Comparison of stimuli used in sound field audiometric testing

Harvey Dillon and Gary Walker

National Acoustic Laboratories, 5 Hickson Road, Millers Point, NSW 2000, Sydney, Australia
(Received 20 November 1980; accepted for publication 11 August 1981)

The problems associated with performing pure-tone threshold measurements in reverberant or diffuse sound fields are illustrated with the use of three-dimensional representations of the sound field within a typical test booth. A microphone mounted on a motorized trolley is used to perform these measurements. A comprehensive comparison is then made of the efficiency with which FM tones, AM tones, damped wave trains, and narrow bands of noise provide a uniform sound field. The conclusion is reached that the bandwidth of the stimulus is the major factor determining the uniformity of the field. A decision about the most appropriate stimulus for sound field work must thus be based on factors other than field uniformity. When the constraints of obtaining suitable spectral distributions, and being able to relate thresholds obtained with complex stimuli to those obtained with pure tones are also considered, FM tones and suitably generated narrow bands of noise appear to be the most suitable stimuli. The selection of suitable parameters is discussed and an Appendix discusses the spectrum and bandwidth of FM signals with different modulation waveforms. The relative accuracy of testing in the direct and reverberant regions in a nonanechoic environment is also discussed.

PACS numbers: 43.66.Yw, 43.66.Sr, 43.66.Cb [JH]

INTRODUCTION

Sound field audiometric procedures have long been used in audiology in order to estimate the thresholds of young children and others who will not tolerate earphones. They are also increasingly used to evaluate the functional characteristics of hearing aid systems. Sound field measurements obtained with the subject wearing the hearing aid with his own custom earmold provide the only accurate indication of the "real ear" performance of the aid.

Although many authors have advocated the use of sound field procedures, relatively little attention has been paid to the associated problems. Pure tones are not considered appropriate stimuli for use in the sound field because the interaction between direct and reflected sound in the test enclosure sets up standing wave patterns with the result that there may be considerable variations in the sound pressure level of the same signal at different locations in the room. This problem is overcome if true anechoic testing conditions are available but this is not practicable in most situations. Several other frequency-specific stimuli have been suggested for use in sound field audiometry. The best known of these are frequency modulated tones (warble tones) and many audiometers include a frequency modulation facility. There seems, however, to be no general consensus among manufacturers as to the most appropriate parameters of the frequency modulation provided (Staab and Rintelmann, 1972). Although several studies have addressed themselves to the question of appropriate parameters for FM stimuli (see Morgan *et al.*, 1979 for a review), the emphasis has been on the effects of the parameters on the threshold obtained. A systematic study on the effects of different parameters (and different stimuli) on the uniformity of the field does not appear to have been made. Other possible stimuli include narrow-band noise, damped wave trains (Victoreen, 1974) and amplitude modulated tones (Goldberg, 1979).

The study to be reported here involves a systematic examination of the performance of these various stimuli in a typical audiometric test enclosure with a view towards making practical recommendations regarding suitable stimulus parameters.

What are the desirable properties of stimuli to be used in sound field audiometry?

(1) The stimulus must be reasonably frequency specific. Frequency selective information is usually required for a satisfactory definition of the hearing status of the patient. The pure tone clearly meets this requirement whereas white noise clearly does not. Other stimuli lie somewhere between these extremes and a decision has to be made about what constitutes adequate specificity. This may vary from frequency to frequency. For example, a $\frac{1}{3}$-oct. band of noise centered on 500 Hz covers the frequency band 448 to 565 Hz, which might be considered adequately narrow for most audiometric purposes. However, a $\frac{1}{3}$-oct. band of noise centered on 4 kHz covers the frequency region 3580 to 4520 Hz. Since considerable and important changes in the threshold of hearing could occur within this range, this stimulus may not be adequately frequency specific in many instances.

(2) The sound pressure level generated in the field must be stable over the whole region which the head could occupy during testing. With the trunk immobile an adult human can move his ear through a distance of about 20 cm by forward and backwards tilt of the head and through about 8–10 cm by sideways tilt of the head. Even when a head rest is employed head movements of several centimeters can occur. With young children it is difficult even to keep the ear within 20 cm of its original position. Thus, the sound field should not vary excessively within a region of this size.

(3) The sound pressure level generated in the field should be stable for small shifts in frequency. Any particular room/loudspeaker/listening position combi-

 0001-4966/82/010161-12$00.80

nation will have a unique frequency response characteristic. Although this is compensated for during calibration of the room, small drifts in frequency may occur between calibrations of the audiometer and it is important that such drifts, which may be up to 5%, do not cause marked changes in the sound pressure level at the measuring point in the room.

(4) It is desirable that measurements obtained in the sound field are able to be related to similar measures obtained with pure tones under earphones.

I. THE EQUIPMENT

All of the experiments were carried out in an audiometric test room (I.A.C. model 404A) having internal dimensions of $2.92 \times 2.75 \times 1.98$ m high. The room is of standard construction but is probably somewhat larger than the average booth used in audiological clinics. A 12-in. loudspeaker (Plessey–Foster model 300K60) was placed at the center of one of the long walls of the room and facing into the room so that its cone was 1.04 m above the floor. A variety of stimuli was generated outside the room and used to excite the loudspeaker via a Technics model 7300 amplifier. For one series of measurements a high-frequency speaker (Philips model AD0160/T8 "tweeter") replaced the 12-in. speaker.

A $\frac{1}{4}$-in. condenser microphone (Bruel & Kjaer type 4135) was mounted on a carriage which itself was mounted on an extruded aluminum rail. This equipment is shown in Fig. 1. The carriage was designed to move along the rail in either direction driven by a variable speed dc electric motor and pulley system. Carriage movement was controlled from outside the room. The rail was elevated by supports so that the microphone mounted on the carriage was the same distance above the floor of the room as the center of the speaker cone. The effects of the rail and supports on the sound field were checked by performing some measurements with the microphone on its carriage and suspended by a cord from the ceiling. The differences were less than 1 dB for all frequencies up to 10 kHz. The output of the microphone was passed through a microphone amplifier (Bruel & Kjaer 2603) thence through a bandpass filter for the constant frequency experiments (Bruel & Kjaer type 2121 frequency analyzer), thence through a square law device and low-pass filter to a level recorder (Bruel & Kjaer type 2305). The role of the square law device and low-pass filter is described briefly later in this paper and more extensively in Dillon and Walker (1980).

It is possible with the above equipment to measure the sound field for a single frequency signal over a continuous range of positions along a straight line, or to measure a continuous range of frequencies at one position.

II. EXPERIMENTS WITH PURE TONES

A series of measurements was carried out with pure tones. The stimuli were generated by a beat frequency oscillator (Bruel & Kjaer type 1022) with the frequency monitored on a frequency counter (Fluke 1900A). By measuring the sound field in the room for one frequency over the full path of travel of the microphone along the rail, and repeating this a number of times with the rail progressively displaced to the left and right of the speaker axis in 5-cm steps, it was possible to build up a three-dimensional plot of the sound field in the horizontal plane passing through the speaker axis (for this particular frequency). The results of one such plot for a frequency of 1200 Hz are presented in Fig. 2. The three dimensions are width of room × length of room × intensity (or SPL) of sound field.

The sound field generated within the room can be divided into two regions, the direct and reverberant fields. In the immediate vicinity of the speaker the sound field is dominated by sound radiated directly from the speaker. This region is not influenced by room effects, and intensity within the field varies according to the inverse square law. However, at some distance from the speaker, reverberant energy begins to have an influence on the field and at greater distances from the speaker the field is dominated by reverberant energy. The extent of the direct and reverberant fields will depend on a number of factors including the directivity of the speaker, the dimensions of the room, and the absorptivity of the walls. For the field depicted in Fig. 2 the region dominated by the direct field is represented by the pronounced peak in the cen-

FIG. 1. Microphone mounted on a cable-drawn trolley used to determine sound field distribution within the test room.

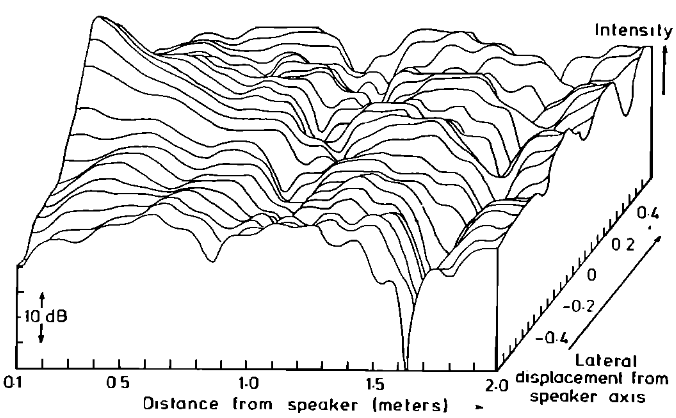

FIG. 2. Sound intensity variations within the test room at a frequency of 1200 Hz. The loudspeaker is located 0.1 m to the left of the center of the left-hand side "wall."

ter of the left-hand edge of the figure. It extends for about 60 cm in front of the speaker and is about 30 cm wide at a point 10 cm in front of the cone.

In the reverberant field, sound pressure level does not appear to vary with increasing distance from the source in any systematic way. This field is characterized by a complex series of peaks and troughs with two pronounced troughs running across the room more or less at right angles to the speaker axis. The troughs on the right-hand face of the figure are directly related to the wavelength of the signal, being approximately 29 cm apart. The other peaks and troughs, however, show no simple relationship to signal wavelength.

As the frequency is changed, the distribution of intensity within the room changes. Clearly, it is not practicable to present a picture like Fig. 2 for all possible frequencies. Instead, the next series of figures was obtained with the microphone track on the speaker axis. With the microphone stationary, the oscillator was swept through a range of frequencies and the microphone's response recorded. This procedure was repeated a number of times with the microphone progressively moved away from the speaker in 5-cm steps. This provided a three-dimensional plot of frequency × distance × intensity. The automatic scanning facility of the oscillator was used to sweep through the chosen frequency range. After amplification, the microphone output was recorded on the level recorder. Figures 3–5 show the results of the measurements for the low-(250–500 Hz), medium-(1–2 kHz), and high-(3–6 kHz) frequency ranges.

For low frequencies (from about 250 to 350 Hz) the field along the path traversed by the microphone is virtually free from peaks and troughs. The intensity of the field falls smoothly with increasing distance from the speaker, and there is little change in field strength from frequency to frequency. For higher frequencies, pronounced peaks and troughs begin to appear in the field. The variations in field strength can be quite large, for example, 28-dB peak to trough at 1.12 kHz at the furthermost measuring position (see Fig. 4). What is more, in Figs. 3 and 4 the position of the peaks and troughs in relation to frequency vary systematically with distance from the speaker. In Fig. 5, however, it is clear that the peaks and troughs retain constant frequencies as distance from the speaker is varied. This indicates that for these higher frequencies, the irregularities in the response are being dominated by the frequency response of the loudspeaker rather than by the room acoustics. Further evidence for this comes from the fact that the peaks and troughs are just as prominent in the direct field as in the reverberant field. To separate the room and speaker effects, the measurement for this frequency range was repeated with a high-frequency speaker (Phillips AD0160/T8). These results are shown in Fig. 6.

This figure is characterized by many sharply tuned troughs. Again, the position of these troughs in relation to frequency appears to be dependent on distance from the speaker. Because of the smaller wavelengths at these frequencies, much fine detail is lost when 5-cm

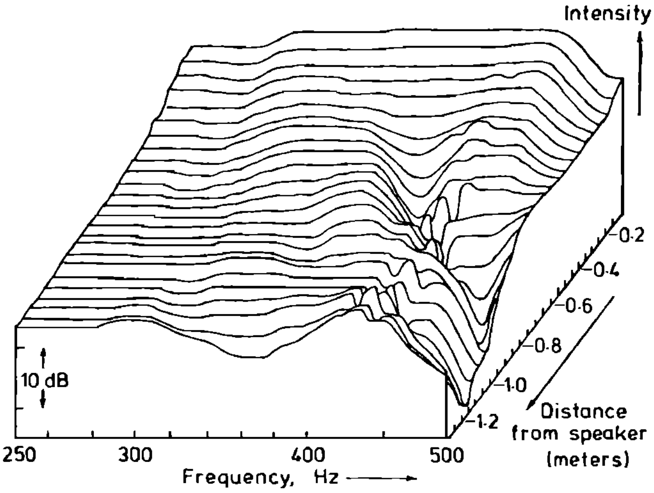

FIG. 3. Sound intensity variations along the speaker axis for frequencies between 250 and 500 Hz. The results were obtained by performing frequency sweeps at successive microphone positions.

steps are used.

The frequency–distance relationship can be seen more clearly in Fig. 7 which shows a frequency sweep between 4 and 5 kHz over 1-cm steps. This represents a more detailed "enlargement" of the rectangular area in the center of Fig. 6.

Clearly pure tones are not appropriate signals for sound field testing, at least in the reverberant field. Small shifts in the head position and/or drifts in frequency can result in significant changes in field strength at the subject's ear.

III. EXPERIMENTS WITH COMPLEX SIGNALS

A. Description of signals

The complex signals which we have investigated include frequency modulated tones, amplitude modulated

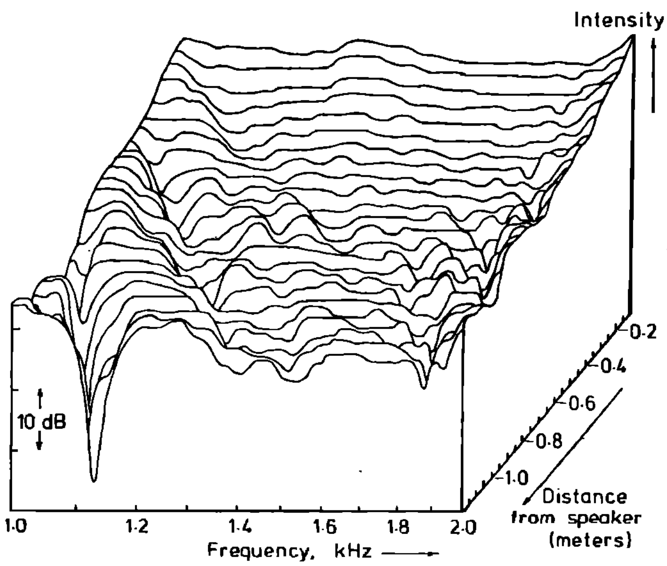

FIG. 4. Same as for Fig. 3 but for the 1- to 2-kHz frequency range. Note the clear frequency-distance inter-relationship of the sound field "troughs."

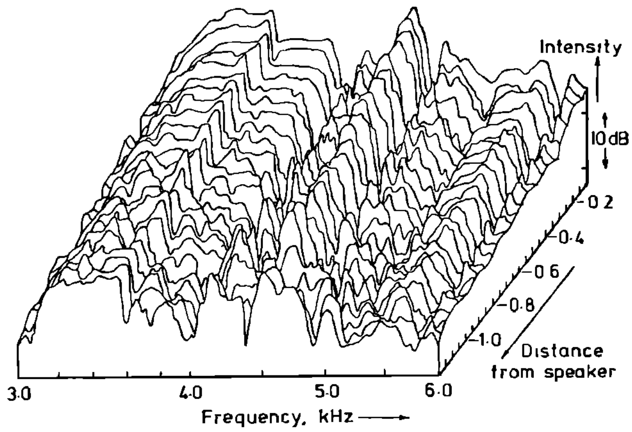

FIG. 5. Same as for Fig. 3 but for the 3- to 6-kHz frequency range. The parallel troughs extending into the direct field indicate a poor loudspeaker response, not the effect of room modes.

tones, damped wave trains, and filtered noise. For convenience, continuous signals were used throughout this study. Although gated signals are normally used in audiometry, the gated signals have a duration far greater than the usual reverberation times for sound treated rooms. Steady-state conditions will thus be set up in the sound field and the results reported here will be directly applicable.

1. Frequency modulated (FM) signals

The carrier frequency may be modulated by a variety of waveforms (e.g., a square wave, a triangular wave, a sine wave, or a ramp). Staab and Rintelmann (1972) reported a survey which indicated that, at that time, audiometer manufacturers were almost equally divided between using sinusoidal and rectangular modulation. We know of no more recent published figures. We have chosen to use ramp modulated tones in our program because this is all that is available on the beat frequency oscillator used, and we particularly wished to utilize the frequency sweeping facilities of this instrument. Two other parameters of the signal need to be specified, namely modulation rate and frequency deviation. The

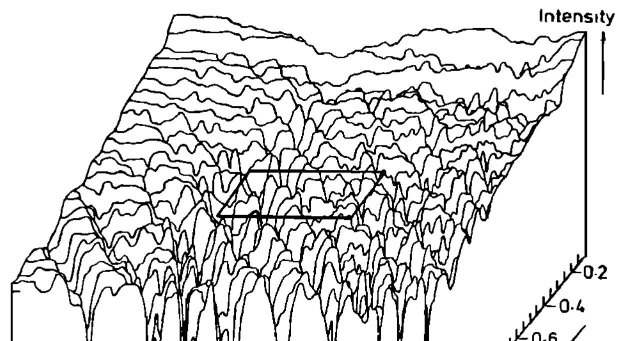

FIG. 6. Same as for Fig. 3 but for the 3- to 6-kHz frequency range and using a high-frequency loudspeaker.

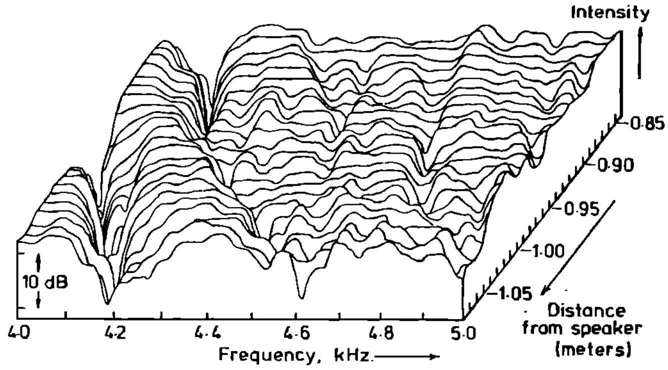

FIG. 7. An expanded version of the sound intensity variations within the small rectangle shown in Fig. 6.

bandwidth of the signal will be almost independent of the modulation rate provided the modulation rate is moderately less than the frequency deviation. Under this same condition, the bandwidth increases in proportion to the frequency deviation. (See the Appendix for a discussion of the bandwidth of FM tones.) Frequency deviations of between 16 and 250 Hz and a modulation rate of 10 Hz (except where stated otherwise) were used.

2. Damped wave trains

Damped wave trains (DWT's) are exponentially decaying sinusoidal tone bursts presented at a repetition rate of approximately 2/s (Victoreen, 1974). The particular ones used by Victoreen had a decrement of 0.9 (that is the intensity of each peak is 0.9 times the intensity of the preceding peak). This corresponds to a Q of 30 or a relative bandwidth of 3.4%.

Damped wave trains are most readily generated by stimulating a bandpass filter with either a narrow pulse train, or a square wave. The energy distribution of the stimulus can thus be easily controlled by varying the shape of the filter response. In this respect the DWT is superior to other forms of tone burst. The spectrum of the DWT is simply equal to the spectrum of the input waveform multiplied by the filter response. It thus consists of components spaced apart by the frequency of the square wave or pulse train, but with amplitudes given by the bandpass filter response characteristic.

In DWT's the number of cycles in each burst remains constant (provided relative bandwidth is held constant) and is independent of frequency. Thus the duration of a 100-Hz DWT is 60 times that of a 6-kHz DWT. The other point to note is that for practical bandwidths, DWT's are very brief duration signals. Even for very low frequencies, the stimulus duration is shorter than the typical integration time of the normal human ear (approximately 200 ms).

The damped wave trains used here were generated by passing a 1-Hz square wave through the bandpass filter of a frequency analyzer (B & K type 2121). This is different from the method employed by Victoreen (1974) but very similar to that used by Wit (1978). Bandwidths of 1%, 3%, 10%, and $\frac{1}{3}$ oct. were used.

3. Narrow-band noise

Narrow-band noise (NBN) was generated by passing the output of a white noise generator (Grason Staedler model 1285) through the bandpass filter of the same frequency analyzer used to generate the DWT's. The bandwidths and skirt slopes were thus also the same as for the DWT's.

4. Amplitude modulated tones

A simple (sinusoidally) amplitude modulated pure tone is seen spectrally as a three tone complex. The two extra tones (sidebands) are positioned on each side of the nominal frequency and their spacing from the nominal frequency is equal to the modulation frequency. Of course, the sidebands will only be resolved if the analyzing instrument has a bandwidth less than the modulation frequency, and if the signal duration is sufficiently in excess of the modulation period. Goldberg (1979) has proposed the use of signal composed of a pure-tone amplitude modulated by two low-frequency incoherent signals.

The amplitude modulated signal used here was generated by passing two pure tones through an unbalanced multiplier (Analog Devices model 426A). A modulation depth of 100% was used. The bandwidth is altered by varying the frequency of the modulating tone. Five modulation rates were chosen to provide bandwidths equivalent to those of the other stimuli.

B. Effect on field uniformity

The position of nulls in the room is dependent on frequency. Thus, if one were to measure the field for a number of adjacent frequencies and average the results, the effect of a null at any one frequency would be largely negated. This is the principle underlying the use of complex signals to overcome room effects. A complex signal will be relatively unaffected by a null if it has significant energy within a frequency band which is wider than the null. We may thus expect that the efficacy with which a complex signal removes the irregularities from a room's response will be largely determined by the spectral bandwidth of the signal. The distribution of energy within the band may also have some effect if it is markedly different from a uniform distribution.

The transmission through a room of any stimulus other than a sine wave results in a signal whose amplitude varies over time, even for a fixed measuring position. How are such fluctuations to be handled when these stimuli are measured? Clearly, some sort of temporal averaging is required, so that a single value for the SPL may be assigned to the stimulus. To enable accurate threshold determination, the averaging should correspond to the averaging performed by the human ear. In this way, the minimum audible signal can be directly related to the subject's pure-tone threshold. Thus, for all the signals being evaluated, the output of the measuring microphone was passed through a system comprising a square law device and low-pass filter with a 200-ms time constant. This system integrates energy over time according to a loudness and/or threshold model of the human ear (Plomp and Bouman, 1959). The output of the device therefore represents that which a human subject would perceive in terms of loudness fluctuations. A full description of the threshold model used is presented elsewhere (Dillon and Walker, 1980).

1. Fixed frequency data

Two types of data were collected. For the first set of measurements the measuring microphone was moved along the speaker axis from a point just in front of the speaker to a point just in front of the opposite wall. This provides a continuous record of the sound pressure level along this axis. One center frequency was tested, namely 1350 Hz. This frequency was chosen on the basis of pilot work and results in a medium amount of field variability for the microphone positions employed. The records obtained were divided into nine equal segments, each of width 20 cm. This distance was chosen to be representative of the distance which an unrestrained head may move during a sound field evaluation. By noting the falloff in intensity as the microphone moved away from the speaker, it was possible to classify the field in the back five segments as being predominantly reverberant, and the field in the front three segments as being predominantly direct. A typical example is shown in Fig. 8. The maximum deviation within each of the five segments furthermost from the speaker were determined and averaged to provide a single index of variability for each stimulus in the reverberant field. The results of this procedure are shown in Fig. 9. Two trends in the data are evident. Firstly, as expected, variability decreases as the stimulus bandwidth increases. That is, a more uniform field is produced in the room by wider bandwidth stimuli. Note that as the bandwidth increases from 1% to 23%, the variability decreases from a value close to that appropriate for a pure tone to a value close to that appropriate for white noise. (The frequency specificity of the latter stimulus, of course, leaves much to be desired.) The other trend apparent in the figure is that for a given bandwidth, some stimu-

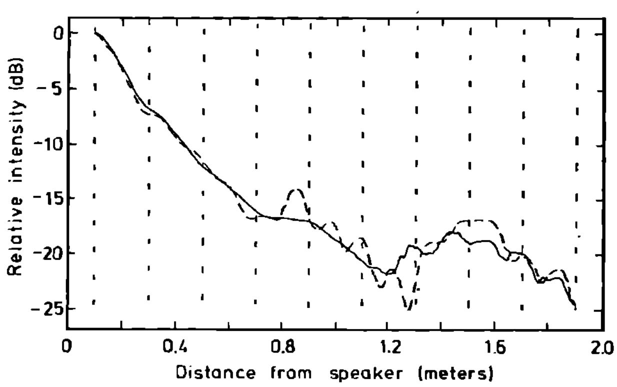

FIG. 8. Sound intensity variations occurring along the speaker axis at a center frequency of 1350 Hz. The dashed curve shows the results for a pure tone and the solid curve the results for a ±5% frequency modulated tone. The dotted vertical lines show the 20-cm "windows" within which the variability was calculated.

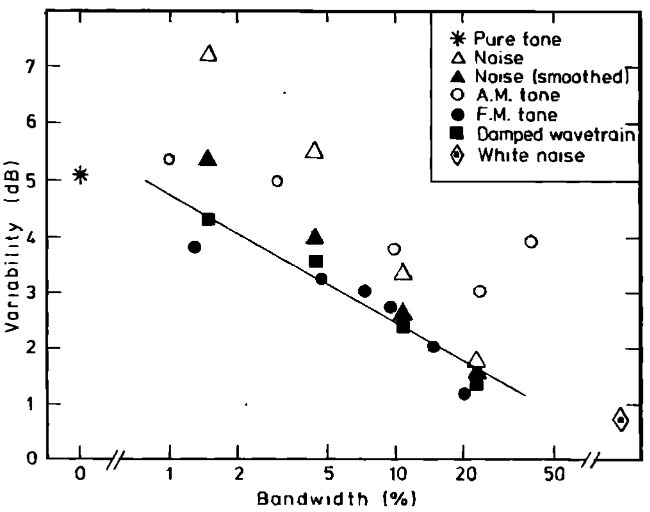

FIG. 9. Average variability within the 20-cm windows in the reverberant field as a function of the effective bandwidth for several types of stimuli.

li produce a more uniform field than others. Specifically, AM tones and NBN seem to produce a more variable field than FM tones, DWT's and "smoothed" NBN. These last three (shown as closed symbols in Fig. 9) can be seen to be equally efficient at reducing the field variability. "Noise" and "smoothed noise" results actually originate from the same stimuli, but differ in the method of measurement, as follows. The "noise" figures were obtained by passing the signal through the averaging system described above which deals with intensity fluctuations in the same way as would a human subject. The "smoothed noise" figures were obtained by passing the signal through the averaging device but with the time constant of the averager set to 1.0 s. This latter data is included because it shows the "pure" effects of the room modes unaffected by random fluctuations in signal level, and is therefore more appropriate for investigating the relationship between signal bandwidth and room effects. For "smoothed noise" the averaging is roughly equivalent to that provided by the "slow" response mode of a standard sound level meter.

Unsmoothed NBN is inferior to all other stimuli for bandwidths up to and including 5%. This can be explained in terms of the random amplitude fluctuations which are a characteristic of noise. These fluctuations are relatively slow at the smallest bandwidths, and consequently are not smoothed out by the low-pass filter to the same extent that more rapid fluctuations are. For higher center frequencies, the absolute bandwidth of a 5% noise band will be larger. The temporal fluctuations in intensity will therefore be faster and so will increasingly be smoothed by both the ear and the calibrating device. Thus for the higher audio frequencies, even the narrower bands of noise can be expected to produce variability similar to the other stimuli.

The relatively large variability observed for the AM tones is readily explained by the distribution of spectral energy which this stimulus possesses. Since a sinusoidally amplitude modulated signal consists of only three tones, the averaging effect (which is achieved by using a stimulus with many components) is not very effective. The type of AM proposed by Golberg would be expected to have a variability between the AM figures presented here and those for FM and DWT's. The AM stimulus was not included in the next measurement procedure.

2. *Fixed position data*

For the second set of measurements, the measuring microphone remained fixed at 1 m from the speaker and the frequencies were swept between 1 and 2 kHz. These measurements provide a frequency response curve over the limited frequency range for this one position. The frequency range and position were selected on the basis of a pilot investigation. They were again chosen to provide a moderate amount of field variability (for small changes in frequency). The frequency response curves obtained were divided into ten equal segments (each having a width of 7% of its center frequency) and the maximum deviation within each segment was determined.

These ten figures were then averaged to provide a single index of variability for each type of stimulus for each bandwidth. An example of two typical curves (for a pure tone and for 10% bandwidth DWT) are shown in Fig. 10. Processing of the results for the FM stimulus was slightly more complex than for the others since for this stimulus, absolute rather than percentage bandwidth remained constant as the frequency swept. The figures from sweeps with different FM parameters were combined in such a way that equal percentage bandwidths were combined at all times in the averaging process. Results for all stimuli are shown in Fig. 11.

Trends very similar to those noted in the distance domain are again observed. For this particular measurement location, field uniformity (with respect to small changes in frequency) improved as the stimulus bandwidth increased. Unsmoothed noise was again more variable, increasingly so for the narrower bandwidths.

IV. DISCUSSION

A. Choice of signal type

For all signals there is a clear trade-off between bandwidth and variability. Thus, comparisons between

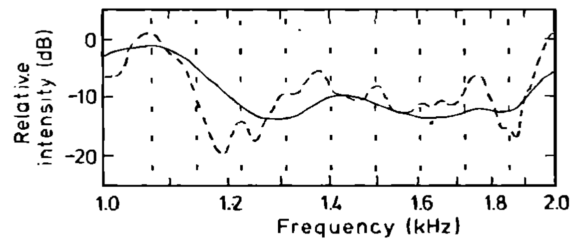

FIG. 10. A typical variation of sound intensity with frequency (for a point 1 m in front of speaker). The dashed curve shows the result for a pure tone and the solid curve for a 10% bandwidth damped wave train. The dotted vertical lines show the 7% wide frequency regions within which the variability was calculated.

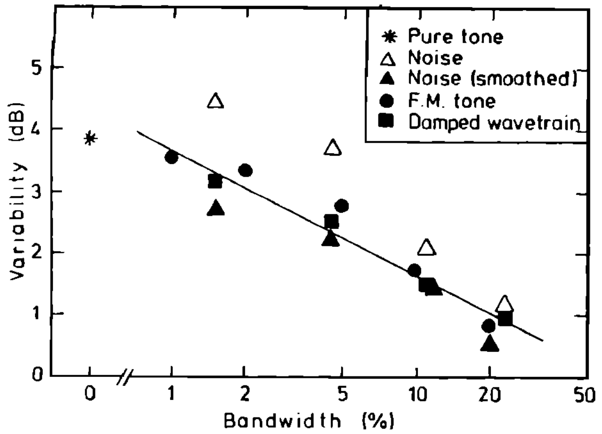

FIG. 11. Average variability within the 7% frequency windows shown in Fig. 10 as a function of the effective bandwidth for several types of stimuli.

different signals must clearly be made at the same bandwidth. When Morgan *et al.* (1979) for example, found that narrow-band noise was superior from the point of view of field uniformity, they were comparing a 10% bandwidth FM tone with a 23% bandwidth noise band, and so invalidly concluded that noise was superior.

What then are the choices for sound field testing? Pure tones and sinusoidally modulated AM tones can be excluded on the ground that they provide a more variable reverberant sound field than either DWT's, noise, or FM. If very narrow bandwidths are required, especially at low and medium frequencies, filtered noise can be likewise excluded on the grounds that perceived variability is greater than for DWT's or FM. That is, although a slow response (long averaging time) can be selected when the signal is being calibrated, a long averaging time *cannot* be selected in the perceptual apparatus of a human subject. A signal whose intensity is varying randomly and slowly (such as very narrow-band noise) will thus be perceived differently from one moment to the next (de Boer, 1966). In particular, the average and peak intensities of a 1- or 2-s burst of such a signal will be somewhat different for different presentations to a subject. These fluctuations will therefore increase the variability of the threshold determining process. Whether the increase will be clinically significant (or even measurable) will depend on the subject's time constant of integration, and the bandwidth and duration of the signal.

For the larger (and probably more useful) bandwidths tested, the choice between DWT's, FM, and NBN must be made on the basis of factors other than the variability which they produce in the test room. For clinical use, two such factors are evident.

1. Shape of signal spectrum

Hearing loss, especially in the high-frequency region can change rapidly with only small changes in frequency. It is thus desirable that a stimulus not only have as narrow a bandwidth as practicable, but that energy outside the bandwidth limits falls off as rapidly as possible. To achieve this for the DWT stimuli, the filter used in generating the stimuli must have fairly steep skirts. Steep filter skirts are also a requirement if NBN is generated by filtering wideband noise as we have done here. For such a stimulus, Orchik and Mosher (1975) have shown that insufficiently steep filter slopes can have a large effect on the measured threshold. While sufficient selectivity can be achieved at low frequencies with fairly simple low-order filters, higher frequency stimuli may well require the use of more complex higher order filters. Such filters normally require very careful alignment—a disadvantage in a clinical instrument.

Narrow-band noise can also be generated by other methods, thus removing the need for high-order filters. A modulation technique has been proposed by Zwislocki (1951) in which a low-pass band of noise is used to amplitude modulate a sinusoid at the desired center frequency. The resulting band of noise has a bandwidth twice that of the low-pass noise, and skirts much steeper than that of the low-pass noise. Alternatively, NBN can be synthesized by adding together a number of closely spaced discrete components (Schafer *et al.*, 1950). For such a stimulus, the skirt slope is limited only by the energy splatter that arises when noncontinuous tones are used. (This is minimized in audiometers by the signal durations and rise/fall times selected.)

If either of these two generating techniques are used, then NBN has a suitable spectral distribution for audiometric purposes, since both the in-band and out-of-band distributions can be readily controlled.

For the generation of FM stimuli no filter is required. The spectral shape of an FM signal depends only on the modulation rate, waveform, and frequency deviation. Thus, provided these are fixed, and the carrier frequency is moderately greater than the frequency deviation, the spectral bandwidth and shape also remain fixed as the carrier (or center) frequency is varied. Because of this, very steep slopes can be readily generated, especially at the higher frequencies where they are most required. At a center frequency of 2 kHz for example, a modulation rate of 10 Hz and a sinusoidal frequency deviation of ±160 Hz leads to a skirt slope of 400 dB per octave! The spectral properties associated with different modulation waveforms are discussed in the Appendix.

In summary, with respect to the stimulus skirt slope (out-of-band component rejection), FM tones and suitably generated narrow bands of noise are superior to both DWT's and NBN obtained by direct filtering. With respect to the in-band spectral distribution, sinusoidally frequency-modulated tones are somewhat less suitable than DWT's, NBN, or FM tones with a triangular modulation waveform.

2. Relationship to other audiometric data

Hearing loss has traditionally been measured with sinusoidal stimuli having a duration sufficiently long that their actual duration has no effect on the results obtained. *All* of the stimuli being evaluated here contain some temporal fluctuation. Periodic fluctuations are inherent in the generation of AM and DWT stimuli, as are random fluctuations for the NBN. The FM tones contain no fluctuations when generated, but do so after passing through the room. As the instantaneous frequency sweeps over the various room modes, the intensity is modulated depending on the transfer function for that particular loudspeaker/microphone setup. The relationship between FM stimulus thresholds and pure-tone thresholds has been investigated elsewhere by Morgan *et al.* (1979), and that for DWT's by Wit (1978). A more extensive comparison for the stimuli compared here is the subject of another paper (Dillon and Walker, 1980). As a general rule, thresholds for pure tones can be most accurately inferred from measurements made using other narrow-band stimuli if the other stimuli contain a minimal amount of temporal fluctuation. There are several studies which show that the well known "law of temporal integration" of short stimuli, which is well established for normal hearers, is grossly changed for many types of hearing loss (e.g., Elliott, 1963; Gengel and Watson, 1971; Pederson and Elberling, 1973; Sanders *et al.*, 1971; Wright, 1968). For practicable bandwidths, DWT's are quite brief stimuli. The thresholds obtained by using these stimuli could thus not be used to accurately predict the thresholds that would have been obtained had sinusoids been used. Furthermore, the thresholds obtained with DWT's will depend on the duration (and hence bandwidth) of the stimulus. As arbitrary choice must thus be made of a suitable duration. By comparison, both FM stimuli and noise have much smaller temporal fluctuations, and a more accurate prediction of the threshold for a continuous sinusoid is thus possible. In this respect then, DWT stimuli are inferior to either NBN or FM stimuli.

A proviso must be added here that the above argument applies only if one wishes to express hearing loss in terms of long duration stimuli. It has been argued by Victoreen (1974) that thresholds for brief stimuli may be more indicative of hearing performance for speech stimuli, since many consonants are of short duration. The value of brief tone audiometry in the diagnosis of the underlying pathology of hearing defects has also been shown (Pederson, 1976). If, for these reasons or others, one wishes to perform brief tone audiometry, DWT's are an excellent choice of stimuli (except for the measurement of sharply sloping losses). Bandwidths larger than the 3% suggested by Victoreen (1974) would seem to be necessary if the inaccuracies caused by standing waves are to be avoided.

3. Summary

If one wishes to obtain a "conventional" audiogram it would appear that either frequency modulated tones or suitably generated narrow bands of noise are the most suitable stimuli for sound field testing. For the NBN stimuli, there may be difficulties associated with the slow (and therefore perceptible) fluctuations that arise when very narrow bandwidths are used at low or mid frequencies. For the FM stimuli, there may be difficulties associated with the nonuniform in-band spectral distribution when sinusoidal modulation is used, or with the more gradual out-of-band rejection rate when triangular modulation is used. We plan to address these problems more quantitatively in a later paper.

B. Choice of signal parameters

1. Bandwidth

The relationship between smoothness of response and bandwidth is shown in Figs. 9 and 11. (For FM tones, the Appendix shows how bandwidth is related to frequency deviation.) Note that although the general trend between bandwidth and variability will hold for all situations, the particular relationship depicted here will be specific to the particular room, source and measurement positions, and frequencies tested. In general, a bandwidth is required that is wider than the major peaks and dips in the transfer function from the loudspeaker to the listener. When this condition is satisfied, the peaks and troughs are largely removed by the averaging that occurs across frequencies. Some room/loudspeaker combinations will obviously be better than others in this regard so it is not possible to stipulate a universal minimum necessary bandwidth, even for one frequency. A comparison of Figs. 3, 4, and 7 will reveal that the minimum bandwidth required to span the troughs in each of the frequency ranges is not a constant fraction of the center frequency of the tone. For our room it would appear that the commonly used frequency deviation of ±5% provides sufficient averaging across frequencies only for the highest frequency range tested. An upper limit to the bandwidth comes from the need to test people with hearing losses that vary rapidly with frequency. Changes in thresholds of up to 70 dB between adjacent audiometric frequencies (e.g., 1 and 2 kHz) are not uncommon. Where the possibility of such an audiometric configuration exists it may be necessary to use narrower bandwidths at the expense of some imprecision caused by standing wave effects, at least in the frequency region of concern. The accurate measurement of a sloping hearing loss in an ill suited room does not appear to be an easy matter. We expect that the bandwidth that is optimum for most purposes will prove to be neither a constant proportion of center frequency nor a constant absolute bandwidth. Further work is required in this area. Further work is also required so that room design, and loudspeaker and listener position can be optimized to provide the most uniform sound field.

2. Modulation rate

If FM tones are chosen for sound field work, the modulation rate also needs to be chosen. The Appendix shows that the bandwidth of the signal will not be altered by changes in modulation rate provided this rate is moderately less than the frequency deviation. It follows that as long as this condition is observed, field variability will not be influenced by modulation rate. To test this prediction a series of sweeps of the microphone across the room was made with FM tones which were constant in center frequency (1350 Hz) and frequency deviation (±125 Hz) but which varied in modulation rate between 0.5 and 200 Hz. An index of variability with changes in position was calculated for each stimulus in the manner previously described. The results are presented in Fig. 12.

The prediction is supported by the data. There is negligible change in variability in the field for modulation rates between 2 and 70 Hz. However, as the modulation rate increases beyond 70 Hz towards the frequency deviation (125 Hz) there is a sharp increase in variability. This occurs because the bandwidth of the signal was held constant at 250 Hz (i.e., ±125 Hz) by a filter, and as the prefiltered bandwidth approaches and exceeds this bandwidth, fewer and fewer frequency components will appear within the filter envelope. (The filter was employed for this experiment to ensure that all comparisons made were between stimuli of the same bandwidth.) The lower limit is determined by the averaging employed on the signal (i.e., that suggested by the temporal integration characteristics of the ear). At very slow modulation rates (less than approximately 5 Hz), as the instantaneous frequency sweeps past the various peaks and troughs in the room both the ear and the room will be able to respond to each intensity fluctuation in turn. The threshold for the signal will thus depend only on the highest peak during the modulation cycle. As the height of this peak can vary markedly from position to position within the room, the averaging effect of a broader spectrum signal is thus lost, and variability approaches that for a pure tone.

Thus, the modulation rate should be chosen such that it is higher than the inverse of the time constant of the averaging employed and considerably less than the frequency deviation. The exact range of suitable modulation rates will depend on the time constant and the frequency deviation, but the rate of 5 Hz which is commonly employed would not seem to be sufficiently fast to allow much smoothing of temporal intensity fluctuations in the signal by hearing impaired subjects.

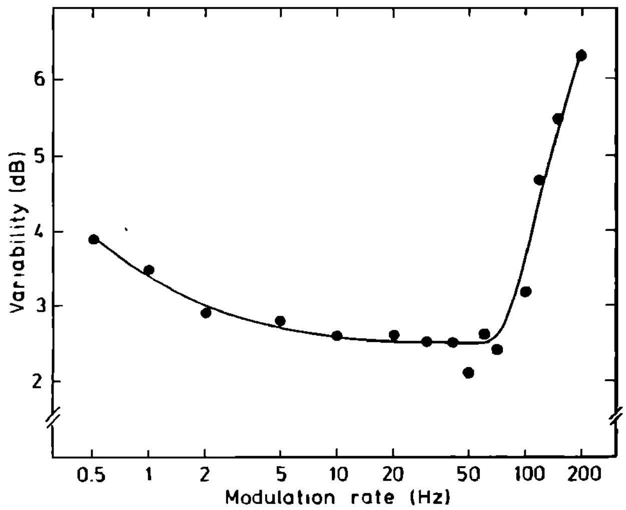

FIG. 12. Average variability within the 20-cm windows within the reverberant field as a function of the FM tone modulation rate. The FM tone had a center frequency of 1350 Hz and a frequency deviation of ±125 Hz.

V. DIRECT VERSUS REVERBERANT FIELD TESTING

One other option needs to be considered. Sound field testing may be carried out in the direct field. In the case of nonanechoic conditions this requires that the loudspeaker be placed very close to the subject's head. Since the direct field is not affected by room acoustics, all of the stimuli discussed above (with the exception of very narrow-band noise) would be equally suitable from the point of view of stimulus variability. In the direct field region, stimulus variability (with respect to distance) is dominated by the inverse-square-law behavior of the field. Widening the bandwidth of the stimulus will thus have little or no effect on variability since all components in the complex signal will be affected by a change of position in the same manner. This effect is shown in Fig. 13 for FM stimuli of various bandwidths. The points indicated correspond to the average deviation within the three 20-cm windows closest to the speaker, (in effect, the difference between the nearest and furthest points within each window). The microphone commenced its traverse at a point 10 cm from the front of the speaker. The lower smooth line shown in Fig. 13 corresponds to the average results presented earlier for the FM tones in the reverberant field. It can be seen that, for the room tested, variability can be minimized by testing in the reverberant field with a stimulus of sufficient bandwidth.

Had the testing been performed at different center frequency, the exact shape of the lower curve would probably have been different. However, the same principles would apply and the conclusions reached would be the same. Stimulus variability in the reverberant field can always be reduced by increasing the stimulus bandwidth while the effects of variability in the direct field can only be removed by restraining the subject's head in some way, which is not possible in all situations. If testing in the direct field is required for some reason, variability is minimized when the measuring point is as far from the speaker as possible (but still within the direct field region). Since the sound pressure in the direct field is not affected by the properties of the particular test room, a general curve for

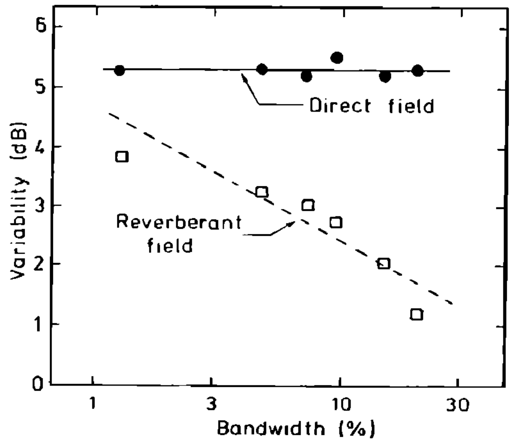

FIG. 13. Average variability within the direct and reverberant fields as a function of stimulus bandwidth for an FM tone of center frequency 1350 Hz.

the variability in this region can be calculated by assuming that the intensity falls off as the square of the distance from the speaker. The calculated variability curve (assuming that the reverberant field component is much less than the direct field component) is shown in Fig. 14. It can be seen that if the field is free from reverberant effects for as much as 1 m from the speaker, the variability in a 20-cm window centered 1 m from the speaker will be less than 2 dB. The direct field will extend the furthest when the room is as large as possible, and the walls are as absorbent as possible, since the crossover from direct to reverberant fields is furthest from the speaker under these conditions (Beranek, 1954, p. 314).

As the substitution of a head (and body) for the microphone used throughout this experiment could result in additional standing waves, the conclusions reached regarding stimulus choice in the reverberant field will still apply in the direct field, although the necessity that the stimulus parameters be properly optimized may not be so great.

In some instances, the choice of direct or reverberant field testing may be dictated by factors other than the relative inaccuracies caused by sound field variability. Due to different head diffraction effects in the two cases, the real ear gain of the hearing aid will also be different for the two types of fields (Kuhn, 1980). Most listening indoors will be done in the reverberant field but listening out of doors or in very large rooms may be done in the direct field. Clearly, for a complete picture of the aided performance of the subject, both direct and reverberant conditions should be tested. However, if a choice has to be made the reverberant field would be appropriate to the majority of commonly encountered listening situations.

VI. SUMMARY

One of the problems of audiometric testing in a reverberant or diffuse sound field is the strong dependence of SPL on both frequency and subject position. Variations (of up to 30 dB) that occur in a typical test enclosure when pure tones are used are illustrated with three-dimensional representations of the field. It is

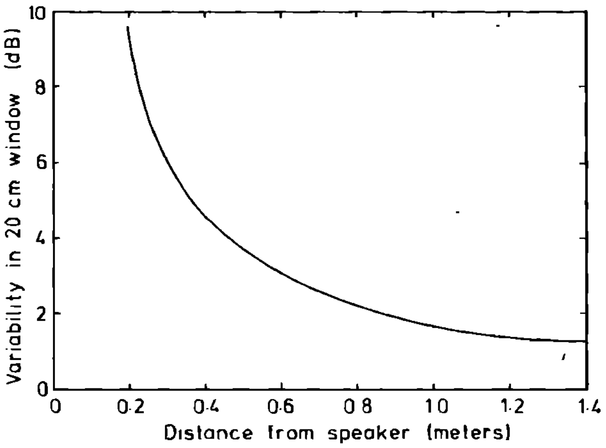

FIG. 14. Theoretical variability within a 20-cm window within the direct field as a function of distance from the speaker.

shown that when more complex stimuli are used the increased bandwidth leads to a more uniform field. Within certain limits, different types of stimuli will produce equally uniform fields provided they are equal in bandwidth. Specifically, these limits are that all the stimuli being compared must have a sufficient number of components within the band. In this respect, sinusoidally amplitude modulated tones and very rapidly frequency modulated tones are found to be inferior at reducing field nonuniformities. Since narrow-band noise, damped wave trains, and frequency modulated tones are found to be equally efficient at producing a uniform field, a choice between them has to be made on other grounds. Frequency modulated tones or suitably generated narrow bands of noise are suggested as the most appropriate sound field stimuli because they can be more directly related to pure-tone thresholds than can damped wave trains (owing to the brief duration of the latter). Some difficulties that might arise when using the recommended stimuli are discussed. These difficulties are associated with the spectral uniformity within the band and rejection rate outside the band, or with the presence of slow intensity fluctuations.

Finally it was shown that while the effects of field nonuniformities in the reverberant field can be greatly reduced by using a stimulus of sufficiently large bandwidth, accurate testing in the direct field can be performed only if the subject's head can be sufficiently immobilized, or if the field extending 1 m or more into the test room is not influenced by the reverberant energy in the test room.

ACKNOWLEDGMENTS

We would like to thank Keith Keen, John Macrae, and Denis Byrne for their helpful comments on an early draft of this paper.

APPENDIX: THE BANDWIDTH OF FM SIGNALS

If we consider a frequency modulated (FM) signal in the time domain, we observe a signal of center frequency f_0 but which deviates between the frequencies $f_0 - f_d$ and $f_0 + f_d$. On this basis, we would expect the signal to have a bandwidth of $2f_d$, and that the bandwidth would be independent of the modulation rate f_m. Physical measurements of the spectrum, however, show that some power does spread outside these limits, and that the distribution of power within the band is very dependent on the modulation rate. By analyzing the signal into its Fourier components, it can be shown that it may be represented by a series of components with a separation between adjacent frequencies equal to f_m (Carson, 1922; Bennett, 1970). For sinusoidal modulation, the amplitude of each component can be calculated from Bessel functions of the first kind once the deviation and modulation rates are known. The argument of the function corresponds to the ratio f_d/f_m, and the order of the function to the order of the particular component in question. Analytically, the amplitude A of each component can be expressed as follows: $A(f_0 \pm nf_m) = J_n(f_d/f_m)$. The results for the bandwidth

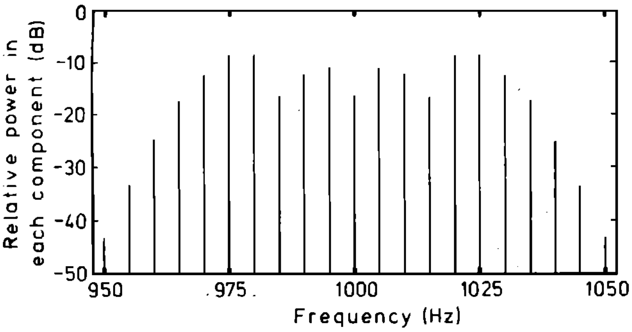

FIG. A1. An example of the spectrum of a 1000-Hz tone, frequency modulated by ±30 Hz at a 5-Hz modulation rate.

given here are based on this equation and assume a sinusoidal modulation, except where stated.

The problems involved in unambiguously specifying a bandwidth for FM signals may be appreciated by considering the typical spectrum shown in Fig. A1. The spectrum shown corresponds to a signal having a center frequency of 1 kHz, a deviation of ±30 Hz, and a modulation rate of 5 Hz. The 0-dB reference corresponds to the total power in the signal. One conventional way of specifying the bandwidth of a signal is to measure the frequency separation between points that have a spectral density 3 dB or 10 dB less than the spectral density within the band. The problem with this method is that the spectral density within the band is not constant for FM signals. An arbitrary decision must be made as to whether the highest, lowest, or average component within the band is used as the reference. (Certainly the lowest will not be appropriate as individual components can at times be absent.) Also, since it is the bandwidth that is being determined, when does a component cease to be "within the band?"

An alternative way of specifying the frequency limits is via the "effective bandwidth" method. This is the bandwidth which is possessed by a signal of constant spectral density in-band and zero power out-of-band and which has the same total power as the signal in question. Again, the in-band power of the FM signal needs to be determined. This method was used in this experiment since the effective bandwidths of the filters used for generating the other stimuli were already known.

A third method which seems appropriate for the present purposes is to specify the frequency limits which encompass, say, 99% of the total power. This method thus avoids the necessity of arbitrarily selecting a method by which the in-band spectral density can be determined. By the use of Bessel functions, the number of components (counting outwards from the carrier frequency) that are required to account for 99% of the total power can be determined. Since these components have frequency separation equal to the modulation frequency, the bandwidth is then readily obtained.

Figure A2 shows the resulting bandwidth as a function of the frequency deviation, for different modulation frequencies. It can be seen that the modulation rate has almost no effect on the bandwidth provided it is much less than the frequency deviation. As the modulation

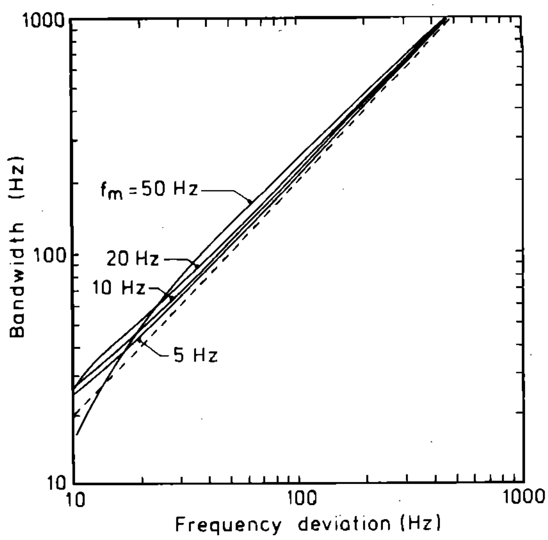

FIG. A2. The expected bandwidth of an FM signal as a function of frequency deviation with modulation rate as a parameter. The bandwidth limits shown are those that encompass 99% of the signal power.

rate approaches the deviation, the bandwidth increases. For further increases in the modulation rate the number of components contributing becomes very small until eventually for $f_m > f_d/7$, 99% of the power is contained in the one central component. Under the definition of bandwidth being used here, such a signal has a bandwidth of zero. Consequently the 20- and 50-Hz curves in Fig. A2 begin to dip for low-frequency deviations.

The broken line shows the estimate that would be obtained if one simply took the bandwidth to be twice the frequency deviation, i.e., the total instantaneous frequency sweep. It can be seen that, for most applications, this estimate is sufficiently precise. (For $f_m \ll f_d$, this estimate can be shown to represent the frequency limits which encompass 95% of the total power.)

The remainder of this Appendix examines, by way of examples, the effect that different modulating waveforms have on the spectrum. Since the generation of an FM stimulus is a nonlinear process, the calculation of the spectrum for modulating waveforms other than sinusoidal is difficult. Consequently, spectra for triangular and square-wave modulating waveforms were measured on a Spectral Dynamics 360 real time analyzer. The waveforms were generated on a Hewlett Packard 3312A function generator.

Figure A3 shows the calculated spectrum for a sinusoidally modulated waveform with a center frequency of 1 kHz, a modulation rate of 10 Hz and a frequency deviation of ±40 Hz. The experimentally determined spectrum is shown in Fig. A3(b). As is to be expected, agreement is quite close. Figures A3(c) and (d) show the experimentally determined spectra for the triangular and square-wave modulated signals, respectively. All signals had the same peak frequency deviation. The square-wave modulation waveform is clearly not well suited for audiometric purposes. Outside the main-frequency band, the falloff in power is much more gradual than for the other two waveforms, and there

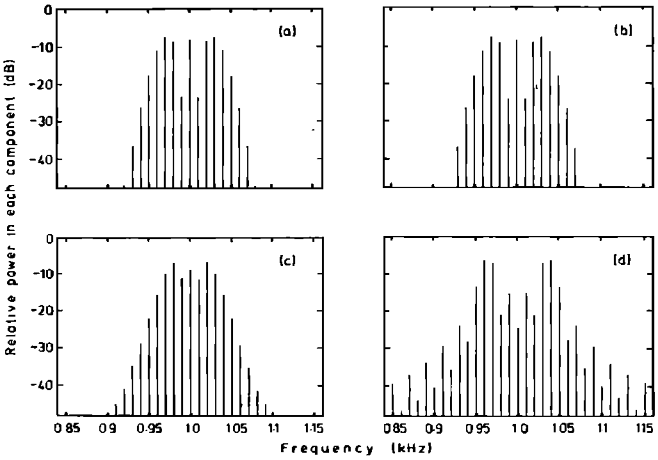

FIG. A3. The effect of modulation waveform on the spectrum of FM signals. Theoretical and experimental spectra for sinusoidal modulation are shown in (a) and (b), respectively. Experimental results for triangular and rectangular (square wave) modulation waveform are shown in (c) and (d), respectively. In all cases the tone had a center frequency of 1 kHz, a frequency deviation of ±40 Hz, and a modulation rate of 10 Hz.

is a concentration of power at the two frequencies on the edges of the band. These results are to be expected since we would expect the power to be greatest at those locations where the instantaneous frequency remains stationary for 50% of the modulation cycle. Furthermore, the rapid change from one frequency to the other creates energy "splatter" well outside the desired frequency range. Despite these facts, Staab and Rintlemann (1972) found that 50% of manufacturers surveyed incorporated rectangular modulation in their audiometers.

A similar, but less marked effect is seen for the sinusoidally modulated signal. Since the instantaneous frequency spends a greater proportion of its time near the band edges than in the center of the band, the components near the band edge usually are of greater amplitude than those near the center. By comparison, the triangular modulation waveform leads to a more uniform distribution within the band, but at the expense of a more gradual out-of-band falloff rate.

These results also hold for other frequency deviations and modulation rates. Ramp modulation (with a rapid return from the higher to the lower frequency or vice versa), although not tested, could be expected to lead to an energy splatter similar to that for the square wave.

On these bases we recommend that ramp or rectangular modulation not be used.

Bennett, W. R. (1970). *Introduction to Signal Transmission* (McGraw-Hill, New York).

Beranek, L. L. (1954). *Acoustics* (McGraw-Hill, New York).

de Boer, E. (1966). "Intensity discrimination of fluctuating signals," J. Acoust. Soc. Am. 40, 552–560.

Carson, J. R. (1922). "Notes on the theory of modulation," Proc. I. R. E. 10, 57ff.

Dillon, H., and Walker, G. (1980). "The perception by normal hearing persons of intensity fluctuation in narrow-band stimuli and its implications for sound field calibration procedures," Aust. J. Audiol. 2, 72–82.

Elliott, L. L. (1963). "Tonal thresholds for short-duration stimuli as related to subject hearing level," J. Acoust. Soc. Am. 35, 578–580.

Gengel, R. W., and Watson, C. S. (1971). "Temporal integration: I. Clinical implications of a laboratory study. II. Additional data from hearing-impaired subjects," J. Speech Hear. Disord. 36, 213–224.

Goldberg, H. (1979). "Discrete frequency sound field audiometry," Hear. Aid J. 32, 7-ff.

Harris, J. D., Haines, H. L., and Myers, C. K. (1958). "Brief tone audiometry," Arch. Otolaryngol. 67, 699–713.

Kuhn, G. F. (1980). "Some effects of microphone location, signal bandwidth, and incident wave field on the hearing aid input signal," in *Acoustical Factors in Hearing Aid Design*, edited by G. Studebaker and I. Hochberg (University Park, Baltimore).

Morgan, D. E., Dirks, D. D., and Bower, D. R. (1979). "Suggested threshold sound pressure levels for frequency modulated (warble) tones in the sound field," J. Speech Hear. Disord. 44, 37–54.

Orchik, D. J., and Moser, N. L. (1975). "Narrow band noise audiometry: The effect of filter slope," J. Am. Aud. Soc. 1, 50–53.

Pedersen, C. B. (1976). "Brief-tone-audiometry," Scand. Audiol. 5, 27–33.

Pedersen, C. B., and Elberling, C. (1973). "Temporal integration of acoustic energy in patients with presbyacusis," Acta Oto-Laryngol. 75, 32–37.

Plomp, R., and Bouman, M. A. (1959). "Relation between hearing threshold and duration for tone pulses," J. Acoust. Soc. Am. 31, 749–758.

Sanders, J. W., Josey, A. F., and Kemker, F. J. (1971). "Brief tone audiometry in patients with VIIIth nerve tumor," J. Speech Hear. Res. 14, 172–178.

Schafer, T. H., Gales, R. S., Shewmaker, C. A., and Thompson, P. O. (1950). "The frequency selectivity of the ear as determined by masking experiments," J. Acoust. Soc. Am. 22, 490–496.

Staab, W. J., and Rintelmann, W. F. (1972). "Status of warble-tone in audiometers," Audiology 11, 244–255.

Victoreen, J. A. (1974). "Equal loudness pressure determined with a decaying oscillatory waveform," J. Acoust. Soc. Am. 55, 309–312.

Wit, H. P. (1978). "Hearing threshold for a decaying oscillatory waveform," J. Aud. Res. 18, 69–77.

Wright, H. N. (1968). "The effect of sensorineural hearing loss on threshold-duration functions," J. Speech Hear. Res. 11, 842–852.

Zwislocki, J. (1951). "Eine verbesserte vertäubungsmethode für die audiometrie," Acta Oto-Laryngol. 39, 338–356.