U.D.C. 534.322, 534.75, 612.858.4, 159.938.37, 159.932

Effect of Phase Relations on the Timbre of Harmonic Tones

Susumu TACHI

Abstract

Monaural phase effects (MPE's) in steady-state musical tones were experimentally studied for two- and three-component tones. The two types of stimulus pairs used were:

$$\begin{cases} \sin 2\pi\cdot ft + A_2/A_1 \sin 2\pi\cdot 2ft, \\ \sin 2\pi\cdot ft + A_2/A_1 \sin (2\pi\cdot 2ft + \theta); \text{ and} \end{cases}$$

$$\begin{cases} \sin 2\pi\cdot ft + A_2/A_1 \sin 2\pi\cdot 2ft + A_3/A_1 \sin 2\pi\cdot 3ft, \\ \sin 2\pi\cdot ft + A_2/A_1 \cos 2\pi\cdot 2ft + A_3/A_1 \sin 2\pi\cdot 3ft. \end{cases}$$

Timbre discrimination between the stimuli of each pair was investigated as a function of f, A_2/A_1, A_3/A_1, and θ at a constant sensation level (about 30 dB). The AB method was applied, and the information rates were calculated and used as a measure of discriminability.

The most important findings are:

(1) MPE's can be perceived for two- and three-component steady state musical tones under certain conditions of f, A_2/A_1, A_3/A_1, and θ.

(2) For stimulus pairs with a frequency beyond 1248 Hz, no timbre difference can be heard.

(3) The maximal discriminability of two-component tones with f of 193 Hz is obtained when A_2/A_1 is about 0.5. For three-component tones maximal discriminability is reached at a certain value of A_2/A_1, which increases with A_3/A_1. And a great difference is observed between the stimulus pairs with $A_3/A_1>1$ and with $A_3/A_1<1$.

The experimental results suggest that MPE's are related to differences in the waveform of the vibration patterns of the basilar membrane, and are particularly related to the corresponding differences in the temporal patterns of the nerve impulses near the peak point of the envelope of the basilar membrane vibration caused by the higher component of the stimulus.

I. Introduction

Recently monaural phase effects (MPE's) in steady state musical tones have come to be argued again. The first controversy waged over the phase rule established by von Helmholtz[1] began when Koenig[2] studied MPE with the aid of a "wave siren" designed by himself. The argument continued for about thirty years concerning an artifact of the wave siren, but it did not solve the problem; in the light of present-day insights, we may say that the effect of phase on timbre is too small to be studied successfully by means of the wave siren. It was useful to neglect MPE in order to establish an auditory theory to a first approximation, which was mainly a pitch theory especially explaining the experimental data which were obtained by using simple sinusoidal stimuli. With the development of experimental equipment and signal processing techniques, it has been made clear that the ear perceives monaural phase cues contained in steady-state musical tones. Chapin and Firestone[3], Trimmer and Firestone[4], Licklider[5], Schroeder[6], and Craig and Jeffress[7] have all demonstrated monaurally detectable changes which are related to phase changes in total stimuli.

The next problems to be solved are as follows.

(1) To investigate relations between the components of a tone under which one can perceive the MPE, and to quantitatively investigate how much the MPE can be perceived.

(2) To integrate MPE's into auditory theory and bring auditory theory to a better approximation.

Plomp and Steeneken[8] quantitatively investigated the effect of phase on timbre and found the extent to which timbre is determined by maximal effect of phase. And Raiford and Schubert[9] succeeded in showing that changes in the phase relation of simple octave waveforms can be discriminated monaurally and the discrimination can be made even by subjects with little or no laboratory listening experience.

In this paper an experimental work is reported, which is designed to provide perception of monaural phase cues data for two and three-component steady-state musical tones as a function of fundamental frequency and amplitude relations between the components with constant sensation level (SL). The results are compared with other psychological experimental data and physiological evidence, and an auditory mechanism to discriminate phase is considered on the basis of experimental results.

II. Experimental Method

1. APPARATUS

In order to investigate MPE as a function of amplitude relations and phase relations between the components, a signal generator is needed which provides control over the amplitudes and phases of the individual components of a periodic complex tone. Two prerequisites of the apparatus designed and used in this study are:

(1) Each component should be synchronized completely so that the phase relation between components

will not change with time.

(2) Phase relations between components should be easily and continuously adjusted.

Figure 1 shows the functional circuit layout for the complex wave generator circuits used for these experiments. The outputs of the scale of 3, 4, and 5 counters are applied to a NAND gate, the outputs of the scale of 2, 3, and 5 to another NAND gate, and the outputs of the scale of 4 and 5 are fed to still another NAND gate. Applying the same procedure we can get rectangular waves which have harmonic relations up to the 6th harmonic with the same synchronized phase at the outputs of the six NAND gates. The output of every NAND gate triggers a flip-flop, whose positive and negative outputs trigger another two flip-flops.

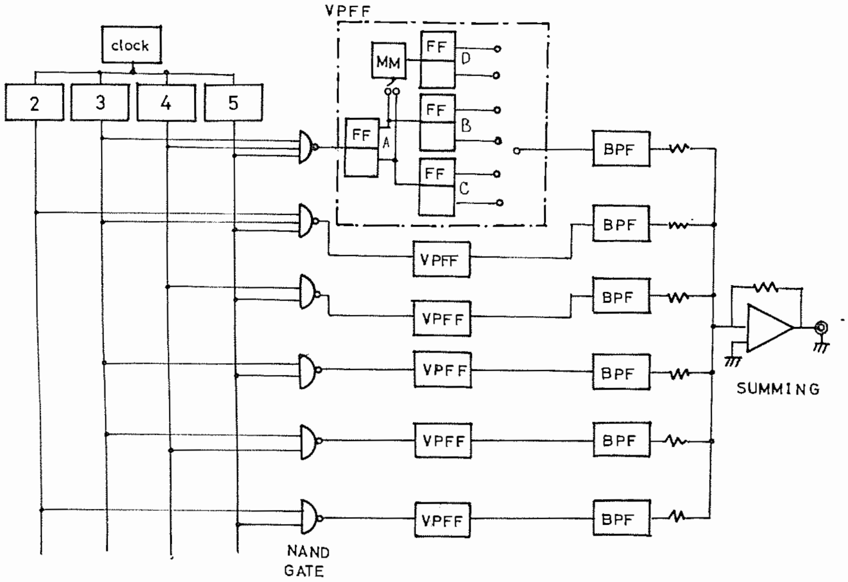

Fig. 1 Details of the Fourier synthesizer. The signs FF, MM, VPFF, and BPF are abbreviations for a flip-flop, monostable multivibrator, variable phase flip-flop, and band pass filter, respectively. A VPFF is the flip-flop the output phase of which can be freely adjusted. Its construction is described in the area enclosed by dotted lines.

The output of each NAND gate triggers each VPFF (variable phase flip-flop), the construction of which is described in the area enclosed by dotted lines in Fig. 1. The outputs of flip-flop A and $\bar{A}$ trigger the other two flip-flops B and C, respectively. The outputs B and C have the phase difference of 90 degree (the phase of the output B differs by 90 degree from the output C). The output of flip-flop D shows an arbitary phase according to the duration time of monostable multivibrator MM, the input of which is the output of flip-flop A. Thus we get rectangular waves which form harmonic relations in their fundamental frequencies and their phase relations can be varied arbitarily. The output of each VPFF is applied to a band-pass filter BPF whose center frequency is tuned to the fundamental frequency of its input signal. By summing up the six outputs with various weights we can get harmonic signals with various amplitude and phase relations between their components.

2. TEST PROCEDURE

Two harmonic tones which have the same amplitude patterns but have different phase patterns are presented to both ears of a subject monaurally by a head receiver (Stax electrostatic ear-speaker SR-3), which has a reasonably flat transmission characteristics up to 3000 Hz. A subject first listens to two tones tentatively named A and B, alternatively, and memorizes which is which. After this he listens to a tone presented according to an m-sequence (quasi-randomly), and judges which of the two original tones he has just heard. He judges about 160 times for one pair of stimuli, and writes A or B according to his judgement. The correct sequence is recorded on a pen-recorder and it is used to check his answer. If we assume that the stimuli A and B are presented equally and randomly, we can get the information transmission rate by the formula,[10]

$$R=1+1/2\left[(p_a+q_b)\log_2\frac{1}{p_a+q_b}+(p_b+q_a)\log_2\frac{1}{p_b+q_a}-p_a\log_2\frac{1}{p_a}-q_a\log_2\frac{1}{q_a}-p_b\log_2\frac{1}{p_b}\right.$$
$$\left.-q_b\log_2\frac{1}{q_b}\right],$$

where p_a is the probability of receiving an A when an A is sent, q_b is the probability of receiving a B when an A is sent, p_b is the probability of receiving a B when a B is sent, and q_b is the probability of receiving an A when a B is sent.

Figure 2 shows a schematic diagram of the apparatus used. A cathode-ray oscilloscope bridged at the point i permits convenient observation of the waveshape being fed to the ear.

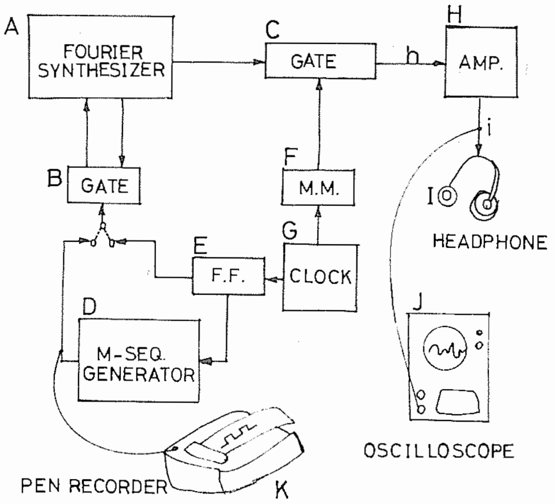

Fig. 2 Schematic diagram of the experimental setup.

III. Experiments

1. EXPERIMENT 1 Control Experiment

In order to confirm that phase cues can be perceived in sustained harmonic tones, the following stimuli were subjected to ten subjects with normal hearing.

(1) $\sin 2\pi\cdot ft+2/5 \sin 2\pi\cdot 2ft+1/5 \sin 2\pi\cdot 3ft$,

(2) $\sin 2\pi\cdot ft+2/5 \cos 2\pi\cdot 2ft+1/5 \sin 2\pi\cdot 3ft$,

where $f=193$ Hz.

Stimuli (1) and (2) are presented alternatively, and all the subjects except one heard definite clear difference between the two. They described the timbre difference as follows:

[1] Stimulus (2) sounds somewhat higher in pitch than stimulus (1).

[2] Stimulus (2) sounds somewhat louder than stimulus (1).

[3] Stimulus (1) sounds like "buh", whereas stimulus (2) sounds like "boh".

Almost all the subject distinguished between the stimuli (1) and (2) by using all of the three terms [1], [2], and [3].

Next the two stimuli were presented pseudo-randomly according to an m-sequence of period 127. Each stimulus was presented for 1.7 sec, and followed by an interval of 0.1 sec to eliminate transient effects. For one testrun about 160 such stimuli were presented, both stimuli appearing almost equally.

If we assume that the subjects cannot distinguish between two stimuli at all and answered randomly, then information rates deviate from the original level of 0 bit. If the transition from 1 bit to 0 bit information rates may be considered as an integrated Gaussian distribution, information rate above 0.03 bits represents a deviation from chance at the 1% level of significance.

Timbre difference due to a small change of amplitude pattern was measured in order to investigate the precision of these experiments and the result is in Table 1, which indicate that amplitude difference below 20% did not affect the results, and the apparatus used has precision of at least 5%.

Table 1 Timbre difference due to a small error of amplitude pattern. The results show that the amplitude errors below 20% are under the 1% level of significance, and can not be said to have effect on timbre.

amplitude patterns of the two stimuli $(A_2/A_1)_1$ vs. $(A_2/A_1)_2$	percent correct	information rate [bit]
0.20 vs. 0.22	50	0.00
0.20 vs. 0.24	52	0.01
0.20 vs. 0.26	64	0.06

Three subjects who can distinguish the timbre difference most clearly were chosen and used as subjects of the following experiments. The data used were the average of the three subjects.

2. EXPERIMENT 2 Experiments with Two-component Tones

In order to investigate the most simple and fundamental case, two-component tones with harmonic relations, which have the same amplitude ratios between the components but differ in their phase relations, were presented to the subjects. The following stimuli were used.

(3) $\sin 2\pi \cdot ft + A_2/A_1 \sin 2\pi \cdot 2ft$,

(4) $\sin 2\pi \cdot ft + A_2/A_1 \cos 2\pi \cdot 2ft$,

(5) $\sin 2\pi \cdot ft - A_2/A_1 \sin 2\pi \cdot 2ft$,

(6) $\sin 2\pi \cdot ft - A_2/A_1 \cos 2\pi \cdot 2ft$, where $f=193$ Hz.

Experiments were carried out under the same condition as in the previous experiment, i. e., stimulus pairs (3) & (4), (3) & (5), and (3) & (6) were used to investigate detectability of timbre difference, and one of the two, for instance, (3) or (4) of (3) & (4) pair, was presented to both ears of the subject monaurally according to an m-sequence of period 127. Each stimulus duration of 1.7 sec was followed by an interval of 0.1 sec. About 160 stimuli were applied in one run of experiment at SL of about 30 db. Figure 3 represents the information rate as a function of A_2/A_1, for the timbre difference between (3) & (4), (3) & (5), and (3) & (6). The marks ○, △, and × indicate results for the stimulus pairs (3) & (4), (3) & (5), and (3) & (6), respectively.

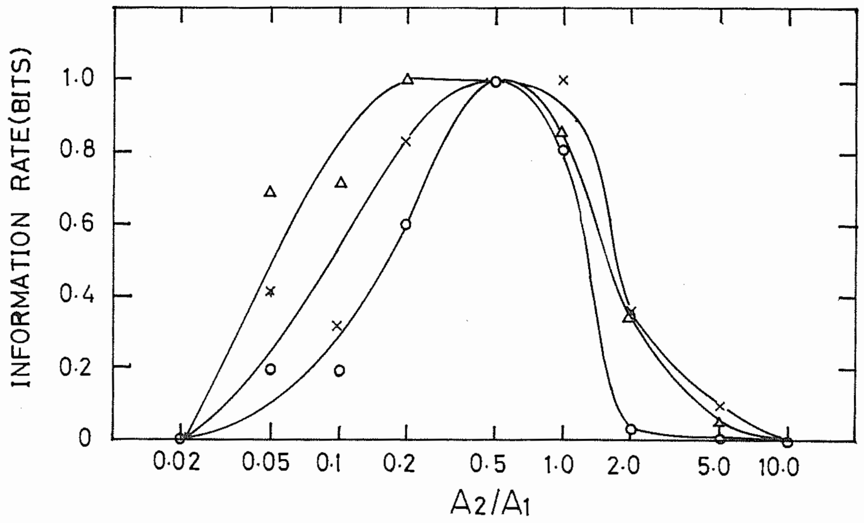

Fig. 3 Timbre discrimination between the stimulus pair:
$$\begin{cases}\sin 2\pi\cdot 193t + A_2/A_1 \sin 2\pi\cdot 386t,\\ \sin 2\pi\cdot 193t + A_2/A_1 \sin (2\pi\cdot 386t+\theta),\end{cases}$$
as a function of A_2/A_1 for each of three values of θ. The marks ○, △, and × indicate results for the stimulus pairs with $\theta=\frac{1}{2}\pi, \pi, \frac{3}{2}\pi$, respectively. The ordinate represents the information rate, the information about the timbre difference (bit) transmitted per stimulus.

3. EXPERIMENT 3 Experiments with Three-component Tones

Experiments were carried out under the same condition as the previous ones using three-component tones as follows:

(7) $\sin 2\pi\cdot ft + A_2/A_1 \sin 2\pi\cdot 2ft + A_3/A_1 \sin 2\pi\cdot 3ft$,

(8) $\sin 2\pi\cdot ft + A_2/A_1 \cos 2\pi\cdot 2ft + A_3/A_1 \sin 2\pi\cdot 3ft$,

where $f=193$ Hz.

Timbre difference between the stimuli (7) and (8) was investigated as a function of A_2/A_1 and A_3/A_1. Figure 4 represents the information rate for each of five values of A_3/A_1=constant as a function of A_2/A_1.

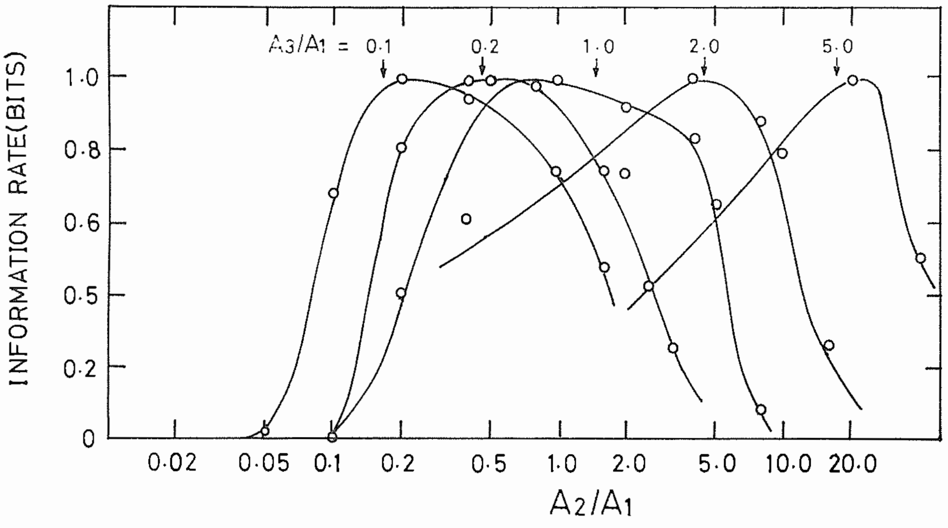

Fig. 4 Timbre discrimination between the stimulus pair:

$$\begin{cases}\sin 2\pi\cdot 193t + A_2/A_1 \sin 2\pi\cdot 386t + A_3/A_1 \sin 2\pi\cdot 579t,\\ \sin 2\pi\cdot 193t + A_2/A_1 \cos 2\pi\cdot 386t + A_3/A_1 \sin 2\pi\cdot 579t,\end{cases}$$

as a function of A_2/A_1 for each of five values of A_3/A_1 (0.1, 0.2, 1.0, 2.0, and 5.0). The ordinate represents the information rate, the information about the timbre difference (bit) transmitted per stimulus.

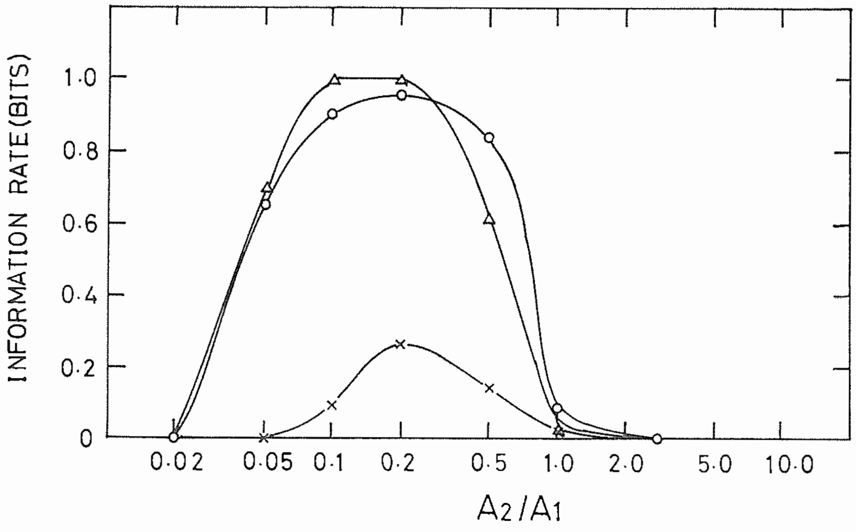

Fig. 5 Timbre discrimination between the stimulus pair:

$$\begin{cases}\sin 2\pi\cdot 416t + A_2/A_1 \sin 2\pi\cdot 832t,\\ \sin 2\pi\cdot 416t + A_2/A_1 \sin (2\pi\cdot 832t + \theta),\end{cases}$$

as a function of A_2/A_1 for each of three values of θ. The ordinate represents the information rate, the information about the timbre difference (bit) transmitted per stimulus. The marks ○, △, and × indicate results for the stimulus pairs with $\theta = \frac{1}{2}\pi, \pi, \frac{3}{2}\pi$, respectively.

4. EXPERIMENT 4 Effect of Fundamental Frequency

For the experiments 1 through 3 the fundamental frequency of the complex-tone stimuli was fixed at 193 Hz. In order to investigate the effect of fundamental frequency on detectability of phase, similar experiments were carried out for fundamental frequencies of 416 Hz, 832 Hz, and 1248 Hz.

4.1. $f=416$ Hz

4.1a. two-component tones

Similar tests as in experiment 2 were carried out under the same conditions, except for the fundamental frequencies of the stimuli:

(3′) $\sin 2\pi\cdot ft+A_2/A_1 \sin 2\pi\cdot 2ft$,
(4′) $\sin 2\pi\cdot ft+A_2/A_1 \cos 2\pi\cdot 2ft$,
(5′) $\sin 2\pi\cdot ft-A_2/A_1 \sin 2\pi\cdot 2ft$,
(6′) $\sin 2\pi\cdot ft-A_2/A_1 \cos 2\pi\cdot 2ft$, where $f=416$ Hz.

Figure 5 represents the results, (a), (b), and (c) indicating the information rates for the difference between (3′) & (4′), (3′) & (5′), and (3′) & (6′), respectively, as a function of relative amplitude A_2/A_1.

4.1b. three-component tones

In order to investigate detectability of the timbre difference between three-component tones which have the same amplitude patterns but different phase patterns, as a function of A_3/A_1 and A_2/A_1 for a fundamental frequency of 416 Hz, an experiment using a method of adjustment was carried out for selected pairs of the following stimuli.

(7′) $\sin 2\pi\cdot ft+A_2/A_1 \sin 2\pi\cdot 2ft+A_3/A_1 \sin 2\pi\cdot 3ft$,
(8′) $\sin 2\pi\cdot ft+A_2/A_1 \cos 2\pi\cdot 2ft+A_3/A_1 \sin 2\pi\cdot 3ft$,
(9′) $\sin 2\pi\cdot ft-A_2/A_1 \sin 2\pi\cdot 2ft+A_3/A_1 \sin 2\pi\cdot 3ft$,
(10′) $\sin 2\pi\cdot ft-A_2/A_1 \cos 2\pi\cdot 2ft+A_3/A_1 \sin 2\pi\cdot 3ft$,
where $f=416$ Hz.

The result is Fig. 6, in which the abscissa represents relative amplitude A_2/A_1 and the ordinate relative amplitude A_3/A_1. The data points are obtained by plotting the detectability threshold achieved by the method of adjustment for the difference (7′) & (8′), (7′) & (9′), and (7′) & (10′), and ○, △, and × represent the data points averaged over five trials, respectively.

4.2. $f=832$ Hz

A similar experiment to that of experiment 2 was carried out for two-component stimuli of fundamental frequency of 832 Hz under the same conditions as in experiment 2, using the following stimuli:

(3″) $\sin 2\pi\cdot ft+A_2/A_1 \sin 2\pi\cdot 2ft$,
(4″) $\sin 2\pi\cdot ft+A_2/A_1 \cos 2\pi\cdot 2ft$,
(5″) $\sin 2\pi\cdot ft-A_2/A_1 \sin 2\pi\cdot 2ft$,
(6″) $\sin 2\pi\cdot ft-A_2/A_1 \cos 2\pi\cdot 2ft$, where $f=832$ Hz.

The timbre difference between the stimuli (3″) & (4″), (3″) & (5″), and (3″) & (6″) were examined. In this case, however, the subjects couldn't notice the difference, and it was hard to memorize which stimulus was supposed to be A and which B, except for one stimulus pair, (3″) & (5″). Figure 7 represents the information rate as a function of A_2/A_1 for that stimulus pair.

For three-component tones,

(7″) $\sin 2\pi\cdot ft+A_2/A_1 \sin 2\pi\cdot 2ft+A_3/A_1 \sin 2\pi\cdot 3ft$,

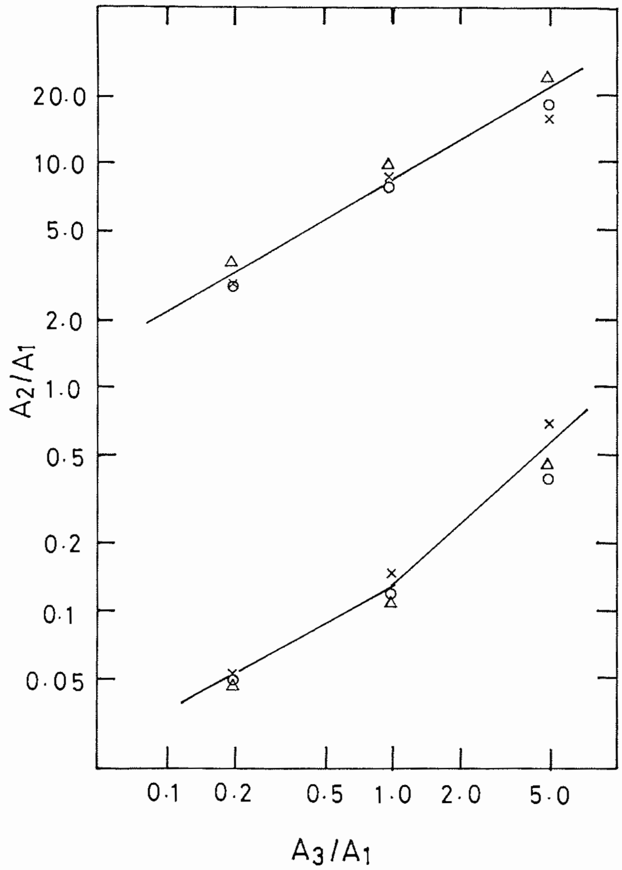

Fig. 6 Upper and lower limits of the distinguishable timbre difference between the stimuli (7)′, (8)′, (9)′, and (10)′. In the intermediate region the stimuli were distinguishable. The marks ○, △, and × indicate the upper and lower limits of the distinguishable timbre difference between (7)′ & (8)′, (7)′ & (9)′, and (7)′ & (10)′, respectively

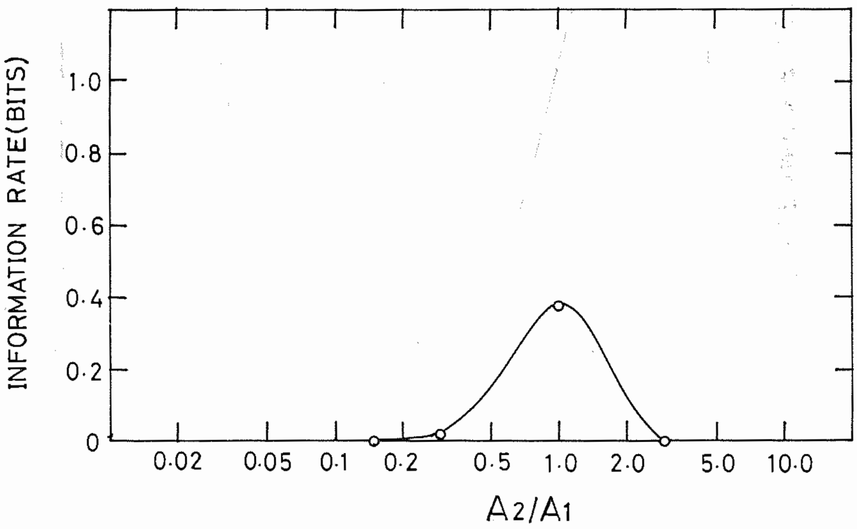

Fig. 7 Timbre discrimination between the stimulus pair:

$$\begin{cases}\sin 2\pi\cdot 832t + A_2/A_1 \sin 2\pi\cdot 1664t,\\ \sin 2\pi\cdot 832t - A_2/A_1 \sin 2\pi\cdot 1664t,\end{cases}$$

as a function of A_2/A_1. The ordinate represents the information rate, the information about the timbre difference (bit) transmitted per stimulus.

(8″) $\sin 2\pi\cdot ft + A_2/A_1 \cos 2\pi\cdot 2ft + A_3/A_1 \sin 2\pi\cdot 3ft$,

(9″) $\sin 2\pi\cdot ft - A_2/A_1 \sin 2\pi\cdot 2ft + A_3/A_1 \sin 2\pi\cdot 3ft$,

(10″) $\sin 2\pi\cdot ft - A_2/A_1 \cos 2\pi\cdot 2ft + A_3/A_1 \sin 2\pi\cdot 3ft$,

where $f=832$ Hz, similar experiments were conducted for $A_3/A_1=0.2$, 1.0, 2.0, and 5.0. For any A_2/A_1 the subjects could hardly notice the timbre differences.

4.3. $f=1248$ Hz

For two-component tones an experiment like 1 or 2 was carried out, but no timbre difference could be heard, between all the combinations of the stimuli, and for all values of A_2/A_1. For timbre difference between $\sin 2\pi\cdot ft+0.3 \sin 2\pi\cdot 2ft$, and $\sin 2\pi\cdot ft-0.3 \sin 2\pi\cdot 2ft$, $f=1248$ Hz, which one subject felt a subtle difference, the information rate was 0.01, which is under the level of significance, so we can't say there is a difference.

5. EXPERIMENT 5 Discrimination of relative phase shift

To investigate the effect of phase angle shift on the detectability of the timbre difference, the information rate was measured as a function of a phase shift of the second harmonics. The timbre difference between the two stimuli,

$$\begin{cases} \sin 2\pi\cdot ft+0.5 \sin 2\pi\cdot 2ft, \\ \sin 2\pi\cdot ft+0.5 \sin (2\pi\cdot ft\pm\theta), \quad \text{where } f=193 \text{ Hz}, \end{cases}$$

was examined. The result is as in Fig. 8, in which the abscissa represents phase shift of the second harmonics (radian) and the ordinate information rate (bits).

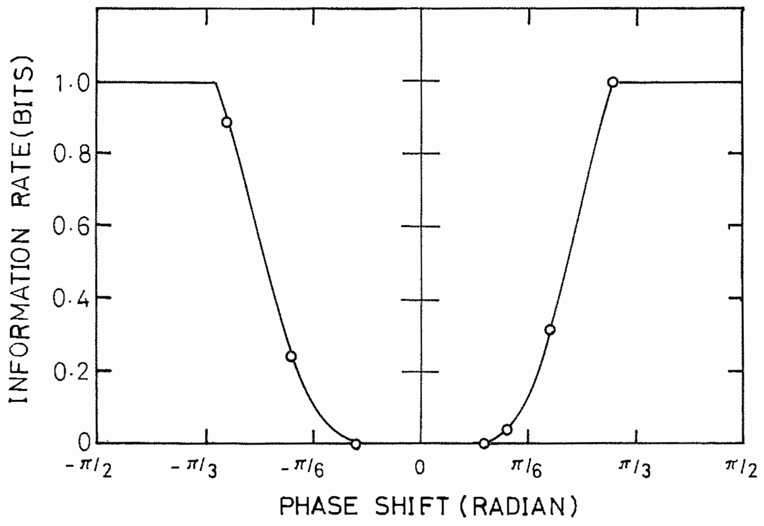

Fig. 8 Timbre discrimination between the stimulus pair:

$$\begin{cases} \sin 2\pi\cdot 193t+0.5 \sin 2\pi\cdot 386t \\ \sin 2\pi\cdot 193t+0.5 \sin (2\pi\cdot 386t+\theta), \end{cases}$$

as a function of θ. The ordinate represents the information rate, the information about the timbre difference (bit) transmitted per stimulus.

IV. Discussion

(1) The fact that no timbre difference due to phase cue can be heard for the two-component stimulus pairs with the fundamental frequency of 1248 Hz and higher coincides in frequency limit with the fact that the pitch phenomena related to the waveform periodicity[11,12], including pitch of the "residue", have the frequency limit of about 1000 Hz. According to the periodicity pitch phenomena, pitch sensation is due to the periodicity of the sound for those sounds with a fundamental frequency of up about to 1000 Hz, whereas for a higher tone nerve impulses can not follow the tone, therefore we hear a pitch corresponding to the place of maximal movement along the basilar membrane. And the periodicity pitch phenomena strongly suggest the existence of a time-related processing of the nerve impulses.

The results of the discriminability of phase cues also support this view. In experiment 4, for a fundamental frequency of 416 Hz and 832 Hz discriminability of two-component stimulus pair can be seen to depend on how much the phase between the stimuli separates (the differences between the curves ○, △, and × in Fig. 5), i. e. the pair with the phase difference of π radian can be distinguished most clearly.

For a fundamental frequency of 193 Hz the dependency of discriminability on phase shift is obtained in experiment 5. (Since the phase shift in experiment 4 is every 90 degree, it is too large to cause the obvious difference between the curves ○, △, and × for the frequency of 193 Hz.) For 832 Hz tone the pair with π radian-shift of second harmonics (i. e., stimuli (3″) and (5″)) is the most easy to descriminate. The information rate (0.4 bits) is about 0.7 radians for 193 Hz tone according to Fig. 8 of experiment 5. The ratio of the phase shift is $\pi/0.7=4.5$, while the ratio of the frequency is $832/193=4.3$. This suggests an existence of a linear relation between the fundamental frequency and the necessary phase shift to transmit a certain amount of information.

According to Fig. 8 the necessary phase shift to discriminate the two-component stimuli of frequency 193 Hz (i. e., to get the information rate of 1.0 bit) is about $\pi/3$ radians. Phase shift of $\pi/3$ radians corresponds, in this case, to about 1 ms, which plays a roll of a unit for phase discrimination. The fundamental frequency of the stimulus that has this unit as its period is 1000 Hz, which corresponds to the frequency limit of the periodicity pitch phenomena.

Those facts suggest that the MPE's are perceived by some kind of time processing of the nerve impulse pattern that corresponds to the vibration pattern along the basilar membrane.

(2) The curves presented in Fig. 3, Fig. 5, and Fig. 7, which are the results for two-component tones, are similar in shape, though their values are somewhat different. A typical shape may be described as follows: It has a maximum at a particular amplitude relation of A_2/A_1, or plateau for a particular region of A_2/A_1, and the information rate decreases to zero as A_2/A_1 becomes either very small or very large. It seems reasonable to have this kind of shape, because the more one component of two-component tone becomes dominant, the more the two-component tone can be granted as a one-component tone, therefore the less the timbre difference between the tones which have the same amplitude pattern but differ their phase relations can be heard out. And for a medium amplitude relation of A_2/A_1 or a medium region of A_2/A_1

the timbre difference can be heard clearly. Then what does the particular A_2/A_1 which gives the maximal information rate mean?

As is seen in Fig. 4, for three-component stimulus pairs maximal discriminability is reached at a certain value of A_2/A_1, which increases with A_3/A_1. And a great difference is observed between the stimulus pair with $A_3/A_1<1$ and with $A_3/A_1>1$: when $A_3/A_1<1$, an obtuse peak is obtained at such a value of A_2/A_1 that $A_2/A_1<1$; on the other hand when $A_3/A_1>1$, a sharp peak is obtained at such a value of A_2/A_1 that $A_2/A_1>1$. This tend is more easily seen in Fig. 9, which is got by replotting the points that transmit 0.9 bits information as a function of A_2/A_1, and A_3/A_1. How are these results interpreted?

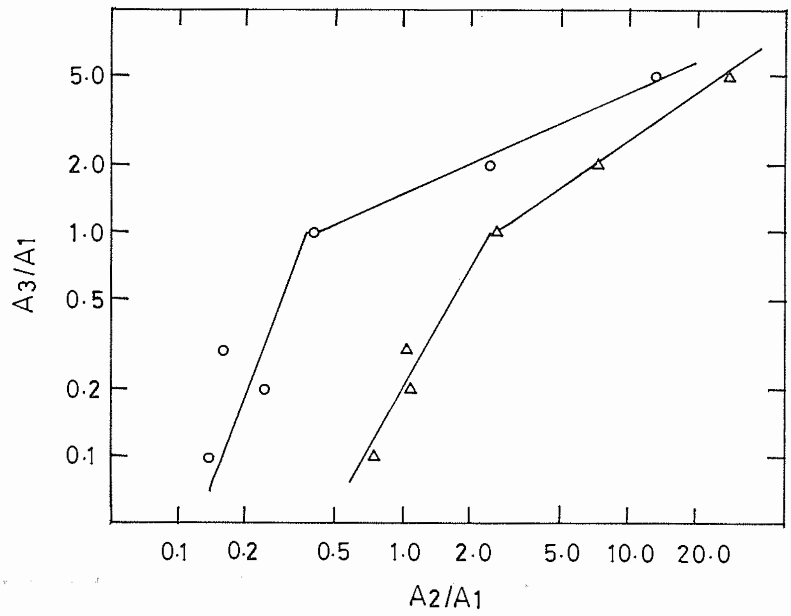

Fig. 9 Replotted data from the curves in Fig. 4 as a function of A_2/A_1, and A_3/A_1. The marks ○ and × represent the points where information rate of 0.9 bits are obtained.

Difference in phase patterns will result in difference in the vibration patterns at various points along the basilar membrane. Figure 10 shows the response of the basilar membrane to sinusoidal stimulation.[13] The amplitudes of vibration of the basilar membrane increase almost linearly with the distance from the stapes up to the maximum and thereafter drop suddenly. If the observation is not of the basilar membrane but of vibration near the tips of the hair cells near the tectorial membrane, the drop in amplitude close to the helicotrema is even sharper, because there is a sharpening effect produced by the mechanical traveling wave that serves to cut off the response in the apical portion of the cochlea. In nervous system there is an added sharpening of the response in the apical portion of the cochlea by the inhibition associated with a time delay, and a funneling effect similar to that responsible for the Mach band (von Békésy[14]).

The mechanical action of the cochlea is assumed to be linear over a reasonably wide range of stimulus intensities, in which the stimuli used were. A two-component stimulus gives rise to the movement of the basilar membrane that is the superposition of two movements, each of which corresponds to the simple stimulus. The envelope of the vibration of the basilar membrane has ordinarily two peaks, which correspond to the higher component and the lower component of the stimulus.

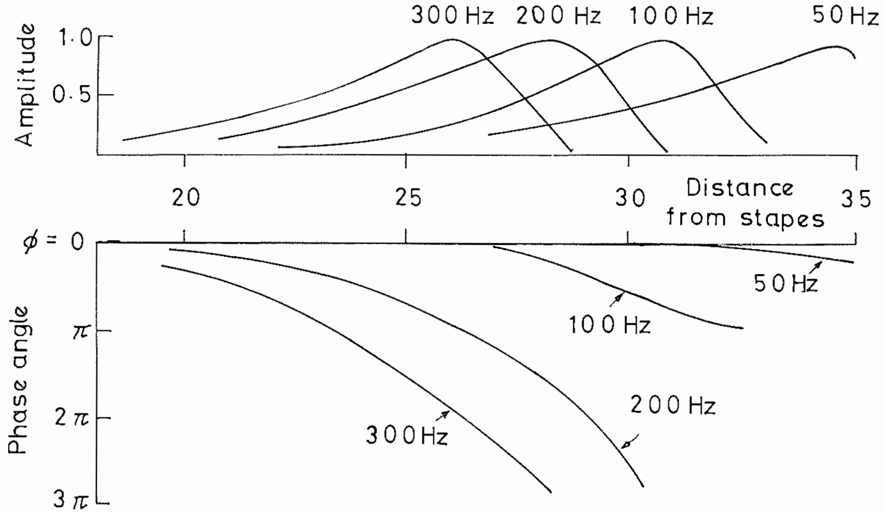

Fig. 10 Phase displacement and resonance curves for four low tones, from von Békésy [13].

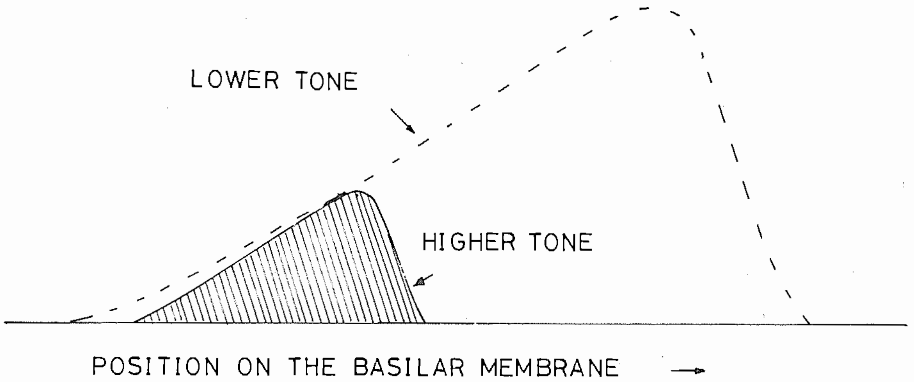

Fig. 11 The vibration patterns of two tones along the basilar membrane. This shows the presumed condition under which maximal information about the interference of the two tones is transmitted.

It can be reasoned that the portion that transmits as nerve impulses the information about the waveform of the original stimulus is the peak corresponding to the higher component of the stimulus. The reason is as follows: The basal turn (near stapes) responds to practically all frequencies in the audible range, while the upper parts of the cochlea (near helicotrema) respond only to the stimuli of low frequencies. This is evident from the tuning curve of Fig. 10, and from the previous discussion in inhibition. This is also apparent from the microphonic responses of guinea pig recorded from basal turn and from third turn of the cochlea (Tasaki^{15}).

So the portion in question must be from the peak corresponding to the higher component of the stimulus to the stapes. It is, however, unlikely to take the whole portion; because Mach-type inhibition is active around the peak, and by funneling nerve activities around the peak may be reduced. Therefore, if the place where the waveform of the stimulus is preserved and transmitted toward the brain should be seeked, although it is somewhat transformed in its amplitude and phase according to the tuning curves of Fig. 10, it must be through the nerve fibers near the peak of the higher component of that stimulus.

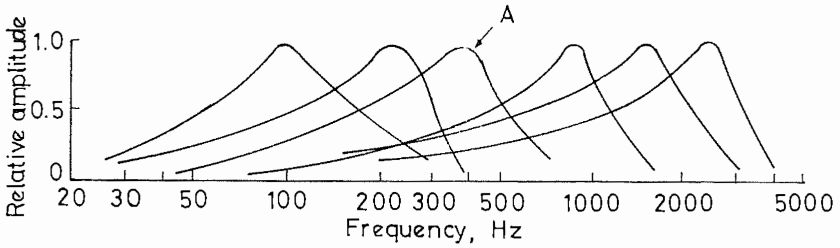

Fig. 12 Forms of resonance curves for six positions along the cochlear partition determined by von Békésy. Each curve shows the relative amplitude of vibration of a point on the cochlear partition as the stapes of the cochlea is driven sinusoidally at various frequencies with constant rms displacement. The six curves are for points "tuned" to 100, 200, 400, 800, 1600 and 2400 Hz. The amplitude for 200 Hz of the curve A, which is the resonance curve for the partition tuned to 400 Hz, gives an estimated value of A_2/A_1 for the fundamental frequency of 200 Hz-tone to meet the condition of Fig. 11. From von Békésy [13].

The results of phase discriminability as a function of relative amplitude can be explained by presupposing that the information about phase is transmitted most effectively when the two components whose phase difference is transmitted have the same amplitude. This situation is presented in Fig. 11, where the information about phase difference is transmitted through the channel corresponding to the higher component.

Figure 12 shows the forms of resonance curves for six positions along the cochlear partition determined by von Békésy[13]. Each curve shows the relative amplitude of vibration of a point on the cochlear partition as the stapes of the cochlea is driven sinusoidally at various frequencies with constant rms displacement. The six curves are for points "tuned" to 100, 200, 400, 800, 1600 and 2400 Hz. The amplitude for 200 Hz of the curve A, which is the resonance curve for the partition tuned to 400 Hz, gives an estimated value of A_2/A_1 for the fundamental frequency of 200 Hz-tone to meet the condition of Fig. 1.1, This value is estimated to be about 0.5. In Fig. 3 we can see the value of A_2/A_1 which gives the maximal discriminability is estimated to be about 0.5. Thus a rough coincidence can be observed.

(3) The difference of the discriminability curve in their shapes between the stimulus pair with $A_3/A_1>1$ and with $A_3/A_1<1$, and the values of A_2/A_1 where the information rate has its maximum can also be explained along the same line. In experiment 3 the discriminability of phase is investigated by changing the phase of the second harmonic. So when the third harmonic is less than the fundamental ($A_3/A_1<1$), the phase difference between the fundamental and the second harmonic may be dominant. At that case the tendency of the curve is natural to be similar to that of two-component stimuli.

On the contrary, when the third is more than the fundamental ($A_3/A_1>1$), the interference between the second and the third may be dominantly observed. In that case the information about the interference between the second and the third will be transmitted by the nerves near the maximal vibration corresponding to the third on the basilar membrane. The value of A_2/A_1 which meets the condition of Fig. 11, therefore becomes larger compared with that for the former case. The distance of the peaks caused by the second harmonic and the third harmonic is shorter than that of the peaks caused by the fundamental and the second. So if one of them becomes larger, then the other is more likely to be buried. The condition of Fig. 11 with some allowance is obtained, therefore, for rather narrow range of the value of A_2/A_1 for the case that $A_3/A_1>1$ than for the case such that $A_3/A_1<1$. This can be the reason why the discriminability curve has sharper maximum for $A_3/A_1>1$.

V. Conclusions

(1) MPE's can be perceived for two- and three-component steady state musical tones under certain conditions of their fundamental frequencies, their amplitude patterns and phase patterns.

(2) For stimulus pairs with a frequency beyond about 1000 Hz, no timbre difference can be perceived.

(3) The maximal discriminability of two-component tones with fundamental frequency of 193 Hz is obtained for those stimulus pairs with $A_2/A_1=0.5$ (relative amplitude ratio between the second harmonic and fundamental). For those tones with $A_2/A_1=0.5$, the two tones are fully distinguished if the phase difference of second harmonics is greater than $\pi/3$ radians.

(4) For three-component tone pairs the maximal discriminability is reached at a certain value of A_2/A_1, which increases with A_3/A_1. When $A_3/A_1<1$, the maximal discriminability is reached at such a value of A_2/A_1 that $A_2/A_1<1$. Whereas when $A_3/A_1>1$, it is reached at such a value of A_2/A_1 that $A_2/A_1>1$, and it becomes suddenly low after A_2/A_1 becomes greater than the value of A_2/A_1 which gives the maximum.

(5) The experimental results suggest that MPE's are related to differences in the waveform of the vibration patterns of the basilar membrane, and are particularly related to the corresponding differences in the temporal patterns of the nerve impulses near the peak point of the envelope of the basilar membrane vibration caused by the higher component of the stimulus.

References

1) H. L. F. von Helmholtz: On the Sensations of Tone, Dover Publications, Inc., New York (1951) Reprint of A. J. Ellis' 1885 English translation of the 4th German edition (and last) of 1877. von Helmholtz completed the 1st edition of his work in 1862.

2) R. König: Bemerkungen über die Klangfarbe, Ann. Physik. Chem., **14**, 369/393 (1881)

3) E. K. Chapin & F. A. Firestone: The influence of phase on tone quality and loudness; the interference of subjective harmonics, J. Acoust. Soc. Am., **5**, 173/180 (1934)

4) J. D. Trimmer & F. A. Firestone: An investigation of subjective tones by means of the steady tone phase effect, J. Acoust. Soc. Am., **9**, 24/29 (1937)

5) J. C. R. Licklider: Effects of changes in the phase pattern upon the sound of a 16-harmonic tone, J. Acoust. Soc. Am., **29**, 780 (A) (1957)

6) M. R. Schroeder: New results concerning monaural phase sensitivity, J. Acoust. Soc. Am., **31**, 1579 (B) (1959)

7) J. H. Craig & L. A. Jeffress: Effect of phase on the quality of a two-component tone, J. Acoust. Soc. Am., **34**, 1752/1760 (1962)

8) R. Plomp & H. J. M. Steeneken: Effect of phase on the timbre of complex tones, J. Acoust. Soc. Am., **46**, 409/421 (1969)

9) C. A. Raiford & E. D. Schubert: Recognition of Phase Changes in Octave Complexes, J. Acoust. Soc. Am., **50**, 559/567 (1971)

10) W. A. Munson & J. E. Karlin: The measurement of human channel transmission characteristics, J. Acoust. Soc. Am., **26**, 542/553 (1954)

11) J. F. Schouten, R. J. Ritsma & B. Lopes Cardozo: Pitch of the residue, J. Acoust. Soc. Am., **34**, 1418/1424 (1962)

12) R. Plomp: Pitch of complex tones, J. Acoust. Soc. Am., **41**, 1526/1533 (1967)

13) G. von Békésy: Experiments in Hearing, McGraw-Hill (1960)

14) G. von Békésy: Sensory Inhibition, Princeton University Press, p. 188 (1967)

15) I. Tasaki: Nerve impulses in individual auditory nerve fibers of guinea pig, J. Neurophysiol., **17**, 97/122 (1954)

16) S. Tachi & T. Isobe: Effect of Phase Relations on the Timbre of Harmonic Tones, Japanese Journal of Medical Electronics and Biological Engineering, **11**, 108/116 (1973) (in Japanese)