Effects of Relative Phases on Pitch and Timbre in the Piano Bass Range

Alexander Galembo
Russian Academy of Sciences

Anders Askenfelt
Royal Institute of Technology

Lola L. Cuddy
Queen's University

Frank A. Russo
Queen's University

digital.library.ryerson.ca/object/124

Please Cite:
Galembo, A., Askenfelt, A., Cuddy, L. L., & Russo, F. A. (2001). Effects of relative phases on pitch and timbre in the piano bass range. The Journal of the Acoustical Society of America, 110(3), 1649–1666. doi:10.1121/1.1391246
Effects of relative phases on pitch and timbre in the piano bass range

Alexander Galembo

Setchenov Institute of Evolutionary Physiology and Biochemistry, Russian Academy of Sciences, Thorez Prospect 44, St. Petersburg 194223, Russia

Anders Askenfelt

Department of Speech, Music and Hearing, Royal Institute of Technology, SE-100 44 Stockholm, Sweden

Lola L. Cuddy and Frank A. Russo

Department of Psychology, Queen’s University, Kingston, Ontario K7L 3N6, Canada

(Received 15 December 1999; accepted for publication 21 June 2001)

Piano bass tones raise questions related to the perception of multicomponent, inharmonic tones. In this study, the influence of the relative phases among partials on pitch and timbre was investigated for synthesized bass tones with piano-like inharmonicity. Three sets of bass tones (A0 = 27.5 Hz, 100 partials, flat spectral envelope) were generated; harmonic, low inharmonic, and high inharmonic. For each set, five starting phase relations among partials were applied; sine phases, alternate (sine/cosine) phases, random phases, Schroeder phases, and negative Schroeder phases. The pitch and timbre of the tones were influenced markedly by the starting phases. Listening tests showed that listeners are able to discriminate between tones having different starting phase relations, and also that the pitch could be changed by manipulating the relative phases (octave, fifth, major third). A piano-like inharmonicity gives a characteristic randomizing effect of the phase relations over time in tones starting with nonrandom phase relations. A measure of the regularity of the phase differences between adjacent partials is suggested for quantifying this randomization process. The observed phase effects might be of importance in synthesizing, recording, and reproducing piano music.

© 2001 Acoustical Society of America. [DOI: 10.1121/1.1391246]

PACS numbers: 43.75.Cd, 43.75.Mn, 43.66.Hg [ADP]

I. INTRODUCTION

Piano bass tones connect instrument design and psychoacoustics in a challenging way. In the bass, the piano design is forced to its limits, and the piano engineer encounters a variety of problems. The produced sound defines a stimulus of extraordinary complexity, of which many aspects have not yet been explored by the psychoacoustician.

A. The piano designer’s problem

Differences in tone quality between pianos of different size, model, and make are due to differences in design, material, and the production process. This truism is best illustrated in the bass range. Above the mid-range (say C4), stringing scales of most pianos are rather similar. The speaking length of the strings is derived iteratively from the length of the C8 string (which always is close to 5 cm), via a scaling factor, which always is close to the 12th root of two [see, e.g., Conklin(1996)]. Design limitations, however, become clearly exposed in the bass range, in particular, in the two lowest octaves A0–A1 covering 27.5–110 Hz [Galembo and Cuddy (1997); Galembo, Askenfelt, and Cuddy (1998)]. Some of the most important difficulties which the piano designer has to tackle in the bass are: (1) the perceptual importance of inharmonicity due to string stiffness increases; (2) the longitudinal string modes become a significant factor of the timbre [Conklin (1996)]; (3) due to physical space limitations it is no longer possible to use plain strings and wrapping is introduced that influences the tone quality; (4) an inevitable discontinuity in the string scale and corresponding change in timbre occurs at the transition from the treble bridge to the bass bridge; and (5) the radiation efficiency of the soundboard decreases rapidly in the lowest bass.

Mainly by trial and error, piano designers have developed a variety of strategies to overcome or compensate for these limitations in order to achieve the best tone quality. Further refinements of the design may still be possible. A deeper understanding of the physical processes involved in tone generation in the bass range, as well as their perceptual relevance, are thus of immediate interest to the piano manufacturer.

B. Psychoacoustic perspective

From a psychoacoustician’s point of view the piano bass range brings several questions to the surface, in particular, how perception is influenced by the temporal evolution in magnitude and phase over the duration of an inharmonic tone.

A piano bass tone is an intriguing stimulus [see Figs. 1(a) and (b)]. Viewed in the frequency domain, it is very rich. The spectra of the lowest bass tones may contain more than 100 partials, and generally extend up to 4–5 kHz within a 60-dB amplitude drop. The magnitude spectrum has a characteristic formant-like envelope with groups of six to eight partials between spectral minima, determined by the striking
position of the hammer on the string. The partials are progressively stretched due to the inherent inharmonicity of the stiff strings [Rayleigh (1894); Fletcher (1964)], as well as a motion of the string terminations [Schuck and Young (1943); Rimski-Korsakov and Diakonov (1952); Exley (1969); Benade (1976)]. The decay rate generally increases toward higher partials [Martin (1947); Hundley, Benioff and Martin (1978)].

In the time domain, the inharmonicity is manifested as aperiodicity, giving a marked evolution in the waveform when viewed over several periods. Among other things, this means that the phase relations among partials (“relative phases”) change continuously. The presence of aperiodicity suggests that temporal processes might be of perceptual importance in the bass range. Psychoacoustical findings support this view; for complex tones in the bass range, no partials are resolved by critical bands [see, e.g., Gelfand (1998)]. As a consequence, temporal pitch coding can be supposed to dominate. Further, as we deal with a high number of audible partials, perceptual effects on pitch and timbre from the relative phases among partials seem quite probable. In particular, it is natural to suppose that audible effects due to the dynamically changing relative phases in piano bass tones may occur.

This is the topic addressed in this study. The perception of synthesized multicomponent bass tones with a piano-like inharmonicity is explored in three experiments. The design of the experiments focused on two main control parameters: (1) the relative phases at note onset (starting phases), and (2) the level of inharmonicity, and its influence on stimulus waveforms, pitch and timbre. The starting phases will control the initial waveform (peak structure), while the inharmonicity will influence the temporal evolution of the phase relations, and hence also the evolution of the waveform.

II. BACKGROUND

A. Phase effects

The classic statement [Helmholtz (1877)] that the auditory system is insensitive to the phase relations among partials is known to be a first-order approximation only. Several studies have demonstrated an influence of relative phases of partials on timbre and pitch perception in experimental situations [see, e.g., Schroeder (1975); Nordmark (1978); Patterson (1987); Moore (1997); Hartmann (1998)]. However, the reported experiments usually involve synthesized complex tones consisting of a few harmonics, restricted to a specific range of fundamental frequencies. In contrast, tones produced by real musical instruments usually contain numerous partials and show a complex dynamic evolution of the spectral envelope. In many cases, such as the piano, musical tones also have a noticeable inharmonicity. In view of these differences, it is not straightforward to apply the psychoacoustical findings on the perception of musical tones. Phase
B. Inharmonicity

Inharmonicity is a common term in psychoacoustics referring to spectra in which one or several partials are shifted away from their harmonic positions. The purpose of most experiments with such controlled inharmonicity has been to gain a better understanding of the perception of harmonic complex tones [see, e.g., Schouten, Ritsma, and Cardozo (1962); Terhardt (1972); Goldstein (1973); Wightman (1973a); Moore, Glasberg, and Peters (1985)]. Moore, Peters, and Glasberg (1985) connected the results of psychoacoustical experiments on inharmonicity with the perception of piano tones.

The main feature of an inharmonic musical tone is that it lacks strict periodicity. This property may affect both pitch and timbre. In particular, the inharmonicity may influence the pitch strength. A harmonic spectrum is known to evoke a more pronounced sensation of pitch than does an inharmonic spectrum [Wightman (1973b)].

The inharmonicity in piano tones is caused by several physical phenomenon, influencing all partial frequencies. A major contribution comes from the stiffness of the piano string, which causes its transverse modes of vibration to be higher than those of an ideal string (“stretched” partials). The classical formula describing the inharmonicity in piano strings is

\[
f_n = n f_1^0 \sqrt{1 + n^2 B},
\]

where \(f_n\) is frequency of the \(n\)th partial, \(n\) is partial number, \(f_1^0\) is fundamental frequency of the ideal, flexible string, and \(B\) is an inharmonicity coefficient, set by the string dimensions, material, and design. Fletcher (1964) provides a good survey of earlier work together with derivations of Eq. (1) for different types of boundary conditions and experimental validations.

In the time domain, the inharmonicity in piano tones is manifested as dispersion, the high-frequency components propagating on the string with a higher velocity than the lower components. This leads to characteristic ripples in the string waveform, preceding each main fundamental pulse, “pitch glides” [Podlesak and Lee (1989)] or “precursors” [Askenfelt and Jansson (1993)]. The precursors are particularly prominent in the bass where the excitation pulse is short compared to the fundamental period. The influence of inharmonicity in piano tones on timbre is not at all fully explored. Perceptual studies of synthesized piano-like tones have suggested that the inharmonicity adds “warmth” or “liveness” to the decay [Fletcher, Blackham, and Stratton (1962)], as well as a “bite” to the attack, enhancing the percussive character [Reinhold, Jansson, and Askenfelt (1987)].

C. Overlooked factors

A serious limitation of earlier works on the perception of piano tones is that two important factors influencing the timbre in combination with the inharmonicity have been overlooked in the experiments. The first factor is the spectral bandwidth of a tone, which covaries with the degree of inharmonicity and influences the brightness (sharpness) markedly [Galembo and Cuddy (1997)]. For example, a threefold increase of the inharmonicity coefficient (a factor that would may be observed between two grand pianos of different size) may increase the spectral bandwidth of a bass tone by as much as 50%.

The second overlooked factor influencing piano timbre as well as pitch is the phase relations among partials. This factor is the focus of the three experiments included in this study. In Experiment I, the listeners’ ability to discriminate between tones with five different starting phase relations under three levels of inharmonicity is investigated. In Experiment II, we test specific pitch effects in harmonic tones (ambiguous or alternative pitches). In Experiment III, temporary timbral effects in inharmonic tones (“squeals”) are examined.
III. STIMULI

A total of 15 stimulus tones were synthesized, comprising five different starting phase relations and three levels of inharmonicity. Different subsets of these tones were used in Experiments I–III.

The tones were computed as

\[
\text{tone}(n,t) = \sum_{n=1}^{N} A \sin\left[2 \pi f_{10}^0 n (1 + B n^2)^{0.5} t + \text{Ph}(n,0)\right].
\]

where \(N\) is the number of partials in the spectrum, \(\text{Ph}(n,0)\) is the starting phase of the \(n\)th partial, and \(B\) is the inharmonicity coefficient (\(B = 0, 60 \times 10^{-6}, \text{or } 600 \times 10^{-6}\)).

All tones had a steady spectrum with a flat envelope, and a common frequency of the unstretched fundamental, \(f_{10}^0 = 27.5\) or \(55.0\) Hz, corresponding to the nominal frequency of the lowest note on the piano (A0), or an octave above (A1). Each tone consisted of 100 partials. This number is representative of a real piano bass tone at a relatively loud dynamic level. The tones were generated by additive synthesis on a Pentium PC at a sampling frequency of 16 kHz.

The inharmonicity was implemented according to Eq. (1), mimicking the partial frequencies of real piano tones. Three levels of inharmonicity were used, defining three sets of tones: “harmonic” (\(B = 0\)), “low inharmonic” (\(B = 60 \times 10^{-6}\)), and “high inharmonic” (\(B = 600 \times 10^{-6}\)). For the low-inharmonic set, the inharmonicity coefficient \(B\) was chosen close to the minimum value in the bass range across many observed pianos, [Galembo and Askenfelt (1999); Galembo and Cuddy (1997)]. For the high-inharmonic set, \(B\) was set close to the maximum observed value in the bass range.

For each level of inharmonicity, five versions of starting phases \(\text{Ph}(n,0)\) of the partials were generated. The resulting waveforms for the harmonic set are shown in Fig. 2.

- **Sine phases:**
  \[
  \text{Ph}(n,0) = 0.
  \]
  Sine (zero) starting phase for all partials.

- **Alternate phases:**
  \[
  \text{Ph}(n,0) = \pi / 4 ( (-1)^n + 1 ).
  \]
  Sine phases for odd partials and cosine for even partials.

- **Random phases:**
  \[
  \text{Ph}(n,0) = \text{random}[0, 2 \pi].
  \]
  Randomized starting phases.

- **Schroeder phases:**
  \[
  \text{Ph}(n,0) = - \pi n^2 / N.
  \]
  Sine phases for all partials, but with a consecutive delay \((n)\) ranging from zero to half the fundamental period.

- **Negative Schroeder phases:**
  \[
  \text{Ph}(n,0) = \pi n^2 / N.
  \]
  Sine phases for all partials, but with a consecutive lead ranging from zero to half the fundamental period.

The Schroeder phases originate from a derivation of the phase relations, which minimize the peak factor of a periodic signal with a flat power spectrum [Schroeder (1970)]. Partial \(n\) of a Schroeder tone is given by

FIG. 2. Computed waveforms of five harmonic A0 tones (\(f_{10}^0 = 27.5\) Hz) consisting of 100 consecutive harmonics of equal and constant amplitude, but with different starting phases.
\[
\text{partial}(n)_{\text{SCH}} = \sin \left[ 2\pi f(n)t - \pi n^2/N \right] = \sin \left[ 2\pi f_0(n)(t - \text{delay}(n)) \right],
\]
where
\[
\text{delay}(n) = n(l(2Nf_0)) = (n/N)T_1^0/2,
\]
and \(T_1^0\) the fundamental period.

The starting phase relations in a Schroeder tone can thus be viewed as sine phases for all partials, but with a consecutive delay ranging from zero to half the fundamental period. Negative Schroeder tones show a corresponding consecutive lead. A (positive) Schroeder tone will resemble a chirped FM signal with monotonically increasing instantaneous frequency, while a negative Schroeder tone will give decreasing frequency within the period.

IV. EXPERIMENT I: DISCRIMINATION BETWEEN TONES WITH FIVE DIFFERENT STARTING PHASE RELATIONS ASSESSED UNDER THREE LEVELS OF INHARMONICITY

A. Overview of the stimulus waveforms

In this section we provide an overview of the stimulus waveforms and some characteristics of pitch and timbre, compiled from the spontaneous first impressions of a few listeners. The discrimination listening test is reported below in Sec. IV B.

1. Harmonic tones

The set of harmonic tones gives a clear demonstration of some features that are also of relevance for the more realistic inharmonic tones. As seen in Fig. 2, the five different relative phase conditions are clearly manifested in the waveforms. The sine-phase condition emphasizes the peak structure of the waveform and gives a chugging (“motorcycle”) timbre with a strong sensation of the fundamental periodicity. The alternate phase results in a quasi-halving of the waveform period, giving an impression of an upward octave shift in pitch compared to sine-phase tones. Random phases provide a waveform with less clear indications of the fundamental period, and a smooth steady timbre. Schroeder phases result in upward pitch glides within each fundamental period, while with negative Schroeder phases the glides are downward. The Schroeder cases will be discussed in detail in connection with the inharmonic examples in the following section.

2. Inharmonic tones

The general effect of inharmonicity on the waveform is illustrated in Fig. 3, which compares the initial three periods of three tones, one from each set (harmonic, low-inharmonic and high-inharmonic), all with sine starting phases.

In the harmonic tone, the starting phases among the partials repeat each period, while in the inharmonic tones the phases will be disordered over time (“dephasing”). As seen
in the figure, the initial pulse in the two inharmonic tones is severely distorted after only a few periods, the energy contained in the initial pulse being smeared out over a substantial fraction of the period. During the first quasiperiods, this effect is manifested as characteristic downward frequency glides prior to the nominal position of the "main pulse." These forerunners, or precursors, occupy about half of the first period of the low-inharmonicity tone, and about three-quarters of the first period of the high-inharmonicity tone. After a few periods, these glides disappear and the waveforms appear more stochastic.

For inharmonic tones it is thus important to distinguish between the starting phase conditions $Ph(n,0)$ and the instantaneous phase $Ph(n,t)$. The dephasing process will successively bring the starting phases out of order, randomizing the phases from period to period. The rate of the dephasing increases with increasing inharmonicity. As inharmonic tones are not periodic, the phase relations will never return to the starting conditions.

The first 200 ms of the waveforms of all ten tones included in the low-inharmonic and high-inharmonic sets are shown in Fig. 4 (the two inharmonic cases in Fig. 3 reap-

---

**FIG. 4.** Computed waveforms of the first 200 ms of ten A0 tones, consisting of 100 consecutive inharmonic partials of equal and constant amplitude. The tones differ in starting phases (rows) and inharmonicity (columns), low-inharmonic tones, $B = 60 \times 10^{-6}$ (left), and high-inharmonic tones, $B = 600 \times 10^{-6}$ (right). The arrows in (g) and (h) indicate moments of maximal in-phase relations in the Schroeder tones (see the text).
parring at the top). A significant difference between random-phase tones [Figs. 4(e) and (f)] and the other tones can be seen. For tones with random starting phases, no characteristic changes in waveform (and timbre) with time can be observed. Due to the inharmonicity, the phases of the other tones (which all start with nonrandom phases) are successively randomized, and after some time it is difficult to attribute the waveform (and timbre) to a specific starting phase.

In the Schroeder case, the inherent upward pitch glides generate an interesting conflict with the downward glides due to inharmonicity. Comparing the waveforms in Fig. 4 it can be observed that the sine starting phases are randomizing gradually from one period to the next. Negative-Schroeder phases are also randomizing monotonically, only faster. In contrast, Schroeder tones start with an upward glide, but the cumulative effect of inharmonicity overrides this tendency with time. Inharmonicity finally overrides, resulting in a downward glide. The turning point between these two conflicting mechanisms corresponds to a distinctive peak in the waveform, separating the initial upward glide and the following downward glide. In Fig. 4(g) this event takes place at about 45 ms after the onset of the tone (indicated by an arrow). With higher inharmonicity, the turning point occurs sooner, after about 4 ms in Fig. 4(h). The presence of a peak in the waveform indicates a “phase concentration” at this moment, meaning that the partials are more in phase at this moment than before or after.

B. Listening test 1

In Listening test 1, we investigated whether listeners were able to discriminate between tones with five different starting phase relations (Sine, Alternate, Random, Schroeder, and Negative Schroeder) under three levels of inharmonicity. Three sets of tones: (1) harmonic (2) low inharmonic, and (3) high inharmonic were independently examined in separate blocks of trials, using an ABX procedure. We expected that tones of the harmonic set would be easiest to discriminate because phase relations of these tones remain constant across duration.

1. Method

(a) Listeners. Eight listeners were recruited from the Queen’s University community, 3 men and 5 women ranging in age from 19 to 36. All listeners were paid for their participation. All had received a minimum of seven years formal training on a musical instrument. All reported normal hearing.

(b) Materials. The stimuli consisted of the three complete sets of tones (harmonic, low inharmonic, and high inharmonic) described in Sec. III, each set including five starting phase relations. All tones had a pitch of A0 \((f_0 = 27.5\) Hz) and duration of 1 s. A pause of 10 ms separated all tones in a trial.

(c) Apparatus. The stimulus presentation and collection of responses were controlled by a PC Pentium computer running dedicated listening test software. Tones were presented using a Sound Blaster acoustical output (32 bit), driving Sennheiser HD580 precision headphones with magnitude response within \(\pm 3\) dB across 16–30 000 Hz. Before the actual collection of data started, the listeners were asked if the loudness was comfortable or needed adjustment, but no one exercised this option. In effect, the loudness level was the same for all listeners. Consideration of the phase response of the headphones is addressed in the Appendix.

(d) Procedure. Listeners heard 80 trials in each block, two presentations each of the 10 possible tone combinations in each of the four possible ABX permutations. The order of trials within blocks was independently randomized for each listener. In any given trial, listeners heard three tones; the third tone was always a repetition of the first or second tone (i.e., ABX procedure). The task was to verbally state whether the third tone (X) matched the first tone (A) or the second tone (B). Within each block, all possible ABX combinations and permutations of the five tones were tested. Listeners were able to hear the trial as often as necessary before submitting a response. A cash prize was awarded to the listener achieving the highest percentage correct; thus listeners were inclined to only submit a response once they were confident. Once a response was submitted, feedback was provided in the form of “correct” or “incorrect.” Prior to commencing each block, listeners were given approximately 15 min to familiarize themselves with the tones used in the block and to practice the discrimination task with those tones. The order of blocks was identical for all listeners: (1) harmonic, (2) low inharmonic, and (3) high-inharmonic. This block order was designed to follow the expected progression from easy to difficult discrimination.

| Table I. Mean discrimination for tone combinations within each tone set in Listening test 1. Note: Mean discrimination was collapsed across all ABX permutations. |
| --- |
| Harmonic \((B = 0)\) | Sine | Random | Schroeder | Negative Schroeder | Alternate |
| Sine | ... | 100% | 100% | 98.43% | 100% |
| Random | ... | 100% | 100% | 100% | 100% |
| Schroeder | ... | 96.88% | 100% | ... | ... |
| Negative Schroeder | ... | 100% | ... | ... | ... |
| Alternate | ... | ... | ... | ... | ... |
| Low inharmonic \((B = 60 \times 10^{-6})\) | Sine | ... | 100% | 93.75% | 90.63% | 98.44% |
| Random | ... | 100% | 100% | 100% | 100% |
| Schroeder | ... | 100% | 100% | ... | ... |
| Negative Schroeder | ... | 100% | ... | ... | ... |
| Alternate | ... | ... | ... | ... | ... |
| High inharmonic \((B = 600 \times 10^{-6})\) | Sine | ... | 100% | 89.06% | 76.56% | 98.44% |
| Random | ... | 100% | 100% | 100% | 100% |
| Schroeder | ... | 93.75% | 100% | ... | ... |
| Negative Schroeder | ... | 100% | ... | ... | ... |
| Alternate | ... | 92.19% | ... | ... | ... |
2. Results and discussion

The overall mean percentage of correct discrimination was well above chance (i.e., 50%) at 97.6%. Variance in mean percentage correct across tone sets was small but significant with harmonic at 99.53% (SD = 0.65), low inharmonic at 98.28% (SD = 1.88), and high-inharmonic at 95% (SD = 5.9), F(2,14) = 2.8, p < 0.05. This variance indicates that ability to discriminate between tones having different phase is related inversely to the level of inharmonicity in those tones. Table I lists for each tone set the mean percentage of correct discrimination for each of the AB tone combinations collapsed across all permutations. The most difficult discriminations were Sine-phase and Negative Schroeder-phase combinations, particularly in the high-inharmonic tone set. However, even for difficult tone combinations the mean discrimination was well above chance. In conclusion, listeners are able to discriminate between tones having different starting phase relations, even when those tones possess high levels of inharmonicity.

V. EXPERIMENT II: PITCH OF HARMONIC ALTERNATE-PHASE TONES

The listeners’ spontaneous reactions to the stimulus waveforms (Sec. IV A) included three kinds of responses when comparing harmonic sine and alternate-phase tones: (1) The pitch of the alternate-phase tone was an octave higher than that of the sine-phase tone; (2) the alternate-phase tone was brighter than the sine-phase tone; (3) the alternate-phase tone had ambiguous pitch, differing by one octave. A dedicated study of the reason for these differing responses was undertaken in Experiment II. The term pitch structure is used as a common term for all possible pitches associated with a particular tone.

A. Pitch and timbre interactions

Response (1) reporting an octave shift is not surprising. Alternate phase relationships are known to generate an octave shift in pitch compared to sine-phase harmonics, when applied to unresolvable higher partials of middle-range fundamentals [Shackleton and Carylon (1994)]. In our bass tones, all partials are unresolvable or close to that, so the octave-shift response could be expected.

The appearance of response (2) reporting differences in brightness as an alternative to the pitch-shift (1) could be explained by the well-known interaction between pitch and timbre [Singh (1987); Crowder (1989); Iverson and Krumhansl (1989, 1993); Hesse (1982)]. In particular, the “brightness” or “sharpness” [Lichte (1941)] dimension of timbre, known to reflect the “position of energized spectral region” [Von Bismark (1974)], or the “spectral center of gravity” [Wessel, Bristow, and Settel (1987)], has been shown to interact with pitch. If other pitch cues are weak, a change in sharpness (a property that can be scaled from low to high as can pitch) might be confused with a change in pitch [Singh and Hirsh (1992); Pierce (1992)]. The pitch-shift response in (1) is thus not categorically different from the brightness-change response in (2). No clear perceptual border between changes in pitch and changes in brightness does exist.

Response (3) about the ambiguous pitch of the alternate-phase tone can be supported by rearranging the expression for the tone in even and odd terms:

FIG. 5. Computed waveforms of the first three periods of three A1 tones (f1 = 55 Hz) consisting of 100 consecutive harmonic partials of equal amplitude. The tones have the same magnitude spectrum, but differ in phase spectrum. In (a) all partials have sine starting phases; in (b) every fifth partial has a phase shift of π, and in (c), every third partial has a phase shift of π/2 and every fifth partial a phase shift of π. The pitch structure is different for the three tones (see the text).
In other words, an alternate-phase tone with 100 harmonics should sound the same as the sum of a sine-phase tone with 100 harmonics and a tone one octave higher with 50 harmonics having \( \sqrt{2} \) times larger amplitude and different phase.

Any change of the phase of some particular partial(s) in a complex tone can be interpreted as an addition (or subtraction) of another complex tone made up of these manipulated partials. This gives a selective attention to the “added” or “subtracted” tone when comparing two consecutive tones with different phase relations. A striking demonstration of this effect was given by Schroeder (1975), who reported that it was possible to play “Mary had a Little Lamb” using sounds consisting of the 31 lowest harmonics of a complex tone with a fundamental of 100 Hz and differing in phase spectrum only.

**B. Alternate-phase tones with additional pitches**

In order to be able to discriminate between “changes in pitch” and “changes in brightness” when comparing sine and alternate-phase tones in listening tests, we decided to try to produce phase effects which would give pitch shifts other than an octave. Preliminary experiments showed that the perceived differences between sine and alternate-phase tones were more audible for tones one octave above A0. The most probable explanation is that A1 gives a stronger perception of the pitch corresponding to the fundamental. For this reason A1 tones \( f_1^{\text{A1}} = 55.0 \text{ Hz} \) were used in Experiment II.

Two A1 tones were synthesized, each consisting of 100 harmonics. In one tone the phases of every third partial were shifted by \( \pi \), and in the other tone the phases of every fifth partial were shifted by the same amount. In this way, a pitch corresponding to the third and fifth harmonic, respectively, was emphasized. The tones were generated as

\[
tone_{\text{ALT}}(n) = \sum_{n=1}^{N/2} \sin(2 \pi n f_0^{\text{A1}} t) + \sin(2 \pi (n+\frac{2}{3}) f_0^{\text{A1}} t + \frac{\pi}{4}) + \sin(2 \pi (n+\frac{5}{6}) f_0^{\text{A1}} t + \frac{\pi}{4}).
\]

(10)

For \( \text{tone3} \) and \( \text{tone5} \), waveforms were obtained in which the period was divided by rather prominent peaks in three and five equal parts, respectively. In a pilot test it was found that when these tones were presented over headphones, a “fundamental pitch” corresponding to A1 was perceived, as expected. In addition, a distinctive pitch at an octave plus fifth, E3 (or just a fifth, E2, due to the common octave ambiguity) was perceived for \( \text{tone3} \). For \( \text{tone5} \), the additional pitch was two octaves plus a major third, C\(^{\#} 4 \), with octave ambiguities. Depending on the listening conditions (headphones, loudspeakers), the C\(^{\#} 4 \) pitch could sometimes be perceived as stronger than the A1 pitch. In Listening test 2 we tested whether a sequence of the phase-manipulated tones (\( \text{tone3} \), \( \text{tone5} \)) gave rise to changes in pitch or in timbre (see Sec. V C).

In further experiments with complexes consisting of 100 equal-amplitude harmonics with a fundamental frequency of 55 Hz (A1), the phases of partials were adjusted so as to emphasize the pitches of the third and fifth harmonics simultaneously. The effect on the waveform is illustrated in Fig. 5. The figure shows the detailed waveforms of (a) an A1 tone with all partials starting with sine phases, (b) an A1 tone with every fifth harmonic phase shifted by \( \pi \), and (c) an A1 tone with every third harmonic phase shifted by \( \pi/2 \) and every fifth harmonic phase shifted by \( \pi \). The amplitude spectrum is the same throughout the sequence. The period of the second tone is divided by peaks into five parts, while the third tone is divided into 15 parts.

In a pilot test, the response of some listeners when listening over headphones was that the sequence consisted of the tone A1 followed by a major third interval (A1–C\(^{\#} 4 \)) spread over one or two octaves, and then by a (weak) major chord (A–C\(^{\#} – E \)). Other listeners reported a major arpeggio in a wide disposition, following the most distinctive pitch changes of the sequence. When listening over loudspeakers in a room with normal reverberation the three tones of the sequence were often perceived as differing in timbre only.

**C. Listening test 2: Pitch effects in harmonic alternate-phase tones**

In Listening test 2, we assessed whether listeners could perceive pitches other than the pitch corresponding to the fundamental frequency in harmonic tones with shifted phases of selected sets of harmonics.

To assess this, three-tone sequences were presented as a short melody. Each tone of a sequence had an identical harmonic spectrum, but different phase relations: (1) all sine phases; (2) all sine phases except every third harmonic phase shifted by \( \pi/2 \); and (3) all sine phases except every fifth harmonic phase shifted by \( \pi/2 \). All three tones could be predicted to possess the pitch A corresponding to the fundamental, supplemented by the additional pitches E and C in (2) and (3), respectively. Listeners were asked to match each three-tone sequence to one of 27 possible three-tone sequence alternatives. Each three-tone sequence was presented...
in each of three reverberation conditions: (1) headphones, (2) speaker near, and (3) speaker far. We expected that pitch effects would be clearest under low-reverberation presentation conditions (i.e., headphones and speaker near).

1. Method

(a) Listeners. Ten listeners, five men and five women, were recruited from the Kingston and Queen’s University communities. Listeners ranged in age from 19 to 51. Five listeners were paid for participation; of these, three were recruited from responses to posters on campus and two were known to the experimenters. Three listeners were recruited from a first-year psychology subject pool and received course credit for their participation. The remaining two listeners were members of the Acoustics Laboratory and did not receive any remuneration. All had received training on a musical instrument to a minimum performance level of Grade VI Royal Conservatory (or equivalent). All reported normal hearing.

(b) Materials. The stimuli were sequences of three synthesized tones, each tone with different phase relations: (1) all sine phases, (2) all sine phases except every third harmonic phase shifted by \( \pi/2 \), and (3) all sine phases except every fifth harmonic phase shifted by \( \pi/2 \). All tones were harmonic with a pitch of A1 (\( f_1 = 55 \) Hz), flat, steady magnitude spectrum, and a duration of 0.5 s.

(c) Apparatus. The tones were presented over headphones and loudspeakers. Headphone presentation was identical to that described in Experiment I (Sec. IV B 1). Speaker presentation was realized using a small computer loudspeaker (Creative CS100), specified to be within \( \pm 3 \) dB across 20–20 000 Hz. The loudspeaker was used in two conditions; (1) Speaker near: The listeners were asked to hold the loudspeaker within 2–3 cm from their right ear; (2) Speaker far: The loudspeaker was positioned at a distance of about 90 cm from the listener with a reverse orientation so that the listener was not exposed to the direct sound. The listening tests were conducted in a small office room (2 × 3 m) with concrete walls. Consideration of the phase response of the loudspeaker presentation is addressed in the Appendix.

(d) Procedure. Each listener heard each of six tone sequences in each of three presentation conditions. The sequences were the six possible permutations of the three tones presented with a 10-ms pause between each tone. The order of presentation conditions was identical for all listeners: (1) headphones, (2) speaker near, and (3) speaker far. In each presentation condition, each of the six tone sequences was presented three times, for a total of 18 trials. The order in which trials were presented was randomized independently for each presentation condition and each listener.

Listeners were told that the three tones could be heard either as a sequence of single pitches, or as a sequence containing some, or all, of the pitches A, C\(^4\), and E. Thus, there were 27 possible response alternatives. A sample of the possible response alternatives was demonstrated on the piano keyboard.

Listeners were asked to enter the letter names of the pitches that they heard, in order, on the computer keyboard (e.g., AAA, AC\(^4\)E, C\(^4\)EA, etc). Listeners were instructed to ignore tone height and to attempt to focus on the similarities or differences among the three sounds. After completing all 18 trials, they were asked to listen to each trial again and to verify their judgements. Listeners were allowed to use the piano keyboard for assistance. They were also permitted to replay each trial as many times as they liked.

2. Results and discussion

A response was considered “expected” if it conformed to the theoretical predictions. A strict criterion was used; each pitch had to be labeled in the predicted order. For example, when the prediction was that the pitches of the three sounds was A E C\(^4\), only that order and no other permutation of the pitches, was considered “expected.”

The majority of responses (72.4%) contained three different pitches. The mean number of expected responses across listeners, out of a possible maximum of 18, was 9.33 for the headphone condition, 8.67 for the speaker-near condition, and 3.00 for the speaker-far condition. Analysis of variance revealed a significant effect of mode of presentation, \( F(2,18) = 10.35, \ p < 0.001. \) Orthogonal contrasts revealed that the number of expected responses in the headphone condition was not significantly different from the speaker-near condition, \( F(1,9) = 0.98, \ p = 0.35, \) but that the combined mean for the headphone and speaker-near conditions surpassed that for the speaker-far condition, \( F(1,9) = 12.54, \ p < 0.001. \)

In conclusion, the listening test showed that the perception of alternative pitches is a relatively difficult task, in particular in the reverberant speaker-far condition. Music training may have played a role in fostering responses that conformed to our expectations. A “high-conformance” group with four listeners (min 17, max 17.5 expected responses) contrasted distinctively against a “low-conformance” group with two listeners (min 1.5, max 3 expected responses). The high-conformance group included listeners with recent or current music training, and activities such as conducting and composing.

VI. EXPERIMENT III. PHASE CONCENTRATIONS (“SQUEALS”)

In general, a correspondence between the peak structure of the waveform and timbre could be observed, as discussed above (see Sec. IV A). Sometimes, however, some significant exceptions occurred. In particular, when listening over headphones to tones with high inharmonicity, two distinctive acoustical “events” or “inclusions” of short duration were distinguished within the one-second duration of the tone (except for the random-phase tones). These events, which had the character of “squeals,” were perceptually separated from the tone and are not visible in the waveforms. The squeals occurred at about 0.4 and 0.8 s from the onset of the tones. After such a long time, the phases of partials are normally randomized so much that the tone would sound as a random-phase tone with steady timbre. The most plausible explanation for the squeals would be that they are generated by a
group of partials, temporarily fused by their phase relations ("phase concentrations"). An analysis that evaluates this hypothesis is given in Sec. VII B. First, listeners' ability to detect squeals was tested.

A. Listening test 3: Timbral effects of phase concentrations (squeals)

In Listening test 3, we assessed the consequences of random and nonrandom starting phase relations on the perception of inharmonic tones. Our prediction was that listeners would only perceive squeals in tones possessing nonrandom starting phase relations (e.g., sine, alternate, Schroeder, or negative Schroeder). As in Listening test 2, tones were presented in each of three reverberation conditions: (1) headphones, (2) speaker near, and (3) speaker far. We anticipated that the detection of squeals would be clearest in low-reverberation presentation conditions, where phase is relatively well preserved.

1. Method

(a) Listeners. Nine listeners, four men and five women, were recruited from the Kingston and Queen's University communities. Six of the nine listeners had participated in Listening test 2. Listeners ranged in age from 19 to 51. Six listeners were paid for participation; of these, five were recruited from responses to posters on campus and one was known to the experimenters. Two listeners were recruited from a first-year psychology subject pool and received course credit for their participation. The remaining listener was a member of the Acoustics Laboratory and did not receive any remuneration. All had received training on a musical instrument to a minimum performance level of Grade VI Royal Conservatory (or equivalent). All reported normal hearing.

(b) Materials. The stimuli consisted of the five tones of the high-inharmonicity set (see Sec. III), including four tones with nonrandom starting phase relations (i.e., sine, alternate, Schroeder, or negative Schroeder) and one tone with random starting phase relations (random). All tones had $B = 600 \times 10^{-6}$, flat, steady magnitude spectrum, pitch of A0 ($f_{0} = 27.5$ Hz) and a duration of 1 s. The four tones with nonrandom phase relations each occurred twice and the tone with random phase relations occurred 8 times, for a total of 16 trials.

(c) Apparatus. The tones were presented over headphones and loudspeakers, as described in Experiment II (Sec. V C 1).

(d) Procedure. The order of trials within each presentation condition was independently randomized for each listener. The order of presentation conditions was identical for all listeners: (1) headphones, (2) speaker near, and (3) speaker far. Listeners were instructed that each trial contained one of two types of sound. They were asked to classify the sound as to whether it contained squeals (i.e., if they heard "events" in the timbre), or not (i.e., they heard a "smooth" timbre). They recorded judgements on the computer keyboard, "1" for "squeals" and "0" for "no squeals." Listeners were allowed to replay each trial as many times as they wished before recording a judgement.

2. Results and discussion

For purposes of scoring, a judgment was considered "expected" if it corresponded to predictions. Given the prediction that only nonrandom starting phase relations would

![Phase images of the starting phases (t=0) of all 100 partials for (a) alternate, (b) random, (c) Schroeder, and (d) negative Schroeder tones. The partial number is shown along the perimeter of the circle, and the phase angle along the radius (0 at the center and $2\pi$ at the perimeter). Grid line circles correspond to steps of $\pi/2$.](image)
contain squeals, an assignment of “1” on trials with nonrandom phases and an assignment of “0” on trials with random phases was scored, as expected. The mean number of expected responses across listeners, out of a possible maximum of 16, was 15.89 for the headphone condition, 15.67 for the speaker-near condition, and 11.44 for the speaker-far condition. Analysis of variance revealed a significant effect of mode of presentation, $F(2,16) = 21.28, p < 0.001$. Orthogonal contrasts revealed that the number of expected responses in the headphone condition was not significantly different from the speaker-near condition, $F(1,8) = 0.64, p = 0.457$, but that the combined mean for the headphone and speaker-near conditions surpassed that for the speaker-far condition, $F(1,8) = 22.72, p < 0.001$.

In conclusion, the listening test showed that listeners were able to detect phase concentrations corresponding to squeals in the tones with nonrandom starting phases with high accuracy. This was true even in the reverberant speaker-far condition, where performance was better than chance, $X^2(8, n = 9) = 21.63, p < 0.01$.

**VII. INDICATORS OF PHASE REGULARITY**

Inspired by the results in Listening test 3, which indicated that temporary phase concentrations can be perceived by a majority of listeners, different ways of illustrating and quantifying the instantaneous phase relations were developed.

**A. Instantaneous phase images**

A visual indicator of the regularity of the relative phases among partials, coined *instantaneous phase image*, was found to be useful when discussing the regularity of phase relations and the evolution of the phase relations over time. Such an image represents the phases of the partials at a particular moment of time as a contour in a radar diagram. The phase angle is shown as the distance from the center $\sim n\pi$, not as an angle $\theta$ with zero at the center and $2\pi$ at the perimeter. The partials (1–100) are arranged clockwise around the circle.

In Fig. 6 four phase images representing the starting phases for (a) alternate, (b) random, (c) Schroeder, and (d) negative Schroeder tones are shown. The sine-phase tone is not included as the phases are all concentrated at the center of the circle in that case. It is easy to distinguish between the three cases with regular patterns of the starting phases (periodicity, symmetry), and the irregularity of the random-phase case. Using the phase image representation, the influence of
inharmonicity (or other dephasing factors) on the phase evolution can easily be traced over time.

In Fig. 7, phase images for the five tones with low inharmonicity in Fig. 4 (left column) are shown at the moment of the peak in the Schroeder tone \( t = 43.8 \text{ ms} \) after the onset, Fig. 4(g)]. Unlike the other tones, the phase image of the Schroeder tone shows a clear regularity [see Fig. 7(d)]. The majority of the partials, including large groups of adjacent partials, have phases below \( \pi \), thus forming a positive peak in the waveform. Also, within these groups there is very little change in phase between adjacent partials. This “in-phase” property between adjacent partials becomes even more evident when examining images of the phase differences between adjacent partials, wrapped to the 0–2\( \pi \) range [see Figs. 8(a)–(e)]. For the Schroeder tone in (d), almost all phase differences between adjacent partials are close to zero (2\( \pi \)) at the moment of the peak in the waveform. The development of the Schroeder tone after that point is consequently close to that of the sine-phase tone.

**B. Quantitative approach**

Listening tests 1 and 3 suggested that the pitch and timbre of the stimuli were consistent with the waveform features, in particular, the presence or absence of a clear peak structure, which in turn are dependent on the phase relations among the partials. The perceptual effects related to this observation can be summarized as follows.

1. If the starting phases among partials are random, there is no evident peak structure in the waveform and the timbre of the tone is only dependent upon the degree of inharmonicity.

2. If the starting phases are not random, the timbre of the tone changes with time because of the dephasing driven by inharmonicity. After some time when the phases have been randomized enough, the initial peak structure of the waveform has disappeared and the tone has reached the timbre of a random-phase tone with the same inharmonicity.

Some instantaneous measure of the regularity of the phases might thus serve as an indicator of the timbre. There is no single mathematical definition of the regularity or randomness of an arbitrary finite series of numbers, but various approaches have been suggested [for a survey, see Pincus (1991)]. Generally, a series of values within a certain interval is considered more random than another series within the same interval, if (1) the values are more uniformly distributed over the range, and/or (2) the order of values is less
predictable. As a reasonable working hypothesis it was assumed that criterion (1)—the uniformity of the distribution of the phase differences between adjacent partials—alone could serve as a first-order estimate of the peak structure in the waveform, and hence also as an indicator of timbre. Some support for this suggestion might be found in the Schroeder tones, which show a low peak factor in spite of highly ordered and predictable phase relations.

An additional assumption was made: The phase relations between adjacent partials were considered more important for pitch structure and timbre than the relations between other combinations of partials. As before, the term "pitch structure" is used as a common term for all possible pitches associated with a particular tone. The justification for this assumption is that the audible interaction between partials such as difference tones, beats, etc. are known to be stronger for close partials.

In a spectrum with \(N\) partials there will be \(N^2 - 1\) phase differences \(Ph(n,t)\) between adjacent partials with instantaneous phase \(Ph(n,t)\):

\[
Pd(n,t) = Ph(n,t) - Ph(n-1,t),
\]

where \(n = 1,2,3,...,N\) and

\[
Ph(1,t) = Ph(1,t) - 0.
\]

The phase difference range \([0, 2\pi]\) is divided into \(M\) equal subranges and the number of phase differences in each subrange is calculated, creating a sequence of occurrences \(o_i\):

\[
\{O(t)\} = o_1(t), o_2(t)\cdots o_M(t),
\]

where

\[
\sum_{i=1}^{M} o_i(t) = N \quad \text{for any } t.
\]

It is further assumed that \(M = N\). This means that an ideally uniform distribution would give one occurrence in each subrange. A calculation of how uniform the distribution of phase differences is across \([0, 2\pi]\) is made, using the redundancy concept [Attnave (1959)]. The redundancy of the phase differences (RPD factor) in percent is calculated as

\[
RPD(t) = 100\left(1 - \frac{H(t)}{H_{\text{max}}}\right),
\]

where \(H(t)\) is the entropy,

\[
H(t) = -\sum_{i=1}^{N} \frac{O_i(t)}{N} \log \left(\frac{O_i(t)}{N}\right),
\]

and

\[
\sum_{i=1}^{M} O_i(t) = N \quad \text{for any } t.
\]
\[ H_{\text{max}} = \log N; \]

\( RPD = 0 \) corresponds to a case where the phase differences are distributed uniformly, and \( RPD = 100\% \) to a case where all phase difference values are concentrated in the same subrange (for example, as for the sine-phase tone).

In Fig. 9 the waveforms and the corresponding RPD trajectories for the two Schroeder tones in Figs. 4(g) and (h) are compared. It is evident that the RPD function gives a good description of the waveform envelope of both tones, the peaks of the waveform being perfectly matched in time with the phase concentrations indicated in the RPD curve. The RPD curves for random-phase tones included in Fig. 9(c) do not differ significantly between the two inharmonicity cases; neither do they change with time. A detailed analysis of all synthesized tones showed that the RPD curves mapped the peak structure of the waveforms accurately.

The RPD trajectory allows a rough estimation of the time elapsed before the timbre has reached its quasisteady state. According to the assumptions above, this should occur when the RPD curve had dropped to the level of the random-phase tone. This would occur after about 150 ms for the tone with low inharmonicity and much earlier for the tone with high inharmonicity [see Fig. 9(c)].

Figure 10 shows a fragment of a negative Schroeder-phase tone containing two short squeals. The waveform does not show any significant changes at the places corresponding to the squeals, but the RPD trajectory gives clear indications of the events. Our assumption that temporary phase concentrations were the reason for the squeals was thus verified.

**VIII. DISCUSSION AND CONCLUSIONS**

**A. Survey of the results**

In this study, we have explored some effects of phase on pitch and timbre that may be of relevance for musical tones in the bass range. The evolution of the waveform and phase relations in synthesized bass tones consisting of 100 partials was analyzed (A0 = 27.5 Hz and A1 = 55 Hz), including harmonic tones and inharmonic tones with string-like inharmonicity. Listening tests showed a clear effect of the relative phases of partials on both pitch structure and timbre.

For harmonic tones with flat steady magnitude spectra, it was shown that it was possible to transform a chugging motor cycle-like sounding sequence of impulses (sine phases) into a smooth timbre with weaker pitch strength (random phases), or into a sequence of periodic downward or upward pitch glides (Schroeder phases), only by changing the phase relations. Further, by changing the phase of selected groups of partials, it was possible to change the brightness of the tone. Only random starting phases of the partials provide steady timbre. If the starting phases are not random, the inharmonicity disorders them (“dephasing”), resulting in randomized phase conditions after some time. An ongoing process of randomization is audible only in the attack, as downward pitch glides, differently depending on the starting phase relations and the degree of inharmonicity. Higher inharmonicity gives a faster dephasing process.

A measure of the redundancy of the phase differences between adjacent partials (RPD) was calculated in order to estimate the phase regularity. High peaks in the RPD curve were shown to be consistent with the peak structure of the waveform. RPD peaks of small amplitude could be used to detect temporary phase concentrations, invisible in the waveform but audible as sudden changes in timbre (“squeals”).

The perceptual effect of different phase relations depends on listening conditions. When the tones were presented over headphones, very clear responses were reported for a majority of the listeners. When presented over loudspeakers in a normal room, the effect decreased with increasing distance to the loudspeakers. In a large and moderately reverberant room most listeners do not perceive any differences between the tones, not even in the simplest task with discrimination between harmonic tones with five different phase conditions. The reason for the large influence of the presentation conditions is that the phase relations are well preserved in the headphone presentation, but more or less deteriorated when listening over loudspeakers due to the interfering reflections in the room. This raises doubts about the significance of phase effects when listening in concert halls.
or in other relatively reverberant conditions. Some specific cases in which an audible effect of the phase relations can be expected are discussed below in connection with the piano.

Regarding a possible effect of audible phase distortion introduced by the equipment in the listening test (primarily by the headphones and loudspeakers), it is unlikely that such distortions would provide cues that were both reliable and distinctive enough to enable perceptual discrimination. According to the evaluation in the Appendix there were large differences in the amount of phase deterioration for three different presentation conditions (headphones, speaker near, speaker far). For the headphone presentation it seems safe to conclude that any phase distortion introduced by the test equipment was by far outweighed by the phase differences in the computed stimuli. For the loudspeaker presentations (speaker near, speaker far) the room reverberation introduced a progressive deterioration in the stimuli with increasing listening distance, which is in concordance with the results of the listening tests.

B. Future studies and applications to the piano

In this study, only A0 and A1 tones were studied in combination with five well-defined phase relations. We consider our conclusions valid for at least the bass range below A1. The prominence of the perceived effects are dependent on the conditions of the presentation (headphones versus loudspeakers, room reverberation), and in some cases, musical training. Despite these limitations the experiments have pointed out several important perceptual phenomena, which need further research. A continuation focusing on perception should include the determination of a parameter space in which relative phases have audible effects on timbre and pitch structure. The main parameters would be fundamental frequency range, number of partials, degree of inharmonicity, and spectral envelope and its dynamic evolution. In this connection it would be appropriate to evaluate how useful the crest factor of a waveform is as a first-order approximation of the in-phase/out-of phase relations among partials [Schroeder (1970)].

Regarding applications to real piano tones, the dephasing processes at successive stages in the tone generation are of particular interest. There are several potential sources of phase randomization after the initial string excitation by the hammer. The first is inharmonicity due to string stiffness, as mentioned. The larger the inharmonicity coefficient and the higher the pitch, the faster the phases will be randomized. A second source of phase randomization is the soundboard, a distributed radiator. An additional phase randomization might be expected in the internal volume of the case of the upright piano. A third phase randomizer is the room, due to multiple reflections. The listener is exposed to all three randomizing processes, while the pianist (tuner, voicer), being closer to the soundboard than the soundboard dimensions, is in a better position to experience a less dephased sound.3

There are few data available on the initial phase relations between the vibration modes of a struck piano string and the relative phases of the partials in the radiated tone. An indication that the string motion signal starts with nearly ordered partials, while the radiated piano tone starts with phases already disordered to some extent, can be obtained from the waveforms in Fig. 11 [Podlesak and Lee (1989)]. Audible pitch glides in the signal reported by Podlesak and Lee, as well as a characteristic evolution in waveform within the period comparable to the inharmonic tones in Fig. 3, illustrate the dephasing effect from the string inharmonicity in a real piano.

A tempting question for further research is the extent to which a phase-randomization process (“from order to disorder”) contributes to the character of real piano bass tones. First, it would be desirable to study phase effects in more realistic synthesized tones with piano-like spectral dynamics. Also, it would be desirable to conduct measurements on various pianos to determine the stages of phase transformations from string vibrations to radiated tones. Preliminary experiments indicate that some of the phase information in signals with a pronounced peak structure (like the sine-phase tone) can survive surprisingly well in reverberant conditions when radiated from loudspeakers. Thus, room reverberation alone may be insufficient to sweep out any ordered phase information in radiated piano tones.

The question then remains how large the dephasing effect is due to the radiation from the soundboard. An indication that the soundboard probably plays a major role in the dephasing process comes from experiences of preparing a
grand piano with a contact microphone for amplification or recording. The position of the pick-up relative to the bridge is then highly important. As the phase randomization due to the soundboard is missing, pitch glides indicating the dephasing due to string inharmonicity can be recognized in bass notes for certain positions of the contact microphone, in particular, when placed close to the bass bridge. A similar phenomenon may be of importance when listening to normally recorded piano music. If the microphone has been placed at a small distance from the soundboard, the signal may contain ordered phase relations among partials. This may explain in part the experience that recorded piano music may sound quite different when listening over loudspeakers (even of high quality) compared to headphones. A headphone presentation offers much better phase characteristics [Blaeurt and Laws (1978)] and avoids the reverberation of the listening room.

Whatever may be the relations between the contributions from the different parts of the piano to the phase randomization, the very attack portion of the tone would be the most significant part for communicating the “ordered” phase information. The listener needs information about the initial, more or less ordered, phases in order to be able to follow the continuous evolution toward a randomized state. The importance of the attack is further underlined by (1) the known fact that if focuses the attention on the new tone [Moore (1997)], and (2) the presence of a direct sound component not contaminated by multiple room reflections.

Regarding the presence of “squeals” that were attributed to temporary phase concentrations, it might be appropriate to revisit the “wolf tones” in the piano bass range observed by Knoblaugh (1929, 1944), and attributed by him to the longitudinal string modes. In view of later experimental evidence that the longitudinal modes decay relatively fast, phase concentrations deserve to be investigated as an alternative explanation of the phenomenon.

In summary, our study indicates that the randomization of the relative phases among partials due to inharmonicity is an important factor for the evolution of the timbre in bass tones with wide spectrums. This effect probably needs to be taken into account in a detailed description of the sound generation in the piano bass range. An immediate application could be foreseen in high-quality piano synthesis.

ACKNOWLEDGMENTS

This project was supported by a NATO Postdoctoral Science Fellowship awarded to the first author and a grant from the Natural Sciences and Engineering Research Council of Canada to the third author. We are grateful to Sonya Dal Cin and Jung-Kyong Kim, Queen’s University, for running participants and data preparation.

APPENDIX: PHASE RESPONSE

The amplitude and phase responses of the headphones and loudspeakers used in the listening tests will influence the sound presented to the listeners. While the amplitude responses are specified by the manufacturers, data on the phase responses are not readily available. A determination of the phase responses is possible by measuring the group delays [see, e.g., Blauert and Laws (1978)]. For headphones this needs some kind of (standardized) coupler that simulate the enclosed air volume in the ear. The results of such tests indicate that the differences in group delay characteristics (group delay versus frequency) among different manufacturers and models are relatively large [Blaeurt and Laws (1978); Hirara and Ueda (1990); Hirara (1997)]. Listening tests have shown, however, that in most cases the phase distortion will be below perceptibility [Blaeurt and Laws (1978)].

A determination of the group delays does not allow an easy interpretation of how severely the computed stimuli are distorted by imperfections in the phase responses. For this reason another approach was taken. The resulting sound pressure waveform for one of the stimuli tones (harmonic Schroeder) was recorded with a high-quality electret miniaturized microphone (diam. 8 mm, height 6 mm) at the position of the listener’s ear in each of the three presentation conditions in the listening tests; (inside) the headphones, speaker near and speaker far (see Fig. A1). In this way the combined effect of amplitude and phase distortions could be examined, as well as the influence of the room. An additional case was included, showing the sound pressure at a distance of about 4 m from a high-quality hi-fi loudspeaker (Audio Pro A4–14 Mk II) in a moderately reverberant lecture room with a reverberation radius of 1.1 m (reverberation time $T_{60}=0.76$ s, room volume $V=300\,\text{m}^3$).

According to the waveforms in Figs. A1(a)–(d), it is evident that there is progressive deterioration in the presented sounds when going from presentation over headphones (b) to speaker near (c) and speaker far (d), as compared to the computed original signal in (a). In the headphone presentation the sound pressure is essentially a replica of the computed signal and all characteristic features of the stimulus are present, including the characteristic upward frequency sweep. In the speaker-near condition the sweep is still relatively well preserved, although modified by large-amplitude fluctuations. In the speaker-far condition the sweep is barely discernible, being severely deformed by high-frequency reverberation ripples. In the lecture room presentation the stimulus is still more deteriorated. Even at a moderate distance of 4 m from the hi-fi loudspeaker, the sound has no resemblance with the original signal [see Fig. A1(e)], the reason being that the reverberant sound dominates.

1SOUNDSWELL™ Signal Workstation, Nyvalla DSP AB, Åldermansvägen 19-21, SE-17148 Solna, Sweden.
2SPRUCE™ Listening Test 2.0, Nyvalla DSP AB, Älvdalsvägen 19–21, SE-17148 Solna, Sweden.
3The phase spectrum was probably an overlooked covariant factor in the Fletcher et al. (1962) experiments on the timbral importance of inharmonicity in pianos [see Galembo and Cuddy (1997)].
4According to Podlesak and Lee (1987), the decay rate of the longitudinal string component in the lowest bass notes is about 100 dB/s.

Askenfelt, A., and Jansson, E. V. (1993). “From touch to string vibrations. III: String motion and spectra,” J. Acoust. Soc. Am. 93, 2181–2196.
Attneave, F. (1959). Applications of Information Theory to Psychology: A Summary of Basic Concepts, Methods and Results (Holt-Dryden, New York).

J. Acoust. Soc. Am., Vol. 110, No. 3, Pt. 1, Sep. 2001 Galembo et al.: Effects of relative phases on pitch 1665
Galembo et al.: Effects of relative phases on pitch