Theory of frequency modulation detection for low modulation frequencies

W. M. Hartmann and M. A. Klein

Physics Department, Michigan State University, East Lansing, Michigan 48824
(Received 3 May 1979; accepted for publication 3 December 1979)

There is general agreement that the frequency difference limen measured in a two-sine-tone frequency discrimination experiment is smaller than that measured in a frequency modulation (FM) experiment. We present a model of frequency modulation detection for low modulation frequencies, within the framework of signal detection theory, which accounts well for the observed difference between frequency discrimination experiments and FM detection experiments. The FM detection model also predicts psychometric functions for detection of FM with different modulation waveforms. FM detection experiments with square, sine, trapezoid, and triangle FM are in reasonable agreement with the model predictions.

PACS numbers: 43.66.Fe, 43.66.Ba [BS]

INTRODUCTION

FM detection and pulsed-sines discrimination

The minimum detectable change ΔF in the frequency of a sine tone is an important quantity in the description of human hearing. This quantity indicates the limits of the frequency analyzing ability of the ear and has played a key role in the development of auditory theory.

There is a long history of efforts by psychoacousticians to measure ΔF. In 1931 Shower and Biddulph attempted to measure ΔF by finding the minumum detectable *frequency modulation* of a sine tone. Their study was comprehensive, spanning a wide range of frequencies and intensity levels.

In 1952 Harris used a two-interval, forced-choice method to determine frequency discrimination for sine tones which had different frequencies in the two intervals. We follow Wier *et al.* (1977) by referring to this type of experiment as a "pulsed-sines" experiment. For frequency ranges below 2000 Hz Harris found frequency difference limens which were notably smaller than those found in the FM experiment of Shower and Biddulph. The discrepancies were very large for frequency ranges below 500 Hz.

In 1954 Stevens compared a number of pulsed-sines frequency discrimination experiments with the Shower and Biddulph data. Although there was considerable variation among experiments, the pulsed-sines data were roughly consistent with the rule, 1 JND=1 mel. The Shower and Biddulph FM data were not consistent with this rule; their data showed frequency difference limens which were too high to fit the rule at low frequencies.

Nordmark (1968) has compared the Shower and Biddulph FM data with FM detection data found by other experimenters. He argued persuasively that the Shower and Biddulph data are not representative. In particular, the FM detection level in the Shower and Biddulph data rises too rapidly with decreasing center frequency compared with other FM studies.

Even if one discounts the low-frequency Shower and Biddulph data there is still a marked tendency for FM detection data, from all modern studies, to show higher-frequency difference limens than are found using pulsed sines (Wier *et al.* 1977). The discrepancy between the frequency difference limens determined by FM and pulsed sine techniques is best observed in these experiments in which the two techniques are directly compared, using the same subjects under similar conditions. Such studies have been performed by Verschurre and van Meeteren (1975), Jesteadt and Sims (1975), Moore (1976), and Fastl (1978). These studies all agree. Frequency difference limens measured with FM detection are higher, by about a factor of two, than those measured using pulsed sines.

It is not really surprising that the FM detection and pulsed-sines techniques give different results. As noted by Harris (1952), and by Jesteadt and Sims (1975), the two experiments involve perceptual tasks which are quite different. The pulsed-sines experiment is easy to analyze from the point of view of signal detection theory. However, there is presently no theory of FM detection within the framework of signal detection theory, which would enable one to make a theoretical comparison between the two kinds of experiments. In this paper we present a theory of FM detection which allows a direct comparison. The FM detection theory also makes predictions for the relative detectability of FM with different modulation wave shapes. The theory is tested by experiments which compare frequency discrimination of pulsed sines with FM detection and which compare FM detection for four different modulation wave shapes.

Alternative models for FM detection

In the long history of FM perception studies much has been written about the spectrum of a frequency modulated tone. The spectrum is calculated by performing a Fourier transform of the modulated signal on the entire time axis from minus to plus infinity. The spectrum has a number of sidebands on either side of the carrier, separated from one another by the modulation frequency. The number of significant sidebands is determined by the modulation index $\beta = \Delta f/f_m$, where Δf is the frequency excursion and f_m is the modulation frequency. Shower and Biddulph referred to the sidebands as extraneous frequencies and considered them to be a kind of distortion.

In 1937 Koch presented a theory of FM detection based upon the nature of the FM sidebands. Koch noted that when the modulation index is increased from zero to a value between 1.5 and 2, the envelope of the spectral components becomes bimodal. Koch argued that modulation should be detected when Δf is large enough to make this two-peaked structure appear. In fact, he argued that this detection procedure represents a theoretical limit, and that performance for human observers should never be better than performance in this theoretical limit, $\beta = 1.5-2.0$.

The most persuasive argument against Koch's model is simply to compare its predictions with modern FM detection data. For example, Kay and Matthews (1972) found, for a carrier of 250 Hz, that the detectable modulation index at $f_m = 1$ Hz was 0.4 and that at $f_m = 2$ Hz it was 0.2. For larger values of f_m the modulation index at the detection level was never greater than 0.13. Thus actual detection performance is better than Koch's theoretical limit by a factor of 10.

Feth *et al.* (1969) examined the difficult question of FM detection as a function of modulating frequency. Like Koch, these authors focused on the envelope of the spectrum. They made measurements of the physical spectrum of frequency modulated tones and produced a table of correction factors to be used in converting FM detection data to frequency difference limens. Unfortunately their physical measurements and correction factors do not seem to be relevant to FM detection because the physical measurements were made at modulation indices much greater than detection levels. The spectral envelopes shown by Feth *et al.*, and particularly their "intuitive" spectrum, do not resemble the spectra at detection levels. The basic problem with the spectral view of FM detection at low values of f_m, as noted, in fact, by Feth *et al.*, is that the ear does not integrate the stimulus signal for a long enough time to create a valid representation of the spectrum.

In our work we take the point of view that FM detection (and amplitude modulation detection as well) can be rather simply understood in two regimes: a high modulation frequency regime, where f_m approaches or exceeds the critical bandwidth (Zwicker, 1952), and a low modulation frequency regime, $f_m \leq 6$ Hz (provided that the carrier frequency is not too low). According to this view, in the high-frequency regime the spectral representation does provide a psychologically relevant characterization of the signal (Hartmann, 1978, 1979). In the low modulation frequency regime the relevant characterization of FM (or AM) is a temporal one. For intermediate modulation frequencies the detection and the perception of modulation is, quite likely, a rather complicated process. There is hope of understanding it, though, if a proper understanding of high and low f_m limits is first achieved.

The present paper considers FM detection in the low-frequency regime. In this regime a human subject can, with some success, attempt consciously to track the temporal variations caused by the modulation. Viewed as an integrator the ear integrates for a long enough time to establish a sensation of pitch, but for a short enough time to notice the variations in frequency due to FM. Our model for FM detection then is based upon successive samplings of the pitch of the stimulus. This model is presented in Sec. II. Section I presents a straightforward model for pulsed-sines discrimination for comparison.

I. THEORY: FREQUENCY DISCRIMINATION FOR PULSED SINES

We consider a two-interval forced-choice experiment intended to determine a frequency difference limen. Each interval contains a sine tone; in one interval the frequency is higher than in the other. The subject's task is to decide which interval contains the higher-frequency tone. The analysis of this experiment in terms of signal detection theory is straightforward.

The frequencies of the two tones are f_1 and f_2 where, by definition, $f_2 > f_1$. The two tones are coded in the auditory system on a tonotopic coordinate x. Because our theory is concerned only with detecting small differences, $\Delta F = f_2 - f_1$, in a single fixed-frequency range, the exact relationship between x and f is unimportant. Therefore, we take the internal representation of the two frequencies to be given by

$$x_1 = f_1 + N_1(t)$$

and (1)

$$x_2 = f_2 + N_2(t),$$

where N_1 and N_2 represent time-dependent noise internal to the auditory system, associated with the lower- and higher-frequency tones.

Our expression for the tonotopic variable x has the same form as the expression for the internal representation of signal amplitude for the familiar study of the detection of a signal in noise (Green and Swets, 1966). It is not difficult to justify this form for the tonotopic variable. The neural encoding of the amplitude of an auditory signal is a stochastic process, in which the primary neural elements fire with a certain amount of randomness. Frequency information contained in the signal, in particular pitch information, is extracted from a neural firing pattern by some transformation of the pattern. The transformation may be based upon neural synchrony or it may be based upon the distribution of excitation with respect to the characteristic frequencies of the primary fibers. In either case the noise present in the original encoding of the stimulus will result in a noisy representation of frequency information on the tonotopic coordinate. This idea is the basis of Siebert's (1968) theoretical comparison of intensity and frequency discrimination and for Henning's (1967, 1970) experimental comparison.

We assume that the internal noise has a Gaussian probability density p_N with zero mean and variance σ^2, i.e.,

$$p_N(x)=(\sigma\sqrt{2\pi})^{-1}\exp(-x^2/2\sigma^2). \tag{2}$$

Let $p_{1N}(x)$ and $p_{2N}(x)$ be the probability densities for x for the lower- and higher-frequency intervals. These densities have mean values f_1 and f_2, respectively. According to the theory of ideal observers the subject makes a correct decision if his sample x_2 from distribution p_{2N} is higher than his sample x_1 from distribution p_{1N}. Therefore, the percent correct responses in a 2IFC experiment is given by the familiar result,

$$P_c=\int_{-\infty}^{\infty} dx\, p_{2N}(x)\int_{-\infty}^{x} dx'\, p_{1N}(x'). \tag{3}$$

For the case of independent Gaussians with the same variance, function P_c is the cumulative normal probability from $-\infty$ to a value of $d'/\sqrt{2}$ where $d'=\Delta F/\sigma$.

II. THEORY: MODULATION DETECTION

Modulation detection is quite different from frequency discrimination for pulsed sines. The most important difference is that in the modulation detection experiment the subject must detect a frequency variation within a single experimental interval, whereas in a pulsed-sines experiment he must detect a frequency difference between two intervals.

A. Sampling–differencing

In the low modulation frequency regime a listener's image of an FM signal is that a pitch is moving in time. Therefore, in our model of FM detection the listener takes successive pitch samples and compares them in an attempt to detect a change. Experiments on pitch discrimination for short tones indicate that it is possible for subjects to operate in a sampling mode. As the duration of a short tone increases pitch acuity improves. According to the data of Henning (1970), asymptotically good performance is achieved when tone durations are about 25 ms. Therefore, at a modulation rate of 4 Hz, where many FM detection experiments have been done, a subject could take ten pitch samples during the course of a single modulation cycle. The subject could take as many as 80 samples during a 2-s stimulus tone. The subject can take more samples in a given length of time if he is willing to sacrifice resolution on each sample.

The frequency modulated signal is represented on the internal tonotopic coordinate according to the formula,

$$x(t)=f_0+\Delta f_M(t)+N(t), \tag{4}$$

where f_0 is the center frequency, $\Delta f_M(t)$ is the modulation waveform, and $N(t)$ is internal noise. Because we are concerned with frequency differences the center frequency is of no importance. It is mathematically convenient to shift the tonotopic scale so that $f_0=0$. We denote the maximum value of the modulation excursion by the symbol x_M. Index M is 0 for the case of no modulation, and has values S, T, Z, and Q for sine, triangle, 50% trapezoid, and square modulation, respectively. For example, the frequency excursion for sine modulation with angular frequency ω_m is given by

$$\Delta f_S(t)=x_S\sin(\omega_m t). \tag{5}$$

A frequency modulated signal can be represented statistically by a frequency probability density $p(x)$, which indicates the probability that a pitch sample has value between x and $x+dx$.

The frequency probability density plays an important role in our theory. It is different from the spectrum of a modulated signal. A frequency probability density is obtained when the sampling time is short compared with the modulation period. A spectrum is obtained from measurements in which the sampling time is long compared with the modulation period.

In the absence of noise the density $p(x)$ is simply proportional to the amount of time that the modulated signal has a frequency which maps to tonotopic coordinate x. We refer to this density as p_M where M indicates the modulation waveform. Densities for various common waveforms are given below.

(1) No modulation:

$$p_0(x)=\delta(x). \tag{6}$$

(2) Sine modulation:

$$p_S(x)=\pi^{-1}(x_S^2-x^2)^{-1/2},\quad 0\leq|x|\leq x_S,\qquad =0,\quad |x|>x_S. \tag{7}$$

(3) Triangle modulation:

$$p_T(x)=(2x_T)^{-1},\quad 0\leq|x|\leq x_T,\qquad =0,\quad |x|>x_T. \tag{8}$$

(4) Trapezoidal (50%) modulation:

$$p_Z(x)=\tfrac{1}{4}[x_Z^{-1}+\delta(x+x_Z)+\delta(x-x_Z)],\quad 0\leq|x|\leq x_Z,\qquad =0,\quad |x|\geq x_Z. \tag{9}$$

(5) Square modulation:

$$p_Q(x)=\tfrac{1}{2}[\delta(x-x_Q)+\delta(x+x_Q)]. \tag{10}$$

There are several ways in which a listener could use successive samples of $p(x)$ to detect FM. For instance, he could compare pitch values obtained on two successive samples and, if the difference is large enough, decide that FM is present. A more efficient process, however, is to perform a cross correlation between the sampled pitches and a modulation waveform template stored in memory. Mathematically this process would be represented as a correlation sum with a contribution to the sum from each sample interval. In this section we develop a general sampling–differencing model based upon only two samples giving a single difference. We defer, until Sec. B, the questions of number of samples, correlation between samples, and correlation of samples with a modulation template.

The sampling and differencing operations are described as follows. The probability density for x at the first sample time is

$$p_{1MN}(x_1)=\int_{-\infty}^{\infty} dx_1'\, p_{M_1}(x_1')\, p_{N_1}(x_1-x_1'). \tag{11}$$

Here p_{N_1} is the probability density of the noise in the first sample, and $p_{M_1}(x_1')\,dx_1'$ is the probability that the stimulus modulation has frequency excursion between

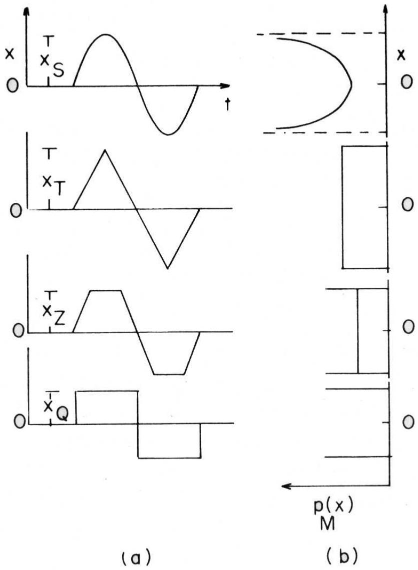

FIG. 1. (a) Frequency modulation waveforms. The waveforms are scaled to have equal rms amplitude. (b) Probability for the waveforms $p_M(x)$. The densities are in the correct scale except for the delta functions.

x_1' and $x_1'+dx_1'$. Note if the first sample is uncorrelated with the modulation then all values of x_1' are *a priori* equally probable, and the density p_{M_1} simply becomes equal to the frequency probability density of the modulation p_M. For the present, however, we keep the formalism general and refer to the density for the first sample as p_{M_1}.

Similarly, the probability density at the second sample time is

$$p_{2MN}(x_2)=\int_{-\infty}^{\infty} dx_2' p_{M_2}(x_2') p_{N_2}(x_2-x_2'), \tag{12}$$

where p_{N_2} is the noise density for the second sample, and p_{M_2} may or may not be correlated with sample 1.

According to the model the subject takes the difference between the values obtained in the two samples,

$$z=x_2-x_1. \tag{13}$$

The probability density for the difference is given by

$$p_{\Delta MN}(z)=\int_{-\infty}^{\infty} dx_1 p_{1MN}(x_1) p_{2MN}(x_1+z). \tag{14}$$

We assume that the noise has the same variance for both samples, i.e.,

$$p_{N_1}=p_{N_2}\equiv p_N, \tag{15}$$

where p_N is given by Eq. (2). Then it is easy to show that

$$p_{\Delta MN}(z)=\int_{-\infty}^{\infty} dx_1' \int_{-\infty}^{\infty} dx_2' p_{M_1}(x_1') p_{M_2}(x_2') r(z+x_1'-x_2'), \tag{16}$$

where r is the convolution of p_N with itself, i.e.,

$$r(x)=(2\sigma\sqrt{\pi})^{-1}\exp(-x^2/4\sigma^2). \tag{17}$$

The special case of no modulation is particularly important. The difference density is

$$p_{\Delta 0N}(z)=r(z). \tag{18}$$

This is the difference distribution obtained by the subject in a 2IFC experiment during the interval in which there is no modulation. The values of the difference z may be positive or negative. When asked to choose which of two intervals contains modulation the ideal observer chooses the interval with the largest absolute value of z. The density for the absolute value of z is given by

$$\begin{aligned} p_{A\Delta MN}(z)&=2p_{\Delta MN}(z), \quad z\geq 0,\\ &=0, \quad z<0. \end{aligned} \tag{19}$$

The expression for the percent correct responses in a 2IFC experiment, based upon a single differencing operation, is then,

$$P_c=\int_{-\infty}^{\infty} dz\, p_{A\Delta MN}(z)\int_{-\infty}^{z} dz'\, p_{A\Delta 0N}(z'). \tag{20}$$

Substituting from the above equations we find that

$$P_c=2\int_{-\infty}^{\infty} dx_1' \int_{-\infty}^{\infty} dx_2' p_{M_1}(x_1') p_{M_2}(x_2') \times \int_{-\infty}^{\infty} dz\, r(z+x_1'-x_2')\,\mathrm{erf}(z/2\sigma), \tag{21}$$

where erf is the error function (Abramowitz and Stegun, 1964, p. 297).

A convenient form for P_c is obtained by writing the modulation frequency probability densities in the form

$$p_M(x)=x_M^{-1}\rho_M(x/x_M). \tag{22}$$

Then with appropriate changes of integration variables we obtain,

$$P_c=\frac{2}{\sqrt{\pi}}\int_{-\infty}^{\infty} dv\, \rho_{M_1}(v)\int_{-\infty}^{\infty} dy\, \rho_{M_2}(y) \times \int_0^{\infty} du \exp\{-[u+\tfrac{1}{2}D(v-y)]^2\}\,\mathrm{erf}(u), \tag{23}$$

where D is the discrimination parameter, the ratio of the maximum frequency excursion to the noise error

$$D=x_M/\sigma. \tag{24}$$

It is interesting to evaluate Eq. (23) in the limit of small D. For a number of important cases the linear term in a Taylor expansion of P_c vanishes, and the leading term is proportional to D^2. This occurs if p_{M_1} and p_{M_2} are the same functions, as would be the case for uncorrelated random sampling. It also occurs if p_{M_1} and p_{M_2} are both even functions of their arguments. The unsampled densities p_M are even functions for the waveforms described in Eqs. (6)–(10). Unless correlations in the sampling operation favor one direction of frequency excursion over the other the sampled densities (p_{M_1} and p_{M_2}) will also be even functions. In these cases psychometric functions for modulation detection should rise from 50% correct more slowly than linearly. Pyschometric functions for pulsed-sines discrimination, by contrast, should rise linearly with the discrimination parameter.

The sampling and differencing operation described above is a rather general model of the way in which two samples can be used to produce a decision variable for FM detection. Each of the sampling operations is accompanied by internal noise. The above model then relates FM detection to pulsed-sines discrimination if one assumes that the internal noise associated with the pitch sampling in FM is statistically similar to the noise associated with pitch perception in pulsed-sines discrimination. To evaluate the above expression for P_c requires some assumptions about the nature and number of the sampling operations. We consider a special and simple case in Sec. B.

B. Correlated-differencing model

The correlated–differencing model for FM detection is a special case of the general sampling–differencing model described in Sec. A. It includes assumptions which make it possible to evaluate the general expression in Eq. (23). We believe that these assumptions are plausible when the modulation is barely detectable, but that the assumptions probably are in error when the modulation is so large that detection performance is close to 100%. Below we list the assumptions of the correlated–differencing model and discuss them.

1. Crude cross correlation

The optimum detection process is a cross correlation between the sampled modulation and a modulation waveform template stored in memory. A subject can operate in this optimum mode in an experimental situation in which the signal is specified exactly (Green and Swets, 1966, p. 162). For FM detection this corresponds to a situation in which the modulating waveform, frequency, and initial phase angle are kept constant during the course of the experiment. In practice, however, not all this information is available to a subject. When modulation widths are so small that FM is barely detectable it is impossible to distinguish one modulation waveform from another (Klein and Hartmann, 1979). A just-detectable trill (square wave FM) and a just-detectable vibrato (sine wave FM) sound identical. For this reason the cross correlation between stimulus and memory waveforms may not be a very detailed one. A subject is mainly aware of the periodicity of the modulation. He expects that the tone pitch will go up, and then, half a cycle later, will go down. Therefore, we assume that only periodicity information is used to correlate the first and second samples.

2. Random initial phase

We assume that the modulation is barely detectable so that the subject does not know the initial phase angle of the modulation. Therefore, he cannot take his first sample at an optimum time, when the frequency excursion is at a positive or negative maximum. Instead the phase angle of the modulation at the first sample time is random and, therefore,

$$p_{M_1} = p_M. \tag{25}$$

3. Half-period correlation

We suppose that the subject does use his memory of the modulation period to help him in the detection task. All the waveforms of interest to us have the symmetry relation

$$\Delta f(t) = -\Delta f(t+T/2), \tag{26}$$

where T is the modulation period.

It is not hard to show that, on the average, the subject's best strategy is to take the difference between two samples separated in time by $T/2$. Therefore, we assume that the second sample is tightly correlated with the first so that

$$p_{M_2}(x_2') = \delta(x_2' + x_1'), \tag{27}$$

where δ is the Dirac delta function and x_1' is the frequency excursion at the time of the first sample.

With these assumptions, the probability density for the difference, Eq. (16), becomes

$$p_{\Delta MN}(z) = \int_{-\infty}^{\infty} dx_1 \, p_{1MN}(x_1) \, p_{2MN}(x_1 + z). \tag{28}$$

Difference densities $P_{\Delta SN}$, for sine modulation, and $P_{\Delta 0N}$ for no modulation are shown for illustration in Fig. 2. Because P_M is symmetric about zero for $M = 0, S, T, Z$, and Q, $P_{\Delta MN}$ is also symmetric.

The final expression for the percent correct responses in a 2IFC experiment, Eq. (23), becomes

$$P_c = \frac{2}{\sqrt{\pi}} \int_0^{\infty} du \int_{-\infty}^{\infty} dv \, \rho_M(v) \exp[-(u + Dv)^2] \, \mathrm{erf}(u), \tag{29}$$

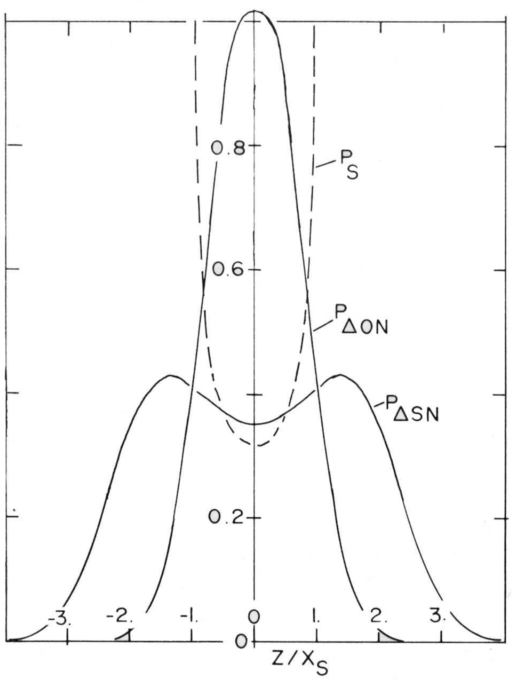

FIG. 2. The figure shows $p_{\Delta SN}$, the probability density for the difference distribution given both sine modulation and internal noise. In this figure, $D = x_S/\sigma = 2$. When there is no modulation the difference density is given by $p_{\Delta 0N}$. The correlated–differencing theory predicts that for these densities a 1-cycle FM 2IFC experiment would find 79% correct responses. The dashed curve shows the probability density p_S for sine modulation alone. The divergences at $z = \pm x_S$ are responsible for the peaks near $z = \pm 2x_S$ in the difference density. All probabilities are normalized to unity.

where D is given by Eq. (24). This expression predicts that P_c should be proportional to D^2 for small values of D.

4. Independent multiple sampling

Equation (29) gives the percent correct responses predicted by the correlated–differencing model if the subject performs only a single differencing operation. However, the stimuli in FM detection experiments are typically long enough for the subject to take many differences. A complete theory of FM detection must deal in some way with the possibilities for multiple sampling. The most convenient assumption to make in the present context is that the subject, who is presumed to be aware of the modulation periodicity, performs a differencing operation on each modulation cycle. As the subject takes additional differences his performance in the FM detection task improves. The simplest statistical description of the improvement is obtained by assuming that successive differencing operations are stochastically identical and are independent. With this assumption our final expression for the percent correct responses in an FM detection experiment is given by Eq. (29) with D given by the equation

$$D=x_M\sqrt{N}/\sigma, \tag{30}$$

where N is the number of differencing operations performed.

Ideally a subject's performance should improve without bound if the number of modulation cycles presented becomes sufficiently large. In practice there is a limit to the number of differencing operations which can be usefully performed, analogous to the maximum integration time of standard signal detection theory. To find the maximum value for N we performed an experiment.

III. EXPERIMENT 1

The goal of experiment 1 was to determine the maximum number of modulation cycles, and hence, according to our simple sampling model, the maximum number of differencing operations which a subject can use in performing an FM detection task. The experimental approach was to monitor FM detection performance in a series of experimental runs in which the number of modulation cycles presented N_c varied from run to run. We expected that as the number of cycles was increased performance should improve, roughly proportional to $\sqrt{N_c}$, until the maximum number N_m was reached. No further improvement was expected when N_c became greater than N_m.

Three subjects participated in the experiment. Subjects had at least a year's experience performing psychoacoustical tasks in pitch perception. The signals were 800-Hz sine tones presented binaurally at 75 dB SPL through Beyer DT-48 headphones. The experiment used a two-interval, forced-choice procedure. On one interval the signal was frequency modulated with a 4-Hz sine wave. Signals began and ended with the modulation at a positive-going zero crossing. On the other interval there was no modulation. A 250-ms silent interval separated the two observation intervals. A 1-s gap, following the subject's response, separated the trials. The experiment used a staircase procedure (Levitt, 1971). Following an incorrect response the FM peak excursion was increased by 0.16 Hz; following two correct responses the peak excursion was decreased by 0.16 Hz. On each experimental run the subject reversed the direction of the staircase 26 times. The first four reversal points were discarded and the average of the remaining 22 was interpreted as a measurement of the FM peak excursion at the 71% correct point. On a given experimental run the number N_c was fixed. Runs with different values of N_c were performed in haphazard order and experimenting continued until the subject's response stabilized. The average of the final four runs for each value of N_c was used to determine the just-detectable excursions plotted in Fig. 3.

The data in Fig. 3 apply to the theory of Sec. II in two ways. First, they provide the value of N_m, the maximum number of usable cycles. Detection performance improves for increasing N_c for $N_c<4$, but there is no statistically significant improvement for $N_c>4$. Therefore, it appears that N_m is about 4. Second, the data test the plausibility of the assumption that detection performance should improve according to Eq. (30), i.e., $x_M \sim N_c^{-1/2}$ for $N_c<N_m$. In Fig. 3 a dashed line has been fitted to the experimental points at $N_c=2, 3$, and 4. This line passes through the point $(\Delta f=0, N_c=\infty)$. Ideally all the experimental data should lie on this line for $N_c<N_m$.

Because N_m is only 4 there are not many data points to test the assumption, but it appears as though performance deteriorates more rapidly with decreasing N_c than the assumption predicts. The data can be interpreted, within the context of the correlated–differencing model, as indicating that successive cycles of the modulation should not be considered statistically equivalent

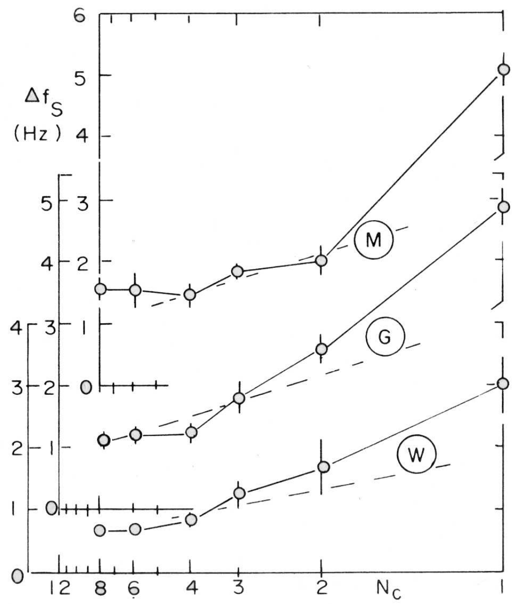

FIG. 3. Just-noticeable sine FM frequency excursion for three subjects as an experimental function of N_c, the number of modulation cycles in the stimulus. The horizontal scale is proportional to $N_c^{-1/2}$.

or independent. With that interpretation our justification for the assumption of independent differencing operations is only one of mathematical simplicity. Alternatively one might note that the stimuli with only a single modulation cycle are brief, 250 ms. It may be that there is some other effect, unknown to us, which makes detection performance unexpectedly poor for these brief tones.

IV. THEORY AND EXPERIMENT

Equation (3) predicts the psychometric function for pulsed-sines discrimination as a function of parameter σ. Equations (29) and (30) predict psychometric functions for modulation detection as functions of σ and N. These equations can be used as follows:

If one wishes to compare FM detection psychometric functions for different modulation waveforms, Eq. (29) provides a model which is entirely parameter free. If one wishes to predict FM detection psychometric functions on an absolute frequency scale one needs parameter σ, the measure of the internal noise on the tonotopic coordinate. However, parameter σ enters the model only as a scale factor, which may be different for different subjects, and it is not of particular interest in the present context. If one wishes to compare FM detection performance with pulsed-sines discrimination performance one needs parameter N. From the results of experiment 1, we take $N=N_m$ equal to 4.

The psychometric functions for FM detection predicted by the correlated-differencing model are shown by curves Q, Z, S, and T in Fig. 4. The upper horizontal scale is x_M/σ, the ratio of the modulation peak frequency excursion to the square root of the noise variance. The lower horizontal scale is X_M/σ, where X_M is the modulation peak-to-peak excursion, i.e., $X_M = 2x_M$. Figure 4 also shows the predictions of Eq. (3) for the pulsed-sines discrimination experiment. The ratio of the frequency difference $\Delta F = f_2 - f_1$ to the square root of the noise variance is given on the lower horizontal scale. Parameter N enters the calculations only as a scale factor which relates the position of the modulation detection psychometric functions to the position of the discrimination psychometric function. Parameter N does not affect the shapes of the psychometric functions. As expected, the computed modulation detection psychometric functions rise quadratically from 50% correct, whereas the discrimination psychometric function rises linearly.

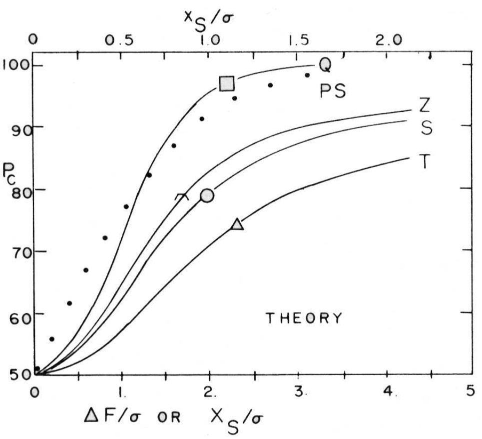

FIG. 4. Percent correct responses in a 2IFC experiment predicted by the theory of Secs. I and II. Dots, PS, are for frequency discrimination of pulsed sines. Curves Q, Z, S, and T are for modulation detection for square, trapezoid, sine, and triangle waveforms. The lower horizontal scale is the ratio of frequency difference (or FM peak-to-peak excursion) to the square root of the noise variance. The upper horizontal scale is for FM peak excursion.

V. EXPERIMENTS 2 AND 3

A. General

To test the predictions of the models, illustrated in Fig. 4, we performed experiments 2 and 3. Experiment 2 found the psychometric function for frequency discrimination for pulsed sines. Experiment 3 measured psychometric functions for FM detection for different modulation waveshapes.

Attempts were made to make the pulsed-sines and FM detection experiments similar. The three subjects from experiment 1 participated in experiments 2 and 3. The signals were 800-Hz sine wave tones, presented binaurally through Beyer DT 48 headphones at 75 dB SPL. Experiments 2 and 3 used the method of constant stimuli and a two-interval, forced-choice procedure. The experimental variable was the frequency difference in the pulsed sines experiment and was the frequency excursion in the modulation detection experiments. Each experimental run consisted of 100 judgements, ten pairs of tones in each of 10 cycles. Within each cycle the experimental variable took on six different values in a random order. The same method of data collection and analysis was followed for each experiment. During the initial runs the performance of the subjects improved with time. After several experimental sessions performance seemed to stabilize. We continued experimental runs until the final five successive runs exhibited no large differences. In all cases the psychometric functions reported here were computed by averaging the final five runs. Although feedback was given during some of the training runs no feedback was given on any of the data runs. The runs of experiments 2 and 3 were interleaved in a random way during the weeks of experimentation.

B. Experiment 2 stimuli

In experiment 2, frequency discrimination was measured using pulsed sines. The stimuli were 500-ms tone bursts with a 500-ms silent interval between observation intervals. A 1-s gap, following the subject's response, separated the experimental trials. Except for the diotic presentation our stimuli were identical to those of Wier *et al.* (1977). The stimulus tones were generated by a Wavetek voltage—controlled oscillator controlled by a microcomputer, which also collected the response data. The results of experiment 2 are given by the dots in Figs. 5–7.

C. Experiment 3 stimuli

Experiment 3 measured a subject's ability to detect frequency modulation for four different modulation waveforms, square (Q), trapezoidal (50%) (Z), sine (S), and triangle (T). The stimuli were 2-s tone bursts with a 250-ms silent interval between observation intervals. A 1-s gap following the subject's response separated the trials. One of the two intervals contained modulation, the other interval did not. The modulation frequency was 4 Hz and the stimulus tones began and ended with the modulation waveform at a positive-going zero crossing. Modulation waveforms were generated by a 12-bit DAC run by the microcomputer. The results of experiment 3 are given in Figs. 5–7.

Note that the horizontal scale plots frequency difference for the pulsed-sines experiment and *peak-to-peak* modulation excursion for the FM detection experiments. These are the two quantities which, historically, have always been compared.

VI. DISCUSSION OF RESULTS

A. Theory versus experiment

Experiments 2 and 3 led to psychometric functions which are plotted in Figs. 5–7 for three subjects. These psychometric functions can be compared with the psychometric functions in Fig. 4 predicted by the correlated–differencing model. The correspondence between theory and experiment seems quite good, especially when it is recalled that the model contains only one parameter N which provides the relative scale between the FM detection prediction and the pulsed-sines discrimination prediction. This one parameter was not adjusted; it was taken from the results of experiment 1.

The comparison of theory and experiment can be made in detail as follows:

Pulsed sines versus FM. The most useful single comparison between pulsed-sines discrimination and FM detection is the ratio $X_S/\Delta F$ at 75% correct point. The

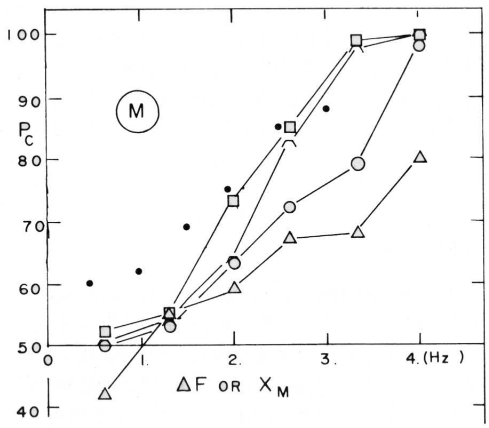

FIG. 5. Pulsed sine frequency discrimination and FM detection experiments for subject M. Percent correct responses are plotted versus ΔF for pulsed sine (•) and against the peak-to-peak frequency excursion x_M for FM with different waveforms: square □, trapezoid ⌒, sine ○, and triangle △. Lines connect the FM points to guide the eye.

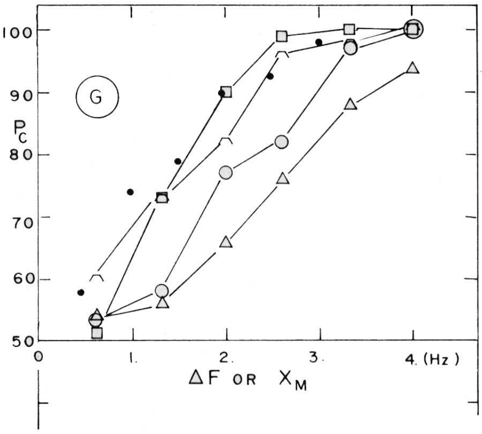

FIG. 6. Same as Fig. 5, for subject G.

theoretical ratio is 1.8. The experimentally determined ratios for subjects M, G, and W are, respectively, 1.6, 1.9, and 1.4 with possible error of ±20%.

Crossing. The model predicts that the pulsed-sines function should cross the FM detection function for square wave modulation. A crossing does occur in the data for two of three subjects and almost occurs for the third subject.

Slope at 50%. According to theory the slope of the pulsed-sines psychometric function should be finite at 50% but the slope of the FM detection psychometric functions should be zero at 50%. The experimental psychometric functions for subject M appear to be entirely consistent with the theoretical slope prediction. The experimental data for subject G also agree with the theoretical slope prediction except for the case of trapezoidal modulation. The data of subject W, with unusually low thresholds, do not appear to agree with the slope prediction, but the disagreement may be due to the particular choice of values of ΔF and X_M used in the experiment.

Ordering. The model makes parameter-independent predictions for the relative detectabilities of the various FM waveforms. The experimental psychometric functions appear to be ordered as the model predicts. For instance, at the 75% correct point there is complete

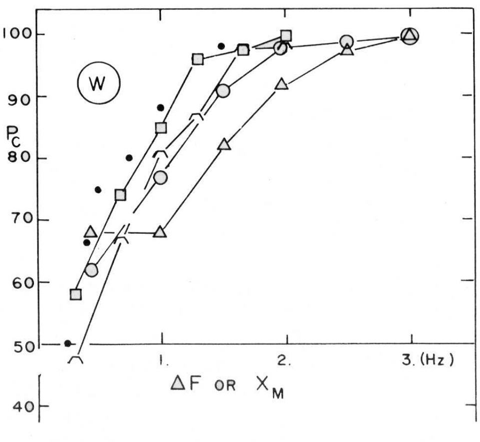

FIG. 7. Same as Fig. 5, for subject W.

agreement between theory and the ordering of the psychometric functions for the three observers. There are occasional crossings among the experimental psychometric functions for the three observers. These crossings show no systematic behavior; they seem to be associated only with irregularities in the experimental curves due to statistical fluctuations.

Psychometric function shape. The main difference between the model predictions and the experiment occurs for large values of X_M, where the modulation is readily detected. In that region the theoretical psychometric function slopes tend to be too small to agree well with experiment. We believe that the failure of the model in this region is caused by the assumptions that each differencing operation is independent and begins with a sample taken at a random modulation phase angle. We believe that these assumptions are reasonable near the detection level, but when the frequency excursion is large, the individual differencing operations are not independent but are, almost certainly, phase locked. For large frequency excursions the subject can probably choose an initial sampling phase angle in a way which is superior to the random procedure assumed in the correlated–differencing model. Both of these changes should lead to improved detection performance for large frequency excursions and to steeper psychometric functions near 100% correct.

B. Alternative models

The correlated–differencing model of Sec. II is in the general framework of signal detection theory. The fact that it correctly predicts the ordering of the psychometric functions for different modulation waveforms is encouraging. The first column of Table I shows the predictions of this model for the relative frequency excursions of different modulation waveforms at the 75% correct point. However, there are other models (Hartmann, 1977) which also predict the observed ordering of detection thresholds. These models are not complete. They do not include noise, they make no prediction for the difference between modulation detection and frequency discrimination for pulsed sines, and they do not, in their present form, predict a psychometric function for modulation detection. Nonetheless it is worth considering these models because they are simple and because they might contain some element of truth.

1. *Root-mean-square model (RMS)*

The rms model assumes that the ear takes the root-mean-square value of the modulation waveform, averaging the modulation just as an ac voltmeter measures an alternating voltage. In this simple model we assume that there is some threshold for the rms value so that FM with any waveform is detected when its rms value is equal to that threshold. The predictions of this model are given in the second column of Table I.

2. *Fundamental extraction model (FEM)*

According to the fundamental extraction model the ear detects modulation by detecting only the fundamental Fourier component of the modulation waveform. This model has its origins in the FM adaptation experiments reported by Kay and Matthews (1972). Those experiments suggested that the auditory system contains channels tuned to specific modulation frequencies. The channels are about 1 octave in width. Thus the adaptation experiments suggest that the ear makes a crude Fourier analysis of the modulating waveform. The process is analogous to the spectral analysis performed in the cochlea for the (much higher) audible frequencies. We assume that the various tuned channels have approximately equal sensitivity. Because the Fourier components of the modulation waveforms used experimentally decrease monotonically with increasing harmonic number, only the channel corresponding to the fundamental frequency of the modulation is excited when the modulation is just detectable. Because all the modulation waveforms in our experimental study have only odd harmonics, no harmonic other than the fundamental contributes to the excitation in the fundamental channel. Therefore, according to the fem, frequency modulation is detected by detection of only the fundamental Fourier component of the modulation (Klein and Hartmann, 1979). Modulations by different waveforms are equally detectable when they have equal fundamental amplitudes. The predictions of this model are given in the third column of Table I.

Unfortunately the experimental ratios in the last three columns of Table I cannot be used to make a reliable choice between the models. The error in the ratios is a combination of the errors in two experimental psychometric functions; we estimate that the ratios could be in error by 20%.

C. Comparison with previous experiments

The psychometric functions of Fig. 4 can be used to compare the model predictions with the results of previous experimental studies. According to the model, the 71% correct points, as would be measured in a staircase experiment (Levitt, 1971), for pulsed sines and modulation occur when $\Delta F/\sigma = 0.78$ and when $X_S/\sigma = 1.4$. Therefore, at the detection point $X_S = 1.8\Delta F$. This result is our model prediction for the comparison of FM

TABLE I. The table shows the frequency excursions required for 75% correct responses in a 2IFC experiment for pulsed-sines discrimination (PS) and for modulation detection for square (Q), trapezoid (Z), sine (S), and triangle (T) modulations, all relative to the excursion for sine modulation. The left three columns show the predictions of the correlated–differencing model, the root-mean-square model and fundamental extraction model. The right three columns were calculated from experimental psychometric functions for three subjects. The experimental columns are not accurate to better than 20%.

	Prediction			Experiment		
M	cdm	rms	fem	M	G	W
PS	0.57			0.62	0.54	0.69
Q	0.62	0.71	0.79	0.70	0.70	0.77
Z	0.87	0.87	0.87	0.77	0.70	0.94
S	1	1	1	1	1	1
T	1.42	1.23	1.23	1.23	1.27	1.39

TABLE II. Ratio of X_s in different modulation detection experiments to the pooled ΔF data from Wier *et al.* (1977) compared with the CDM prediction.

f (kHz)	CDM	Shower Biddulph	Zwicker	Groen Versteegh	Kay Matthews	Jesteadt Sims	Moore	Fastl
0.125	1.8	5.1						
0.25	1.8	3.4		2.9	1.1	2.9		
0.35	1.8							4.2
0.5	1.8	2.4		2.1				3.2
1	1.8	2	2	1.4		2.2	2.7	3.3
f_m(Hz)	≤ 6	2	4	4	4	8	4	4
SL(dB)	>20	40	40	30–50	40	60	50	70
Method	2IFC	A	Y/N	A	Y/N	2IFC	2IFC	A
Date		1931	1952	1957	1972	1975	1976	1977

detection and pulsed-sines discrimination experiments. Of particular interest is the model prediction that this ratio should be independent of frequency range. Although the internal noise σ may depend upon frequency range, it cancels in the above experimental ratio. Generally, the model predicts that if $\log(X_M)$ and $\log\Delta F$ are both plotted against f (or against $\log f$) then the two curves should be parallel.

The theoretical prediction can be compared with the observed difference between the modulation detection data obtained in numerous FM studies and the pooled pulsed-sines data collected by Wier *et al.* (1977). Wier *et al.* have fitted a straight line to the pulsed-sines frequency discrimination data of six different studies. The frequency discrimination studies were monaural at 30–50 dB SL.

There are several factors which complicate the comparison of the modulation data with the pulsed-sines data. First, at high center frequencies, FM induces amplitude modulation because of the rapidly varying frequency dependence of the transfer function of the earphone–ear system. The AM artifact would be expected to produce artificially low FM thresholds. This objection has been noted by Harris (1952) and by Henning (1966). We therefore compare only modulation detection and pulsed-sines experiments performed below 2000 Hz.

Other problems in the comparison are that different workers use different modulation frequencies f_m, different sound levels, and different psychophysical methods.

Table II gives the ratio $X_S/\Delta F$, where X_S is the threshold peak-to-peak excursion for seven different monaural FM detection studies and ΔF is the frequency difference limen from the pooled data of Wier *et al.* In this table, f gives the frequency range and f_m gives the FM modulation frequencies. Other FM experimental parameters include the sound level (SL), and the method: A = method of adjustment, Y/N = yes/no, 2IFC = two-interval forced-choice. The Jesteadt and Sims FM data are divided by a factor of two, to compensate for the large value of f_m (cf. Kay and Matthews, 1972).

The values in Table II can be compared with the prediction of the theory $X_S/\Delta F = 1.8$, independent of the frequency range f. The worst agreement is with the Shower and Biddulph data at 125 Hz. As noted by Nordmark (1968) the Shower and Biddulph data seem anomalously high at low frequencies.

Apart from the Shower and Biddulph data there still seems to be a tendency for the experimental ratio in Table II to rise with decreasing range, contrary to our model. We note however, that if the data of Wier *et al.* rather than the pooled data, are used for ΔF the tendency for the ratio to rise with decreasing f tends to disappear. We do not believe that there is good evidence against the contention that the ratio is constant.

Three of the studies listed in Table II made comparisons of FM detection with frequency discrimination for pulsed sines. In these cases it makes much more sense to compare the two experiments within a study than to make the comparison in Table II. The intrastudy ratios are given in Table III. The Jesteadt–Sims data have been corrected as in Table II. In his study Moore used 20 unpracticed observers. Some of these observers exhibited very high values of ΔF. To find the ratio in Table III we averaged only the ten best observers from Moore's study.

The data in Table III can be compared with the theoretical value, 1.8. As expected there is less scatter among the data in Table III, within studies, than there is in Table II, different FM and pulsed-sines studies. Nevertheless the experimental errors can be large. For instance, the largest deviation from the theoretical value of 1.8 occurs in Fastl's data. However, within the experimental error bars Fastl's data are consistent with a ratio of 1.8.

Finally, as noted in Sec. II, a clear prediction of the theory is that psychometric functions for FM detection should rise quadratically from 50% with increasing frequency excursion, whereas psychometric functions for

TABLE III. Ratio of X_s to ΔF within studies.

f (kHz)	Jesteadt Sims	Moore	Fastl
0.25	2.6		
0.35			3.3
0.5			2
1.	2.0	2.1	3

pulsed-sines discrimination should rise linearly with increasing frequency difference. We are aware of only one other study in which psychometric functions for FM detection and pulsed-sines discrimination were compared. This is the study of Jesteadt and Sims (1975). Jesteadt and Sims fitted straight lines to their experimental psychometric functions. They noted that the major difference between the psychometric functions for the two experiments was that the intercepts at 50% correct responses in FM detection occurred for values of X_S markedly greater than zero. By contrast the intercepts for 50% pulsed-sines discrimination occurred at $\Delta F = 0$. They noted that no matter what scales were used in plotting the FM data at least two of their three subjects would show chance performance for the smallest FM range. This experimental observation is in agreement with the shapes of the psychometric functions predicted by the correlated–differencing model.

VII. CONCLUSION-SUMMARY

We began our study by noting that frequency difference limens found in FM detection experiments are larger than those found in pulsed-sines frequency discrimination experiments. An understanding of this common observation requires a theory in which both FM detection and pulsed-sines discrimination are treated on the same basis. We fulfilled this requirement by proposing a model for FM detection within the general framework of signal detection theory.

For low modulation frequencies, FM is perceived as a time-varying pitch. Therefore, our model of FM detection involves a sampling of the pitch at different times. The absolute value of the difference between the sampled pitch values forms a decision variable used in modulation detection. Pitch is represented quantitatively on an internal tonotopic coordinate which includes an admixture of internal noise. The sampling–differencing model for FM detection relates FM detection to pulsed sines discrimination if the noise is statistically the same in both kinds of experiments. A specific application of the sampling–differencing model, called the correlated–differencing model, presumes that the subject is aware of the modulation period and uses this information in performing the detection task. In a single differencing operation optimum use is made of the periodicity information by acquiring pitch samples at half-period intervals. The model assumes that the subject cannot use modulation phase information.

In the present application multiple sampling is included in the simplest possible way, by assuming that successive differencing operations are statistically identical and independent.

The correlated–differencing model of FM detection makes a parameter-free prediction of the relative detectabilities of modulations with different waveforms. The predictions were successfully tested against experiments with four different waveforms. The model requires one parameter to make a prediction for the discrepancy between frequency DL experiments employing FM detection and frequency DL experiments employing pulsed-sines discrimination. This parameter, the number of useful differencing operations, was found by a separate experiment. The resulting prediction for the DL discrepancy was in general agreement with the discrepancy observed in DL experiments done in recent decades, and with our own comparative experiments. It was particularly encouraging to find that the theory predicts different shapes for the psychometric functions for FM detection and pulsed-sines discrimination which are, in fact, the experimentally observed shapes. We believe that the correlated–differencing model provides an attractive approach to the understanding of FM detection near the detection level.

The most serious shortcoming of the correlated–differencing model for FM detection appears to be the assumption that successive differencing operations are statistically independent. We are quite sure that this assumption is incorrect when the modulation excursion is so wide that detection performance is significantly above the detection level. Intuitively one feels that near 100% detection the successive sampling operations are phase locked, and our experimental psychometric functions, steeper near 100% correct than the model predicts, corroborate one's intuition. The assumption of uncorrelated–differencing operations may even be incorrect at the detection level, as suggested by the most straightforward interpretation of our experiment 1, which monitors detection performance as a function of the number of stimulus modulation cycles. However, the assumption of uncorrelated–differencing operations has the virtue of mathematical simplicity. Because the theory of this paper seems to be the first realistic attempt to understand FM detection at low modulation frequencies the mathematical simplification seems warranted. Improvement of the model to include correlated–differencing is an interesting challenge.

There are two other research directions suggested by the general sampling–differencing formalism. It would be interesting to try to extend the model beyond the regime of low modulation frequencies. For modulation frequencies greater than about 6 Hz the auditory sampling time is probably not short compared with the modulation period. In that event the effective frequency probability density is different from the ideal given in Eqs. (7)–(10). In any reasonable sampling model the probability density must shrink in width for increasing modulation frequency resulting in decreased detection performance.

Finally the general formalism of the sampling–differencing model is not limited to FM detection but could be applied to AM detection or to the detection of any periodic stimulus variation in the low-frequency regime.

ACKNOWLEDGMENTS

We are grateful to Professor David M. Green for bringing the discrepancy between pulsed-sines frequency discrimination and FM detection to our attention. Professor Larry Feth and Dr. Walt Jesteadt provided helpful information and suggestions. Supported by NSF grants BNS 76-20225 and BNS 79-14155. One of the authors (M. A. Klein) was supported in part by NIMH National Research Service Award 5T 32 MH14622.

Abramowitz, M., and Stegun, L. A. (1964). *Handbook of Mathematical Functions* (National Bureau of Standards, Government Printing Office, Washington, DC).
Fastl, H. (1978). "Frequency Discrimination for Pulsed vs. Modulated Tones," J. Acoust. Soc. Am. 63, 275–277.
Feth, L. L., Wolf, R. V., and Bilger, R. C. (1969). "Frequency Modulation and the Difference Limen for Frequency," J. Acoust. Soc. Am. 45, 1430–1437.
Green, D. M., and Swets, J. A. (1966). *Signal Detection Theory and Psychophysics* (Wiley, New York).
Groen, J. J., and Versteegh, R. M. (1957). "Frequency Modulation and the Human Ear," Acta Otolaryngol. 47, 421–430.
Harris, J.D. (1952). "Pitch discrimination," J. Acoust. Soc. Am. 24, 750–755.
Hartmann, W. M. (1977). "Five Experiments on Frequency Modulation Width Perception," J. Acoust. Soc. Am. Suppl. 1 61, S50(A).
Hartmann, W. M. (1978). "Detection of Frequency Modulation at Large Modulation Frequencies," J. Acoust. Soc. Suppl. 1 63, S66(A).
Hartmann, W. M. (1979). "Detection of Amplitude Modulation." J. Acoust. Soc. Am. Suppl. 1 65, S58(A).
Henning, G. B. (1966). "Frequency Discrimination of Random-Amplitude Tones," J. Acoust. Soc. Am. 39, 336–339.
Henning, G. B. (1967). "A Model for Auditory Discrimination and Detection," J. Acoust. Soc. Am. 42, 1325–1334.
Henning, G. B. (1970). "A comparison of the effects of signal duration on frequency and amplitude discrimination," in *Frequency Analysis and Periodicity Detection in Hearing*, edited by R. Plomp and G. F. Smoorenburg (Sijthoff, Leiden).
Jesteadt, W., and Sims, S. L. (1975). "Decision processes in frequency discrimination," J. Acoust. Soc. Am. 57, 1161–1168.
Kay, R. H., and Matthews, D. R. (1972). "On the Existence in Human Auditory Pathways of Channels Selectively Tuned to the Modulation Present in Frequency-Modulated Tones," J. Physiol. 225, 657–677.
Klein, M. A., and Hartmann, W. M. (1979). "Perception of Vibrato Width," in *Proceedings of the Research Symposium on the Psychology and Acoustics of Music*, Lawrence, Kansas (to be published).
Kock, W. E. (1937). "A new interpretation of the results of experiments on the differential pitch sensitivity of the ear," J. Acoust. Soc. Am. 9, 129–134.
Levitt, H. (1971). "Transformed up–down methods in psychoacoustics," J. Acoust. Soc. Am. 49, 467–477.
Moore, B. C. J. (1976). "Comparison of frequency DL's for pulsed tones and modulated tones," Br. J. Audiol. 10, 17–20.
Nordmark, J. O. (1968). "Mechanisms of frequency discrimination," J. Acoust. Soc. Am. 44, 1533–1540.
Siebert, W. M. (1968). "Stimulus transformations in the peripheral auditory system," in *Recognizing Patterns: Studies in Living and Automatic Systems*, edited by P. A. Kolers and M. Eden (MIT, Cambridge, MA).
Shower, E. G., and Biddulph, R. (1931). "Differential pitch sensitivity of the ear," J. Acoust. Soc. Am. 3, 275–287.
Stevens, S. S. (1954). "Pitch discrimination, Mels and Kock's contention," J. Acoust. Soc. Am. 26, 1075–1077.
Verschurre, J., and van Meeteren, A. A. (1975). "The effect of intensity on pitch," Acustica 32, 33–44.
Wier, C. C., Jesteadt, W., and Green, D. M. (1977). "Frequency Discrimination as a Function of Frequency and Sensation Level," J. Acoust. Soc. Am. 61, 178–184.
Zwicker, E. (1952). "Die Grenzen der Horbarkeit der Amplitudenmodulation und der Frequenzmodulation Eines Tones," Acustica 2, AB125–133.