

Periodicity and pitch perception

John R. Pierce

Center for Computer Research in Music and Acoustics, Department of Music, Stanford University, Stanford, California 94305

(Received 11 December 1990; revised 20 April 1991; accepted 4 June 1991)

There has been experimental evidence pointing to at least two pitch mechanisms, the first involving low-order harmonics that are resolved along the basilar membrane, and the second a periodicity mechanism that depends only on the repetition rate of the time waveform on the basilar membrane. If this time waveform is derived from repeated bursts of sinusoidal tone, the second mechanism might be the sole pitch mechanism. It is found that this can be so up to rates as high as 250 bursts of 4978-Hz tone per second. The stimuli used are periodic patterns of equally spaced tone bursts, with either successive tone bursts in the same phase, or every fourth tone burst 180° out of phase with respect to the rest. Up to a critical transitional rate of tone bursts a second, the two sequences sound exactly the same, despite their different fundamental frequencies and frequency separation of harmonics. Critical rate data are given for sinusoidal bursts of seven different frequencies. Critical rates appear to be closely related to the critical bandwidth. Pitch matching appears to be consistent with these observations; it is on rate below the critical rate and can be on fundamental frequency above the critical rate.

PACS numbers: 43.66.Mk, 43.66.Hg [WAY]

INTRODUCTION

Investigations of just noticeable differences (jnd's) of pitch continue to indicate the plausibility of two "pitch mechanisms," the first operating on resolved harmonics, and the second "periodicity pitch" mechanism on unresolved clusters of harmonics (Houtsma and Smurzynski, 1990), as discussed earlier by de Boer (1976).

The second mechanism is exemplified by the jnd's (or DLs) obtained by Nordmark (1968) for periodic short pulses, beginning at rates so low that there is no pitch, but extending into the range of musical pitch. His jnd's as fractions of the period (and hence, of the pulse rate) are about 1% at 1 pulse per second, and about 0.07% at 4000 pulses a second. The shape of the curve of jnd (DL) versus pulse rate suggests a transition between two mechanisms between 62.5 and 500 pulses per second.

Such a transition is supported by experiments on matches between periodic all-positive pulses and periodic patterns of positive and negative pulses, carried out by Flanagan and Guttman (1960), Guttman *et al.* (1964), and Rosenberg (1966). At low frequencies the match is on pulse rate; at higher frequencies the match is on fundamental frequency.

Can the mechanism depending on periodicity be completely disentangled from the other mechanism and investigated separately in the range of musical pitch?

Guttman *et al.* (1964) took a step in this direction by high-pass filtering the pulse trains, thus removing components easily resolved by the ear. Rosenberg (1965) used low-pass noise to mask the fundamental and lower harmonics. These approaches only partially resolved the mechanisms. Moreover, the observations were pitch matches, which could reveal dominance of one mechanism but not the absence of the other.

In demonstrating that pitch can be conveyed by higher harmonics, Houtsma and Smurzynski (1990) used sequences of equal-amplitude higher harmonics. In order to avoid gross effects of place, they used different sets of an equal number of successive harmonics. This is helpful, but their observations could not assure against a mixture of mechanisms.

In endeavoring to measure the frequency resolution of the basilar membrane, Gold and Pumphrey (1948) used a quite different approach. Their stimuli were made up of short bursts of a sinusoidal tone n cycles long separated by m cycles of silence. In one stimulus the sinusoidal tones repeated regularly in the same initial phase, and in the other alternate 180° in phase (that is, the sign of the tone burst changed at each repetition).

Gold and Pumphrey did not make matches of pitch or rate. Rather, for a given frequency of the sinusoid in the burst of tone, they asked, how long must the m cycles of silence be if one stimulus is to be indistinguishable from the other? This gives a critical rate of tone bursts per second below which the ear can have no clue that Fourier analyses of the two stimuli would reveal line spectra with different fundamental frequencies and different harmonic spacings.

Their argument is that if, at the place corresponding to the frequency of the sinusoid in the tone burst, the vibration of the basilar membrane due to one tone burst had died out before the next tone burst came along, the phase of the subsequent tone burst could have no effect on the quality of the sound. However, if the vibration due to one tone burst persisted past the time the next came along, the pattern of vibration along the basilar membrane induced by the two sequences would be different, giving a different sound quality. In essence, something of a line spectra would be sensed because two or more tone bursts would "fall into the analysis window."

Gold and Pumphrey give only fragmentary data concerning experimental results.

The work of Gold and Pumphrey, and some of their conclusions, have been questioned by others (Hiesey and Schubert, 1972; Green *et al.*, 1975). The questions raised are concerned with why or how the two stimuli sound different at sufficiently low values of m , rather than with the fact that they do *not* sound different at high enough values of m (low enough tone burst rates). There seems to be no clear, plausible alternative explanation for the observation that the difference does disappear for large enough values of m .

If there is no audible difference between sequences of tone bursts of the same sign and sequences with alternating signs, the ear is unable to hear any effect of fundamental frequency or of harmonic spacing. Any pitch mechanism functioning must be one that does not depend on the resolution of harmonics by the ear. It must be a purely periodicity pitch mechanism.

The experiments to be described here indicate that the tone burst rate up to which an uncontaminated periodicity pitch mechanism can be maintained increases with the frequency of the sinusoid that constitutes the tone burst, the carrier frequency, here called the burst frequency. At 4978 Hz (E8) the critical tone burst rate is around 250 Hz, corresponding to B3.

I. STIMULI AND PRESENTATIONS

The stimuli used were like those of Duifhuis (1973), short burst of sinusoidal tone with a raised cosine envelope.

Such tone bursts are characterized by Bf , the burst or carrier frequency (the frequency of the sinusoidal wave that is modulated by the raised cosine envelope), and by N , the integer number of cycles in each tone burst.

The width of the spectrum decreases with the number of cycles per tone burst. Figure 1 shows the computed spectra for two, four, and eight cycles per tone burst. For two cycles per tone burst, the spectrum spans about two octaves (at 10 dB down). A spectrum so wide would tend to smooth or obscure the variation of the critical rate with frequency. For eight cycles per tone burst, the spectral width is narrower, about a half an octave, but the duration of the tone burst is too long to allow rapid tone burst rates. As a compromise, a length of four cycles was chosen, which gives a bandwidth of about an octave between 10-dB points.

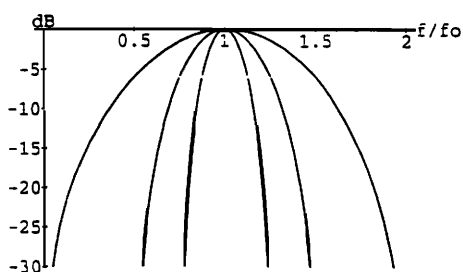


FIG. 1. Spectra of tone bursts with number N of cycles per tone burst equal to two, four, and eight cycles, in dB as a function of ratio of frequency f to center frequency f_0 . Four-cycle tone bursts were used in the experiments.

The tone burst rate R is the number of tone bursts per second. The time T between the beginning of one tone burst and the beginning of the next is the reciprocal of R ; $T = 1/R$.

Different stimuli with different patterns of tone bursts are used. As shown in Fig. 2, in stimulus a each tone burst is identical. In stimulus b every fourth tone burst differs 180° in phase from the others. This is equivalent to saying that the modulating function has been made negative in every fourth tone burst.

For pattern a the fundamental frequency of the tone burst sequence is the tone burst rate R .

For pattern b the fundamental frequency of the tone burst sequence is $R/4$. There are four tone bursts per period.

In the actual experiment, a stimulus c was also presented. The stimulus c is like a but has twice the amplitude. It is presented at a quarter the tone burst rate of a or b . Stimulus c has the same power spectrum as stimulus b . At high enough rates, c sounds like b (Pierce, 1990).

In the experiments described here neither the number of cycles per tone burst N nor the burst frequency (carrier frequency) Bf changed as rate R was changed. The experiments were performed using the NeXT computer, with software written by David Jaffe using the NeXT Music Kit.

In the experiments, N and Bf were specified. The program produced repeatedly, at an adjustable tone burst rate, 1 s of a , $\frac{1}{2}$ s of silence, 1 s of b , $\frac{1}{2}$ s of silence, 1 s of c (at one-fourth the tone burst rate of a and b), $\frac{1}{2}$ s of silence, back to 1 s of a , and so on. If a match was based on rate, a and b would sound the same. If a match was based on fundamental frequency, b and c would sound the same (Pierce, 1990). Sequence c had been included for historical reasons. It played no part in the present experiments and was disregarded by the subjects.

In the experiments, the tone bursts were heard diotically over headphones. For each tone burst frequency the signal level was adjusted to a level that could barely be heard (sensation level) at a rate R of ten tone bursts per second. Then a preselected value of dB was removed from the listening path. Thereafter energy per tone burst was maintained constant as

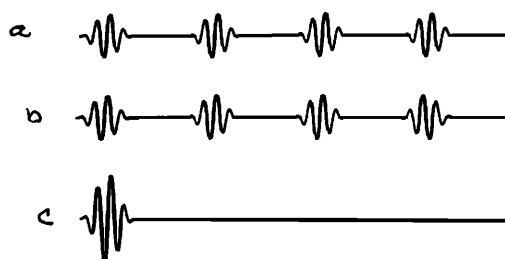


FIG. 2. The three tone burst sequences used: a , b , and c . All have the same power. In a all tone bursts have the same phase. In b every fourth tone burst is shifted in phase 180° with respect to the others. As presented, a and b have the same tone burst rate, and c has one-fourth that rate. Here, b and c have the same fundamental frequency, and a has a fundamental frequency four times as great; c has the same power as a and b . The critical rate up to which b sounds like a is shown as a fraction of the burst frequency as (critical rate)/ Bf . The ratio of half the critical bandwidth to the burst frequency is shown as $ECB/2f$.

the rate R was changed. Thus constancy of tone burst amplitude rather than of sensation level was maintained as the rate was changed. It was observed that the stimuli sounded somewhat louder as the rate R was increased.

While maintaining a given tone burst level (set as noted above), burst or carrier frequency Bf , and number of cycles per tone burst N , the subject was asked to increase the rate R by means of a slider until b sounded just perceptibly different from a . This transitional rate will be referred to as the critical rate. As noted earlier, it is the rate up to which the ear has no clue as to differences in line spectrum and fundamental frequency between stimuli a and b .

II. EXPERIMENT I: EFFECT OF TONE BURST LEVEL

One might expect that the time it takes for the vibration due to a tone burst to effectively die out at its place along the basilar membrane should depend on the tone burst level, that the vibration caused by a just-perceptible tone burst would become undetectable sooner than that induced by an intense tone burst. Therefore, in experiment I the rate at which b sounds just noticeably different from a was determined at several different SL[10] (sensation levels at ten tone bursts per second). Tables I and II give data for four subjects on the reciprocal T of this rate in milliseconds as a function of SL[10] burst frequencies of 2489 and 4978 Hz.

It will be noted that that T increases more rapidly with SL[10] (with stimulus level) for 4978-Hz tone bursts than for 2489-Hz tone bursts. It should also be noted that hearing is more acute at 2479 Hz than at 4978 Hz. Thus the process of level adjustment results in a higher tone burst intensity at 4978 Hz than at 2479 Hz.

III. EXPERIMENT II

Experiment II determined the critical rate as a function of tone burst frequency. For each tone burst frequency the sensation level was set at 20 dB at ten tone bursts a second and a constant tone burst intensity was maintained thereafter. A low sensation level was chosen to avoid as much as possible the increase of T with level that had been found in experiment I.

The critical rates at which b becomes just distinguishably different from a are given as fractions of the tone burst or carrier frequency for four subjects in Table III as (critical rate)/ Bf . Also, the ratio of half of the critical bandwidth

TABLE I. For 2489-Hz burst tone burst frequency, interval T in milliseconds between tone bursts below which tone burst pattern b sounds different from pattern a for the same tone burst rate, for four subjects. SL[10] is the sensation level at ten tone bursts per second. For other rates, the tone burst level was held constant.

| SL[10] dB | $T = 1/(\text{Maximum rate for } b \text{ to sound like } a), \text{ for subject}$ | | | |
|--------------|--|-----|-----|------|
| | jrp | dvo | cc | f ll |
| 10 | 4.7 | 4.9 | 6.5 | 5.0 |
| 20 | 4.6 | 5.3 | 8.1 | 6.5 |
| 30 | 4.9 | 5.7 | 8.1 | 6.6 |
| 40 | 5.7 | 7.2 | 8.5 | 8.7 |

TABLE II. For 4978-Hz tone burst frequency, interval T in milliseconds between tone bursts below which tone burst pattern b sounds different from pattern a for the same tone burst rate, for four subjects. SL[10] is the sensation level at ten tone bursts per second. For other rates, the tone burst level was held constant.

| SL[10] dB | $T = 1/(\text{Maximum rate for } b \text{ to sound like } a), \text{ for subject}$ | | | |
|--------------|--|-----|------|------|
| | jrp | dvo | cc | f ll |
| 10 | 3.3 | 3.2 | 3.1 | 3.0 |
| 20 | 4.4 | 4.3 | 4.3 | 3.4 |
| 30 | 8.5 | 5.0 | 7.4 | 6.1 |
| 40 | 11.9 | 6.3 | 10.3 | 10.0 |

(ERB, or effective rectangular bandwidth, as given by Moore and Glasberg, 1983) to the tone burst frequency (here written as f), that is, $ERB/2f$, is given for comparison. In going from a tone burst frequency of 622–4978 Hz the decline in critical rate as a fraction of burst or carrier frequency can be seen to be somewhat greater than the decline in critical bandwidth.

Despite the fact that the critical rates as fractions of the tone burst frequency differ somewhat among subjects, and that declines in this ratio with increasing tone burst frequency are not strictly monotonic, the trend with frequency and the agreement among subjects seem clear.

IV. DISCUSSION

The experiments performed involved tone burst rates, or, the separation in time of short bursts of tone. It was assumed that for the tone burst lengths in cycles and the tone burst frequencies used, at low enough rates the ear would make no distinction between a tone burst and its negative. This proved to be so. At rates high enough so that a difference is heard, the difference might depend in part on the fact that a tone burst and its negative peak at slightly different times.

The aim of these experiments was to find the critical rate or frequency of repetition below which the sound of the stimuli depended only on rate and was not influenced by differ-

TABLE III. (Critical rate)/ Bf , the ratio to the tone burst frequency of the highest tone burst rate at which pattern b sounds just noticeably different from pattern a , and $ERB/2f$, the ratio of half the critical bandwidth to the tone burst frequency, for various tone burst frequencies and four subjects. For each Bf , the sensation level was set at 20 dB at ten tone bursts per second and the tone burst level was held constant as the rate was changed.

| Note | Bf , Burst frequency, Hz | $ERB/2f$ | (Critical rate)/ Bf , for subject | | | |
|------|-------------------------------|----------|-------------------------------------|------|------|------|
| | | | jrp | dvo | cc | f ll |
| E5 | 622 | 0.07 | 0.14 | 0.13 | 0.10 | 0.14 |
| A5 | 880 | 0.07 | 0.11 | 0.10 | 0.07 | 0.11 |
| E6 | 1244 | 0.06 | 0.10 | 0.09 | 0.08 | 0.08 |
| A6 | 1760 | 0.06 | 0.09 | 0.08 | 0.08 | 0.07 |
| E7 | 2489 | 0.06 | 0.08 | 0.06 | 0.06 | 0.07 |
| A7 | 3520 | 0.06 | 0.06 | 0.04 | 0.06 | 0.08 |
| E8 | 4978 | 0.07 | 0.04 | 0.04 | 0.05 | 0.06 |

ences in line spectrum and in fundamental frequency. This gives the rate up to which the periodicity mechanism is uncontaminated by the effects of low-order harmonics, by any place or related mechanism, if you will. For bursts of a 4978-Hz tone, this critical rate is about 250 Hz, corresponding to B3. At lower rates, experiments can be done to explore the periodicity mechanism in isolation.

The purpose of the tone burst experiment of Gold and Pumphrey (1948) was to measure the Q of the basilar membrane as a function of frequency. We could interpret their Q as a measure of the critical bandwidth. Thus it seems appropriate in Table III to compare the critical rate with the critical bandwidth. A brief argument given as an appendix indicates that the appropriate comparison may be to compare critical rate with half the critical bandwidth. It seems remarkable to the author that the agreement is as close as it is.

The nature of the barely perceptible difference at the critical rate was not determined. One subject heard it as a slight difference in timbre near the critical rate and as a slight difference in pitch at higher rates. In some cases the difference appeared not to change entirely monotonically with rate, but to become more and less as the rate was increased.

The time it takes for the oscillation set up by a tone burst to become negligible increases with the amplitude of the oscillation that has been set up. Experiment I investigates this effect. No analysis of the nature of the nonlinearity involved is attempted. It is merely noted that investigations involving the periodicity mechanism can best be made at low signal levels.

Experiments I and II do not involve pitch directly, but they do give a basis for predicting the outcome of pitch-matching experiments. It seemed desirable to verify that pitch matches are in accord with the observations and interpretations given above. To this end an informal pitch-matching experiment was undertaken. The stimuli used were (1) a tonal stimulus consisting of the first six harmonic partials with equal intensities; (2) tone burst stimulus a (successive tone bursts same phase); and (3) stimulus b (every fourth tone burst 180° out of phase with the rest—fundamental frequency one-fourth that of a for the same tone burst rate).

The three tasks posed all involved adjusting the pitch of the tonal stimulus with respect to that of a tone burst stimulus. The tasks were: TI, to match the tonal stimulus to the pitch of a at prescribed rates; TII, to match the pitch of the tonal stimulus to that of b at prescribed rates; and TIII, to adjust the pitch of the tonal stimulus to one-fifth above that of a at prescribed rates. The results for subject PRC, a musician and engineer with an excellent sense of pitch, are summarized in Table IV.

The response expected in TI if the match is on periodicity is zero semitones. Errors were small below 110 tone bursts per second, and absent for higher rates.

The response expected in TII is zero semitones below the critical rate (about 150 tone bursts per second from Table III) and either 0 (match on rate) or 24 semitones (match on fundamental frequency) above the critical rate.

The response expected for TIII if an interval judgment can be made between tone burst sequence and the adjustable harmonic tone is -7 semitones.

TABLE IV. Pitch matches and offset outcomes in semitones for various rates in tone bursts per second. Adjustable tonal stimulus is the first six harmonic partials with equal intensities. Task I is to adjust this to match stimulus a ; the expected outcome is zero semitones error. Task II is to match to stimulus b ; the expected outcome is zero semitones below the critical rate and 0 or 24 above the critical rate. Task III is to set a pitch of adjustable stimulus one-fifth above that of stimulus a ; the expected outcome is -7 semitones.

| Tone bursts per second | Task I match to a | Task II match to b | Task III, set one-fifth above a |
|------------------------|---------------------|----------------------|-----------------------------------|
| 29.1 | -2 | 1 | -7 |
| 43.7 | 0 | -1 | -7 |
| 73.4 | 1 | 1 | -7 |
| 110 | 0 | 1 | -5 |
| 164.8 | 0 | 0 | -7 |
| 261.6 | 0 | 25 | -7 |
| 370.0 | 0 | 24 | -7 |
| 493.9 | 0 | 24 | -7 |

The data in Table IV are in accord with the outcomes of experiments I and II.

APPENDIX

Think of the tone burst patterns as produced by double-sideband modulation of the tone burst or carrier frequency. If this is so, the bandwidth of a tone burst sequence will be twice the bandwidth of the envelope, the modulating signal.

Consider the period occupied by four successive tone bursts. Let us ask as a sort of minimum criterion for successive tone bursts not overlapping that the amplitudes at the peaks of the modulating signal be equal and that the waveform go to zero between tone bursts.

To accomplish this we must fit the modulating waveform at at least eight points per four tone bursts. We can do this with pairs of sines and cosines at frequencies one, two, three, and four times the reciprocal of the period of the pattern of four tone bursts. Thus the top envelope frequency required will be the number of tone bursts per second. The bandwidth necessary to pass the tone bursts themselves will be twice this. Consequently, the maximum tone burst rate at which we can have the tone burst amplitudes equal and zero amplitude between tone bursts will be half of the bandwidth. If the critical bandwidth at the burst frequency is the bandwidth that determines tone burst overlap, the critical bandwidth should be at least twice the rate at which b sounds like a . This is roughly in agreement with the outcome of experiment II.

de Boer, E. (1976). "On the "Residue" and Auditory Pitch Perception," in *Handbook of Sensory Physiology*, edited by W. D. Keidel and W. D. Neff (Springer-Verlag, Berlin, 1976), Vol. V.

Dufuis, H. (1973). "Consequences of peripheral frequency selectivity for nonsimultaneous masking," *J. Acoust. Soc. Am.* **54**, 1471-1488.

Flanagan, J. L., and Guttman, N. (1960). "On the pitch of periodic pulses," *J. Acoust. Soc. Am.* **32**, 1308-1319.

Gold, T., and Pumphrey, R. J. (1948). "Hearing I, The cochlea as a frequency analyzer," *Proc. R. Soc. London Ser. B* **135**, 462-491.

Green, D. M., Wier, C. C., and Wightman, F. L. (1975). "Pumphrey and gold revisited—again," *J. Acoust. Soc. Am.* **57**, 935-938.

- Guttman, N., and Flanagan, J. L. (1964). "Pitch of high-pass-filtered pulse trains," *J. Acoust. Soc. Am.* **36**, 757-765.
- Hiesey, R. W., and Schubert, E. D. (1972). "Cochlear resonance and phase-reversed signals," *J. Acoust. Soc. Am.* **57**, 935-938.
- Houtsma, A. J. M., and Smurzynski, J. (1990). "Pitch identification and discrimination for complex tones with many harmonics," *J. Acoust. Soc. Am.* **87**, 304-310.
- Moore, B. C. J., and Glasberg, B. R. (1983). "Suggested formulae for calculating auditory-filter bandwidths and excitation patterns," *J. Acoust. Soc. Am.* **74**, 750-753.
- Pierce, J. R. (1990). "Rate, place, and pitch with tonebursts," *Music Percept.* **7**, 205-212.
- Nordmark, J. O. (1968). "Mechanisms of frequency discrimination," *J. Acoust. Soc. Am.* **44**, 1533-1540.
- Rosenberg, A. E. (1965). "Effect of masking on the pitch of periodic pulses," *J. Acoust. Soc. Am.* **38**, 747-758.