Evidence for a general template in central optimal processing for pitch of complex tones

A. Gerson and J. L. Goldstein

Bioengineering Program, School of Engineering, Tel Aviv University, Tel Aviv, Israel
(Received 2 February 1977; revised 9 August 1977)

Optimum prcessor theory sucessfully accounts for earlier pitch data by including the constraint that component tones in a complex stimulus are estimated as successive harmonics. This constraint gives the paradoxical prediction that a periodic complex tone comprising nonsuccessive harmonics cannot evoke periodicity pitch corresponding to its period. Most published data from pitch-shift experiments imply the necessity for this constraint. New periodicity pitch experiments on pitch shift and musical interval recognition were performed which prove that the theoretical constraint is not generally true. New and old data are reconciled by replacing the maximum likelihood estimation of the theory with maximum posterior probability estimation and removing the successive harmonic constraint. Periodicity pitch is estimated by optimizing the match between the aurally measured frequencies of stimulus components and a general harmonic template over some *a priori* expected pitch range. The new, more general, formulation reduces in many experimental situations to the successive harmonic constraint as a special case.

PACS numbers: 43.66.Hg, 43.66.Ba, 43.66.Fe

INTRODUCTION

The optimum processor theory of auditory perception of periodicity pitch accounts comprehensively for pitch perception of complex tones (Goldstein, 1973; Goldstein *et al.*, preceding paper). Basically the theory states that there exists a central optimal processor that operates on a noisy internal measurement of the frequencies of component tones analyzed by both ears to optimally measure the period of the complex acoustical stimulus. The central processor presumes the stimulus is periodic, comprising a set of successive harmonics of some fundamental frequency. The unknown fundamental frequency that is estimated by the processor is the pitch that is attributed to the stimulus.

Although fully agreeing with available data on pitch perception, this formulation reveals a paradox: The pitch attributed to a stimulus comprising successive harmonics of a fundamental frequency should be different from the pitch attributed to the same stimulus after removal of one intermediate component; or, the performance in pitch perception should deteriorate when a noncontiguous harmonic is added to a stimulus comprising successive harmonics (Goldstein, 1973, 1975). Other rules than estimation as successive harmonics are possible in estimating fundamental frequency from a set of components (Goldstein *et al.*, preceding paper); thus no fundamental limitation of optimum processor theory is implied. However, the successive harmonic constraint gives the best quantitative account of available experimental data. The basic issue addressed by this paper is whether stimuli comprising successive harmonics are *always* assumed by the optimum central processor for periodicity pitch. If not, the reason why all previously reported data agree with the constraint of successive harmonic estimates lies in the particular conditions imposed by the design of the earlier experiments, and we must extend the original theory.

The classical pitch-shift experiment (Schouten, 1940; Schouten, Ritsma, and Cardozo, 1962; de Boer, 1956) provided the main empirical basis for the successive harmonic constraint assumed in the original formulation of optimum processor theory. In this experiment subjects were required to match a periodic stimulus comprising successive harmonics to a test stimulus comprising component tones with frequencies that are uniformly shifted from a successive harmonic series, i.e., $nf_0+\Delta$, $(n+1)f_0+\Delta$, $(n+2)f_0+\Delta$, ... The extensive experimental data reported by de Boer (1956) and Schouten, Ritsma, and Cardozo (1962) imply central estimation of the component frequencies as successive harmonics. It was recognized that at a frequency shift equal to half the original fundamental, i.e., $\Delta = f_0/2$, the test stimulus is periodic comprising successive odd harmonics, yet Ritsma (1967) concluded that this stimulus is always heard with pitch values corresponding to successive harmonic estimation of the stimulus frequencies. The major dissenting data were reported by Flanagan and Guttman (1960a, 1960b) who found that in some circumstances, i.e., at fundamentals below 200 Hz, subjects heard a pitch corresponding to the true period of a test stimulus comprising only successive odd harmonics. More recently Patterson (1973) and Patterson and Wightman (1976) reported pitch-shift data which also suggest that stimulus components are not always estimated as successive harmonics.

As the classical pitch-shift experiment was the main basis for the successive harmonic constraint, new experiments of this type were performed. The most objective demonstration of the presence or absence of periodicity pitch has been given by musical interval recognition experiments (Houtsma and Goldstein, 1972); therefore this experiment was used to inquire whether energy at successive harmonics is required in a stimulus to perceive its true periodicity. Our new findings led us to propose a modification of the original formulation. Section I presents the paradigms and results of these new experiments. The formal statement of the extended formulation required for the new experimental data is given in Sec. II. Theory and data are compared

 0001-4966/78/6302-0498$00.80

in Sec. III, while the final Sec. IV provides a summary of key points and implications of optimum processor theory.

I. DESCRIPTION OF EXPERIMENTS

A. Apparatus

All the experiments were designed around a Varian 620-L-100 minicomputer. Dichotic sound stimuli were computed using a real-time procedure,[1] output through a two-channel digital-to-analog interface, then low-pass filtered by a Rockland type 1520 dual analog filter, and dichotically delivered via TDH 39 earphones to the subject sitting in an IAC soundproof room. Communication from the subject to the computer was provided by a set of pushbuttons, whose status could be sensed by the computer. The subject could communicate his responses as well as choose some stimulus parameters in the free-choice experiments. Otherwise, the sequencing of the sounds to be delivered and the changes in the sound parameters from trial to trial were under computer control. Due to the real-time procedure used, the frequencies of the elementary sine waves had to be multiples of 2 Hz and when four components were to be delivered, had an upper limit of 1.66 kHz, as the computation time for a sample of a four-component stimulus is slightly less than 300 μs. Background noise in the stimuli was measured with a Hewlett–Packard 141T spectrum analyzer and found to have a flat spectrum with a total level about 55 dB below the desired signal.

B. Experimental paradigms

Two types of experiments were performed; a free-choice experiment, referred to as the pitch-shift experiment, and a forced-choice experiment, referred to as the musical interval recognition experiment.

1. Pitch-shift experiment

The subject's task in a pitch-shift experiment is to match the pitch of a harmonic *matching stimulus* (sound B) to the pitch of a complex test stimulus (sound A) comprising tones with fixed spacing in frequency. Sound A comprised four tones with equispaced frequencies f_l, $f_l+\Delta$, $f_l+2\Delta$, $f_l+3\Delta$, where Δ was equal to 200 Hz. Experiments were also performed with other spacings, but are not reported here. Sound A can be considered as derived from a harmonic sound by shifting all its components by the same frequency. Depending upon the experiment, harmonic sound B comprised the four frequencies 2, 3, 4, 5 times f_0 or 3, 4, 5, 6 times f_0 or 4, 5, 6, 7 times f_0, with f_0 being integral multiples of 2 Hz. These choices for the sound B show almost no pitch ambiguity, and their pitch will almost always correspond to f_0. It has been found to be more reliable to use a harmonic complex sound as the matching stimulus rather than a sine wave of frequency f_0. The most obvious reason was that, in the latter case, the considerable difference of timbre of sounds A and B added significantly to the difficulty of the experiment; moreover, it is uncertain whether the pitch of the complex sound and the pitch of the pure sine wave are measured by the same processor in all cases.

The experiment is performed as follows. The subject is presented with the sounds A and B of duration half-second each, separated by a half-second silence. The spectral components of these sounds are divided between the two ears so as to minimize the influence of combination tones (components 1 and 3 to one randomly selected ear, and 2 and 4 to the other ear). The subject then has control over the fundamental frequency f_0 of the matching sound B, and his task is to adjust it until he judges that the test and matching stimuli, A and B, have equal or nearly equal pitch. Obviously, in order to do so, the subject may listen to the sequence A–B as many times as needed until a pitch match is achieved. When this is done, the computer goes to the next trial by modifying the test sound A through choosing a new lowest frequency f_l differing from its previous value by a random multiple of 2 Hz (random walk process). All frequencies in sound A are shifted by the same random amount, maintaining the difference of 200 Hz between frequencies.

2. Musical interval recognition experiment

The subject's task in a musical interval recognition experiment is to report the musical interval formed by two successive complex tones. Two harmonic sounds A and B were used, the nominal fundamental frequency of A being always 98 Hz (G), while the nominal fundamental frequency of B being chosen among the four values 87.3 Hz (F), 92.5 Hz (F#), 103.8 Hz (G#), and 110.0 Hz (A). These fundamentals are nominal because the components actually presented were the closest 2-Hz multiples to the nominal harmonics. The four possible intervals were thus an ascending or descending semitone or tone. Two different experiments were done, one using as components four successive harmonics of the fundamental frequency, the other using four successive *odd* harmonics of the fundamental frequency. The harmonic numbers were randomly chosen among the sets (4, 5, 6, 7), (5, 6, 7, 8), (6, 7, 8, 9), and (7, 8, 9, 10) in the first case, among the sets (1, 3, 5, 7), (3, 5, 7, 9), (5, 7, 9, 11), and (7, 9, 11, 13) in the second, the average harmonic number being 7 in both cases. These random selections were made independently for the first and second sounds.

The experiment is performed as follows. The subject is presented with the sounds A and B of duration half-second each, separated by a half-second of silence. Here again, the four components are divided dichotically between the two ears, the first and third being directed to one randomly selected ear, the second and fourth to the other ear. Each session begins with a training phase, in which the subject requests a typical stimulus corresponding to each of the four possible intervals. When the subject feels ready, the experiment begins, in which an interval is randomly chosen, a stimulus corresponding to this interval is delivered, and the subject must report the interval he heard. He is not given any feedback or opportunity to correct himself; after having recorded the answer, the computer selects and immediately delivers a new musical interval.

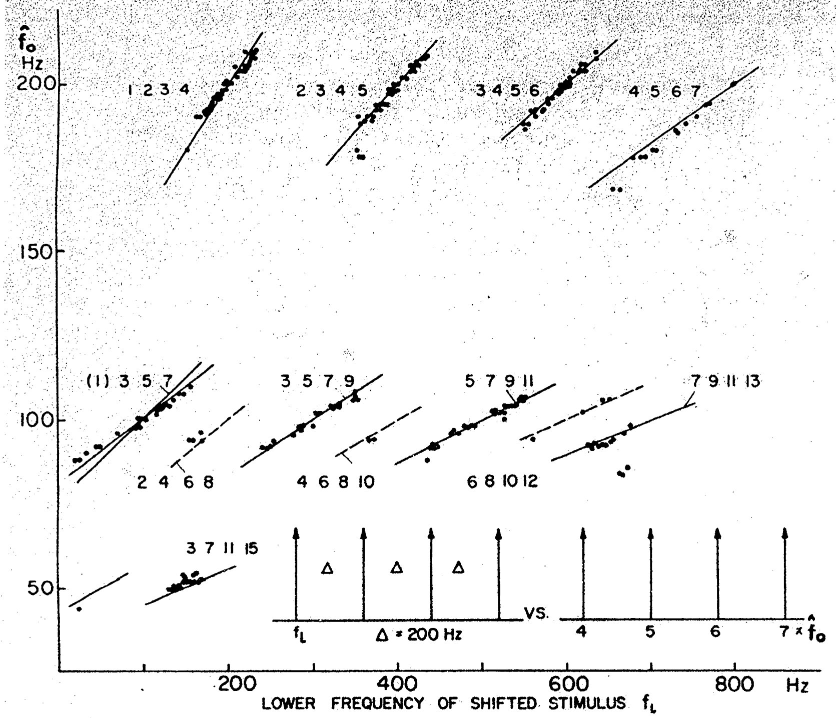

FIG. 1. Experimental results of the pitch-shift experiment for subject AG. The subject adjusted the fundamental frequency f_0 of the harmonic sound to match the pitch of the "shifted" sound. The lower frequency f_l of the shifted sound was chosen in a random sequence. The value $\hat{f}_0$ of f_0 at which pitch match is achieved is plotted against f_l for each experimental point. The lines describe the different theoretical modes of pitch, where for each mode the stimulus components are estimated as having the harmonic ranks indicated. The stimulus was presented dichotically, two components per ear, to minimize combination tones.

C. Results

1. Pitch shift experiment

Systematic data were obtained with two musically skilled subjects, AG and YR, and a more naive subject, EP. Subject AG has the most extensive musical background including conducting experience. The intensity of the sounds was moderate, being 40 dB SPL per component for subjects EP and YR, 50 dB per component to the left ear of subject AG and 35 dB per component to his right ear, subject AG having a known loss of 10–15 dB of known middle-ear origin at the left ear.

Figure 1 presents a typical outcome of the pitch-shift experiment for subject AG. The fundamental frequency $\hat{f}_0$ of the harmonic reference sound B, which matches the pitch of the shifted sound A, has been plotted against the frequency f_l of the lowest component in the sound A. It is easily seen that all the experimental points are grouped along sets of lines. When the frequency f_l is a multiple of $\Delta = 200$ Hz (i.e., when the sound A comprises successive harmonics of a fundamental frequency of 200 Hz), the pitch of A corresponds usually to the fundamental frequency 200 Hz. The upper set of lines can therefore be attributed to the estimation of the complex stimulus (sound A) by a series of successive harmonics, and is similar to numerous previously reported outcomes of pitch-shift experiments. (See the Introduction.)

The important new systematic finding is the additional sets of lines in Fig. 1 that do not correspond to central estimation by a series of successive harmonics. It is evident that when the lowest frequency f_l of the test sound A is an odd multiple of 100 Hz, the pitch heard corresponds usually to an $\hat{f}_0$ of 100 Hz. Therefore, it is reasonable to attribute the set of lines including the $\hat{f}_0 = 100$ Hz matches to the estimation of the complex stimulus by a series of successive odd harmonics. Similarly, the lowest group of points, being located around an f_l of 150 Hz and a $\hat{f}_0$ of 50 Hz, could reasonably be attributed to the estimation of the complex sound by a 3, 7, 11, 15 harmonic series. The remaining points obtained in this experiment, lying between the successive odd harmonic lines, can be seen to be exactly one octave (vertically) lower than points corresponding to the same stimulus (same f_l) on the successive harmonic lines. These points correspond to estimation by a series of successive even harmonics (e.g., 6, 8, 10, 12).

Figure 2 presents the outcome for a similar pitch-shift experiment with subject YR. Though fewer data points were collected, subject YR always performed as did subject AG, making pitch matches with $\hat{f}_0$ in the neighborhoods of both 100 and 200 Hz.

Figure 3 reports the result of the same pitch-shift experiment, performed by subject EP. These results seem to contradict the previous results, for no lines corresponding to nonsuccessive harmonic estimates of the test sound were obtained. In fact, when the lowest frequency f_l was around 500 Hz, so that the stimulus could be interpreted as a successive odd harmonics series (5, 7, 9, 11 times f_0), no such match was made. But neither was there a match to a successive harmonic series (2, 3, 4, 5 or 3, 4, 5, 6 times f_0). A close examination of these "mismatched" data points, circled in the figure, shows that their attributed pitch is almost exactly to one-half of the lowest frequency f_l. Since the reference sound B comprised the harmonic series 2, 3, 4, 5 times f_0, subject EP seemed to have matched the frequencies of the lowest components in both sounds A and

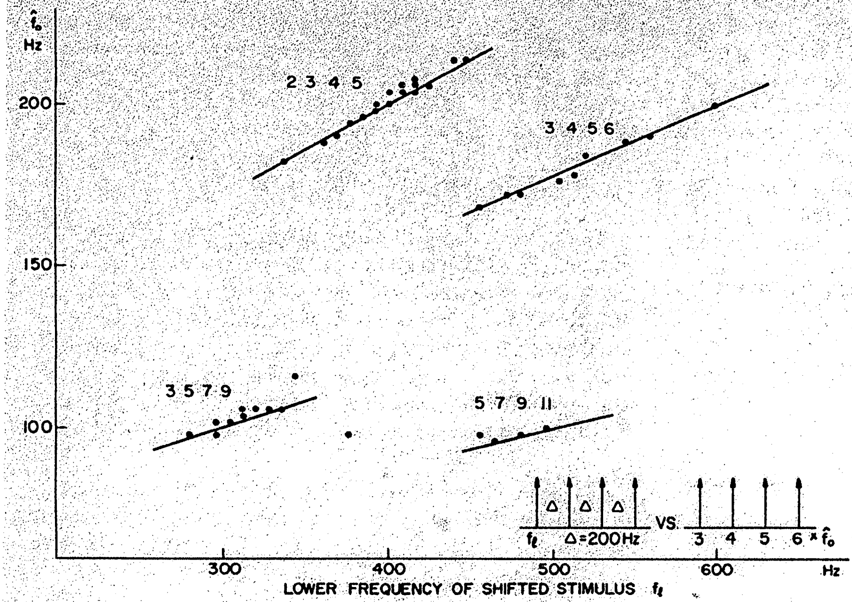

FIG. 2. Experimental results of the pitch-shift experiment for subject YR. These results are comparable with subject AG's data (Fig. 1).

B. Thus instead of measuring their periodicity pitches, the subject switched to an "analytic" mode of perception of the complex sound. (Pitch of complex sounds comes from a "synthetic" mode of perception; both these modes are well known to musicians since Helmholtz articulated the concepts.)

2. *Musical interval recognition experiment*

In the pitch-shift experiment subject EP was unable to hear the periodicity pitch corresponding to the true period of a stimulus comprising only odd harmonics. The results of the experiments on musical interval recognition (with only harmonic sounds), described below, indicate that this inability is not generally true for subject EP.

In the musical interval recognition experiment with successive harmonic stimuli, subject AG scored 99.7% correct answers out of 340 trials, subject YR scored 99.5% correct answers out of 200 trials and subject EP scored 88.4% correct answers out of 1128 trials. In the case of the successive odd harmonic stimuli, the results were 99.7% correct out of 331 trials for subject AG, 97% correct out of 200 trials for subject YR, and 85.5% correct out of 1156 trials for subject EP. If the pitch processor estimated all stimulus components as successive harmonics, then the percentage of correct

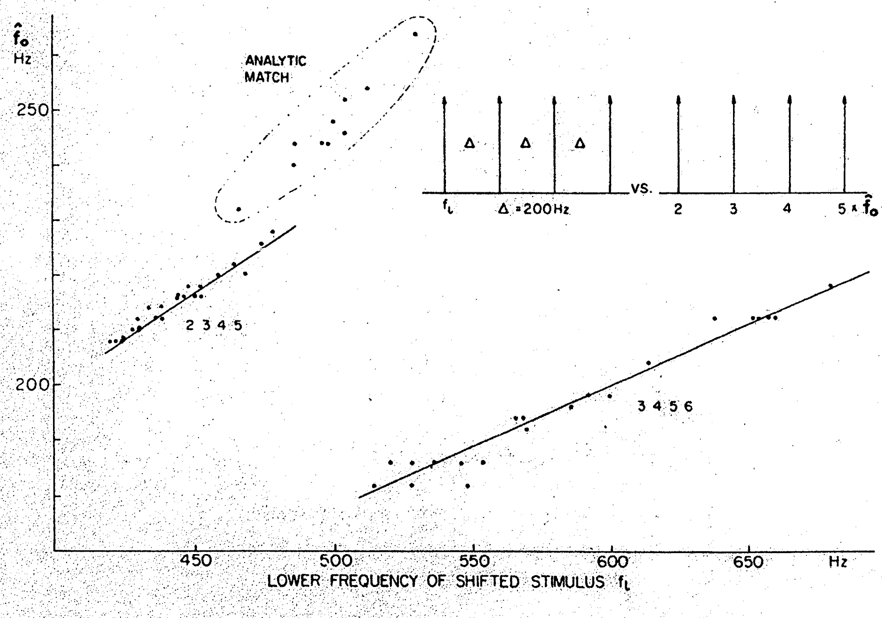

FIG. 3. Experimental results of the pitch-shift experiment for subject EP. In contrast with subject AG, no pitch corresponding to 100 Hz was heard for f_l near 500 Hz. Instead, subject EP switched to an analytic mode giving the circled data.

answers for successive odd harmonic stimuli would not be greater than 49% in this experiment.[2] The performance of subject EP seems to indicate clearly that, in this experiment, she was perfectly able to measure pitch on the basis of a general harmonic series estimate, although she would not do so in the pitch-shift experiment. The basis of this discrepancy will be discussed in Sec. III A. Next we consider how the optimum processor model can be modified to include the new findings.

II. MODIFICATION OF OPTIMUM PROCESSOR THEORY

A. Maximum posterior probability estimation

Earlier pitch-shift experiments (de Boer, 1956; Schouten, Ritsma, and Cardozo, 1962) suggested that the auditory processor of periodicity pitch presumes that all complex tone stimuli are comprised of successive harmonics (see Ritsma, 1967; Goldstein, 1973, Sec. IX, constraint number 4). This implies that in every case a periodic stimulus comprising nonsuccessive harmonics should be heard with a pitch not corresponding to the true period. Our new experimental results prove that the successive harmonic constraint assumed in the original development of optimum processor theory is not generally valid. Depending upon the context provided by the experimental paradigm, the same subject EP either succeeded or failed to hear the pitch corresponding to the true period of an odd harmonic stimulus. Furthermore, in the same experimental pitch-shift paradigm, subject EP failed while subjects AG and YR succeeded in hearing the true periodicity pitch of odd harmonic stimuli.

It was parsimonious in accounting for earlier data to ignore prior probabilities for the stimulus conditions and model the estimation process as maximum likelihood (i.e., Bayesian) with a successive harmonic constraint. The successive harmonic constraint can easily be discarded in the maximum likelihood estimation procedure, and pitch would correspond to that fundamental which together with its 14 or 15 overtones (see Goldstein *et al.*, preceding paper, Sec. III) best matches the aural frequency measurements. An obvious problem here is that we now cannot account for the different performances by subject EP with odd harmonic stimuli in different experimental paradigms. A more subtle difficulty is that the accuracy of earlier theoretical accounts (Goldstein, 1973; Goldstein *et al.*, preceding paper) of musical interval recognition data is poorer without the successive harmonic constraint. It appears as though the successive harmonic constraint is true under certain conditions.

The foregoing considerations lead to a natural generalization of the optimum estimation operation performed by the central processor as maximum posterior probability. This form of estimation provides the maximum number of correct estimates on the average and is better than maximum-likelihood estimation if the prior probabilities of the possible pitch values are not all equal (van Trees, 1968). We will show that we can accommodate the new experimental results by introducing the concept of prior probability into optimum processor theory. Let $\{X_k\}$ be the set of aurally measured component frequencies, and $\hat{f}_0$ be the optimum estimate of the reference fundamental, which we define as "periodicity pitch." Then the posterior and prior probabilities are related according to Eq. (1).

$$P(\hat{f}_0|\{X_k\}) = P(\{X_k\}|\hat{f}_0)\cdot P(f_0)/P(\{X_k\}) . \tag{1}$$

The probabilities in order of appearance from left to right in Eq. (1) are known as the posterior probability, likelihood function, prior probability, and signal probability. The optimum estimated periodicity pitch $\hat{f}_0$ maximizes the posterior probability. If the prior probability is uniform, then maximizing the posterior probability or the likelihood function in estimating periodicity pitch are equivalent procedures.

The form of prior probability that we find adequate to account for the new as well as older data is a rectangular distribution for periodicity pitch. Prior expectations, as established by the conditions of the experiment and the particular subject, are presumed to determine the upper and lower bound of the rectangular distribution. Because the prior probability distribution is rectangular, a maximum-likelihood estimation procedure would again be equivalent to maximum posterior probability, provided the likelihood function is evaluated only within the bounded pitch range.

The mathematical formulation of the estimation procedure proposed here is formally identical to the original presentation by Goldstein (1973), except that the successive harmonic constraint is removed. The central processor receives a set of component frequency estimates $\{x_k\}$, and assumes they are independent Gaussian deviates from $\{n_k f_0\}$. The likelihood of the assumed harmonic stimulus is then

$$L = \prod_k \frac{1}{\sigma_k\sqrt{2\pi}} \exp\left[-\frac{(x_k - n_k f_0)^2}{2\sigma_k^2}\right], \tag{2}$$

where $\sigma_k = \sigma(n_k f_0)$ is the value, known to the processor, of the standard deviation of the measurement of a stimulus component of frequency $n_k f_0$. The estimated pitch $\hat{f}_0$ and the estimated harmonic numbers $\hat{n}_k$ are those that maximize L, or equivalently, its logarithm

$$\Lambda = -\sum_k \ln(\sigma_k\sqrt{2\pi}) - \frac{1}{2}\sum_k\left(\frac{x_k - n_k f_0}{\sigma_k}\right)^2 . \tag{3}$$

An adequate procedure is to assume that the values σ_k are locally constant when differentiating Eq. (3) in search of a maximum for Λ. This is equivalent to saying that $\sigma_k = \sigma(\hat{n}_k \hat{f}_0)$ does not differ significantly from $\sigma(x_k)$, the value of the standard deviation at the received frequency rather than at the estimated harmonic. In this case, maximizing Eq. (3) amounts to a least-squares procedure, where one looks for the minimum of

$$\epsilon^2 = \sum_k \left(\frac{x_k - n_k f_0}{\sigma_k}\right)^2 , \tag{4}$$

where the σ_k are now known values.

Given the harmonic numbers $\{\hat{n}_k\}$ the solution for the pitch $\hat{f}_0$ is

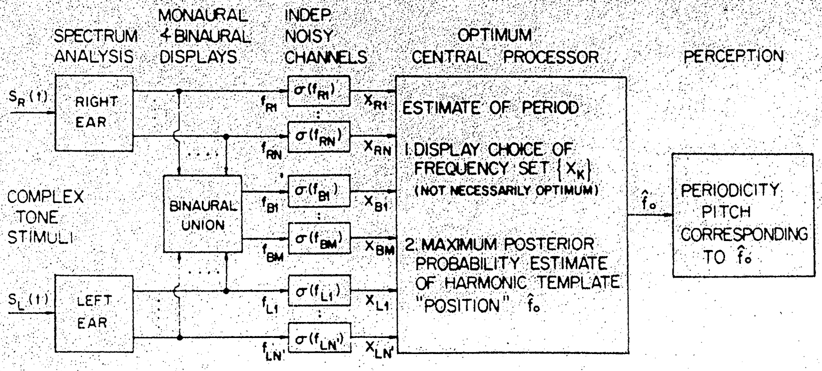

FIG. 4. Modified optimum processor theory, in which maximum posterior probability estimation with a general harmonic frequency template replaces maximum-likelihood estimation with the successive harmonic constraint (see text). The explicit schematic representation of the display options to the central processor has been added for clarity; it is not a change in the theory.

$$\hat{f}_0 = \sum_k \frac{\hat{n}_k x_k}{\sigma_k^2} \Big/ \sum_k \frac{\hat{n}_k^2}{\sigma_k^2} \quad . \tag{5}$$

The value ϵ^2 of the least squares error for this value of the pitch $\hat{f}_0$ is then

$$\epsilon^2 = \sum_k \frac{x_k^2}{\sigma_k^2} - \left[\left(\sum_k \frac{\hat{n}_k x_k}{\sigma_k^2} \right)^2 \Big/ \sum_k \frac{\hat{n}_k^2}{\sigma_k^2} \right] \tag{6}$$

or

$$\epsilon^2 = \left(\sum_k \frac{x_k^2}{\sigma_k^2} \right) \cdot (1 - \rho^2) \, ,$$

where the correlation coefficient is

$$\rho = \sum_k \frac{\hat{n}_k x_k}{\sigma_k^2} \Big/ \left[\left(\sum_k \frac{\hat{n}_k^2}{\sigma_k^2} \right) \left(\sum_k \frac{x_k^2}{\sigma_k^2} \right) \right]^{1/2} \, .$$

Since the harmonic numbers $\hat{n}_k$ must also be estimated the processor computes the value ϵ^2 and looks for the minimum over all possible combinations of integers n_k. This is equivalent to searching for the set of integers $\{\hat{n}_k\}$ that has the greatest correlation with the received frequencies $\{x_k\}$. This set of integers is then used in Eq. (5) to obtain the estimated pitch $\hat{f}_0$ of the acoustical stimulus, where the pitch must lie within some bounded range, $f_{0\,\min} < \hat{f}_0 < f_{0\,\max}$.

A useful conceptualization is to consider the estimation procedure as an harmonic template matching (Goldstein *et al.*, preceding paper, Sec. I, Fig. 3). The component frequencies $\{x_k\}$ received by the optimum processor are plotted along a logarithmic axis, providing the rank ordering $x_{k+1} > x_k$. The processor scans the component frequency axis with an harmonic template, including up to the 15th harmonic, and locates the best matching template position within some bounded range. The criterion for best fit is minimum error in Eq. (4), i.e.,

$$\epsilon_{\min}^2 = \sum_k \left(\frac{x_k - \hat{n}_k \hat{f}_0}{\sigma_k} \right)^2, \quad f_{0\,\min} \leq \hat{f}_0 < f_{0\,\max} \, .$$

We no longer require successive harmonics in the template match. Thus, in terms of the ranking defined above, we replace the requirement $\hat{n}_{k+1} = \hat{n}_k + 1$ with the looser ranking condition $\hat{n}_{k+1} \geq \hat{n}_k + 1$.[3] For a given template "position" corresponding to the reference fundamental f_0, $\hat{n}_k f_0$ is the nearest template harmonic either above or below x_k. From Eq. (6) we observe also that the best-fitting position maximizes the correlation between the aurally measured component frequencies $\{x_k\}$ and the template harmonics $\{\hat{n}_k \hat{f}_0\}$. The modified optimum processor model is shown schematically in Fig. 4. In addition to the introduction of maximum posterior estimation, we explicitly show that the central processor has various monaural and binaural display options in collecting component frequencies for periodicity pitch. The availability of these options is not a new finding, but the rules for their choice by the central processor is a subject requiring systematic study.

B. Probability distribution of the pitch of a complex stimulus

Due to the randomness in transmission of the component frequencies to the central processor, the pitch attributed to a given stimulus will vary in some random fashion. Disregarding for the moment the effect of the pitch range limitation, the Gaussian variations of the transmitted frequencies may yield "best matches" corresponding to different sets of estimated harmonic numbers $\{\hat{n}_k\}$, i.e., different placements of the template along the transmitted frequencies, and hence very different pitches. This is the known basis for the ambiguity of pitch perception. In general, the different modes of the multimodal distribution of the pitch of a given stimulus correspond to different estimates of the set of harmonic numbers $\{\hat{n}_k\}$, and the scatter within each mode (truncated Gaussian distributions) is due to variations in the transmitted frequencies small enough to insure that the best match will still be achieved by the same set of harmonic numbers $\{\hat{n}_k\}$. To accommodate the new experimental facts presented above (Sec. I), the present formulation of the pitch processor does not restrict $\{\hat{n}_k\}$ to successive integers. This modification increases the number of modes, and increases the complexity of the computation of the probability distribution for the pitch of a given stimulus. Even the probabilities of falling within a given mode are analytically untractable. Hence, computer simulation is the only way to evaluate these probabilities. Numerical calculations of the pitch estimate $\hat{f}_0$ were made for many sets of randomly chosen sample frequencies $\{x_k\}$.

Figure 5 shows the outcome of such a simulation for the stimuli used in the pitch-shift experiment (Fig. 1). The four-component "stimuli" are f_l, $f_l + 200$, $f_l + 400$,

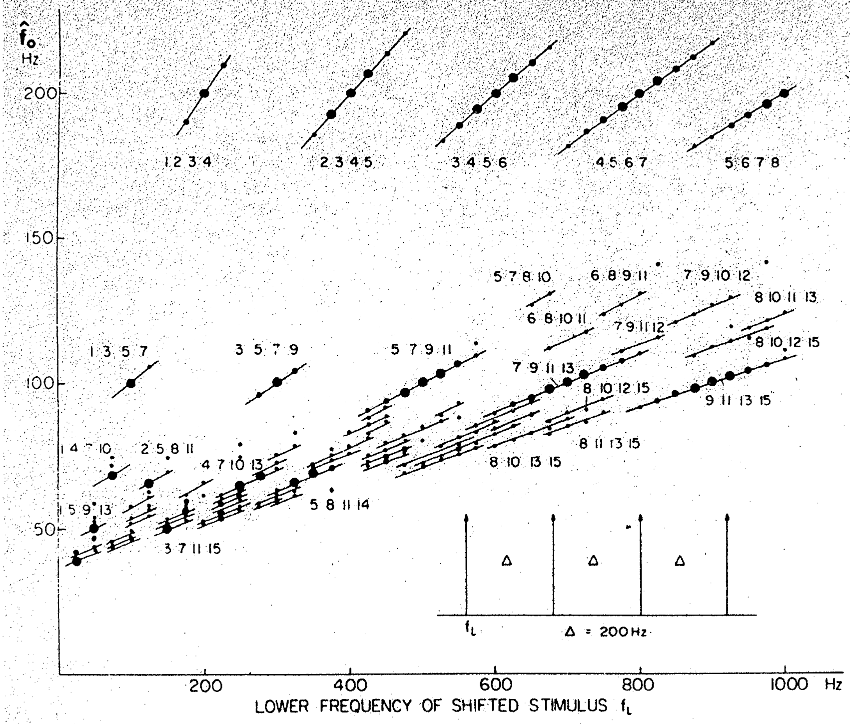

FIG. 5. Theoretical prediction for the pitch-shift experiment obtained by simulation. The measured probabilities of the different modes are schematically represented by the size of the dots. The estimated harmonic numbers are shown near most modes.

and f_l+600 Hz, for all f_l from 25 to 1000 Hz by steps of 25 Hz. For all values of f_l (horizontal axis), the mean estimate of $\hat{f}_0$ for each mode obtained in the simulation (200 trials per "stimulus") is plotted on the vertical axis. The scatter within each mode is too small to be represented in the scale of the figure. The size of each dot is roughly proportional to the probability of occurrence of the corresponding mode. The upper limit of the possible harmonic numbers $\hat{n}_k$ was taken as 15. No range restriction for pitch was introduced in the simulation. Probability distributions for a restricted pitch range can then be readily calculated as a conditional probability on a subspace of the unrestricted range.

The sloping lines in Fig. 5 show how the mean pitch of a given mode changes as the stimulus is changed. The set of estimated harmonic numbers $\{\hat{n}_k\}$ for each of these lines is also shown in Fig. 5. The slope of each line can be obtained as follows: If all transmitted frequencies x_k undergo a uniform shift Δx, the pitch $\hat{f}_0$ undergoes a shift Δf_0 that is easily shown to be

$$\Delta f_0 = \Delta x \sum_k \frac{\hat{n}_k}{\sigma_k^2} \Big/ \sum_k \frac{\hat{n}_k^2}{\sigma_k^2} \quad . \tag{7}$$

In a region where $\sigma(f)$ is roughly proportional to $\sqrt{f}$, such as the region involved in the displayed simulation, this amounts to

$$\Delta f_0 = \Delta x \left(N \Big/ \sum_k \hat{n}_k \right) , \tag{8}$$

where N is the number of resolved components of the stimulus. The reciprocal of the slope of the pitch-shift line is the arithmetic average of the estimated harmonic numbers that characterize this line. This relationship (8) can be easily verified in Fig. 5, at least for the two upper sets of lines corresponding to estimates of the harmonic numbers $\hat{n}_k$ that are successive or successive odd integers. It should be noted in Fig. 5 that nearly all the probability is concentrated on lines where estimates of the harmonic numbers are equispaced, as were the components of the "stimuli" used in the simulation. The shorter lines of low probability all correspond to nonequispaced sets of estimated harmonic numbers.

It is instructive to consider an example and examine in numerical detail the effects of stochastic variability and restricted pitch range on the estimates of pitch $\hat{f}_0$. The effect of stochastic variability on the estimation process of choosing the best template position is illustrated in Fig. 6. The example stimulus comprises the two frequencies 1400 and 1800 Hz, and each panel corresponds to a different random set of aurally transmitted frequencies. The curves show the value of the template error ϵ^2 as a function of template position f_0 as given by Eq. (4). In computing the template error for a given position f_0, each sample component frequency x_k is associated with a different nearest neighboring template component, the criterion for nearest neighbor being minimum template error. Each valley in the error curves is a local minimum, corresponding to a given set of harmonic numbers. The pitch $\hat{f}_0$ is the value of f_0 corresponding to the deepest valley. It can be seen that the valleys have almost the same abscissas in the two panels, but that their relative depths vary greatly between panels, giving rise to the modal structure in the distribution of the random estimates of $\hat{f}_0$ and $\{\hat{n}_k\}$.

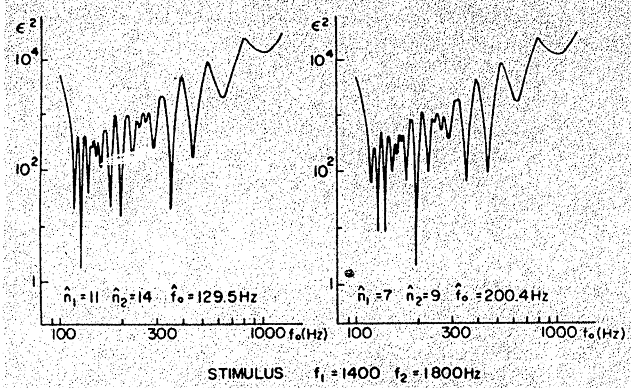

FIG. 6. Squared error of the template match [Eq. (4)] versus template "position" f_0 for two different random sample measurements of the stimulus frequencies. The estimated periodicity pitch $\hat{f}_0$ is given by the template position for minimum error; the corresponding estimates of harmonic number are also indicated.

The effect of the restriction of the range of possible pitches on the probability distribution of the estimated pitch corresponding to the same stimulus (1400 and 1800 Hz) is shown in Fig. 7, where for each panel, the probability of estimating a pitch f_0 has been plotted versus $\hat{f}_0$. The result of the simulation, done with 5000 trials, with no restriction on the range of the pitch, is drawn in the upper-left panel, showing five significant modes with various probabilities. The other panels show the probability distribution of the possible modes, for various ranges of allowed pitch; they comprise conditional probabilities computed from the first panel. It can be noticed that, when the pitch range is one octave around 200 Hz (the true periodicity) then the correct answer is given all the time, and if no pitch is expected below 300 Hz, the matches will be made to successive harmonic patterns only.

Conditional probability calculations provide an analytic approach for accommodating experimental observations (de Boer, 1956; Houtsma and Goldstein, 1971), that although the task of a pitch experiment can be performed perfectly with stimuli comprising few harmonics, the task is made easier by adding harmonics to the stimuli. Adding harmonics generally can provide a more salient pitch. In a given experiment in which the subject uses a restricted pitch range, the increase in saliency or pitch "strength" need not affect the measured behavior. On the other hand, the theory predicts that the probability of a given pitch mode in the unrestricted distribution can be significantly increased by adding harmonics. Thus the unrestricted distribution is useful for characterizing pitch "strength."

III. THEORY AND DATA

A. Pitch modes and expected pitch range

The pitch-shift lines predicted by the theory (Fig. 5) are plotted in Figs. 1 and 2 together with the experimental results of the pitch-shift experiment performed by subjects AG and YR. It can be seen in Fig. 1 that almost all the high-probability modes predicted by theory have been found as experimental data, and that the position and slopes of the experimental lines are in close agreement with the theory. The deviation from the

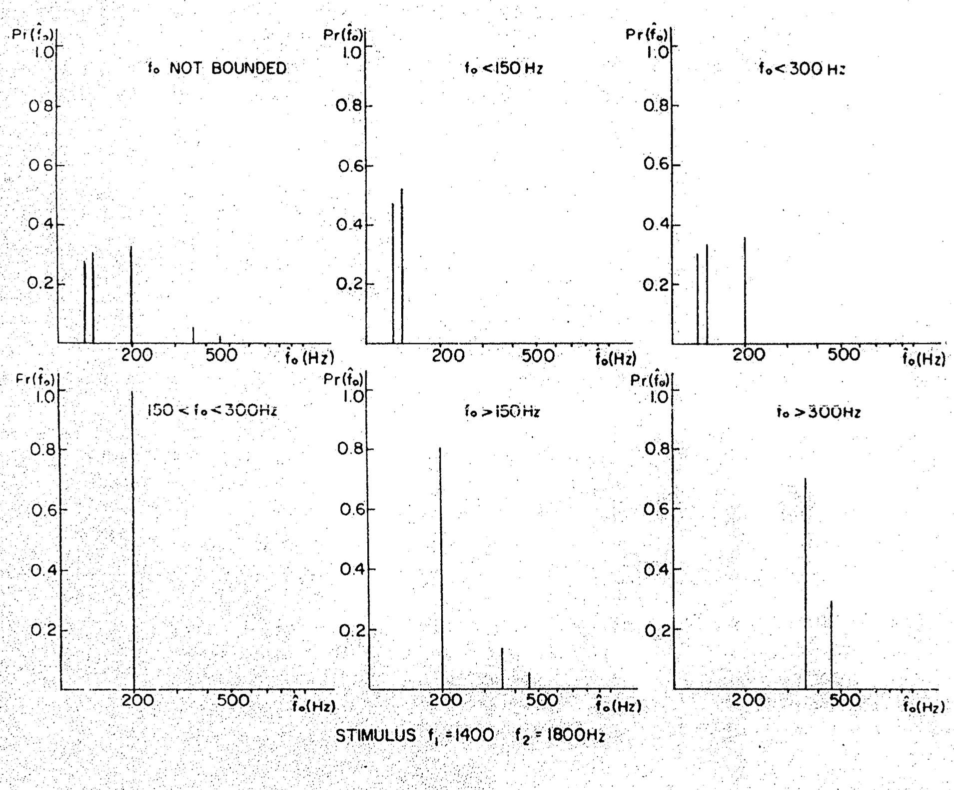

FIG. 7. Discrete probability distribution for estimated periodicity pitch of the stimulus in Fig. 6. Restriction of the template scan range produces distributions that are subsets of the unrestricted distribution (upper left). The probabilities of the modes for the unrestricted distribution provide a measure of the "strength" of the different modes. The means of the estimated pitch $\hat{f}_0$ for each mode and the corresponding harmonic number estimates $\hat{n}_1$ and $\hat{n}_2$ are, respectively, 128 Hz, 11 and 14; 139 Hz, 10 and 13; 200 Hz, 7 and 9; 355 Hz, 4 and 5; 457 Hz, 3 and 4.

slope of the (1, 3, 5, 7) mode (steeper theoretical line) can be explained by the fact that the first component of the stimulus was probably below the auditory threshold; its intensity was only 35 dB SPL. The effective stimulus would therefore comprise only the three upper components. The shallower line, in much better agreement with the data, is the corrected theoretical prediction for a three-component stimulus (3, 5, 7). As there are generally no important differences between theoretically expected pitch lines and best linear fits to data, we do not pursue this comparison. It should be recalled that combination tones were minimized by the dichotic, low-level stimuli.

Low-probability modes of special interest in Fig. 1 are the data corresponding to successive even harmonic estimates of the stimulus components. From the theoretical viewpoint, once the successive harmonic constraint is removed we could expect these modes to occur with high probability. The square error ϵ^2 [Eq. (4)] has the same value if all estimated harmonic numbers $\hat{n}_k$ are replaced by $l\hat{n}_k$ and the pitch $\hat{f}_0$ is replaced by $\hat{f}_0/l$, l being an integer and $l\hat{n}_k$ being no greater than 15. It is highly significant that such subharmonic matches occur infrequently in the experimental data.

We claim that it is in accord with the data that subharmonic pitch matches are avoided in the normal operation of the pitch processor. Consider the three following possible rules in the operation of the pitch processor:

(a) "Subharmonic" matches (i.e., $l > 1$) are strictly forbidden.

(b) "Subharmonic" matches are exactly as likely as "normal" matches ($l = 1$), since ϵ^2 is the same in all cases, and the probability obtained in the simulation has to be distributed amongst all possible subharmonic modes.

(c) "Subharmonic" matches may occur, but only if the "normal" match is outside the expected range of pitches; if not, the "normal" match will be preferred. Thus in general, if two allowed pitch estimates yield the same square error ϵ^2, the processor prefers the higher pitch (in which the estimated harmonic numbers are smallest).

The fact that data corresponding to an "octave ambiguity" ($l = 2$) has been found eliminates rule (a). The fact that most of the data gathers around "normal" matches excludes also rule (b). The last possibility seems therefore the most realistic rule. The few "octave ambiguity" matches that have been obtained can be explained by a restriction in the expected pitch range that occurs naturally in the context of matching the pitches of sounds A and B in the pitch-shift experiment.

This context effect may be understood as follows. In performing the matching task in the pitch-shift experiment the subject initially should have an unrestricted prior expected range for pitch. After the initial estimate of the test sound A, the subject restricts the pitch range to the neighborhood of the initial estimate. In subsequent presentations of sounds A and B, the subject adjusts the fundamental of the harmonic matching sound B to match the pitches of both sounds. The fact that subjects AG and YR were not hopelessly confused with widely varying pitch estimates from successive presentations of sound A is evidence for this pitch focusing ability.

Different abilities of subjects to focus their expected pitch range (i.e., template scan range) appear to underlie the apparently contradictory results obtained from the pitch-shift and musical interval experiments. It was found that subject EP failed in the pitch-shift experiment to correctly estimate as odd harmonic a stimulus that actually comprised successive odd harmonics. On the other hand, in the context of a musical interval recognition experiment performed with sounds built up of successive odd harmonics, the same subject reaches scores unattainable if the processor had a set of successive harmonics as the only possibility for estimating the stimulus. This suggests that the general template matching processor exists for all subjects, but that the design of the pitch-shift experiment prevents some subjects from using it in the "best" way possible. In this way we can understand the conflicting data of Ritsma (1967) versus Flanagan and Guttman (1960a, 1960b), mentioned in the Introduction, for pitch of odd harmonic stimuli. The difference between our two experiments is that in musical interval recognition the subject is given a pitch reference (the first note of the interval) around which to look for the pitch of the second note, while such a reference is missing in the pitch-shift experiment. In other words, the design of the musical interval recognition experiment contains in itself an instruction to the subject to look for a pitch within a restricted range (e.g., one tone above or below the reference note). On the other hand, the random sequence of stimuli given in the pitch-shift experiment leads to unpredictable pitch jumps, that can be greater than one octave.

It seems therefore that the "natural" operation of the pitch processor includes a restriction in the range of expected pitches, the possibility of flexibly extending this range being only reached after considerable training (facilitated by previous musical experience or skill). This corroborates well with centuries of musical history, where one of the basic characterizations of melody is precisely "to be built of small intervals," and the purpose of pitch measurement is (in the musical sphere at least) nothing but an access to melody. Thus a context effect is normal in pitch measurement, the pitch of a note being expected in a close range around a just previously heard note, or sequence of notes. This effect is probably what prevents most subjects from jumping down to odd harmonic estimates of the stimulus when focused on previous higher-frequency pitches. The successive harmonic constraint implied by earlier data can now be understood as being derived from more general constraints and the particular expected pitch range of the experiment and subject.

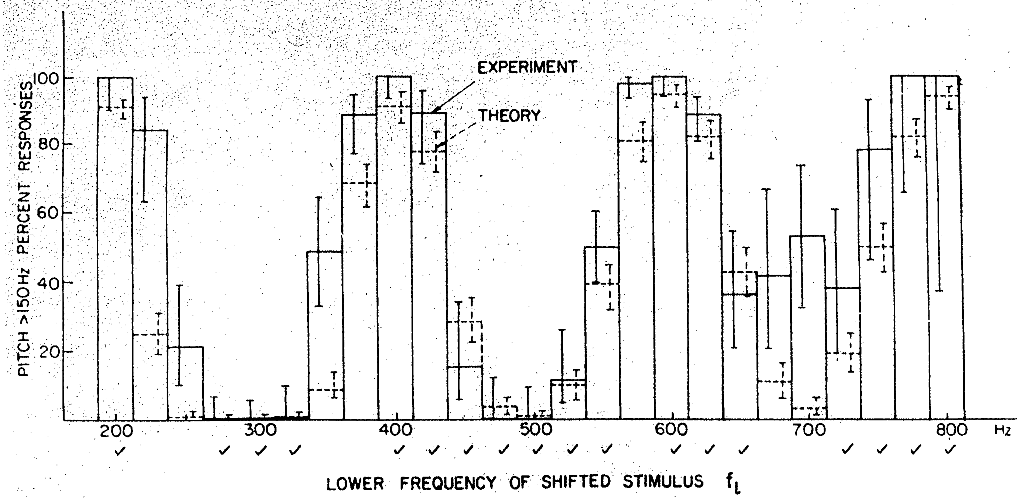

FIG. 8. Percent of high-pitch judgments (>150 Hz) as a function of stimulus in the pitch-shift experiment. Data produced by subject AG (778 points, similar to the 230 points shown in Fig. 1) are compared with the simulated prediction of optimum processor theory. The vertical lines within each histogram bin represent the 95% confidence intervals for experiment and theory. Agreement within the confidence intervals exists for 18 out of 25 bins (indicated by checks below the abscissa).

B. Probabilistic aspects

Given the standard error $\sigma(f)/f$ for aural measurement of component frequencies, and the prior pitch range, the theory predicts the probabilities that a given complex tone stimulus will evoke a pitch from each of several modes (Sec. II B). The multimodal outcome of the pitch-shift experiment for subject AG (Fig. 1) provides an important opportunity for testing some probabilistic predictions of the theory (Fig. 5). Because several repetitions of the test sound A were presented to enable the subject to adjust sound B for a pitch match, the pitch-shift data do not necessarily give a direct measure of the probabilities of each of the possible modes.

We have argued that subject AG makes his initial pitch estimate for a given test sound on the basis of an unrestricted expected pitch range. For subsequent presentations of the same test sound he focuses his expected pitch range to some neighborhood of the initial estimate. If the subject focuses in a very small neighborhood of the initial estimate, then the pitch matches should distribute among the possible modes in direct proportion with the theoretical probabilities. However, the absence of low-probability modes in the experimental data suggests that the focusing was not very fine, thereby causing the pitch matches to be biased toward the high-probability modes lying in the restricted pitch range. On the basis of the clustering of the experimental and theoretical pitch estimates (Figs. 1 and 5) above 170 Hz and below 120 Hz, it is reasonable to assume that the pitch-matching experiment caused subject AG to partition his responses on the basis of high and low expected pitch ranges. Thus it would be meaningful to compare theory and experiment on the basis of the probability of a high pitch for a given stimulus, where for the particular stimuli of the experiment, pitch estimates above 150 Hz would be considered high.

Figure 8 compares experimental and theoretical probabilities for a high pitch in the pitch-shift experiment performed by subject AG. Data obtained from four experimental sessions were similar to the typical case shown in Fig. 1; therefore, the data were pooled to give a total of 778 data points. The data points were grouped, according to the frequency f_l of the lowest component of the shifted stimulus, into frequency bins of width 25 Hz centered on multiples of 25 Hz from 200 to 800 Hz. For each frequency bin, the proportion of responses whose pitch was higher than 150 Hz was computed. Pitch matches above 150 Hz in this experiment corresponds to successive harmonic estimates of the stimulus components. The theoretical probability of having a pitch higher than 150 Hz for stimuli corresponding to the same frequency band was calculated by parabolic interpolation and integration of the results of the simulation described in Fig. 5. A typical standard error function was assumed in the simulation (idealized curve in Fig. 4 of Goldstein *et al.*, preceding paper). Confidence intervals stemming from the limited number of samples in the experimental data and theoretical simulation are also given in Fig. 8.

On comparing theory and experiment in Fig. 8, it can be seen that the basic prediction of the theory is confirmed. In particular, 18 out of 25 frequency bins (indicated by checks below the axis) show agreement within the confidence intervals between theory and experiment. For bins with nonoverlapping confidence intervals, the experimental probability for a high pitch is always greater than the theoretical value. This could be the result of differences between the idealized standard error function $\sigma(f)/f$ used in the simulation and the exact function (not systematically measured) for subject AG, or possibly subject AG may have occasionally restricted the expected pitch range for his initial pitch estimate to the high range.

TABLE I. Percent-correct response in musical interval recognition experiment.

	Experimental data			Theory (simulation)		
Stimulus	Subject	% corr.	95% confidence	f_0 range (Hz)	% corr.	95% confidence
Successive	EP	85.5	83.4 – 87.6	OPEN	88.25	85.4 – 91.5
odd	YR	97.0	94.2 – 98.8	70–140	93.53	90.7 – 96.0
harmonics	AG	99.7	97.4 – 99.9	85–120	97.18	96.0 – 98.5
				135–270	49	44 – 54
Successive	EP	88.4	86.5 – 90.3	OPEN	72.63	67.4 – 77.7
harmonics	YR	99.5	98.2 – 99.8	70–140	90.06	87.3 – 93.3
	AG	99.7	97.4 – 99.9	85–120	92.88	90.2 – 95.6

The outcome of the pitch-shift experiment for subjects EP and YR, in contrast with subject AG, does not provide an opportunity for probabilistic testing. Subject YR's data comprise too few points; and we have argued that subject EP restricted the expected pitch range of her initial pitch estimate to the high-pitch range, thereby systematically excluding the nonsuccessive harmonic modes. Proof that subject EP is capable of perceiving the true periodicity of stimuli comprising nonsuccessive harmonics was obtained in the musical interval recognition experiment (Sec. I C 2). Our argument is supported by the experimental and theoretical probabilities for percent correct interval recognition that are given in Table I.

From the experimental results in Table I it is clear that the three subjects EP, YR, and AG perceived the musical intervals correctly in the vast majority of presentations independently of whether the complex tone stimuli comprised successive harmonics or successive odd harmonics. It is to be recalled that subject EP encountered similar odd harmonic stimuli in the pitch-shift experiment (for f_l at odd harmonics of 100 Hz), and failed to percieve a pitch corresponding to stimulus period. This discrepancy in subject EP's behavior is accounted for by the concept of expected pitch range. The pitch-shift experiment biases most subjects to expect a pitch in the (octave) range around the frequency difference between component tones. If this bias were inappropriately used in the musical interval recognition experiment with odd harmonic stimuli, then the theoretically predicted performance would be significantly poorer than experimental performance. (See Table I.)

On the other hand, if the expected pitch range is centered around the true fundamental frequencies of the stimuli, as it is in the musical interval recognition, theory and experiment are in closer agreement. We do not judge the small, but statistically significant differences between experiment and theory to be of importance. A great deal of further experimental measurement would be required to clarify these second-order effects, including a systematic determination of the exact value of the standard error for each subject. The data presented, however, already provide strong evidence that the central pitch processor does not generally constrain its harmonic pattern recognition by estimating stimulus frequencies as successive harmonics. Instead, new and old data agree with the more general formulation of estimating stimulus frequencies with a general template constrained by an expected pitch range.

IV. SUMMARY AND IMPLICATIONS

Our new findings in pitch-shift and musical interval recognition experiments proved definitively that subjects can estimate stimulus components as nonsuccessive harmonics. Earlier reports that subjects always estimated stimulus components as successive harmonics (de Boer, 1956; Schouten, Ritsma, and Cardozo, 1962; Ritsma, 1967) must be considered as special cases.

To accommodate the new experimental results, optimum processor theory for pitch of complex tones has been made more comprehensive by introducing a more general rule for the central estimation of periodicity pitch. In place of maximum-liklihood estimation of the aurally measured component frequencies as a successive harmonic series, we have maximum posterior estimation of the measured frequencies as a series of harmonics that need not be successive. From the new and old data on periodicity pitch for complex tones we have found that the central processor can assign harmonic estimates to component tones ranging from the fundamental up to the 14th or 15th harmonic (Fig. 1, and Goldstein *et al.*, preceding paper).

Maximum posterior probability estimation introduced prior expectations into the estimation process. Differences in prior probabilities can explain the different results in similar pitch-shift experiments obtained by different subjects and investigators. It is significant that the old and new data imply that variable prior expectations exist only for the estimated periodicity pitch, f_0, but not for the harmonic numbers. Also significant is our finding that the prior expectation for periodicity pitch is adequately described by a uniform distribution over some specific, narrow or broad pitch range.

The conceptually useful interpretation that emerges from the mathematical formalism is that the operation of the central processor for periodicity pitch involves a scanning of the aurally measured frequency pattern by a general harmonic template. The template scan range is set by the prior expected range for periodicity pitch. It would be important to discover whether the physiological implementation of processing for periodicity pitch is homologous with harmonic spectral pattern recognition. The frequency domain appears to be the natural domain for the central processor, because the analysis stage in auditory processing for periodicity pitch (Fig. 4) requires frequency analysis and because auditory tonotopic organization is available for this analysis and further processing. However, on the basis of excisting psychophysics, we cannot assert whether the scanning is performed serially in time or in parallel spatially.

It should be emphasized that while the view of central pitch processing as optimum probabilistic harmonic pattern recognition was evident before (Goldstein, 1973, 1975), the usefulness of the view has been strengthened by the new experimental and theoretical findings. First, the successive harmonic constraint, though providing the best account of earlier data, seemed unnatural for a frequency template matching process. Secondly, the multimodal probability distribution for periodicity pitch measured by Schouten, Ritsma, and Cardozo (1962) was fully compatible with the original form of the theory (Goldstein *et al.*, preceding paper) and implied that no prior expectation exists either for estimates of periodicity pitch or harmonic numbers (Goldstein, 1973, Sec. IX A). This was judged inconsistent with the real existence of a frequency template. Subsequently it was learned (Schouten, Ritsma, and Cardozo, 1975) that the multimodal data were obtained by instructing the subject to report *any* pitch value he could perceive on each of many repeated presentations of the same stimulus. The subject, therefore, was given in instruction which in terms relevant to the original version of the theory, *caused* him to disregard prior estimates of periodicity pitch and harmonic numbers. In terms of the present version of the theory, the subject used the same scan range for each stimulus presentation.

Every mechanism or theory of periodicity pitch must contend with the problem of subharmonics, whether operating in the time or frequency domain. A periodic sound is in principle periodic at the true (smallest) period or any integral multiple of that period. Auditory perception of periodicity pitch does not appear to suffer from this ambiguity. Formally, the period of a waveform can be uniquely chosen by choosing the smallest repetition cycle in the time domain or the largest reference fundamental in the frequency domain.

The original version of the optimum processor theory solved the problem of subharmonics by restricting the estimates of harmonic numbers to successive integers; we now know from the new version of the theory that this is a special, though important, case. According to the new version of the optimum processor theory, the pitch processor chooses the reference fundamental which gives the best match between the general template and the aurally measured component frequencies. If more than one fundamental within the template scan range are equally matched, only the largest is estimated. Thus the theory operates as it logically should with the additional provision that a subharmonic periodicity pitch can be heard if the scan range *excludes* the true fundamental.

An interesting degenerate case of subharmonic periodicity pitch was reported recently by Houtgast (1976), in which periodicity pitch was heard at the subharmonic frequency of a simple tone stimulus. Since in Houtgast's experiment the subject's task was to discriminate between two slightly different values of periodicity pitch, his finding is compatible with periodicity pitch estimation with a very narrow scan range. For a sufficiently narrow scan range a single harmonic stimulus would be adequate for matching the template. Houtgast also found that this subharmonic phenomenon only occurred when the simple tone was accompanied by noise. Possibly the role of the noise was to enable the subject to exclude the pitch of the simple tone from his expected pitch range and focus on the narrow subharmonic range.

Perhaps the most interesting general aspect of periodicity pitch, as shown by the quantitative theory that describes it comprehensively, is its optimum probabilistic basis. In measuring pitch of complex tones the auditory system has been shown to operate very efficiently in extracting the desired information and overcoming the stochastic uncertainty of sensory analysis. We have pointed out elsewhere (Goldstein, 1975) that the mathematical approach in the optimum processor theory of periodicity pitch is based on the same statistical estimation concepts that underlie modern statistical communications theory, in particular Eq. (1) (Shannon and Weaver, 1949; Woodward, 1953). Indeed, Shannon and Weaver speculated that efficient statistical communication may be an appropriate model for human communication in general. Periodicity pitch is one case where this is true. The general potential of the approach for sensory perception of complex stimuli should be examined carefully in future studies.

It was pointed out by Shannon and Weaver (1949) that useful application of statistical communication theory to human communication requires a bridging of the gap between the purely physical aspects of the theory concerning measurements on signals, and the human aspects of communication concerning meaning and effectiveness. This gap was bridged for the psychophysics of periodicity pitch of complex tones by the general template within the central "statistical communication receiver." The "meaning" of the communicated message is defined as the estimated template parameter, the reference fundamental $\hat{f}_0$. The "effectiveness" of the communicated message is measured by the probability distribution of the estimated meaningful parameter. It is clear that the definition of an appropriate internal model of the external signals with meaningful estimated parameters is a key general problem in quantifying human communication.

ACKNOWLEDGMENTS

Dr. A. Gerson's postdoctoral fellowship was supported by the Israel Office for Absorption of Scientists, Office of the Prime Minister, Jerusalem. We are grateful for the cooperation and patience contributed by subjects Edith Polkewitz and Yoash Refael, from the School of Communications Disorders, School of Medicine, Tel Aviv University.

[1]In the real-time procedure employed, each sample of the sound stimulus was computed as a sum of values selected from a 5000-point cosine table, and immediately outputted. A non-real-time procedure would involve instead the computing and storing of a full period of the waveform before outputting.

[2]This value of 49% comes from the following considerations. If the processor worked only with successive harmonic estimates, a (1 3 5 7×100 Hz) stimulus would always be interpreted as (1 2 3 4×155.8 Hz), a (3 5 7 9×100 Hz) stimulus would always be interpreted as (2 3 4 5×170.7 Hz), a (5 7 9 11×100 Hz) stimulus would be interpreted as (2 3 4 5 ×228.6 Hz) with probability 0.15 or as (3 4 5 6×177.6 Hz) with probability 0.85, and a (7 9 11 13×100 Hz) stimulus would be interpreted as (3 4 5 6×221.9 Hz) with probability 0.3 or as (4 5 6 7×181.7 Hz) with probability 0.7. If the subject answers correctly only when the perceived interval between sounds A and B is *precisely* one semitone or one tone, then correct answers would occur only when the two sounds A and B have both the same harmonic construction and the same interpretation, the probability of which is 0.21 (less than random guess). But if the subject classifies as a "semitone" any perceived interval smaller than $\frac{3}{4}$ of a tone, and as a "full tone" any greater interval, then other possibilities could also bear a correct answer. As an example, a (5 7 9 11×98 Hz) sound A interpreted as (3 4 5 6×174.1 Hz) and a (7 9 11 13×92.5 Hz) sound B interpreted as (4 5 6 7 ×168.1 Hz) give a perceived interval of 0.965 that will be "correctly" analyzed as a descending semitone, being very close to the true semitone interval of 0.944. The sum of the probabilities of all such events gives a 49% probability of correct answer.

[3]Under the new ranking condition an integral multiple of a given set of estimated harmonic numbers will be an equally good estimate. Such ambiguities are generally not found experimentally (Sec. III A) and are easily eliminated theoretically by accepting only the largest equally fitting $\hat{f}_0$ in the bounded range.

de Boer, E. (1956). "On the Residue in Hearing," Ph.D thesis (University of Amsterdam).
Flanagan, J. L., and Guttman, N. (1960a). "On the Pitch of Periodic Pulses," J. Acoust. Soc. Am. 32, 1308–1318.
Flanagan, J. L., and Guttman, N. (1960b). "Pitch of Periodic Pulses without Fundamental Component," J. Acoust. Soc. Am. 31, 836 (A).
Goldstein, J. L. (1973). "An optimum processor theory for the central formation of the pitch of complex tones," J. Acoust. Soc. Am. 54, 1496–1516.
Goldstein, J. L. (1975). "Periodicity Pitch: Auditory Pattern Recognition Presumes Natural Stimuli," in *Signal Analysis and Pattern Recognition in Bio-Medical Engineering*, edited by G. Inbar (Wiley, New York), pp. 277–288.
Goldstein, J. L., Gerson, A., Srulovicz, P., and Furst, M. (1978). "Verification of the Optimal Probabilistic Basis of Aural Processing in Pitch of Complex Tones," J. Acoust. Soc. Am. 63, 486–497.
Houtgast, T. (1976). "Subharmonic pitches of a pure tone at low S/N ratio," J. Acoust. Soc. Am. 60, 405–409.
Houtsma, A. J. M., and Goldstein, J. L. (1971). "Perception of Musical Intervals: Evidence for the Central Origin of Musical Pitch," Tech. Rep. 484, (Research Laboratory of Electronics, MIT 1 October 1971).
Houtsma, A. J. M., and Goldstein, J. L. (1972). "The central origin of the pitch of complex tones: Evidence from musical interval recognition," J. Acoust. Soc. Am. 51, 520–529.
Patterson, R. D. (1973). "The effects of relative phase and the number of components on residue pitch," J. Acoust. Soc. Am. 53, 1565–1572.
Patterson, R. D., and Wightman, F. L. (1976). "Residue pitch as a function of component spacing," J. Acoust. Soc. Am. 59, 1450–1459.
Ritsma, R. J. (1967). "Frequencies Dominant in Pitch of Periodic Pulses of Alternating Polarity," IPO Ann. Prog. Rep. 2, 14–24.
Schouten, J. F. (1940). "The Perception of Pitch," Philips Tech. Rev. 5, 286–294.
Schouten, J. F., Ritsma, R. J., and Cardozo, B. L. (1962). "Pitch of the Residue," J. Acoust. Soc. Am. 34, 1418–1424.
Schouten, J. F., Ritsma, R. J., and Cardozo, B. L. (1975). Personal communication during visit to Institute for Perception Research, Eindhoven (Sept. 1975).
Shannon, C. E., and Weaver, W. (1949). *The Mathematical Theory of Communication* (Univ. Illinois P. Urbana).
van Trees, H. L. (1968). *Detection, Estimation and Modulation Theory*, Part I (Wiley, New York).
Woodward, P. (1953). *Probability and Information Theory with Applications to Radar* (Pergamon, London).