NASA Reference Publication 1001

Overall Loudness of Steady Sounds According to Theory and Experiment

Walton L. Howes

(NASA-RP-1001) OVERALL LOUDNESS OF STEADY SOUNDS ACCORDING TO THEORY AND EXPERIMENT (NASA) 150 p MF A01; HC $00 CSCL 05J
N79-25753
Unclas
H1/53 27232

JUN 1979
RECEIVED
NASA STI FACILITY
ACCESS DEPT.

NASA

NASA Reference Publication 1001

Overall Loudness of Steady Sounds According to Theory and Experiment

Walton L. Howes

Lewis Research Center
Cleveland, Ohio

Prepared for
Aerospace Safety Research and Data Institute
NASA Lewis Research Center

NASA

Scientific and Technical Information Branch 1979
National Aeronautics and Space Administration
Washington, DC

Library of Congress Cataloging in Publication Data

Howes, Walton L.
Overall loudness of steady sounds accordng to theory and experiment.

(NASA reference publication ; 1001)
Bibliography: p.
1. Auditory perception. 2. Noise—Physiological effect. 3. Sound-measurement. I. United States. Lewis Research Center, Cleveland. Aerospace Safety Research and Data Institute. II. Title. III. Series: United States. National Aeronautics and Space Administration. NASA reference publication ; 1001.
QP465.H68 612′:85′028 78-24563

For sale by the Superintendent of Documents,
U.S. Government Printing Office, Washington, D. C. 20402
Stock Number 033-000-00690-0

FOREWORD

The growth of technology has been accompanied by numerous consequences harmful to the environment. Among these negative effects is noise. In recent years, as noise sources have become more prevalent and more intense, the general population, increasingly annoyed by noise, particularly from aircraft, trucks, and machinery, has exerted pressures for laws and regulations to establish allowable limits on noise. As a consequence, the reduction of noise has become a significant cost of doing business.

One of the problems of noise abatement and legal restrictions is that of quantitatively defining the level of a noise when it becomes annoying. There is no precise answer to this problem because the individual, his environment, and the exposure circumstances all strongly influence the result. However, annoyance is closely related to loudness, which is much less influenced by nonauditory effects. If two sounds are annoying, then the louder sound tends to be more annoying. Loudness provides a simpler specification than annoyance for noise abatement and legal purposes.

For both abatement and legal purposes it is most important to be able to predict the loudness of sounds. The relationship between the sound and loudness is determined by the operation of the intervening auditory system. Previous loudness calculation schemes are questionable because they have incorporated this operation in a sketchy and dubious fashion.

The purpose of this publication is to present a new quantitative, analytical model for the operation of the human auditory system as it pertains to loudness. The resulting theory provides a practical means for calculating the loudness of any steady sound from given information regarding the sound. It is hoped that this new model of the auditory system will provide important new insights concerning the manner in which the loudness sensation develops and that the resulting loudness calculation scheme will become accepted as a replacement for existing schemes.

CONTENTS

	PAGE
FOREWORD	iii
INTRODUCTION	1
OPERATION OF AUDITORY SYSTEM	2
Filtering	4
Neural Response	5
Beats	7
Critical bandwidth	7
Secondary tones	8
Channels and neuron bandwidths	10
Aural harmonics and combination tones	10
Loudness	11
MODEL OF AUDITORY SYSTEM	11
LOUDNESS OF ANY STEADY SOUND: ANALYSIS	13
Discrete Spectrum	15
Category A: Subcritical bandspread	17
Category B: Supercritical tone spacing	21
Category C: Hypercritical tone spacing	27
Category AB: Subcritically spaced tones spread over supercritical bandwidth	27
Category AC: Subcritically spaced tones spread over hypercritical bandwidth	28
Category BC: Supercritically spaced tones spread over hypercritical bandwidth	28
Continuous Spectrum	28
Subcritical bandwidth	28
Supercritical bandwidth	30
Hypercritical bandwidth	31
Mixed Spectrum	31
LOUDNESS LEVEL	31
TRANSMITTANCE FUNCTIONS	32
SIMPLE FORMULAS FOR LOUDNESS LEVEL	37
TEST CONDITIONS	39

PAGE

LOUDNESS OF SPECIFIC SOUNDS WITH DISCRETE SPECTRA (COMPARISON OF MODEL WITH EXPERIMENT) ---- 40

One Tone ---- 40

Two Tones ---- 42

Von Békésy's experiment (1929) ---- 43

Fletcher and Munson's experiment (1933) ---- 45

Small and Thurlow's experiment (1954) ---- 47

D. H. Howes' experiment (1950) ---- 50

Scharf's experiment (1959) ---- 53

Three Tones ---- 57

Bauch's experiment (1956) ---- 57

D. H. Howes' experiment (1950) ---- 62

Scharf's experiment (1962) ---- 65

Four Tones ---- 67

D. H. Howes' experiment (1950) ---- 67

Zwicker, Flottorp, and Stevens' experiment (1957) ---- 69

Scharf's experiment (1959) ---- 73

Five Tones ---- 75

D. H. Howes' experiment (1950) ---- 75

Zwicker, Flottorp, and Stevens' experiment (1957) ---- 80

Quietzsch's experiment (1955) ---- 80

Six Tones ---- 82

D. H. Howes' experiment (1950) ---- 82

Seven Tones ---- 82

D. H. Howes' experiment (1950) ---- 82

Eight Tones ---- 83

D. H. Howes' experiment (1950) ---- 83

Quietzsch's experiment (1955) ---- 84

Ten Tones ---- 85

Fletcher and Munson's experiment (1933) ---- 85

Eleven Tones ---- 88

D. H. Howes' experiment (1950) ---- 88

Fletcher and Munson's experiment (1933) ---- 89

Twenty Tones or More ---- 91

Quietzsch's experiment (1955) ---- 91

Loudness as Function of Number of Tones Contained Within Fixed Frequency Band ---- 92

Scharf's experiment (1959) ---- 92

LOUDNESS OF SPECIFIC SOUNDS WITH CONTINUOUS SPECTRA (COMPARISON OF MODEL WITH EXPERIMENT) ---- 93

Effect of Bandwidth ---- 93

Zwicker and Feldtkeller's experiment (1955) ---- 93

Niese's experiment (1961) ---- 94

Zwicker, Flottorp, and Stevens' experiment (1957) ---- 95

Octave Bands ---- 96

Robinson and Whittle's experiment (1964) ---- 96

Cremer, Plenge, and Schwarzl's experiment (1959) ---- 99

G. Jahn's experiment (1959) ---- 101

Stevens' experiment (1956) ---- 102

PAGE

Various Bands ---- 103
Zwicker's experiment (1966) ---- 103
Sums of Contiguous Bands ---- 104
Stevens' experiment (1956) ---- 104
Sums of Noncontiguous Bands ---- 106
Stevens' experiment (1956) ---- 106
Effect of Bandwidth Near Loudness Threshold ---- 107
Scharf's experiment (1959) ---- 107
White Noise ---- 108
Feldtkeller's experiment (1956) ---- 108
Pollack's experiment (1951) ---- 109
Miller's experiment (1947) ---- 110
Brittain's experiment (1939) ---- 112
Niese and Köhler's experiment (1958) ---- 112
Broadband Noises ---- 114
Lübcke, Mittag, and Port's experiments (1964) ---- 114
M. Jahn's experiment (1965) ---- 115
Warren's experiment (1973) ---- 116
Quietzsch's experiment (1955) ---- 116
Corliss and Winzer's experiment (1965) ---- 117
Robinson and Bowsher's experiment (1961) ---- 118
Copeland et al. experiment (1960) ---- 119

LOUDNESS OF SPECIFIC SOUNDS WITH MIXED SPECTRA ---- 119
Lübcke, Mittag, and Port's Experiment (1964) ---- 120
M. Jahn's Experiment (1965) ---- 120
Quietzsch's Experiment (1955) ---- 120
Copeland et al. Experiment (1960) ---- 120
RELATION BETWEEN PRESENT THEORY AND A-WEIGHTED, SOUND-PRESSURE LEVEL ---- 121

DISCUSSION AND REMARKS ---- 122

SUMMARY OF RESULTS ---- 124

APPENDIXES
A–SYMBOLS ---- 127
B–SPECTRAL REPRESENTATION OF OUTPUT POWER-SPECTRAL-DENSITY W_q ---- 131
C–LOCATION OF CRITICAL BANDS ---- 133
D–CONCERNING BEATS OF MISTUNED CONSONANCES ---- 134
E–PLANCHEREL'S THEOREM FOR LIMITED INTEGRATION TIME ---- 136
F–EXTERNAL TRANSMITTANCE FUNCTION ---- 137

REFERENCES ---- 139

INTRODUCTION

If the attention of a listener is consciously directed toward perceiving sounds, then he is capable of estimating the relative magnitudes of pairs of sounds. The judged magnitude of a sound heard by a listener is called its loudness. By making quantitative judgments of the relative loudness of pure-tone stimuli at measured sound pressures and by adopting one tone and sound pressure as a reference, a quantitative relation between loudness and sound pressure can be established for pure tones. The resulting loudness scale is a ratio scale because it is derived by judging ratios of loudnesses. Next, by subjectively equating the loudnesses of other sounds to that of the reference tone the loudnesses of the other sounds can be determined.

(Analogous procedures have been adopted to evaluate the relative noisiness, annoyance, or other sensations induced by sounds. However, noisiness and annoyance, for example, involve the emotional state of the listener. This implies that nonacoustic neural processes affect the judgment (ref. 1). These processes have not been evaluated. Moreover, noisiness and annoyance judgments should be quite variable because emotional states vary widely among individuals as well as for any given individual at different times (ref. 2). It is reasonable to expect, however, that a listener in a given emotional state will judge the louder of two similar sounds to be noisier and more annoying as well. Thus, the loudness scale may sometimes serve equally well as a noisiness, or annoyance, scale (ref. 3). For these reasons, noisiness and annoyance will not be considered further.)

This report is concerned with the loudness of steady sounds. Steady sounds are defined as those having statistical properties independent of time. For practical purposes a sound can be regarded as steady if its statistics vary slowly over a few auditory integration periods, say over a second or so. With this relaxed definition the class of steady sounds is very important because it is very common. As will be shown, the loudness of a steady sound is not necessarily steady.

The main purpose of this report is to propose and test a new psychophysiological model for predicting quantitatively the loudness of steady sounds. The model is termed a psychophysiological model because it is constructed from both psychoacoustic and physiological information. The present theory utilizes new ideas, particularly

power summation and frequency interaction, to account for loudness in place of the older concepts of loudness summation and inhibition of discharges. Because of its almost complete departure from previous models (e.g., refs. 4, 5, 6, and 7), it is hoped that this new model will provide important new insights concerning the manner in which the loudness sensation develops in the auditory system.

The proposed model was synthesized from data obtained using specific sound stimuli, that is, the development of the model proceeded from the specific to the general. However, the subsequent presentation will be in reverse order, that is, the general theory is presented and then analyzed to obtain specific results. In the process of analysis an attempt is made to show how the model is related to physiological and psychoacoustic observations other than loudness. Loudness levels predicted using the model are compared, where possible, with available loudness judgment data, which are believed to consist of most of the published data.

OPERATION OF AUDITORY SYSTEM

The complete auditory system consists of a complicated transducer which transmits the information in sound waves acoustically, then, mechanically, then hydrodynamically and, finally, electrodynamically to the brain (refs. 8 and 9). The fundamental equations of acoustics, solid dynamics, hydrodynamics and electrodynamics are well known. However, the geometry of the auditory system is so complicated that a mathematical analysis of the complete physical system has not been achieved. Moreover, the processes involved in the hydrodynamic to electrodynamic conversion are not well understood. Because of these and other difficulties it is more feasible in predicting the loudness of sounds to consider what the auditory system does rather than how it does it. The latter approach involves analysis of the physical system, whereas the former approach, adopted herein, involves, instead, an analysis of the operational characteristics of the system.

A block diagram of the basic auditory system is shown in figure 1. In the outer, middle, and partly in the inner ear the wave is filtered and is physically represented by mechanical and hydrodynamical motions. In hair cells in the inner ear the mechanical motion becomes coded in the form of constant amplitude, impulsive, electric-potential discharges (up to several hundred per second in each active neuron) propagated (at speeds up to 100 m/s) along auditory nerve fibers (approximately 30 000) toward the brain. In the mechanical-to-electrical transformation additional filtering probably occurs. In the auditory nerve the waveform can be depicted by the number of discharges occurring as a function of time. The discharges occur only in association with part of the waveform, so that a partial-wave

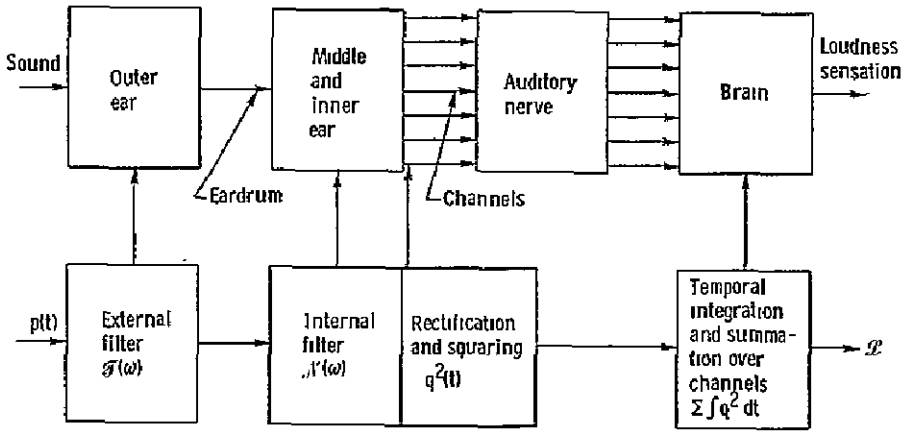


FIGURE 1.—Operational diagram of auditory system.

rectification of the waveform appears in the auditory nerve. In the brain, possibly by means of a metastable chemical reaction, the energy transmitted in the discharges gets summed over a short, running time interval, the auditory integration time (approximately 0.1 s). The auditory integration time is the time taken to build up a steady response to a steady stimulus. The sum determines the intensity of the subjective sensation, that is, the loudness.

Henceforth, for convenience in terminology only, quantities such as the effective sound pressure, physical power, and the neural discharge rate will all be referred to the auditory integration time (0.1 s rather than 1.0 s). Thus, the effective sound pressure is the root-mean-square value over the auditory integration time; the physical power is the energy transmitted during the auditory integration time; the discharge rate is the number of discharges which occur during the auditory integration time, and so forth.

In order to predict the loudness of any steady sound stimulus, consider, first, a previous model of the system (ref. 10) used to predict the loudness of a tone.

Specifically, it was proposed (ref. 10) on the basis of experimental data (ref. 11) that, in response to a tonal stimulus, the number of stimulated auditory nerve fibers is proportional to the effective sound pressure until all available fibers are stimulated. Experimentally, the discharge rate in each fiber is also approximately proportional to the effective sound pressure until the fiber discharge rate becomes saturated (ref. 12). At about 40 decibels above the loudness threshold, discharge rates in some fibers begin to saturate. However, other, previously unstimulated, fibers exceed their stimulation thresholds as the sound pressure is increased, so that the total discharge rate continues to increase until all available fibers are stimulated and, then, saturated. The total discharge rate was expressed (ref. 10) as the product of the total number of stimulated fibers and the dis-

charge rate per fiber averaged over all stimulated fibers. Consequently, as the sound pressure is increased, the number of stimulated fibers increases linearly, but the average discharge rate per fiber increases very little because some fibers have become saturated whereas others have just achieved their thresholds of stimulation. Hence, the total discharge rate is essentially proportional to the effective sound pressure. Since electric potential is proportional to electric energy, and all discharges are identical, the total discharge rate is proportional to the total electric power transmitted by the auditory nerve. Therefore, the total electric power is proportional to the effective sound pressure of the stimulus, and the ear is found to be a nonlinear detector, specifically a squaring detector.

It has long been suspected (e.g., refs. 13, 14, and 15, chap. 12) on the basis of experimental evidence (refs. 16 and 17), that, except near threshold, the magnitude of the loudness sensation in response to a tone stimulus is proportional to the total discharge rate (total electric power) transmitted in the auditory nerve. (The fact that the discharges are summed implies that total current, which is proportional to discharge rate, is proportional to loudness.) As a consequence of this observation, loudness is proportional to the effective sound pressure. Since the mean-square pressure is proportional to the average power of the stimulus, the loudness must also be proportional to the square root of the stimulus power transmitted by the system.

Experimental evidence (ref. 18) indicates that for individual nerve fibers the dependence of discharge rate on sound pressure is the same for broadband noise as for a tone. If it is assumed that the dependence of the number of active fibers on sound pressure is also the same for both sounds, then it may be concluded from an analysis similar to that in reference 10 that, in general, loudness is proportional to the square root of the power transmitted by the auditory system.

In order to predict the loudness of sounds more complicated than a lone tone, a broader description of the system is required.

Filtering

The auditory system is a bandpass filter. Psychoacoustic data (refs. 19 and 20) indicate that the filter is passive, and, of course, nonideal. From the experimental standpoint, it is convenient to analyze the filter into an external filter in series with an internal bandpass filter (ref. 21). The external filtering occurs in the outer ear, which precedes the eardrum (cf. fig. 1). The characteristics of this filter depend on the orientation and extent of the sound source relative to the listener. The internal filter involves the entire auditory

system succeeding the eardrum (cf. fig. 1). This filter is weakly dependent on the effective sound pressure.

The internal filter may be analyzed into a low-pass filter in series with a high-pass filter. The low-pass filter is probably associated primarily with the mechanical part of the auditory system (ref. 8), whereas the high-pass filter may be associated primarily with cochlear processes (particularly excitation of cochlear hair cells) involved in converting the signal to electrochemical form.

Neural Response

Mechanical vibrations of a region of the tectorial membrane within the inner ear (ref. 8) produce excitation of hair cells and, hence, of the auditory neurons. Each neuron is specified by its "best," or "characteristic," frequency, that is, the frequency for which its stimulation threshold is lowest. However, each neuron may be excited by a broad range of stimulus frequencies. In figure 2 nominal bandwidths (adapted from ref. 12) of auditory nerve fibers in squirrel monkeys are shown.

Consider a tonal stimulus. If the tone is sufficiently intense and its frequency lies within the transmission bandwidth of a given fiber, then the tone periodically stimulates discharges in the fiber (ref. 12). (Although a periodicity is commonly believed to exist only for stimulus frequencies less than 5000 Hz, a periodicity which may be associated with pitch actually exists at all frequencies (ref. 22).) At the lower amplitudes of stimulation the periodicity of the discharges tends to be statistically distributed at time intervals which

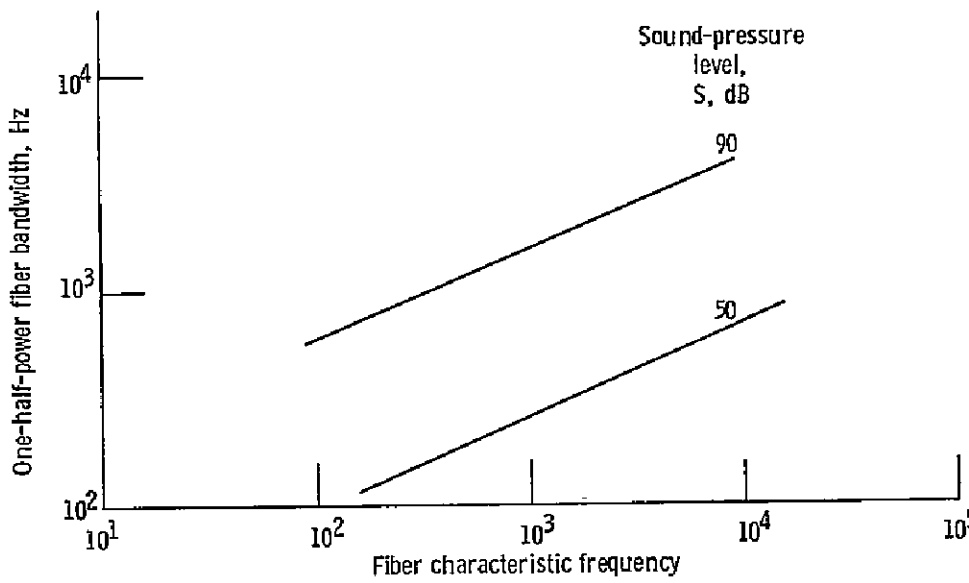


FIGURE 2.—Approximate one-half-power bandwidth of auditory nerve fiber as function of its best frequency with sound-pressure level as parameter. (Adapted from data in ref. 12. Maximum bandwidths are much greater.)

are integer numbers of times the periodicity of the stimulus (ref. 23). At higher stimulus amplitudes a discharge tends to occur more often during successive stimulus periods. For a large number of realizations of the stimulus, the number of discharges as a function of the phase of the stimulus graphically resembles that part of the stimulus waveform which exceeds some constant amplitude (often zero in the auditory nerve, refs. 24 and 12). The number density is effectively zero throughout the rest of the period. Hence, for each fiber the waveform appears to be partial-wave rectified. The preponderance of data indicate that, when all fibers are considered, the waveform may be approximately half-wave rectified. For periodic stimuli the trace of the waveform from discharges in single units over many realizations of the stimulus might be approximated alternatively by discharges over many units for a single realization of the stimulus. This resembles an ergodic hypothesis.

Stimulation by more than one tone yields a distribution of stochastically independent discharges which also tends to duplicate graphically part of the waveform (refs. 12 and 24). Similar results are expected for any other periodic sound. In fact, it seems reasonable to expect that the temporal history of the instantaneous number of discharges over all fibers traces out part of the waveform (suitably filtered) even if it is aperiodic.

In the mechanical part of the auditory system the trace of the waveform is determined by the amplitude of mechanical motions, whereas in the electrical part of the system the waveform is traced by the instantaneous number of discharges, which is proportional to total instantaneous power. Therefore, since power is proportional to the square of the amplitude, the squaring operation in the auditory system results simply from the transformation of the representation of the waveform from mechanical to electrical form!

Additional clues on how to model the auditory system are supplied by various observed psychoacoustic phenomena and their plausible explanation in terms of neural activity. Among possible significant phenomena are audible beats (refs. 25 and 26), critical bandwidths for loudness (refs. 27 and 28) and frequency resolution (ref. 29), secondary tones (ref. 30), harmonic stimuli (ref. 31), and aural harmonics and combination tones (e.g., refs. 15 and 32). The loudness features associated with the beats are claimed herein to be a consequence of squaring and of the limited auditory integration time; whereas those associated with critical bandwidths, secondary tones, and harmonic stimuli are proposed consequences of squaring and of the broad bandwidths of individual neurons.

In order to explain beats and other psychoacoustic phenomena in terms of neural activity, the assumption will be made that *stimuli which excite the same neuron interact whereas stimuli which do not*

excite the same neuron do not interact. This interaction has an important effect on loudness.

Beats.—The loudness of two tones with comparable amplitudes and nearly equal frequencies, or of tones which are nearly harmonics of one another, oscillates because of interference between the tones (refs. 25, 26, and 33 (pp. 14–16)). The loudness oscillations, equal in frequency to physical beats, are referred to as "audible beats." For nearly equal tone frequencies the oscillations are called "beats of imperfect unisons" (ref. 25); whereas for nearly harmonic tone frequencies they are called "beats of mistuned consonances" (ref. 26). Subjectively, since tones which produce beats of imperfect unisons are separated by less than a critical separation, they cannot be resolved. (This defines the critical separation, which is less than one-third octave, except at the lower audible frequencies.) The consequence in the case of nearly equal tones, is an intertone—with frequency near that of the primaries—modulated at the beat frequency. As the separation of the primary tones is increased, the beat frequency increases; but the loudness of the beats vanishes, as will be shown mathematically.

Two, or more, closely spaced tones stimulate the same nerve fibers because the fiber bandwidths are broad. Hence, the beats produced by interference of two, or more, tones are observed as fluctuations of the number of discharges at the beat frequency as these discharges trace out the rectified waveform of the combined stimuli (ref. 34). The period of the beats exceeds the auditory integration time. Loudness is presumed to be the subjective measure of the transmitted electric power corresponding to the discharge rate over all fibers, or, alternatively, of the equivalent transmitted stimulus power (ref. 10). If the auditory averaging time were infinite, the power contribution from the tone interaction would vanish. However, the averaging time is finite (0.1 s), so that a time-dependent interaction exists if the period of the beats is greater than, or approximately equal to, the integration time.

As two closely spaced tones or harmonics are increasingly separated in frequency, the period of the beats becomes less than the auditory integration time. Consequently, the maximum and minimum average power transmitted by the auditory nerve throughout a running integration period tend to the same value, independent of the tone separation, as does the loudness (ref. 35).

Critical bandwidth.—Two critical bandwidths (ref. 36) are significant, namely the critical bandwidth for loudness (ref. 28) and the critical bandwidth for frequency resolution (ref. 29). The critical bandwidth for loudness about any given center frequency is the maximum sound bandwidth for which loudness is independent of bandwidth when the overall sound pressure is held constant. The critical

bandwidth for frequency resolution about any given center frequency is the maximum tone separation for which two, or more, tones cannot be resolved psychoacoustically. The two critical bandwidth functions are similar, as shown in figure 3. As the respective critical bandwidths are exceeded by two, or more, tones, the loudness rises (refs. 37 and 38), and the tones are resolved (ref. 29). The critical bandwidths indicate an important transition in the neural coding.

Secondary tones.—In addition to the loudness increase and tone resolution, secondary tones other than the imposed primary tones may be heard as the critical bandwidth for frequency resolution is exceeded. The most prominent secondary tone is frequently the "periodicity pitch" (ref. 39) with frequency equal to the difference in frequency of the primary tones. If several primary tones are harmonically related, then the fundamental may be heard even if it is not included among the primaries. An audible missing fundamental has been called the "pitch of the residue" (refs. 31 and 40).

The following explanation for these phenomena is proposed. The pattern of neural discharges traces a filtered and rectified version of the acoustic stimulus waveform. This has been demonstrated from many realizations of periodic stimuli in single units (ref. 30). If two tones are closely spaced, the intertone is determined by the periodicity of the envelope of the instantaneous number of discharges. Dis-

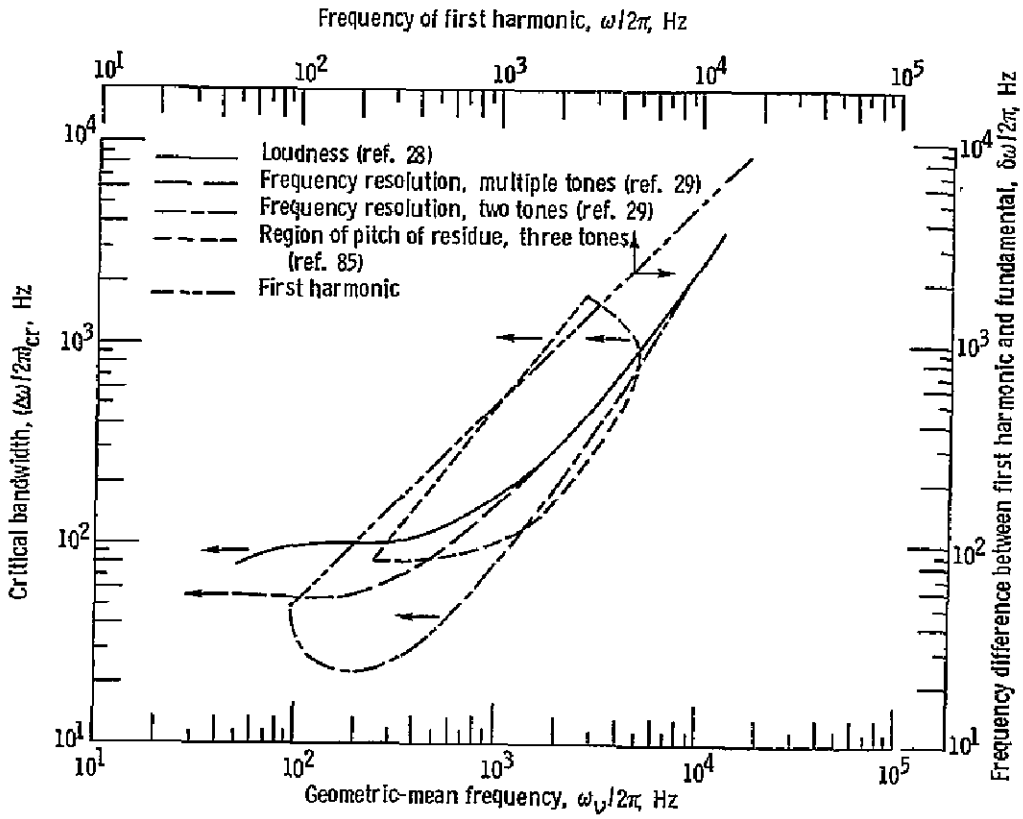


FIGURE 3.—Critical bandwidths and region of pitch of residue.

charges occur in both positive and negative phases of the beat cycle. The beat frequency lies outside the transmission bands of the fibers. As the tone separation approaches the critical value the beat frequency may exceed the lower cutoff frequency of some fibers which also transmit the primaries. Then, fibers will be stimulated at approximately the periodicity of the beat, so that the beat will be transmitted as if it were a real tone and may be heard as a secondary tone. In this regime, due to the variation of frequency pass bands among fibers, a mixture of the primary tones, intertone, beats and secondary tones may combine to produce the sensations of "intermittence" or "roughness" (ref. 25). When the critical separation is exceeded, the primary tones and secondary tone are dominant (ref. 30). All three tones lie within the pass bands of the same fibers and are transmitted as audible tones. This accounts for the periodicity pitch. The pitch of the residue is similarly explained when the stimulus consists of more than two tones. Since the auditory system is a squaring device, the secondary tone is heard as if stimulated by real acoustic power. The detection of the secondary tone should result from an increase in the overall discharge rate and a redistribution of the discharges in time to establish the periodicity of the secondary tone. An increase in loudness should, therefore, also be expected. *According to this argument the critical bandwidth for loudness is equal to a characteristic lower cutoff frequency of those nerve fibers which propagate information on the primary tones.* This comparison is shown in figure 4, where equality of critical bandwidth and fiber cutoff frequency is denoted by the "ideal" curve.

The boundary of the regime of a clear periodicity pitch obtained from a uniformly spaced, three-tone stimulus heard through ear-

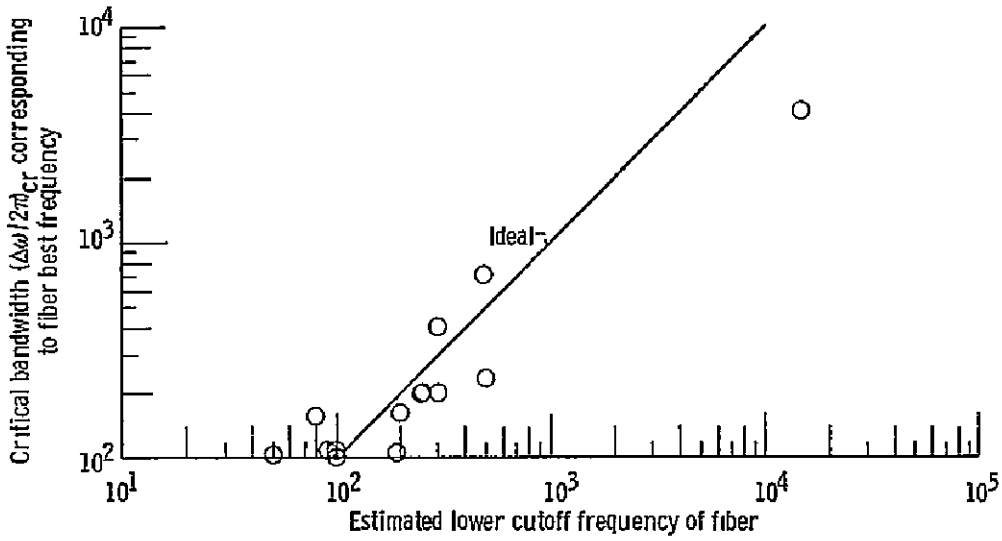


FIGURE 4.—Critical bandwidth as function of complete-lower-cutoff frequency of single auditory nerve fibers. (Data from refs. 12 and 42.)

phones is shown in figure 3 for comparison with the critical bandwidths previously discussed. The lowest audible periodicity pitch corresponds roughly to the critical bandwidth. (Of course, many of the physiologically significant frequencies are not precise because of physiological variations among and within individuals.) The highest audible periodicity pitch is less than the frequencies of the tones which produce it. The frequency range over which the periodicity pitch may be obtained is discussed in the next paragraph.

Channels and neuron bandwidths.—Since the critical bandwidths represent important transition bandwidths in neural coding, it is convenient to evaluate loudness by grouping all neurons into artificial channels having widths equal to the critical bandwidths for loudness. Each channel includes every neuron whose best frequency is within the critical bandwidth for that channel. In most channels the neuron bandwidths tend to be much greater than the channel, or critical, bandwidths (refs. 12 and 28). Thus, each channel may transmit a much broader range of frequencies than the channel bandwidth would imply. All frequencies transmitted in a channel may interact if they excite the same neurons. Thus, if tones transmitted in a channel are subcritically separated, they may generate audible beats; whereas if they are supercritically separated, they may generate secondary tones. The secondary tones can only be generated if the neuron bandwidths (fig. 2) exceed the critical bandwidth (fig. 3). Neuron bandwidths clearly exceed critical bandwidths for best frequencies in the range from 200 to 5000 hertz, approximately, at moderate sound pressures. These frequencies establish the minimum and maximum limits, respectively, for which secondary tones can readily be heard. (See limits of pitch of the residue in fig. 3.)

The preceding channels are "floating" in the sense that their midfrequencies or, alternatively, their cutoff frequencies are determined by the spectrum of the imposed sound. However, the bandwidth of a channel with a given midfrequency is effectively always the same in response to a steady sound.

Aural harmonics and combination tones.—These tones are assumed to be subjective consequences of mechanical nonlinearity, that is, of functional defects, of the auditory system in response to single tones or to multitone stimuli, respectively (ref. 41). This is in contrast with the squaring process, a nonlinear process fundamental in the operation and interpretation of the system. Although these tones may be audible, they are usually very weak relative to the primaries, beats, or secondary tones. Hence, their contribution to overall loudness is likely to be negligible, except possibly at loudnesses near the threshold of feeling (ref. 32, ch. 7). The contribution of these tones to overall loudness will be neglected.

Loudness

The loudness of any sound is numerically expressed relative to that of some precisely defined sound which serves as a reference. Thus, in order to specify the loudness it is not necessary to know how much power is transmitted by the system. Only the ratio of transmitted powers, namely the ratio of the transmitted power of the test sound to that of the reference sound, need be determined. This can be found after the test sound and characteristics of the auditory system are specified mathematically.

MODEL OF AUDITORY SYSTEM

A model of the auditory system previously proposed (ref. 10) for predicting the loudness of a tone will serve as the basic model to be further expanded for predicting the loudness of any steady sound.

It is assumed that the sensation of loudness in response to a sound stimulus is measured by the power (energy per auditory integration time) transmitted to the brain. Physiologically, the power is determined by the product of the total current (proportional to total discharge rate) and the potential, which is the same for each discharge. Since the loudness sensation commences for some discharge rate exceeding zero (ref. 10), the loudness is proportional to the amount by which the electric power exceeds some threshold power. The proportionality constant is a reciprocal power (reference power), so that the loudness is a dimensionless quantity, and the power can always be expressed relative to the reference power in dimensionless form.

In the basic model the electric power is proportional to the effective sound pressure. This results from the analysis of discharge rates in neurons (ref. 10). From the preceding discussion of the operation of the auditory system it is concluded that the extension of the basic model of the system in response to any steady sound stimulus must also include filtering, channeling, squaring, and rectification.

The complete filter, which is separable into an external filter in series with an internal filter, possesses known characteristics derived from experiment. The complete filter is a linear, passive, time-invariant, nonideal, band-pass type. Its input is the sound. Its output is determined from subjective judgments, and therefore, takes into account the entire auditory system.

In the organ of Corti the overall signal is distributed among up to 24 parallel channels spanning the audible frequency range (ref. 28). Each channel has the width of a critical band for loudness in the sense that it includes all auditory neurons with best frequencies lying

within the critical band. Each channel should have a different frequency response. Nevertheless, since the specific response for each channel is unknown, it will be assumed that the response associated with each channel is that of the complete system. This may be the main deficiency of the model. However, this is compensated for by assuming that only those channels within the limits of the sound spectrum transmit appreciable power. Since the transmitted power in each channel is to be specified relative to the same overall reference power, it may be assumed that the entire input is imposed in each channel. Thus, the signal tends to get multiplied by the number of channels which transmit it. In the limiting case of a tone, only one channel, that which includes the tone, is activated. The frequency response is precisely that for the tone. The resulting loudness is found to be that previously predicted for a tone (ref. 10), as will be shown.

Since the neuron bandwidths exceed the associated critical bandwidths over most of the audible frequency range, signals over a wide range of frequencies may interact. In the squaring process involved in the transformation from a mechanical to an electrical signal, both self-interaction of signals at the same frequency and cross-interaction of signals at different frequencies occur in each channel. However, because of the expected decline in response in the channel to frequencies outside the associated critical band, appreciable cross-interaction between different frequencies will be assumed to be limited to frequency pairs where at least one frequency of the pair lies within the critical band.

Since all channels within the limits of the sound spectrum are assumed to transmit appreciable power, whereas the sound spectrum may vanish over entire intermediate critical bandwidths, "empty" channels may occur which transmit no power in the associated critical bands. However, the empty channels may transmit power due to the self-interaction of frequencies outside the empty critical band. As in the case of nonempty bands, it will be assumed that the cross-power contributions due to the cross-interaction of frequencies outside the empty critical bands are negligible.

The pattern of neural discharges is partial-wave rectified. Specifically, it will be assumed that the pattern is half-wave rectified.

This completes the expanded model which now incorporates most of the previously discussed psychoacoustic and physiological phenomena. The subsequent mathematical representation of the model yields the magnitude of the loudness sensation $\mathcal{L}$ as some measurable function of the free-field pressure perturbation p of an arbitrary steady sound.

LOUDNESS OF ANY STEADY SOUND: ANALYSIS

A 1-kilohertz tone imposed as progressive plane waves from the front at a free-field, sound-pressure level S of 40 decibels is defined to have a loudness $\mathcal{L}$ of one sone. (All symbols are defined in appendix A.) A sound λ times as loud as the tone at one sone is said to have a loudness of λ sones. The sound-pressure level S is defined by

$$S=20 \log \mathcal{P} \text{ (decibels)} \tag{1}$$

where

$$\mathcal{P}=[(1/Tp_{ref}^2)\int_{t-T}^{t} p^2(t')\,dt']^{1/2} \tag{2}$$

The quantity $\mathcal{P}$ is the normalized root-mean-square, or normalized "effective," sound pressure over the averaging time T. The time $T\approx 0.1$ second at suprathreshold loudness is the integration time of the auditory system, that is, the time taken to achieve a steady loudness sensation in response to a steady sound. The instantaneous sound pressure p is real, and t is the time. Finally,

$$p_{ref}=20\ \mu\text{Pa}$$

is a reference sound pressure near that at the threshold of loudness of a 1-kilohertz tone.

According to the proposed model, the overall loudness $\mathcal{L}$ of any steady sound is given by

$$\mathcal{L}=\kappa'(r-r_t) \tag{3}$$

where r, the total neural discharge rate (in 0.1 s) over all fibers, is proportional to the total electric power transmitted in the auditory nerve, κ' is a constant, and the subscript t refers to the value at the loudness threshold. Since the electric power is proportional to the square root of the stimulus power distributed over all neural channels,

$$\mathcal{L}=\left[\kappa^2\sum_{k=1}^{M}(\mathcal{R}_q)_k\right]^{1/2}-\left[\kappa^2\sum_{k=1}^{M}(\mathcal{R}_{qt})_k\right]^{1/2} \tag{4}$$

where M is the total number of transmission channels spanning the sound spectrum, $(\mathcal{R}_q)_k$ is the relative power transmitted in each channel k to the brain, and κ is a constant ($\kappa=10^{-2}$ if $\mathcal{L}$ is in sones, as will be shown. Also, see ref. 10). The relative power $\mathcal{R}_q$ in each channel is proportional to the square of the linearly filtered, instantaneous sound pressure integrated over the running auditory integration time T. The filtered output from each neural channel is given by

$$q(t')=\int_{-\infty}^{\infty} h(t'')p(t'-t'')\,dt'' \tag{5}$$

where h is the real impulse response of the auditory system, and q is the output. As a result of the transformation from mechanical to

electrical form, the signal transmitted in the auditory nerve is proportional to q^2. This yields the relative output power

$$\mathcal{R}_q(0,t,T) = (1/Tp_{ref}^2)\int_{t-T}^{t} q^2(t')\,dt' \tag{6}$$

involved in equation (4). In equation (6), $\mathcal{R}_q = \mathcal{R}_q(0,t,T)$ is a specialization of the autocorrelation function $\mathcal{R}_q(\tau,t,T)$, in which τ is a delay time. By virtue of equation (6), equation (4) becomes

$$\begin{aligned}\mathcal{L} = &\left\{(\kappa^2/Tp_{ref}^2)\sum_{k=1}^{M}\int_{t-T}^{t}[q^2(t')]_k\,dt'\right\}^{1/2}\\ &-\left\{(\kappa^2/Tp_{ref}^2)\sum_{k=1}^{M}\int_{t-T}^{t}[q_t^2(t')]_k\,dt'\right\}^{1/2}\end{aligned} \tag{7}$$

where, in concordance with equation (6),

$$\mathcal{R}_{qt} = (1/Tp_{ref}^2)\int_{t-T}^{t} q_t^2(t')\,dt' \tag{8}$$

at the loudness threshold. Equation (7) in combination with equation (5) yields the overall loudness in terms of measurable quantities.

It may be more convenient to express the overall loudness in terms of spectral, rather than temporal, quantities. The output autocorrelation function is defined by

$$\mathcal{R}_q(\tau,t,T) = (1/Tp_{ref}^2)\int_{t-T}^{t} q(t')\,q(t'+\tau)\,dt' \tag{9}$$

and is the Fourier transform of the output power-spectral-density W_q, that is,

$$\mathcal{R}_q(\tau,t,T) = (1/2\pi)\int_{-\infty}^{\infty} W_q(\omega,t,T)\,e^{i\omega\tau}\,d\omega \tag{10}$$

where ω is the angular frequency. Hence,

$$\mathcal{R}_q(0,t,T) = (1/2\pi)\int_{-\infty}^{\infty} W_q(\omega,t,T)\,d\omega \tag{11}$$

If equation (9) is combined with the Fourier inverse of equation (10), then

$$W_q(\omega,t,T) = (1/Tp_{ref}^2)\int_{-\infty}^{\infty} e^{-i\omega\tau}\,d\tau\int_{t-T}^{t} q(t')\,q(t'+\tau)\,dt' \tag{12}$$

This expression can be further transformed into the spectral form

$$\begin{aligned}W_q(\omega,t,T) = &(1/p_{ref}^2)H(\omega)P(\omega)\\ &\times(1/2\pi)\int_{-\infty}^{\infty} H^*(\omega'')P^*(\omega'')\operatorname{sinc}[(\omega-\omega'')T/2\pi]\\ &\times e^{-i(\omega''-\omega)(t-T/2)}\,d\omega''\end{aligned} \tag{13}$$

as shown in appendix B, where

$$H(\omega) = \int_{-\infty}^{\infty} h(t)\,e^{-i\omega t}\,dt \tag{14}$$

is the Fourier transform of the impulse response $h(t)$ and is called the frequency response, or system function;

$$P(\omega)=\int_{-\infty}^{\infty} p(t)e^{-i\omega t}dt \tag{15}$$

is the Fourier transform of the pressure history and is called the pressure spectral density; and the asterisk denotes the complex conjugate. Also, by definition,

$$\operatorname{sinc}\xi=\sin(\pi\xi)/\pi\xi \tag{16}$$

Finally, the combination of equations (4) and (11) yields

$$\begin{aligned}\mathcal{L}=&\left\{(\kappa^2/2\pi)\sum_{k=1}^{M}\int_{-\infty}^{\infty}[W_q(\omega,t,T)]_k d\omega\right\}^{1/2}\\ &-\left\{(\kappa^2/2\pi)\sum_{k=1}^{M}\int_{-\infty}^{\infty}[W_{qt}(\omega,t,T)]_k d\omega\right\}^{1/2}\end{aligned} \tag{17}$$

where the W_q's are given in spectral form by equation (13). This alternative form of equation (7) is the desired spectral representation of the overall loudness.

The threshold, output spectral density W_{qt} is defined by shifting the spectrum, without changing its shape, so that the contribution from every channel is subliminal, except for one or more, which corresponds to the loudness threshold. The threshold terms in equations (7) and (17) are significant only if the overall loudness is near the loudness threshold. These equations are defined only for $\mathcal{L}\geqq 0$. The possibility that $\mathcal{L}>0$ for a sound which is subliminal at every frequency is implicit in equations (7) and (17).

In general, the sound spectrum may be partly discrete and partly continuous. The discrete part consists of tones whereas the continuous part might consist, for example, of random noise. Thus, it is sometimes convenient to treat the output spectral density W_q as if it were either discrete $w_q(\omega_n)$, where n is an integer, or continuous $W_q(\omega)$ depending, respectively, upon whether the discrete or the continuous components of the sound are dominant, but also upon the method of detemining the spectrum. Of course, a sound can be specified in terms of both $w_q(\omega_n)$ and $W_q(\omega)$.

Discrete Spectrum

The discrete part of the sound consists of a number of tones, which can be represented by

$$p(t)=\sum_{n=1}^{\infty}A_n\cos(\omega_n t+\phi_n) \tag{18a}$$

or

$$p(t)=\sum_{n=-\infty}^{\infty}a_n e^{i\omega_n t} \tag{18b}$$

where the amplitude A_n is real, ϕ_n is the phase angle of the tone denoted by the integer n,

$$a_n = (1/2) A_n e^{i\phi_n}$$

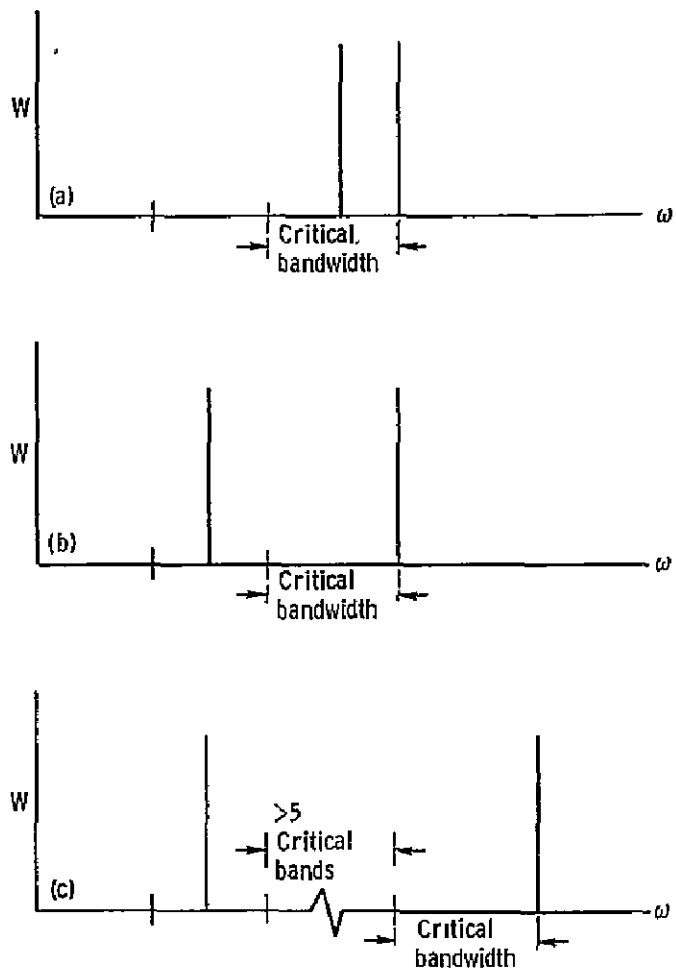
$$a_{-n} = a_n^*, \quad \omega_{-n} = -\omega_n, \quad \phi_{-n} = -\phi_n$$

According to equation (5) the corresponding output is

$$q(t) = \sum_{n=-\infty}^{\infty} H(\omega_n) a_n e^{i\omega_n t} \tag{19}$$

Because the signal transmitted by the auditory nerve is most conveniently analyzed in channels, three categories of tone separation or bandspread appear relevant, as shown in figure 5. These are denoted as

Category A: Subcritical bandspread (fig. 5(a)) in which the overall tone separation is less than, or equal to, a critical bandwidth for loudness (approximately one-third octave, but greater at low frequencies). Beats of imperfect unisons fall in this category.



(a) Two tones subcritically spaced.
(b) Two tones supercritically spaced.
(c) Two tones hypercritically spaced.

FIGURE 5.—Categories of tone spacing relative to channel bandwidths.

Category B: Supercritical tone spacing (fig. 5(b)) in which the overall tone spacing is greater than a critical bandwidth. This category includes beats of mistuned consonances, periodicity pitch, and the pitch of the residue.

Category C: Hypercritical tone spacing (fig. 5(c)) in which the tone spacing is greater than the bandwidths of nerve fibers with characteristic frequencies intermediate to the tones, so that no tone interaction occurs. The theory for category B also includes this category implicitly. Overall loudness judgments of such widely spaced tones are very difficult.

Various combinations of these basic categories occur. The bandspread of subcritically spaced tones may be supercritical (category AB) or hypercritical (category AC), and the bandspread of supercritically spaced tones might be hypercritical (category BC).

The overall loudness $\mathcal{L}$ for each category is determined as follows:

Category A: Subcritical bandspread.—When the tone bandspread is less than a critical bandwidth, equation (7) reduces to

$$\mathcal{L}=[(\kappa^2/Tp_{ref}^2)\int_{t-T}^{t} q^2(t')dt']^{1/2}-(\kappa^2\mathcal{R}_{qt})^{1/2} \tag{20}$$

where $\mathcal{R}_{qt}$ is given by equation (8). By considering equation (19) the integral in equation (20) becomes

$$\begin{aligned}\mathcal{R}_q&=(1/Tp_{ref}^2)\int_{t-T}^{t} q^2(t')dt'\\&=(1/Tp_{ref}^2)\int_{t-T}^{t}\sum_{m=-\infty}^{\infty}\sum_{n=-\infty}^{\infty} H^*(\omega_m)H(\omega_n)a_m^* a_n e^{-i(\omega_m-\omega_n)t'}dt'\\&=(1/4Tp_{ref}^2)\int_{t-T}^{t}\sum_{m=-\infty}^{\infty}\sum_{n=-\infty}^{\infty} H^*(\omega_m)H(\omega_n)A_mA_n\\&\quad\times e^{-i[(\omega_m-\omega_n)t'+\phi_m-\phi_n]}dt'\end{aligned}$$

Define the relative power

$$w(\omega_m,\omega_n)=A_mA_n/2p_{ref}^2 \tag{21}$$

where, in particular,

$$w(\omega_m,\omega_n)\equiv w(\omega_m)=A_m^2/2p_{ref}^2$$

if $\omega_m=\omega_n$. Also, for brevity, let

$$E_{m,n}\equiv e^{-i[(\omega_m-\omega_n)t'+\phi_m-\phi_n]}$$

Then,

$$\begin{aligned}\mathcal{R}_q=(1/2T)\sum_{m=1}^{\infty}\sum_{n=1}^{\infty}\int_{t-T}^{t}[&H^*(\omega_m)H(\omega_n)\,w(\omega_m,\omega_n)E_{m,n}\\&+H^*(\omega_m)H(\omega_{-n})w(\omega_m,\omega_{-n})E_{m,-n}\\&+H^*(\omega_{-m})H(\omega_n)w(\omega_{-m},\omega_n)E_{-m,n}\\&+H^*(\omega_{-m})H(\omega_{-n})w(\omega_{-m},\omega_{-n})E_{-m,-n}]dt'\end{aligned}$$

The frequencies $\omega_m-\omega_n$ and $\omega_m+\omega_n$ implicit in the E's can be generated only in those neurons which transmit periodic discharges corresponding with the frequencies ω_m and ω_n, as well. If $\omega_m-\omega_n$ is less than the critical bandwidth, then $\omega_m-\omega_n$ lies outside the pass band of those neurons which transmit ω_m and ω_n. Discharges do not occur regularly with a periodicity corresponding to $\omega_m-\omega_n$, and this tone is not heard. Rather, $\omega_m-\omega_n$ is the frequency of oscillation of the envelope of the output power due to the interaction of frequencies ω_m and ω_n. On the other hand, except possibly at low frequencies ω_m and ω_n and high stimulus amplitudes, the sum $\omega_m+\omega_n$ tends to lie outside the nominal transmission band of those neurons which best transmit ω_m and ω_n. Thus, $\omega_m+\omega_n$ is transmitted very weakly, if at all. Therefore, those terms which involve $\omega_m+\omega_n$ will be neglected.

Next, the complex, total cross-transmittance Q is defined according to

$$Q(\omega_m,\omega_n)=H^*(\omega_m)H(\omega_n)=Q_r(\omega_m,\omega_n)+iQ_i(\omega_m,\omega_n) \tag{22}$$

where Q_r is the real part, and Q_i is the imaginary part, of Q. If $\omega_m=\omega_n$, $Q(\omega_m, \omega_n)$ reduces to the transmittance

$$Q(\omega_n)=|H(\omega_n)|^2=Q_r(\omega_n)$$

which is real.

The total transmittance Q consists of the product of an external part $\mathcal{T}$ preceding the eardrum and an internal part $\mathcal{N}$ succeeding the eardrum, that is,

$$Q(\omega_m,\omega_n)=\mathcal{T}(\omega_m, \omega_n)\mathcal{N}(\omega_m,\omega_n) \tag{23}$$

or

$$Q(\omega_n)=\mathcal{T}(\omega_n)\mathcal{N}(\omega_n)$$

if $\omega_m=\omega_n$. In parallel with the definition of Q,

$$\mathcal{T}(\omega_m,\omega_n)\equiv\mathcal{H}_e^*(\omega_m)\mathcal{H}_e(\omega_n)=\mathcal{T}_r(\omega_m,\omega_n)+i\mathcal{T}_i(\omega_m,\omega_n) \tag{24}$$

or

$$\mathcal{T}(\omega_n)=|\mathcal{H}_e(\omega_n)|^2=\mathcal{T}_r(\omega_n)$$

if $\omega_m=\omega_n$. Also,

$$\mathcal{N}(\omega_m,\omega_n)=\mathcal{H}^*(\omega_m)\mathcal{H}(\omega_n)=\mathcal{N}_r(\omega_m,\omega_n)+i\mathcal{N}_i(\omega_m,\omega_n) \tag{25}$$

or

$$\mathcal{N}(\omega_n)=|\mathcal{H}(\omega_n)|^2=\mathcal{N}_r(\omega_n)$$

if $\omega_m=\omega_n$. The external and internal frequency responses $\mathcal{H}_e(\omega_n)$ and $\mathcal{H}(\omega_n)$, respectively, which will be explicitly specified subsequently, satisfy the conditions

$$\mathcal{H}_e(\omega_{-n}) = \mathcal{H}_e(-\omega_n) = \mathcal{H}_e(\omega_n) \tag{26}$$

in the case of the external response, and

$$\mathcal{H}(\omega_{-n}) = \mathcal{H}(-\omega_n) = \mathcal{H}^*(\omega_n) \tag{27}$$

in the case of the internal response. Therefore,

$$\mathcal{T}(\omega_{-m},\omega_{-n}) = \mathcal{H}_e^*(\omega_{-m})\mathcal{H}_e(\omega_{-n}) = \mathcal{H}_e^*(\omega_m)\mathcal{H}_e(\omega_n) = \mathcal{T}(\omega_m,\omega_n)$$

whereas

$$\mathcal{N}(\omega_{-m},\omega_{-n}) = \mathcal{H}^*(\omega_{-m})\mathcal{H}(\omega_{-n}) = \mathcal{H}(\omega_m)\mathcal{H}^*(\omega_n) = \mathcal{N}^*(\omega_m,\omega_n)$$

By neglecting the nontransmitted frequencies $\omega_m + \omega_n$ and adopting the newly defined transmittance functions, it follows that

$$\mathcal{R}_q = (1/2T) \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} \mathcal{T}(\omega_m,\omega_n) w(\omega_m,\omega_n) \times \int_{t-T}^{t} [\mathcal{N}(\omega_m,\omega_n) E_{m,n} + \mathcal{N}^*(\omega_m,\omega_n) E^*_{m,n}] dt'$$

where $w(\omega_{-m},\omega_{-n}) = w(\omega_m,\omega_n)$. Because $z + z^* = 2\,\mathcal{R}_E\, z$ for z complex, and because $\mathcal{R}_q$ is real,

$$\mathcal{R}_q = \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} \mathcal{T}_r(\omega_m,\omega_n)\, w(\omega_m,\omega_n) \times (1/T) \int_{t-T}^{t} \mathcal{R}_E[\mathcal{N}(\omega_m,\omega_n) E_{m,n}] dt'$$

or

$$\mathcal{R}_q = \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} \mathcal{T}_r(\omega_m,\omega_n) w(\omega_m,\omega_n) \times (1/T) \int_{t-T}^{t} \Big\{ \mathcal{N}_r(\omega_m,\omega_n) \cos[(\omega_m - \omega_n)t' + \phi_m - \phi_n] + \mathcal{N}_i(\omega_m,\omega_n) \sin[(\omega_m - \omega_n)t' + \phi_m - \phi_n] \Big\} dt' \tag{28}$$

An alternative form of the equation may be written by noting that

$$|\mathcal{N}(\omega_m,\omega_n)| = [\mathcal{N}_r^2(\omega_m,\omega_n) + \mathcal{N}_i^2(\omega_m,\omega_n)]^{1/2} \tag{29}$$

and setting

$$\tan\theta_{mn} = \mathcal{N}_i(\omega_m,\omega_n) / \mathcal{N}_r(\omega_m,\omega_n) \tag{30}$$

Then,

$$\mathcal{R}_q = \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} \mathcal{T}_r(\omega_m,\omega_n) |\mathcal{N}(\omega_m,\omega_n)| w(\omega_m,\omega_n) \times (1/T) \int_{t-T}^{t} \cos[(\omega_m - \omega_n)t' + \phi_m - \phi_n - \theta_{mn}] dt' \tag{31}$$

so that

$$\mathcal{R}_q = \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} \mathcal{T}_r(\omega_m,\omega_n) |\mathcal{N}(\omega_m,\omega_n)| w(\omega_m,\omega_n)$$
$$\times \operatorname{sinc}[(\omega_m-\omega_n)T/2\pi]\cos[(\omega_m-\omega_n)(t-T/2)+\phi_m-\phi_n-\theta_{mn}] \tag{32}$$

Therefore, the overall loudness of tones distributed over a subcritical frequency band is given by

$$\mathcal{L} = \left\{\kappa^2 \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} \mathcal{T}_r(\omega_m,\omega_n)|\mathcal{N}(\omega_m,\omega_n)|w(\omega_m,\omega_n)\right.$$
$$\times \operatorname{sinc}[(\omega_m-\omega_n)T/2\pi]\cos[(\omega_m-\omega_n)(t-T/2)$$
$$\left.+\phi_m-\phi_n-\theta_{mn}]\right\}^{1/2} - (\kappa^2\mathcal{R}_{qt})^{1/2} \tag{33}$$

Since, for subcritical bandspread, loudness is independent of bandspread if the overall sound pressure is held constant, it is reasonable to assume that, at the loudness threshold,

$$\mathcal{R}_{qt} = Q_r(\omega_\mu) w_t(\omega_\mu) \tag{34}$$

which is the lone tone result for the characteristic frequency $\omega_\mu/2\pi$, say the mean frequency of the tones. Because $\mathcal{R}_{qt}$ may be determined by several tones it follows that several subcritically spaced tones, each of which is subliminal, may produce an audible sound. This possibility is reinforced by electrophysiological experiments (ref. 42).

Except possibly at high frequencies, the transmittance functions $\mathcal{T}(\omega_m,\omega_n)$ and $\mathcal{N}(\omega_m,\omega_n)$ tend to remain constant over a critical bandwidth, as will be shown. Hence,

$$\mathcal{T}_r(\omega_m,\omega_n) \approx \mathcal{T}_r(\omega_\mu)$$

and

$$|\mathcal{N}(\omega_m,\omega_n)| \approx \mathcal{N}_r(\omega_\mu)$$

so that

$$\mathcal{L} \approx \left\{\kappa^2 Q_r(\omega_\mu) \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} w(\omega_m,\omega_n)\operatorname{sinc}[(\omega_m-\omega_n)T/2\pi]\right.$$
$$\left.\times \cos[(\omega_m-\omega_n)(t-T/2)+\phi_m-\phi_n]\right\}^{1/2} - (\kappa^2\mathcal{R}_{qt})^{1/2} \tag{35}$$

where

$$Q_r(\omega_\mu) = \mathcal{T}_r(\omega_\mu)\mathcal{N}_r(\omega_\mu)$$

and, as before, $\omega_\mu/2\pi$ is the mean frequency of the tones.

The loudness may fluctuate slowly as audible beats if

$$\gamma\omega_m \approx \epsilon\omega_n$$

where γ and ϵ are integers. Beats obtained when $\gamma=\epsilon$ are called beats of imperfect unisons (ref. 25). These beats are predicted by the preceding loudness equations. The sinc function accounts for the relative amplitude of the loudness oscillations as a function of $\omega_m-\omega_n$ and the sinusoidal dependence on the time t accounts for the time history of the oscillations. It will be shown in connection with two-tone loudness judgment tests by von Békésy that beats might be detectable for $|\omega_m-\omega_n|/2\pi\leqq 100$ hertz, but are strong only if $|\omega_m-\omega_n|/2\pi<10$ hertz. Beats for which $\gamma\neq\epsilon$ will be discussed in connection with more widely separated tones.

Category B: Supercritical tone spacing.—When the tone separation exceeds the critical bandwidth, that is, for supercritical tone spacing, the audible range is divided into adjacent channels. The channels are allowed to shift as a group in accordance with the spectrum of the sound stimulus. A method for locating channels is described in appendix C. The bandwidth of each channel is a critical bandwidth for loudness. The neurons associated with each channel have best frequencies within its associated band. However, except at the lowest and highest audible frequencies, neuron excitation bandwidths exceed the associated channel bandwidth so that the power transmitted in a channel will be affected generally by tones outside the critical band (cf. figs. 3 and 6). The interaction of tones inside the critical band with those outside the critical band accounts for the secondary tones.

Consider a sound stimulus consisting of several tones located in adjacent critical bands. None of the tones lies in the same critical band. Each tone stimulates all channels whose neurons respond to the tone

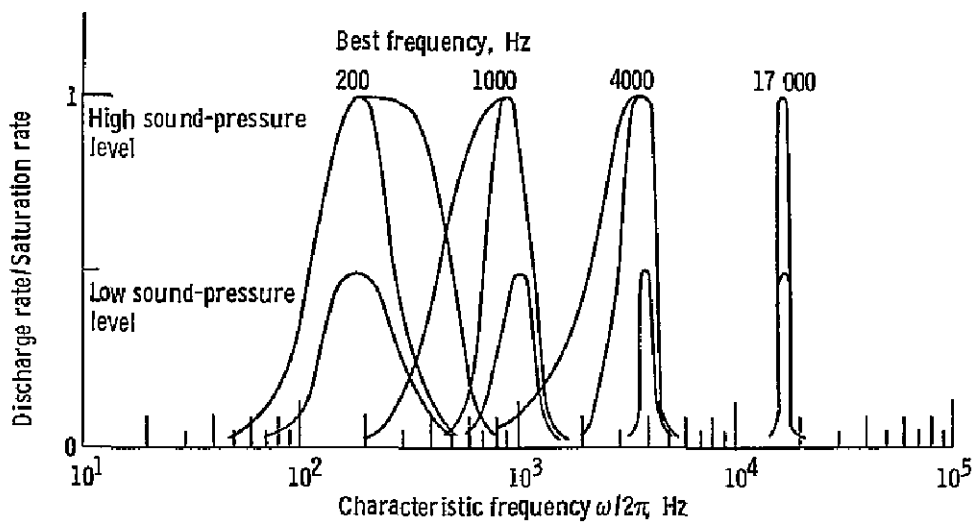


FIGURE 6.—Nominal transmission curves for individual nerve fibers. (Adapted from ref. 12.)

frequency. The output from a typical channel k may be represented by

$$q_k(t) = \sum_{n=-\infty}^{\infty} J_k(\omega_n) a_n e^{i\omega_n t}$$

where $J_k(\omega_n)$ is the frequency response of channel k to rotational frequency ω_n. Since the $J_k(\omega_n)$ are presently indeterminate, they will be replaced by $H(\omega_n)$, as in the treatment of tones spread over a subcritical band (and as in other loudness calculation procedures). The total output from all channels can be approximated by considering only enough channels to encompass the bandwidth of the sound. As mentioned earlier, the replacement of J by H is expected to constitute the main weakness of the present model since it implies that the frequency response is the same in each channel when, in fact, the response in each channel peaks in the neighborhood of the channel midfrequency (ref. 12).

In conformity with the treatment of subcritically spaced tones, it is assumed that frequencies $\omega_m+\omega_n$ are not transmitted appreciably by the nervous system because, generally ω_m, ω_n and $\omega_m+\omega_n$ do not all lie within the nominal transmission band of the same fibers.

In general, in any channel k which includes a stimulus frequency,

$$(\mathcal{R}_q)_k = \left\{ \sum_{m=1}^{N} \sum_{n=1}^{N} \mathcal{J}_r(\omega_m,\omega_n) \,|\mathcal{N}(\omega_m,\omega_n)|\, w(\omega_m,\omega_n) \right.$$
$$\left. \times (1/T) \int_{t-T}^{t} \cos\left[(\omega_m-\omega_n)t' + \phi_m - \phi_n - \theta_{mn}\right] dt' \right\}_k \qquad (36)$$

as in the subcritical case, except that N tones are now spread supercritically over N critical bands. By assuming that only self-interactions of all tones and cross-interactions involving one tone in the channel are significant, equation (36) reduces to

$$(\mathcal{R}_q)_k = \left(\sum_{m=1}^{N} \left\{ Q_r(\omega_m) w(\omega_m) \right.\right.$$
$$+ 2(1-\delta_{mk}) \mathcal{J}_r(\omega_m,\omega_k) \,|\mathcal{N}(\omega_m,\omega_k)|\, w(\omega_m,\omega_k)$$
$$\left.\left. \times (1/T) \int_{t-T}^{t} \cos\left[(\omega_m-\omega_k)t' + \phi_m - \phi_k - \theta_{mk}\right] dt' \right\} \right)_k \qquad (37)$$

It is quite apparent from psychoacoustic experiments that, for supercritical tone spacing, since the secondary tones $\omega_m-\omega_k$ are clearly heard, they must contribute appreciably to the overall loudness. In order to incorporate this contribution to the overall transmitted power, the tone interaction must be evaluated in more detail.

If the difference frequency $\omega_m-\omega_k$ is supercritical, it will be transmitted as if it were a real acoustic signal (ref. 30) and will be audible

if it lies within the passbands of auditory neurons which also transmit ω_m and ω_k. Whereas the period of beats is of the same order as, or greater than, the auditory integration time, the period of this difference tone is much less than the auditory integration time. Since $\omega_m-\omega_k$ is transmitted as an acoustic signal, neural discharges occur in only part, approximately one-half, of each cycle of oscillation (refs. 12, 24, 42, and 30). This effect must now be considered in evaluating the preceding integral of the cosine. Over time T,

$$|\omega_m-\omega_k|T>>2\pi$$

Define

$$\Delta t=2\pi/|\omega_m-\omega_k|<<T$$

The active phase of each cycle of the wave at frequency $|\omega_m-\omega_k|/2\pi$ during which discharges are generated is centered at the maximum value of $\cos[(\omega_m-\omega_k)t'+\phi_m-\phi_k-\theta_{mk}]$, that is, when

$$t=t_{max}=-\frac{\phi_m-\phi_k-\theta_{mk}}{\omega_m-\omega_k}$$

Over each stimulus cycle, during which neural activity occurs in one-half cycle, the normalized value of the integral in equation (37) is

$$\begin{aligned}(1/\Delta t)\int_{t_{max}-\Delta t/4}^{t_{max}+\Delta t/4}\cos[(\omega_m-\omega_k)t'+\phi_m-\phi_k-\theta_{mk}]dt' \\ =(1/2)\operatorname{sinc}[(\omega_m-\omega_k)\Delta t/4\pi] \\ =(1/2)\operatorname{sinc}(1/2) \\ =1/\pi\end{aligned}$$

In order to be consistent with the other terms in equation (37), which are treated as full-wave rectified, the preceding integral should be multiplied by 2. (The reference power is implicitly full-wave rectified.) Since there are $\sigma+\eta$ cycles of $\omega_m-\omega_k$ throughout the period T, where σ is an integer and $0\leqq\eta<1$, it follows that over the auditory integration time

$$\begin{aligned}(1/T)\int_{t-T}^{t}\cos[(\omega_m-\omega_k)t'+\phi_m-\phi_k-\theta_{mk}]dt' \\ \approx(2/\pi)(\sigma+\eta)\Delta t/T=2/\pi\end{aligned}$$

because $\eta<<\sigma$. Therefore, for each channel, equation (37) becomes

$$\begin{aligned}(\mathcal{R}_q)_k=\Bigg[\sum_{m=1}^{N}\Big\{&\mathcal{Q}_r(\omega_m)w(\omega_m) \\ &+(4/\pi)(1-\delta_{mk})\mathcal{T}_r(\omega_m,\omega_k)|\mathcal{N}(\omega_m,\omega_k)|\,w(\omega_m,\omega_k)\Big\}\Bigg]_k\end{aligned}$$

The corresponding overall result for all channels is obtained by appropriately combining the results from each channel. Tones cen-

tered in adjacent critical bands are a critical bandwidth apart. For this condition the preceding equation summed over any two adjacent bands must agree with the corresponding equation for two tones subcritically separated by almost one critical bandwidth. Thus,

$$\begin{aligned}\sum_{k=1}^{N} (\mathcal{R}_q)_k &= (1/2) \sum_{k=1}^{N} \sum_{m=1}^{N} \Big\{ Q_r(\omega_m) w(\omega_m) \\ &\quad + (4/\pi)\beta^2_{mk}(1-\delta_{mk}) \mathcal{T}_r(\omega_m,\omega_k) \,|\mathcal{N}(\omega_m,\omega_k)|\, w(\omega_m,\omega_k) \Big\}_k \\ &= (1/2) \sum_{k=1}^{N} \Big[Q_r(\omega_k) w(\omega_k) + \sum_{m=1}^{N} (1-\delta_{mk}) \Big\{ Q_r(\omega_m) w(\omega_m) \\ &\quad + (4/\pi)\beta^2_{mk} \mathcal{T}_r(\omega_m,\omega_k) \,|\mathcal{N}(\omega_m,\omega_k)|\, w(\omega_m,\omega_k) \Big\} \Big]\end{aligned}$$

where, in the first form, the subscript k denoting the channel has been replaced in the second form by the subscript k denoting the tone in channel k. The factor 1/2 is associated with the fact that two tones separated by a critical bandwidth are also approximately centered in adjacent critical bands. Thus, there is a significant overlapping of neurons appreciably stimulated by the two tones in the sense that the two adjacent channels in which the tones are centered are associated with a power transfer equal to that from one channel bounded by the tones.

The newly introduced empirical quantity β_{mk} is a function of ω_m and a monotonic function of $\omega_m - \omega_k$. Specifically,

$\beta_{mk}=0$	for adjacent tones separated by a critical bandwidth
$\beta_{mk}\rightarrow 1$	as the separation of adjacent tones approaches two critical bandwidths,
$\beta_{mk}=1$	for nonadjacent tones

Thus β_{mk} differs from unity only for adjacent tones separated by one to two critical bandwidths, and then only in interaction terms containing m and k. Therefore, in general, the function β_{mk} can be set equal to unity. The only time its variation is likely to be of practical significance is in the case of sounds with bandwidths slightly supercritical and with nearly equal power in both critical bands.

For adjacent tones the function $\beta_{mk} \neq 0$ is related psychoacoustically to the onset of secondary tones and physiologically to the fraction of the total number of neural pathways which may transmit information on these tones. Thus, as the tone separation increases, the corresponding β_{mk} tends to increase to unity. However, as this separation limit is greatly exceeded, β_{mk} might decline again as the number of fibers which transmit secondary tones declines. Also, the interaction transmittance function $Q(\omega_m,\omega_k)$ must tend to vanish when hyper-

critical separation is reached since both frequencies, ω_m and ω_k, cannot, then, stimulate the same fibers. The loudness of hypercritically spaced tones will not be studied further.

When the preceding equation is introduced in the loudness equation (4), it becomes

$$\begin{aligned}\mathcal{L}=\Bigg\{(\kappa^2/2)\sum_{k=1}^{N}\Bigg[&Q_r(\omega_k)w(\omega_k)\\&+\sum_{m=1}^{N}(1-\delta_{mk})\{Q_r(\omega_m)w(\omega_m)\\&+(4/\pi)\beta_{mk}^2\mathcal{T}_r(\omega_m,\omega_k)|\mathcal{N}(\omega_m,\omega_k)|w(\omega_m,\omega_k)\}\Bigg]\Bigg\}^{1/2}\\&-\Bigg[\kappa^2\sum_{k=1}^{N}(\mathcal{R}_{qt})_k\Bigg]^{1/2}\end{aligned}\tag{38a}$$

which is the overall loudness equation for tones supercritically spaced in adjacent critical bands, and where

$$\begin{aligned}\sum_{k=1}^{N}(\mathcal{R}_{qt})_k=(1/2)\sum_{k=1}^{N}\Bigg[&Q_r(\omega_k)w_t(\omega_k)\\&+\sum_{m=1}^{N}(1-\delta_{mk})\Big\{Q_r(\omega_m)w_t(\omega_m)\\&+(4/\pi)\beta_{mk}^2\mathcal{T}_r(\omega_m,\omega_k)|\mathcal{N}(\omega_m,\omega_k)|w_t(\omega_m,\omega_k)\Big\}\Bigg]\end{aligned}\tag{39}$$

This threshold expression might seem too large since $w_t(\omega_k)$, alone, is a threshold value. In order to justify equation (39) consider the broad range of frequencies for which $Q_r(\omega)$ is constant. Suppose that the threshold term in equation (38a) is $(\kappa^2\mathcal{R}_{qt})^{1/2}$ with no sum, and that $\mathcal{R}_{qt}$ is given by equation (34), which applies for a subcritical band of tones. Then, as the tone separation is made supercritical without changing the overall sound pressure, equation (38a) would indicate that the loudness rises. This is true for suprathreshold loudness. However, experimental data (ref. 43) clearly indicate that, near threshold, loudness does not rise but, in fact, may tend to decline. Both the suprathreshold and threshold loudness results are accounted for if it is assumed that the power $\mathcal{R}_q$ is determined in the same way down to the loudness threshold. Then, equation (39) is valid. Of course, the threshold term is important only for sounds near the loudness threshold.

The preceding loudness equation applies for tones in adjacent critical bands. If the tone separation is supercritical but the tones do not lie in adjacent critical bands, then the equation must be modified to incorporate the fact that neurons characterizing intermediate, "empty" channels are effectively stimulated by the interacting tones. This might be accomplished by multiplying the contributions to

occupied channels bounding the empty channels by a factor greater than unity in order to account for the empty channels. An alternative procedure is to relabel the sums so that they denote critical bands rather than tones. Then, the preceding loudness equation becomes

$$\mathcal{L}=\Bigg((\kappa^2/2)\sum_{\kappa=1}^{M}\Bigg[\mathcal{Q}_r(\omega_\kappa)w(\omega_\kappa)+\sum_{\mu=1}^{M}(1-\delta_{\mu\kappa})\{\mathcal{Q}_r(\omega_\mu)w(\omega_\mu)$$
$$+(4/\pi)\beta_{\mu\kappa}^2\mathcal{J}_r(\omega_\mu,\omega_\kappa)\,|\mathcal{N}(\omega_\mu,\omega_\kappa)|\,w(\omega_\mu,\omega_\kappa)\}\Bigg]\Bigg)^{1/2}$$
$$-\left[\kappa^2\sum_{\kappa=1}^{M}(\mathcal{R}_{qt})_\kappa\right]^{1/2} \tag{40}$$

where the subscripts κ and μ refer to each of the M critical bands spanning the total bandwidth of the tones, $w(\omega_\mu,\omega_\kappa)=0$ if the critical band is empty, and $\beta_{\mu\kappa}$ may differ from unity only in conjunction with interactions between adjacent critical bands. Also, $\sum_\kappa(\mathcal{R}_{qt})_\kappa$ is expressed as in equation (39), with the reinterpretation just stated.

Beats of mistuned consonances (ref. 26) represent a special case in which loudness fluctuations may occur even though $|\omega_m-\omega_n|>>0$. Closely spaced tones produce audible beats. More generally, any tone pair which satisfies $\gamma\omega_m\approx\epsilon\omega_n$ may yield audible beats. Beats of mistuned consonances occur when the integers γ and/or ϵ are greater than unity. The possible values of γ and ϵ which yield subjective beats are limited by the fact that the tones must interact. This implies that both tones must stimulate the same nerve fibers. When the transmittance bandwidths for individual fibers are considered (ref. 12), it is concluded that possible absolute values of γ and ϵ are probably limited to 5, or so. This conclusion is confirmed by psychoacoustic experiments (ref. 44), where the maximum number of audible harmonics was found to be 5 to 7. For γ and/or $\epsilon>1$, in equation (37),

$$(1/T)\int_{t-T}^{t}\cos[(\omega_m-\omega_k)t'+\phi_m-\phi_k-\theta_{mk}]dt'$$
$$\approx(2/\pi)\operatorname{sinc}[(\gamma\omega_m-\epsilon\omega_k)T/2\pi\gamma]\cos\{[(\gamma\omega_m-\epsilon\omega_k)/\gamma](t-T/2)\}$$

as shown in appendix D. Hence,

$$\mathcal{L}=\Bigg\{(\kappa^2/2)\sum_{k=1}^{M}\Bigg[\mathcal{Q}_r(\omega_k)w(\omega_k)+\sum_{m=1}^{M}(1-\delta_{mk})\{\mathcal{Q}_r(\omega_m)w(\omega_m)$$
$$+(4/\pi)\beta_{mk}^2\mathcal{J}_r(\omega_m,\omega_k)\,|\mathcal{N}(\omega_m,\omega_k)|\,w(\omega_m,\omega_k)$$
$$\times\operatorname{sinc}[(\gamma\omega_m-\epsilon\omega_k)T/2\pi\gamma]\cos\{[(\gamma\omega_m-\epsilon\omega_k)/\gamma](t-T/2)\}\}\Bigg]\Bigg\}^{1/2}$$
$$-\left[\kappa^2\sum_{k=1}^{N}(\mathcal{R}_{qt})_k\right]^{1/2} \tag{38b}$$

replaces equation (38a) in the presence of beats of mistuned consonances. Equation (38b) reverts to equation (38a) in the absence of

these beats, that is, if $\gamma\omega_m - \epsilon\omega_k \neq 0$. Beats of mistuned consonances may be significant only if $|\gamma\omega_m - \epsilon\omega_k| < 10$ hertz, as in the case of beats of imperfect unisons. In figure 7 the upper limit of the frequency difference for detectable beats of imperfect unisons (ref. 25) is compared with the critical bandwidth (ref. 28).

Category C: Hypercritical tone spacing.—When the tones are sufficiently separated, $Q(\omega_m, \omega_n) \to 0$, as will be shown; so that the tone interaction, and hence the periodicity pitch, vanishes in equation (40).

Category AB: Subcritically spaced tones spread over supercritical bandwidth.—Subcritically spaced tones cannot be resolved. Supercritically spaced tones can be resolved. Therefore, the output power from subcritically spaced tones spread over a supercritical band may be evaluated from the standpoint of loudness by treating each critical band of tones as if it were a single effective tone and then treating the effective tones from all bands as if they were supercritically spaced tones. Thus, the output power in any given critical band, assuming unity transmittance, is

$$w(\omega_\mu) = \left\{ \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} w(\omega_m, \omega_n) \operatorname{sinc}[(\omega_m - \omega_n) T/2\pi] \times \cos[(\omega_m - \omega_n)(t - T/2) + \phi_m - \phi_n] \right\}_\mu$$

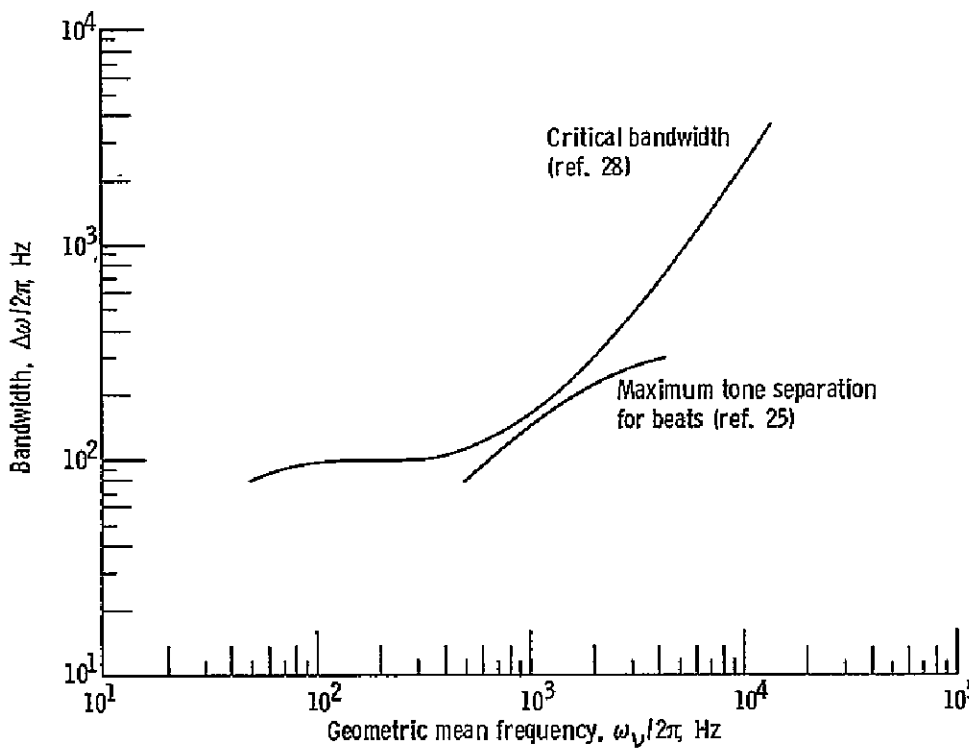


FIGURE 7.—Comparison of critical bandwidth for loudness and maximum tone separation for beats of imperfect unisons.

as in equation (35). However, the subscript μ refers to a critical band. By using this expression to evaluate the effective power in each critical band, the overall loudness can subsequently be computed from equations (38) or (40).

An elementary interpretation of the sums in equation (40) is as follows when each critical band contains several tones. The tones contained in each of the critical bands $\mu \neq \kappa$ interact collectively in each neuron with those within a critical band κ to produce additional relative output power, ascribed to secondary tones, from the neurons characterizing the κth band. All contributions to the κth band from the μ bands act independently because they arise from statistically independent neural discharges. Therefore, the contributions of the μ bands to the transmitted power in the κth band should tend to be additive in all fibers which are dynamically active. The same argument applies with respect to tones in each critical band $\kappa \neq \mu$ interacting collectively with tones within the critical band μ. Since the neuron bandwidths generally exceed the critical bandwidths the same argument can be used to account for power transmitted by fibers characterizing intermediate, empty, critical bands.

Category AC: Subcritically spaced tones spread over hypercritical bandwidth.—The loudness is evaluated in the same manner as in category AB, by means of equations (38) or (40).

Category BC: Supercritically spaced tones spread over hypercritical bandwidth.—The loudness is given by equation (40).

Continuous Spectrum

The loudness of sounds with continuous spectra may be deduced in the same manner as for tones. As in the case of sounds with discrete spectra, the appropriate expression for the loudness depends on the spectrum bandwidth. The bandwidth of the sound may be subcritical, supercritical, or hypercritical.

Subcritical bandwidth.—Sounds with continuous spectra are usually specified in spectral, rather than temporal, form. The loudness of a subcritical band of noise expressed in spectral form is

$$\mathcal{L} = [(\kappa^2/2\pi) \int_{-\infty}^{\infty} W_q(\omega,t,T)\, d\omega]^{1/2} - [(\kappa^2/2\pi) \int_{-\infty}^{\infty} W_{qt}(\omega,t,T)\, d\omega]^{1/2} \tag{41}$$

by virtue of equation (17), where W_q is given by equation (13). In order to evaluate W_q from spectral measurements, equation (13) must be approximated. Specifically, by assuming that $H(\omega)$ and $P(\omega)$ are constant in frequency intervals at least 20 hertz wide and don't vary by orders of magnitude within a few hertz, it is shown in appendix B that

$$W_q(\omega,t,T)=\square\left(\frac{t}{T}-\frac{1}{2}\right)|H(\omega)|^2|P(\omega)|^2/Tp_{ref}^2 \tag{42}$$

where

$$\square\left(\frac{t}{T}-\frac{1}{2}\right)=\begin{cases}1 & 0<t<T\\ 0 & 0>t>T\end{cases}$$

is the rectangular function. Let

$$W(\omega,T)=\square\left(\frac{t}{T}-\frac{1}{2}\right)|P(\omega)|^2/Tp_{ref}^2 \tag{43}$$

define the relative power-spectral-density of the sound, where, without loss of generality, the dependence of W on t has been suppressed. (Note that $W(\omega,T)$ has the dimension [seconds], whereas $w(\omega_n)$ for tones is dimensionless (cf. eq. (21)).) Then, the loudness is given by

$$\begin{aligned}\mathcal{L}=&[(\kappa^2/2\pi)\int_{-\infty}^{\infty}Q(\omega)W(\omega,T)d\omega]^{1/2}\\ &-[(\kappa^2/2\pi)\int_{-\infty}^{\infty}Q(\omega)W_t(\omega,T)d\omega]^{1/2}\end{aligned} \tag{44}$$

where

$$Q(\omega)=|H(\omega)|^2 \tag{45}$$

This formula for loudness can easily be related to experimental measurements. As defined in equations (1) and (2), the overall sound-pressure level is given by

$$S=10\log[(1/Tp_{ref}^2)\int_{t-T}^{t}p^2(t')dt']$$

However, as shown in appendix E,

$$(1/T\,p_{ref}^2)\int_{t-T}^{t}p^2(t')dt'=(1/2\pi)\int_{-\infty}^{\infty}W(\omega,T)d\omega \tag{46}$$

if the temporal integral is stationary. Hence,

$$S=10\log[(1/2\pi)\int_{-\infty}^{\infty}W(\omega,T)d\omega]$$

An audio-spectrum analyzer indicates the sound-pressure level in frequency bands. Ideally, in the band denoted by α,

$$S(\omega_\alpha,(\Delta\omega)_\alpha,T)=10\log[(1/2\pi)\int_{\omega_a}^{\omega_b}W(\omega,T)d\omega] \tag{47}$$

where S is to be measured over an ideal subcritical bandwidth

$$(\Delta\omega)_\alpha=(\omega_b-\omega_a)_\alpha$$

geometrically centered at the rotational frequency

$$\omega_\alpha=\sqrt{\omega_a\omega_b}$$

Thus, the cutoff frequencies are explicitly given by

$$\omega_a = \frac{1}{2}\left\{ -(\Delta\omega)_\alpha + [(\Delta\omega)_\alpha^2 + 4\omega_\alpha^2]^{1/2} \right\}$$

$$\omega_b = \frac{1}{2}\left\{ (\Delta\omega)_\alpha + [(\Delta\omega)_\alpha^2 + 4\omega_\alpha^2]^{1/2} \right\}$$

By inverting equation (47) it follows that the relative power in the bandwidth $(\Delta\omega)_\alpha$ is given by

$$\Omega(\omega_\alpha, (\Delta\omega)_\alpha) \equiv (1/2\pi) \int_{\omega_a}^{\omega_b} W(\omega, T)\, d\omega$$
$$= \text{antilog}[S(\omega_\alpha, (\Delta\omega)_\alpha)/10]$$

where, for brevity, the variable T has been suppressed. Since (neglecting beats) loudness is found to be independent of bandwidth for sounds with subcritical bandwidth, the loudness of a sound with subcritical bandwidth may be determined from measurements over a critical band which contains the subcritical band, that is,

$$\mathcal{L} = [\kappa^2 Q(\omega_\mu)\Omega_{cr}(\omega_\mu)]^{1/2} - \{\kappa^2 Q(\omega_\mu)[\Omega_{cr}(\omega_\mu)]_t\}^{1/2} \tag{48}$$

where $(\Delta\omega)_\mu$ (suppressed) spans the μth critical band, which is assumed to contain all the noise. In equation (48), Ω_{cr} is to be interpreted as the acoustic power imposed in time T in the optical bandwidth (so that $(\Delta\omega)_\mu$ can be suppressed). Also, it has been assumed that Q is constant over the critical band.

Supercritical bandwidth.—If the spectrum bandwidth is supercritical, then the relative output power can be predicted in a manner analogous to that in category AB for tones, except that the relative power is over critical bands rather than for discrete frequencies. The relative output power in any given critical band μ is given by

$$[\Omega_q(\omega_\mu)]_{cr} = Q(\omega_\mu)\Omega_{cr}(\omega_\mu) \tag{49}$$

where Q is assumed to be constant over the critical bandwidth. Equation (49) permits the calculation of $(\Omega_q)_{cr}$ from spectral measurements of S. Since $(\Omega_q)_{cr}$ can be calculated, the loudness can be calculated from

$$\begin{aligned}\mathcal{L} = \Bigg((\kappa^2/2) \sum_{\kappa=1}^{M} \Bigg[& Q_r(\omega_\kappa)\Omega_{cr}(\omega_\kappa) \\ & + \sum_{\mu=1}^{M} (1-\delta_{\mu\kappa}) \{ Q_r(\omega_\mu)\Omega_{cr}(\omega_\mu) \\ & + (4/\pi)\beta_{\mu\kappa}^2 \mathcal{J}_r(\omega_\mu,\omega_\kappa) \,|\mathcal{N}(\omega_\mu,\omega_\kappa)|\, \Omega_{cr}(\omega_\mu,\omega_\kappa) \} \Bigg] \Bigg)^{1/2} \\ & - \left[\kappa^2 \sum_{\kappa=1}^{M} (\mathcal{R}_{qt})_\kappa \right]^{1/2}\end{aligned} \tag{50}$$

which follows from equation (40), and where

$$\Omega_{cr}(\omega_\mu,\omega_\kappa)\equiv[\Omega_{cr}(\omega_\mu)\,\Omega_{cr}(\omega_\kappa)]^{1/2} \tag{51}$$

The interpretation of the sums in equation (50) is similar to that previously given for the sums in equation (40). Signals associated with separate critical bands interact because the neuron bandwidths exceed the critical bandwidths.

Hypercritical bandwidth.—If the spectrum bandwidth is hypercritical, then the output power Ω_q and overall loudness $\mathcal{L}$ are obtained as in the preceding case.

Mixed Spectrum

A mixed spectrum is defined as one consisting of both discrete and continuous components. For a sound with a mixed spectrum the output over a critical band is given by the sum of contributions from the discrete and continuous parts of the sound spectrum. If a broadband spectrum analyzer, say a one-third-octave analyzer, is used to analyze a sound with mixed spectrum, the tones will be unresolved. Then, the spectrum may be treated as if it were continuous. This might be satisfactory if the tones are not closely spaced. However, it should be noted that since one-third-octave bandwidths exceed critical bandwidths over a large portion of the audio-frequency range an instrument with narrower bandwidths would be preferred.

LOUDNESS LEVEL

It is more common to express the subjective intensity as a loudness level in phons rather than loudness in sones. In general, loudness $\mathcal{L}$ and loudness level L are related by

$$L=20\log[(\mathcal{L}/\mathcal{K}_1 p_{ref})+1]$$

where $L=0$ when $\mathcal{L}=0$ (ref. 21). By definition, $\mathcal{L}=1$ sone when $L=40$ phons. Hence, $\mathcal{K}_1 p_{ref}=10^{-2}$, with negligible error. Therefore,

$$L=20\log(10^2\mathcal{L}+1) \tag{52}$$

or, for suprathreshold loudness ($\mathcal{L}\geqq 1$),

$$L\approx 20(\log\mathcal{L}+2) \tag{53}$$

For future use, the loudness-level difference between N tones and any one tone, or group of tones, is expressed by

$$\Delta L=20\log[\mathcal{L}(\omega_1,\omega_2,\ldots,\omega_N)/\mathcal{L}(\omega_\nu,\omega_\rho,\ldots)] \tag{54}$$

so that

$$L(\omega_1,\omega_2,\ldots,\omega_N)=L(\omega_\nu,\omega_\rho,\ldots)+\Delta L \tag{55}$$

TRANSMITTANCE FUNCTIONS

The transmittance function $Q(\omega)$ may be evaluated explicitly from results of psychoacoustic studies (refs. 20, or 19, 45, and 46). These studies show that Q, more particularly $\mathcal{N}$, is a weak function of the sound pressure p, as well as of ω. Specifically,

$$Q(\omega,p)=\mathcal{T}(\omega)\mathcal{N}(\omega,p_e) \tag{56}$$

where $p_e=p_e(p,\mathcal{T})$ is the sound pressure at the eardrum (ref. 21).

It is convenient to express $\mathcal{T}$, $\mathcal{N}$, and p_e in decibel notation. Thus, by definition,

$$T=10 \log \mathcal{T} \tag{57}$$

$$N=10 \log \mathcal{N} \tag{58}$$

where T is the external transmittance level, and N is the internal transmittance level. The sound pressure at the eardrum p_e corresponds to the sound-pressure level at the eardrum S_e, in conformity with equations (1) and (2). Also, S_e is readily evaluated from

$$S_e=S+T$$

The total transmittance Q is derived from judged equal-loudness curves shown in figure 8 (ref. 20). The external transmittance level T has been measured directly (refs. 45 and 46). The curves tabulated in table I and displayed in figure 9 approximate these results. Hence, the internal transmittance $\mathcal{N}$, or its level N, can be determined.

For free-field listening the experimentally derived transmittance $\mathcal{T}$ can be satisfactorily approximated by an empirical formula given in appendix F. The empirical formula and the experimental curves are in satisfactory agreement, as shown in figure 9.

For listening through earphones the selected model of the outer ear consists of a larger diameter outer tube, representing the earphone and pinna, in series with a smaller diameter closed tube, representing the auditory meatus. For an assumed plane wave the external transmittance of this arrangement is independent of frequency (ref. 33, pp. 154–156), so that it is sufficient to set $\mathcal{T}=1$, or $T=0$, in this case.

The internal-transmittance-level function $N(\omega,S_e)$ can be found (ref. 21) from the equal-loudness curves shown in figure 8 by subtracting out the external-transmittance-level function $T(\omega)$ for free-field listening (fig. 9). The resultant dashed curves are in fairly good agreement, especially for suprathreshold loudnesses, with equal loudness curves for earphone listening reported in reference

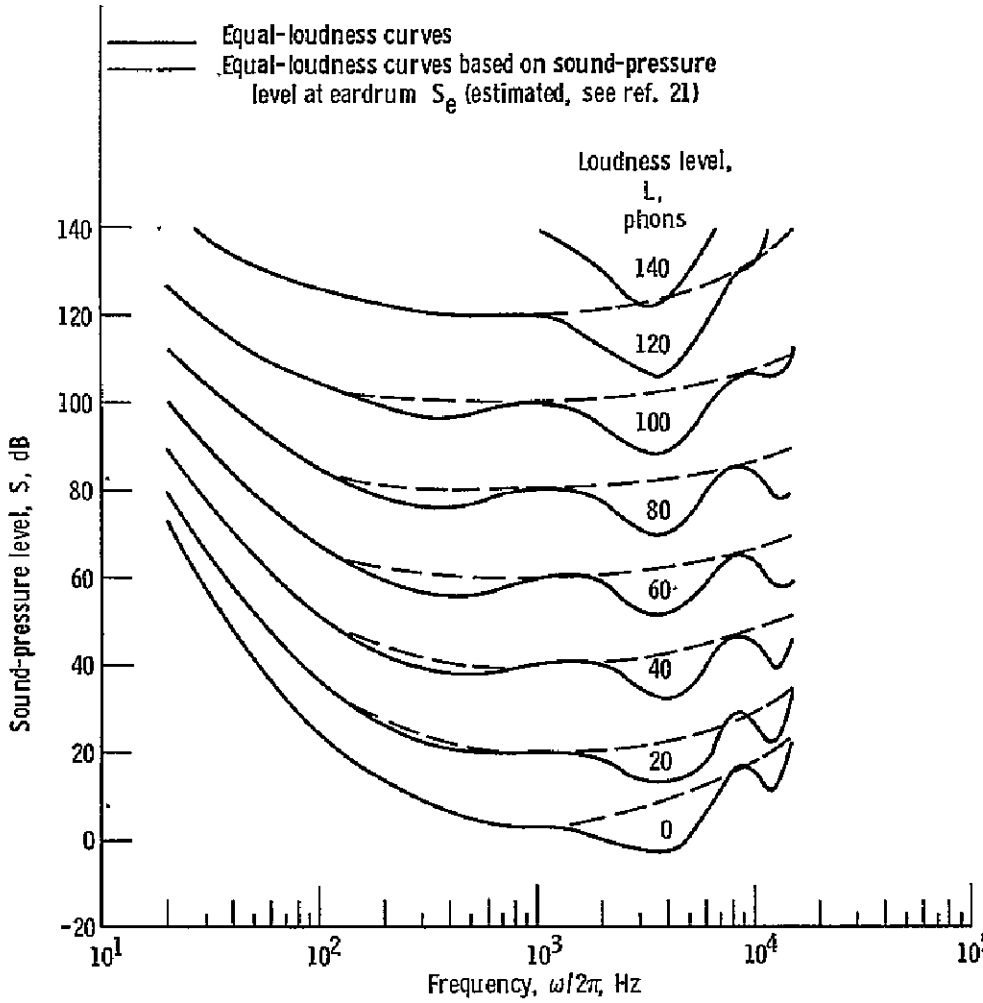


FIGURE 8.—Equal-loudness curves for plane waves incident from front. (Data from ref. 20.)

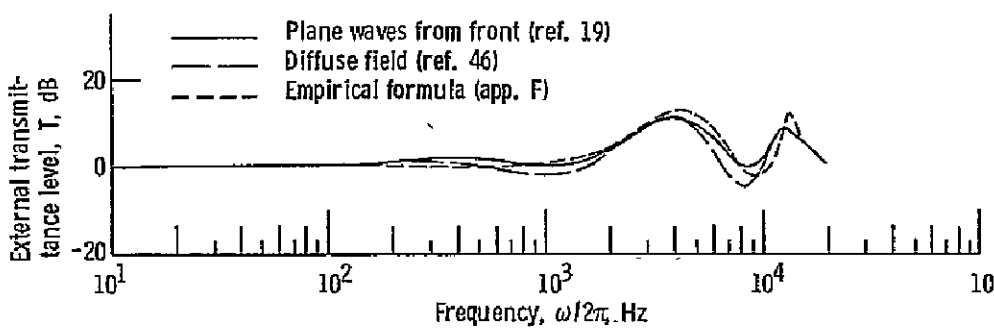


FIGURE 9.—External transmittance level function $T(\omega)$.

27. The experimentally derived internal transmittance $\mathcal{N}$ can be satisfactorily approximated (ref. 21) by the formula

$$\mathcal{N}(\omega,S_e)=\{(\omega/\omega_c)^2\,[1+(\omega_c/\omega_l)^2]\,[1+(\omega_c/\omega_u)^2]\,[1+(\omega/\omega_l)^2]^{-1}[1+(\omega/\omega_u)^2]^{-1}\}^6 \qquad (59)$$

TABLE I.—EXTERNAL TRANSMITTANCE LEVEL FUNCTION T

One-third-octave-band midfrequency, Hz	External transmittance level, T, dB	
	Plane wave incident frontally	Diffuse field
<160	0	0
160	.5	.5
200	1	1
250	1.5	1
315	2	1
400	2	.5
500	2	0
630	1.5	−1
800	.5	−2
1 000	0	−2.5
1 250	0	−2
1 600	1.5	0
2 000	4	3.5
2 500	7	7.5
3 150	10	11
4 000	12	12.5
5 000	10	8.5
6 300	6	2
8 000	0	−5
10 000	2	−1
12 500	9	9
16 000	5	----
20 000	0	----

which originates from the problem of a high-pass filter in series with a low-pass filter (ref. 47). In equation (59),

$$\omega_c/2\pi = 1000 \text{ hertz}$$

is taken as the center frequency;

$$\omega_u/2\pi = 18\,000 \text{ hertz}$$

is the upper cutoff frequency, and

$$\omega_l/2\pi = \text{antilog}_{10}\ (2.1761 - 0.005706\ S_e) \text{ hertz}$$

is the lower cutoff frequency of the auditory system. The lower cutoff frequency is a weak function of S_e. The experimental and analytical transmittance levels $N(\omega,S_e)$ compared in reference 21 are in good agreement. The analytical curves are shown in figure 10. (The standard A-weighting for sound-level meters is also shown in fig. 10 for comparison. Clearly the A-weighting is in poor agreement with the standard transmittance curves for loudness (ref. 20), especially at low frequencies.)

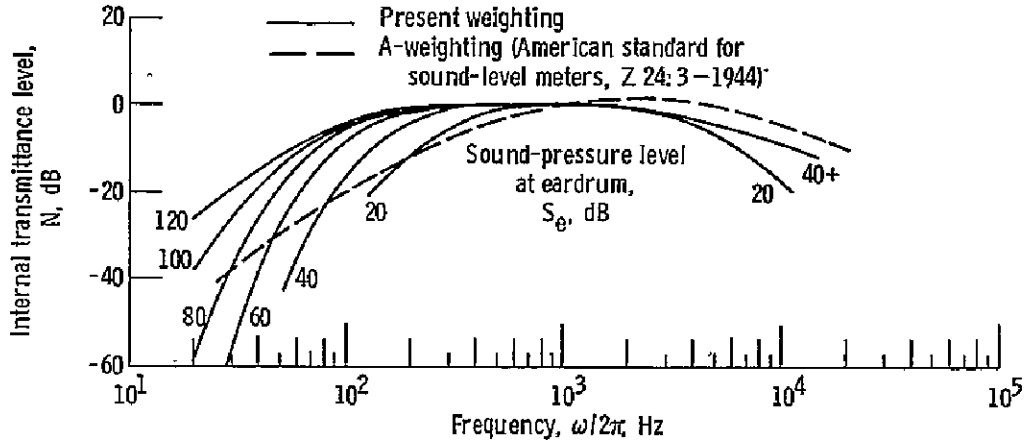


FIGURE 10.—Internal transmittance level function $N(\omega, S_e)$.

In summary, for free-field (without earphones) listening, Q is given by equation (56), and the required formulas for $\mathcal{T}$ and $\mathcal{N}$ are given in appendix F and by equation (59), respectively. For listening through earphones, $\mathcal{T}=1$, so that

$$Q(\omega,S_e)=\mathcal{N}(\omega,S_e) \tag{60}$$

In addition to the total transmittance $Q(\omega)$ (where, for brevity, the dependence on S_e has been suppressed), the cross-transmittance $Q(\omega_m,\omega_n)$ is involved in those terms of the loudness equations which represent spectrum interactions.

For free-field listening, in which case the external transmittance $\mathcal{T}$ must be considered, the required real part of the external cross-transmittance $\mathcal{T}(\omega_m,\omega_n)$, that is, $\mathcal{T}_r(\omega_m,\omega_n)$, can be expressed by an empirical function given in appendix F. The dependence of this function on $|\omega_m-\omega_n|/2\pi$ for various geometric-mean frequencies $\sqrt{\omega_m\omega_n}/2\pi$ is shown in figure 11. It is apparent that

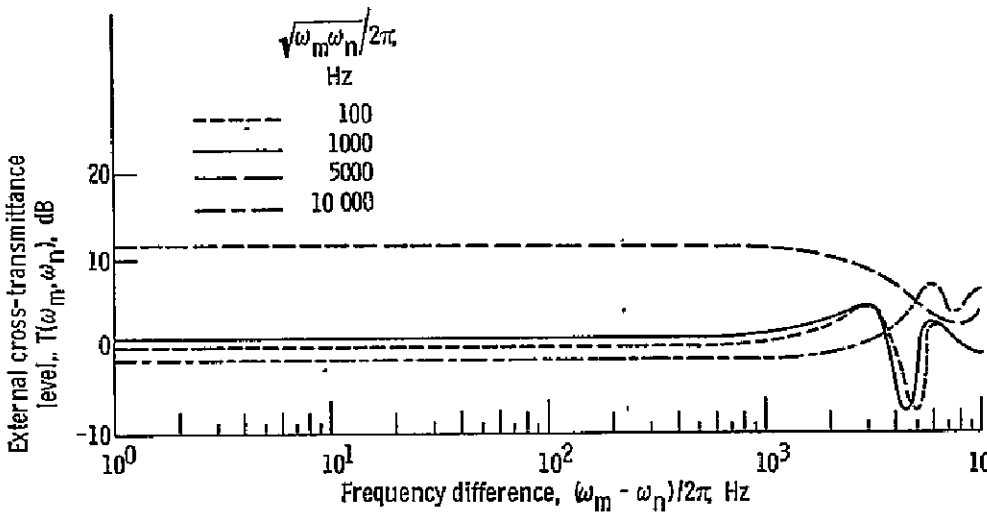


FIGURE 11.—External cross-transmittance-level function $T(\omega_m,\omega_n)$.

$$T_r(\omega_m,\omega_n) \approx T_r(\sqrt{\omega_m\omega_n}) \qquad \text{for } |\omega_m-\omega_n|/2\pi \lesssim 2000 \text{ hertz}$$

The required absolute value (cf. eq. (38) or (40)) of the internal cross-transmittance $\mathcal{N}(\omega_m,\omega_n)$ can be determined by comparing the relation $\mathcal{N}(\omega_m,\omega_n)=\mathcal{H}^*(\omega_m)\mathcal{H}(\omega_n)$ with equation (59) for $\mathcal{N}(\omega,S_e)$. It is apparent that equation (59) results if

$$\mathcal{H}(\omega_n)=\{(i\omega_n/\omega_c)\,[1+(\omega_c/\omega_l)^2]^{1/2}\,[1+(\omega_c/\omega_u)^2]^{1/2} \times(1+i\omega_n/\omega_l)^{-1}(1+i\omega_n/\omega_u)^{-1}\}^6 \tag{61}$$

since $\mathcal{H}(-\omega_n)=\mathcal{H}^*(\omega_n)$. From this it can be shown that

$$|\mathcal{N}(\omega_m,\omega_n)|=(\alpha^2+\beta^2)^3 \tag{62}$$

where

$$\alpha\equiv(c/d)\,[(1+\omega_m\omega_n/\omega_l^2)(1+\omega_m\omega_n/\omega_u^2)-(\omega_m-\omega_n)^2/\omega_l\omega_u] \tag{63}$$

$$\beta\equiv(c/d)\,\{(1+\omega_m\omega_n/\omega_l^2)\,[(\omega_m-\omega_n/\omega_u)]+(1+\omega_m\omega_n/\omega_u^2)\,[(\omega_m-\omega_n)/\omega_l]\} \tag{64}$$

and

$$c\equiv[1+(\omega_c/\omega_l)^2][1+(\omega_c/\omega_u)^2]\,(\omega_m\omega_n/\omega_c^2) \tag{65a}$$

$$d\equiv\{(1+\omega_m\omega_n/\omega_l^2)^2+[(\omega_m-\omega_n)/\omega_l]^2\}\times\{(1+\omega_m\omega_n/\omega_u^2)^2 +[(\omega_m-\omega_n)/\omega_u]^2\} \tag{65b}$$

Also

$$\mathcal{N}_r(\omega_m,\omega_n)=|\mathcal{N}(\omega_m,\omega_n)|\cos 6\psi$$

$$\mathcal{N}_i(\omega_m,\omega_n)=|\mathcal{N}(\omega_m,\omega_n)|\sin 6\psi$$

where

$$\tan\psi=\beta/\alpha$$

To a good approximation,

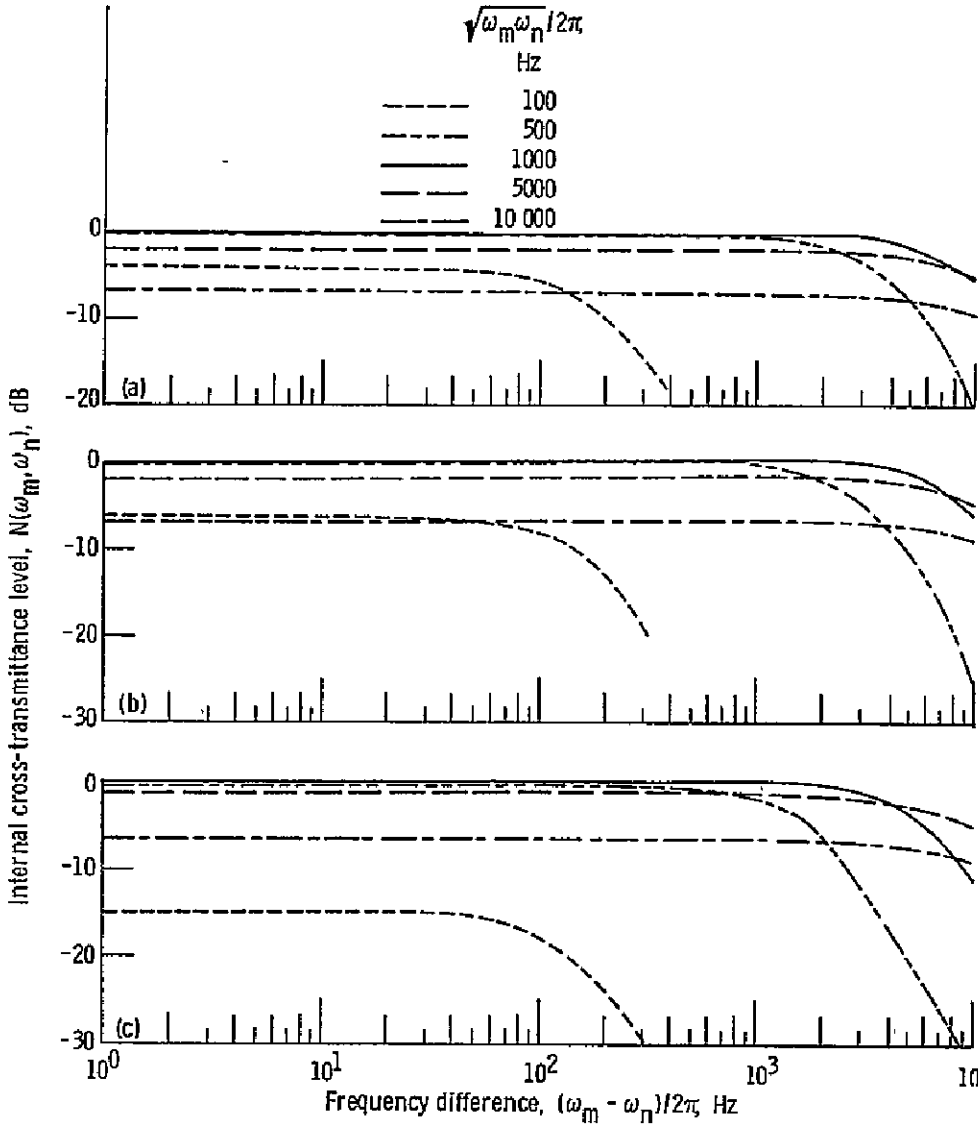
$$|\mathcal{N}(\omega_m,\omega_n)|=|\mathcal{H}(\sqrt{\omega_m\omega_n})|^2\approx\mathcal{N}_r(\sqrt{\omega_m\omega_n})$$

if $|\omega_m-\omega_n| << \omega_l$. This approximation is also valid if $\omega_l << (\omega_m,\omega_n) << \omega_u$. The dependence of $|\mathcal{N}(\omega_m,\omega_n)|$ on $\omega_m-\omega_n$ for various geometric-mean frequencies is shown in figure 12. This function also equals $|Q(\omega_m,\omega_n)|$ for listening with earphones.

Since $\mathcal{T}(\omega_m,\omega_n)$ can often be approximated by $\mathcal{T}(\sqrt{\omega_m\omega_n})$ and $\mathcal{N}(\omega_m,\omega_n)$ can often be approximated by $\mathcal{N}(\sqrt{\omega_m,\omega_n})$, it follows that

$$Q(\omega_m,\omega_n)\approx Q(\sqrt{\omega_m\omega_n}) \tag{66}$$

if $|\omega_m-\omega_n| << \omega_l$ and $|\omega_m-\omega_n|/2\pi \lesssim 2000$ hertz. From figures 11 and 12 it would appear that equation (66) is satisfactory if $|\omega_m-\omega_n|/2\pi \lesssim 2000$ hertz, except at low frequencies.



(a) Sound-pressure level at eardrum S_e=100 db.
(b) Sound-pressure level at eardrum S_e=80 db.
(c) Sound-pressure level at eardrum S_e=40 db.

FIGURE 12.—Internal cross-transmittance-level function $N(\omega_m, \omega_n)$.

SIMPLE FORMULAS FOR LOUDNESS LEVEL

For large numbers of tones or broadband sounds the lengthy summations make the loudness formulas too cumbersome for manual calculation. A simple approximation for the loudness level would be desirable. For broadband, suprathreshold sounds which are not dominated by a signal at one frequency, each term of the sums in equation (40), say, contributes only a small amount to the overall loudness. Morever, the total cross-transmittance $\mathcal{T}_r(\omega_\mu,\omega_\kappa)\,|\mathcal{N}(\omega_\mu,\omega_\kappa)|$ is less than, but generally approximately equal to, the total transmittance $Q(\sqrt{\omega_\mu\omega_\kappa})$ (cf. eq. (66)). Hence, $w_q(\omega_\mu,\omega_\kappa) \approx [w_q(\omega_\mu)w_q(\omega_\kappa)]^{1/2}$

for frequencies not too widely separated. In equation (40) the total value of the self-interaction terms is $M \sum_{\mu}^{M} w_q\ (\omega_\mu)$ when there are no empty channels. The total value of the cross-interaction terms is

$$(4/\pi) \sum_{\kappa}^{M} \sum_{\mu}^{M} (1-\delta_{\mu\kappa})\, w_q(\omega_\mu,\omega_\kappa) \approx (4/\pi) \sum_{\kappa}^{M} \sum_{\mu}^{M} (1-\delta_{\mu\kappa}) [w_q(\omega_\mu) w_q(\omega_\kappa)]^{1/2}$$

There are M^2 self-interaction terms and $M(M-1)$ cross-interaction terms. The cross-interaction terms satisfy

$$w_q(\omega_\mu) + (8/\pi)[w_q(\omega_\mu) w_q(\omega_\kappa)]^{1/2} + w_q(\omega_\kappa) \lesssim 2[w_q(\omega_\mu) + w_q(\omega_\kappa)]$$

which follows from Schwarz's inequality by assuming that $8/\pi \approx 2$. Thus, at suprathreshold loudnesses, equation (40) may be approximated by

$$\begin{aligned} \mathcal{L} &\lesssim \left\{ (\kappa^2/2) \left[\sum_{\mu}^{M} w_q(\omega_\mu) + 2(M-1) \sum_{\mu}^{M} w_q(\omega_\mu) \right] \right\}^{1/2} \\ &= \kappa \left[(M-1/2) \sum_{\mu}^{M} w_q(\omega_\mu) \right]^{1/2} \end{aligned} \tag{67}$$

Therefore, the loudness level is given by the simple formula

$$\begin{aligned} L &\approx 20 \log(10^2 \mathcal{L}) \\ &\lesssim 10 \log \sum_{\mu}^{M} w_q(\omega_\mu) + 10 \log(M-1/2) \end{aligned} \tag{68}$$

which follows using equation (53), and where $\kappa = 10^{-2}$. If there are k empty bands, then

$$\begin{aligned} \mathcal{L} &= \left\{ (\kappa^2/2) \left[(1+k) \sum_{\mu}^{M} w_q(\omega_\mu) + 2(M-1) \sum_{\mu}^{M} w_q(\omega_\mu) \right] \right\}^{1/2} \\ &= \kappa \left[\left(M - \frac{1-k}{2} \right) \sum_{\mu}^{M} w_q(\omega_\mu) \right]^{1/2} \end{aligned} \tag{69}$$

and

$$L = 10 \log \sum_{\mu}^{M} w_q(\omega_\mu) + 10 \log \left(M - \frac{1-k}{2} \right) \tag{70}$$

Similar formulas apply for sounds with continuous spectra, except that w is to be replaced by Ω. These formulas can easily be evaluated by manual calculations. Note that, for each critical band, $w_q = Q_r w$, and that $\sum w_q$ is simply the total transmitted relative power. The quantity $\sum w_q$ can be evaluated from spectral measurements in bandwidths equal to, or less than, critical bandwidths. As a last resort the quantity $10 \log \sum w_q$ can be approximated by the A-weighted, overall, sound-pressure level S_A. Then,

$$L \approx S_A + 10 \log \left(M - \frac{1-k}{2} \right) \tag{71}$$

This may not be a sufficiently good approximation. Specifically, since the A-weighting tends to underestimate the response (cf. figs. 9 and 10), the loudness levels indicated by this formula may be too low.

Other simplifications are possible. For equally intense tones, $w(\omega_n)$ is the same for each tone. On the other hand, for equally loud tones, $Q(\omega_n)w(\omega_n)$ is the same for each tone. Hence, for N equally loud tones spaced supercritically in adjacent critical bands ($k=0$), it follows from equation (38a) that

$$\mathcal{L}(\omega_1,\omega_2,\ldots,\omega_\nu,\ldots,\omega_N)/\mathcal{L}(\omega_\nu)\approx\{(1/2)[N^2+(4/\pi)N(N-1)]\}^{1/2}$$
$$=\{(2N/\pi)[(1+\pi/4)N-1]\}^{1/2} \quad (72)$$

if $N>1$ (ref. 48). If empty critical bands exist between the tones (<5 critical bands between tones), equation (40) yields the corresponding result

$$\mathcal{L}(\omega_1,\omega_2,\ldots,\omega_\nu,\ldots,\omega_N)/\mathcal{L}(\omega_\nu)$$
$$\approx\{(1/2)[(N+k)N+(4/\pi)N(N-1)]\}^{1/2}$$
$$=\{(2N/\pi)[(1+\pi/4)N-1+\pi k/4]\}^{1/2} \quad (73)$$

For equally intense, supercritically spaced tones the same formulas are valid if all the tones lie in the frequency range for which $Q_r\approx1$. For listening through earphones the maximum number of tones which may satisfy this condition is about 15, since a maximum of about 15 supercritically spaced tones will span the frequency range for which $Q_r\approx1$. For free-field listening the maximum number of tones is less.

TEST CONDITIONS

Von Békésy (ref. 49) has listed some requirements for good psychological observations. These conditions include precisely defined stimuli, repeatable experimental conditions, an understanding by the listener of the particular phenomenon to be judged, and inhibition of irrelevant sensations. The listener should be unaware of what the experimenter regards as a "correct" judgment, say, of the loudness (ref. 50). The experiment should be "blind" in the sense that neither the experimenter nor the listener should exercise immediate control over the stimulus. Each test should involve a single comparison between the test and reference sounds (ref. 51). Finally, because of the large errors inherent in psychoacoustic judgments the data must be averaged property (ref. 52). The importance of averaging sound power rather than sound-pressure level is well recognized in reading a meter (ref. 53, p. 181f.) but appears to have been generally overlooked in evaluating loudness judgments.

Some of the preceding requirements have been violated in all psychoacoustic studies. The magnitudes of the resulting biases are often

unknown. However, in comparing experiments with theory, the divergence among the results of similar experiments, when compared with the theory, may disclose significant biases.

LOUDNESS OF SPECIFIC SOUNDS WITH DISCRETE SPECTRA (COMPARISON OF MODEL WITH EXPERIMENT)

In this section the loudness of specific sounds having discrete spectra is calculated and compared with psychoacoustic data reported by various investigators. Sounds with discrete spectra consist of one or more tones.

In the experiments to be discussed a variety of loudness comparison methods were employed and many empirical, loudness-calculation procedures were tested. These will not be discussed herein. This information is best obtained by consulting the original references.

One Tone

The sound consists of a single discrete frequency. Equation (18a) reduces to

$$p(t)=A_1\cos(\omega_1 t+\phi_1)$$

Equation (33), or (35), becomes

$$\mathcal{L}=[\kappa^2 Q_r(\omega_1)w(\omega_1)]^{1/2}-[\kappa^2 Q_r(\omega_1)w_t(\omega_1)]^{1/2} \tag{74}$$

where

$$w(\omega_1)=A_1^2/2p_{ref}^2 \tag{75}$$

according to equation (21). The quantity κ can be evaluated by comparing this loudness equation with equation (29) in reference 10, which is

$$\mathcal{L}=Kp_{ref}(\mathcal{P}-1) \tag{76}$$

in the present notation, where K is a function of the number of dynamically active and saturated nerve fibers, hence of the sound pressure and frequency. Since the loudness of any sound is specified in sones relative to that of a 1-kilohertz tone at one sone, equation (76) may be rewritten for any tone relative to a 1-kilohertz tone at one sone in the form

$$\mathcal{L}(\omega_1)=\mathcal{K}[\mathcal{P}(\omega_1)]\,[\mathcal{P}(\omega_1)-\mathcal{P}_t(\omega_1)] \tag{77}$$

where

$$\begin{aligned}\mathcal{K}[\mathcal{P}(\omega_1)]&\equiv\{K[\mathcal{P}(\omega_1)]/K[\mathcal{P}_{L=1}(\omega/2\pi=1000)]\}\\&\quad\times[1/\mathcal{P}_{L=1}(\omega/2\pi=1000)]\\&=10^{-2}\{K[\mathcal{P}(\omega_1)]/K[\mathcal{P}_{L=1}(\omega/2\pi=1000)]\}\end{aligned} \tag{78}$$

and

$$\mathcal{P}_{L=1}(\omega/2\pi=1000)=10^2$$

is the relative, effective sound pressure of a 1-kilohertz tone when its loudness is one sone. By comparing equations (74) to (78), it is evident that $\mathcal{P}(\omega_1)=[w(\omega_1)]^{1/2}$, and

$$[Q_r(\omega_1)]^{1/2}=K[\mathcal{P}(\omega_1)]/K[\mathcal{P}_{L=1}(\omega/2\pi=1000)] \tag{79}$$

so that

$$\kappa=10^{-2} \tag{80}$$

Therefore,

$$\mathcal{L}=10^{-2}[Q_r(\omega_1)w(\omega_1)]^{1/2}-10^{-2}[Q_r(\omega_1)w_t(\omega_1)]^{1/2} \tag{81}$$

As shown in reference 10, with ω_1 in the neighborhood of 1 kilohertz and with sound-pressure levels in the range 34 dB$<S<$90 dB, $\mathcal{K}=10^{-2}$. Therefore, $Q_r(\omega_1)=1$; and, since $w>>w_t$,

$$\mathcal{L}=10^{-2}\mathcal{P}$$

follows from equation (81). Corresponding formulas for the entire acoustic amplitude range are derived in reference 10. For completeness, the loudness function given by equation (81) is compared in figure 13, as it was in reference 10, with psychoacoustic data reported by Warren. The agreement is excellent.

For suprathreshold loudnesses the special result $\mathcal{L}(\omega_1)=\mathcal{K}(\omega_1)\mathcal{P}(\omega_1)$ (cf. eq. (77)) resembles an input–output relation for a linear system. For more complex steady sounds it is, therefore, tempting to assume that the overall loudness is obtained by summing the loudness contributions associated with the individual frequency components comprising the sound. For example, Stevens, following others (ref. 4), proposed several versions of such a procedure (refs. 5, 6, 54, and 55) in which the individual loudness components were also unequally weighted in order to incorporate empirically the apparent "mutual inhibition" of neural discharges by different frequency components.

A more careful scrutiny of the preceding specialization of equation (77) leads to doubt concerning the validity of loudness summation. Firstly, the nervous system is a physical system and obeys physical laws. The psychoacoustic quantities, specifically the loudness sensation in the present instance, are subjective, nonphysical consequences of physical processes in the nervous system. Input–output relations generally indicate the amplitude or the power transferred by a physical system. However, equation (77) relates sound pressure to loudness in a mixed physical–psychoacoustic system. Thus, equation (77) does not correspond to the usual input–output relation because equation (77) does not relate physical quantities nor does it directly indicate

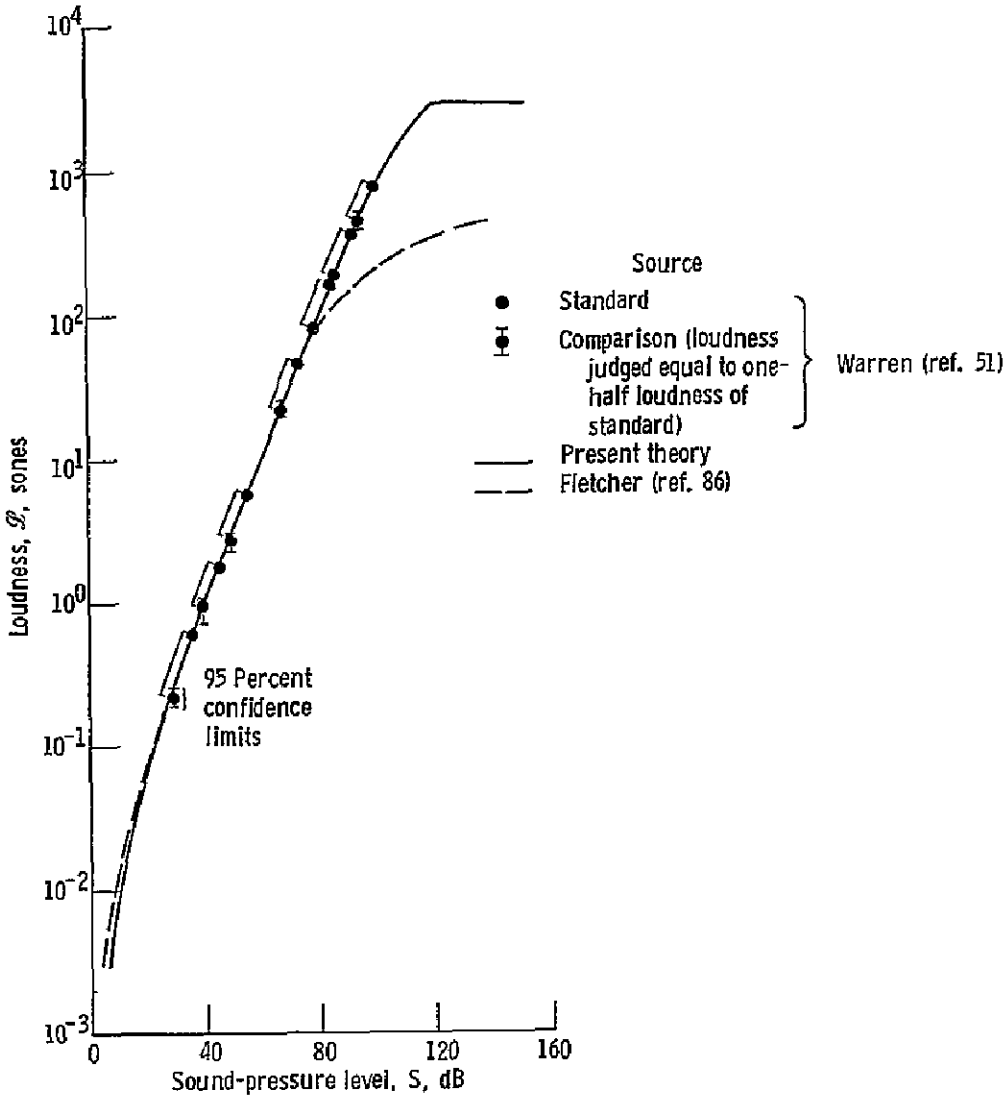


FIGURE 13.—Comparison of judged and calculated loudness for 1-kilohertz tone.

amplitude or power transference. Loudness is an alternative expression of the relative overall output power of the physical system.

Two Tones

The sound consists of two discrete frequencies. Equation 18(a) reduces to

$$p(t) = A_1 \cos(\omega_1 t + \phi_1) + A_2 \cos(\omega_2 t + \phi_2)$$

For two subcritically spaced tones, equation (33) becomes

$$\mathcal{L}=10^{-2}\{Q_r(\omega_1)w(\omega_1)+Q_r(\omega_2)w(\omega_2)$$
$$+2\mathcal{T}_r(\omega_1,\omega_2)|\mathcal{N}(\omega_1,\omega_2)|w(\omega_1,\omega_2)$$
$$\times\mathrm{sinc}[(\omega_2-\omega_1)T/2\pi]\cos[(\omega_2-\omega_1)(t-T/2)+\phi_2-\phi_1]\}^{1/2}$$
$$-10^{-2}\mathcal{R}_{qt}^{1/2} \tag{82}$$

where $\theta_{12}\approx 0$. If the tone frequencies are not too low or too high (200 Hz$\leq(\omega_1,\omega_2)\leq$2000 Hz) or if the two tones are closely spaced (say 100 Hz, or less, apart at high frequencies),

$$\mathcal{L}=10^{-2}\Big[Q_r(\sqrt{\omega_1\omega_2})\{w(\omega_1)+w(\omega_2)+2w(\omega_1,\omega_2)$$
$$\times\mathrm{sinc}[(\omega_2-\omega_1)T/2\pi]\cos[(\omega_2-\omega_1)(t-T/2)+\phi_2-\phi_1]\}\Big]^{1/2}$$
$$-10^{-2}\mathcal{R}_{qt}^{1/2} \tag{83}$$

where, in the former case, $Q_r(\sqrt{\omega_1\omega_2})=1$.

If the tone separation is supercritical or hypercritical, equation (40) becomes

$$\mathcal{L}=10^{-2}\{(1+k/2)[Q_r(\omega_1)w(\omega_1)+Q_r(\omega_2)w(\omega_2)]$$
$$+(4/\pi)\beta_{12}^2\mathcal{T}_r(\omega_1,\omega_2)|\mathcal{N}(\omega_1,\omega_2)|w(\omega_1,\omega_2)\}^{1/2}$$
$$-10^{-2}\Big[\sum_{k=1}^{2}(\mathcal{R}_{qt})_k\Big]^{1/2} \tag{84}$$

where the subscripts refer to tone numbers (rather than critical band numbers); k (not the subscript) denotes the number of empty, intermediate, critical bands; and $\beta_{21}=\beta_{12}$ is unity unless the tone separation is only slightly supercritical.

Von Békésy's experiment (1929).—The predicted relative loudness of two closely spaced tones with equal suprathreshold amplitudes is shown in figure 14 as a function of the spacing for comparison with loudness judgments reported by von Békésy for listening monaurally through an earphone (ref. 35, p. 235). The theoretical curves (to be explained) were derived from equation (83). In the experiment one tone was kept at a frequency $\omega_1/2\pi=500$ hertz, whereas the frequency $\omega_2/2\pi$ of the second tone was adjustable. The loudness of the combined tones was judged relative to the loudness of the 500-hertz tone alone. Von Békésy reported that the observations were difficult. Because the tones were closely spaced, beats were significant, thus the observational difficulty. Periodicity pitch was not important because, at the extreme separatiòn, the tone spacing was only slightly supercritical, in which case $k=0$, and $\beta_{12}\approx 0$, in equation (84). Nearly all the data were for a subcritical separation in the frequency range for uniform transmittance of the auditory system. Specifically, $Q(\omega_1)=Q(\sqrt{\omega_1\omega_2})=Q(\omega_2)=1$. As a consequence of these conditions the loudness of the combined tones relative to the loudness of the 500 hertz tone is given by

$$\mathcal{L}(\omega_1,\omega_2)/\mathcal{L}(\omega_1)=\sqrt{2}\{1+\mathrm{sinc}[(\omega_2-\omega_1)T/2\pi]$$
$$\times\cos[(\omega_2-\omega_1)(t-T/2)+\phi_2-\phi_1]\}^{1/2}$$

which follows from equation (83). For ω_1 and ω_2 given, this function oscillates temporally between the limits

$$\sqrt{2}\{1+\mathrm{sinc}[(\omega_2-\omega_1)T/2\pi]\}^{1/2}$$

and

$$\sqrt{2}\{1-\mathrm{sinc}[(\omega_2-\omega_1)T/2\pi]\}^{1/2}$$

These relative loudness limits for $T=0.1$ second are shown in figure 14. For larger values of T the central peak would be narrower and the fluctuations of the loudness limits would occur for smaller frequency intervals. Figure 14 illustrates that, for narrow spacing of the tones ($\leq$10 Hz), the listener tended to characterize the loudness of the tone combination by its maximum loudness. For greater tone spacing the judged relative loudness either fell within the limits or else, on the average, fell about 1.25 decibels below the limits set by the theory. Von Békésy reported that test results similar to those shown in figure 14 were obtained for $\omega_1/2\pi=200$ and 2000 hertz and that the

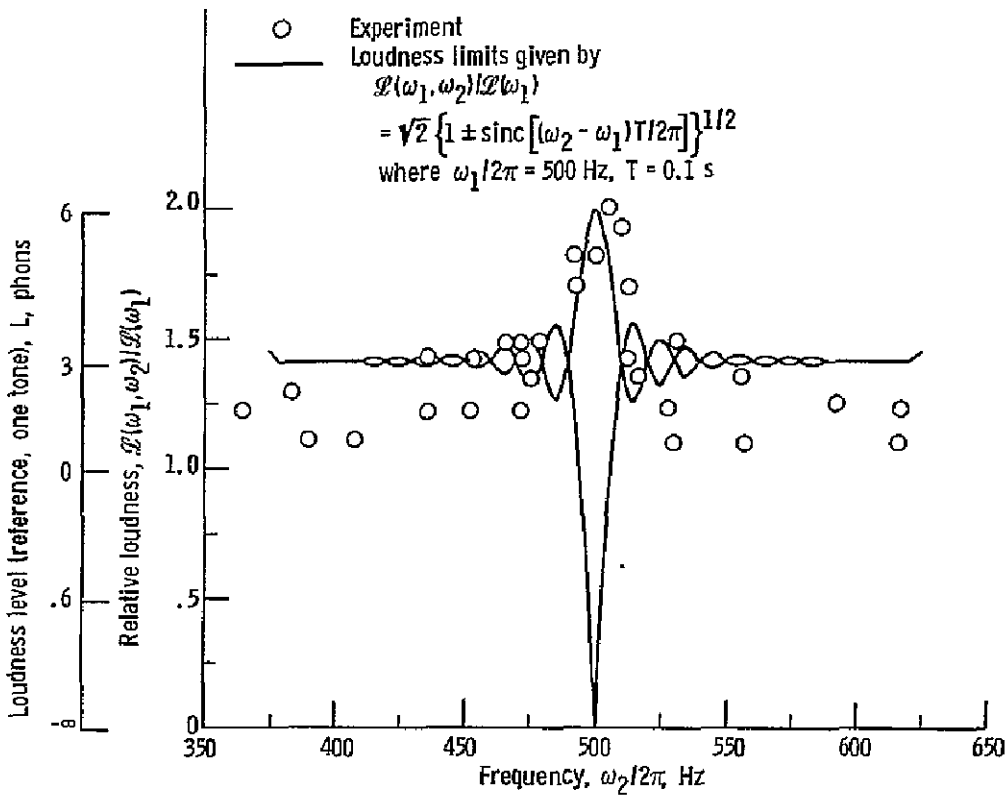


FIGURE 14.—Relative loudness of two equally intense tones as function of tone spacing. (Data from von Békésy (ref. 35).)

loudness decrease depended on the frequency difference, as in the present thory, and not on the relative frequency difference. Contrary to his expectation that the relative loudness should tend to unity as $|\omega_2-\omega_1|$ is increased, the present theory and other tests (ref. 36) suggest that the loudness should remain constant, as in figure 14, and then begin to increase when the tone separation exceeds the critical bandwidth (115 Hz for $\sqrt{\omega_1\omega_2}/2\pi=500$ Hz). Von Békésy's results as $\omega_2\rightarrow\omega_1$ illustrate that a 6-decibel (not the usually accepted 10-dB) reduction in the sound-pressure level of a tone corresponds to halving its loudness (ref. 56).

Fletcher and Munson's experiment (1933).—From data obtained by 11 listeners with normal hearing, who listened through earphones, Fletcher and Munson compared the binaural loudness of pairs of equally loud tones with that of one of the tones alone (ref. 13). The tonal pairs were as follows:

Pair	$\omega_1/2\pi$, Hz	$\omega_2/2\pi$, Hz	k
1	1000	1100	0
2	1000	2000	3
3	1000	125	6

The frequency separation of the first tone pair is somewhat less than the critical bandwidth (see fig. 3 or 7), so that equation (83) should be applicable in computing the loudness. The auditory system is equally sensible to both tones of this, and the second, pair. Therefore, both tones must be incorporated in the threshold term. For the first two pairs $Q(\omega_1)=Q(\omega_2)=1$ (cf. fig. 10), and $Q(\omega_1,\omega_2)=Q(\sqrt{\omega_1,\omega_2})=1$. Therefore, for each tone the given loudness level L equals its sound-pressure level S.

For suprathreshold loudnesses it is convenient to set $w(\omega_1)=w(\omega_2)=w(\omega_1,\omega_2)=1$ when the tones are equally intense and then to compute the ratio of the loudness of the tone combination to that of one of the tones. This procedure eliminates the necessity for calculating $w(\omega_n)$ if the loudness, rather than the loudness ratio, were evaluated. Considering equation (83), the loudness ratio is given by

$$\mathcal{L}(\omega_1,\omega_2)/\mathcal{L}(\omega_1)=\sqrt{2}$$

for two subcritically spaced tones in the absence of beats. In phons this loudness ratio corresponds to a loudness-level difference ΔL given by

$$\Delta L=20\log[\mathcal{L}(\omega_1,\omega_2)/\mathcal{L}(\omega_1)]=3.0 \text{ phons}$$

which is to be added to the loudness level of the lone tone, that is,

$$L(\omega_1,\omega_2)=L(\omega_1)+\Delta L$$

in order to obtain the loudness level $L(\omega_1,\omega_2)$ of the pair.

The judged and calculated loudness levels are compared in figure 15. If the agreement between the judged and calculated values were perfect, the data would fall along the diagonal straight line. Except near threshold ($S_t=3$ dB; see fig. 8), the levels indicated by equation (83) average about 1 phon less than the judged values. It is noteworthy that, for 100 hertz separation of the tones, von Békésy's data (ref. 35) yielded a loudness level about 1.5 phons less than that predicted using equation (83) whereas Fletcher and Munson's data yielded an average judged loudness level 1 phon greater than that predicted using equation (83). The difference between these results seems to imply that the principal difference between theory and experiment may be due to systematic experimental error.

The second pair of tones is harmonically related. The tone separation is supercritical, and the tones are separated by three empty critical bands. Thus, equation (84) applies with $k=3$, and the Q's$=1$. For the results shown in figure 15 the absolute average difference between the calculated and judged loudness levels is about 1.5 phons.

The final tone pair is also harmonically related. The tone separation is supercritical, and the tones are separated by six empty critical bands. For these tones, the transmittances are nonuniform; $Q(\omega_1)=1$, $Q(\omega_2)=0.45$ (cf. fig. 10) and $Q(\omega_1,\omega_2)\approx 1$. Since the tones were equally loud, it follows that $w(\omega_2)=w(\omega_1)/0.45$. The loudness level pre-

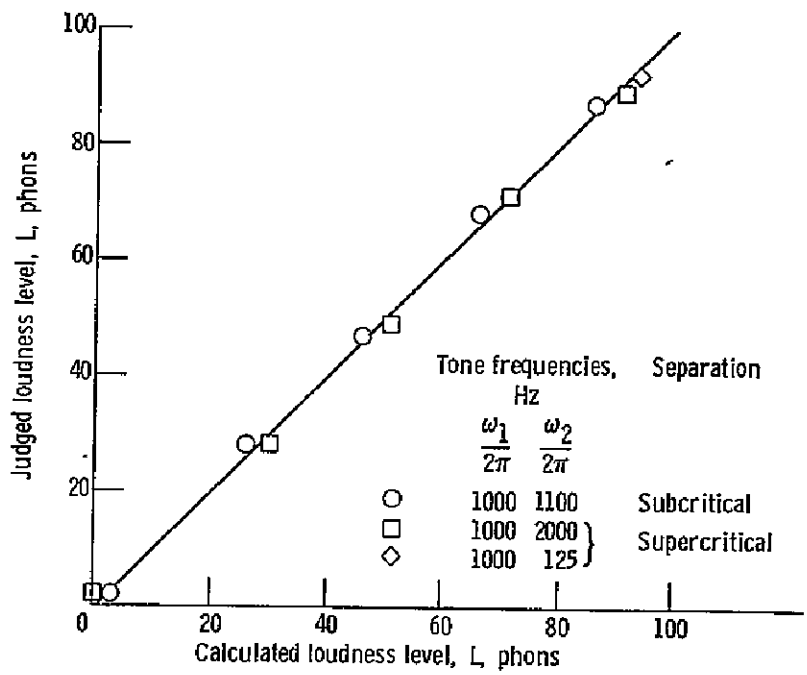


FIGURE 15.—Comparison of judged and calculated loudness levels for two tones. (Data from Fletcher and Munson (ref. 13).)

dicted using equation (84) is compared in figure 15 with the judged value for one sound-pressure level. The difference between the calculated and judged loudness levels is 2 phons.

Small and Thurlow's experiment (1954).—Small and Thurlow, using themselves as listeners, equated the loudness of pairs of tones imposed monaurally through earphones to the loudness of a 2000-hertz comparison tone imposed on the other ear (ref. 57). The following pairs of tones were used:

Pair	$\omega_1/2\pi$, Hz	$\omega_2/2\pi$, Hz	k
1	400	1000	3
2	1900	3100	2
3	1900	4400	4
4	1900	5100	5

The primary tones were imposed at unequal intensities. Moreover, the transmittance function Q is not uniform for all frequencies involved. Finally, the data were displayed in terms of sensation levels $\mathcal{S}$ rather than sound-pressure levels.

The sensation level $\mathcal{S}$ is the level of a sound above its loudness threshold. In general,

$$\mathcal{S}=S-C \tag{85}$$

where

$$\mathcal{S}\equiv 20\log(\mathcal{P}/\mathcal{P}_t)$$

and

$$C\equiv 20\log\mathcal{P}_t$$

is the sound-pressure level of the sound at its loudness threshold.

The sensation level of each tone alone is shown in figure 16. The sensation level of the lower frequency tone was held fixed in each set of tests while the sensation level of the higher frequency tone was adjusted. For all of the fixed-level (lower-frequency) tones, it happens that $\mathcal{S}=S$. The sound-pressure levels of the adjusted (higher-frequency) tones can be determined from the sensation-level data. The sensation level of an equally loud, 1200-hertz tone is indicated by the ordinate in figure 16. Since, for a tone,

$$L=N+S \tag{86}$$

(where N is the internal transmittance level, and the external transmittance level $T=0$ when using earphones), and considering that the ordinate in figure 16 is $\mathcal{S}(\omega/2\pi=1200)$ and that $\mathcal{S}(\omega/2\pi=1200)$

$=\mathcal{S}(\omega/2\pi=1000)\approx S(\omega/2\pi=1000)=L(\omega/2\pi=1000)=L(\omega/2\pi=1200)$, it follows from this approximation and equations (85) and (86) that

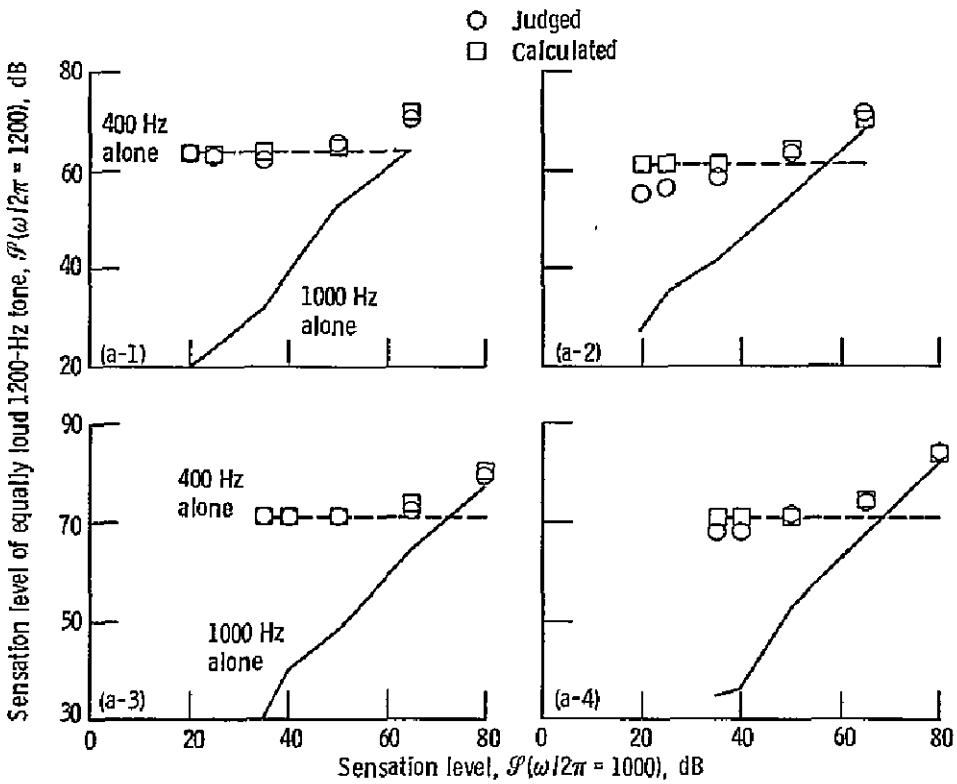
$$C(\omega_n)\approx\mathcal{S}(\omega/2\pi=1200)-N(\omega_n)-\mathcal{S}(\omega_n)$$

Therefore, reconsidering equation (85), there results

$$S(\omega_n)=\mathcal{S}(\omega/2\pi=1200)-N(\omega_n) \tag{87}$$

where the appropriate values of $N(\omega_n)$ (cf. fig. 10) and, hence, of $Q(\omega_n)=\mathcal{N}(\omega_n)$ are as follows:

$\omega_n/2\pi$, Hz	$N(\omega_n)$	$Q(\omega_n)$
1000	0	1.00
3100	−1.5	.71
4400	−3.0	.50
5100	−3.5	.45



(a-1) Reference 57, figure 7(a); listener A.M.S.

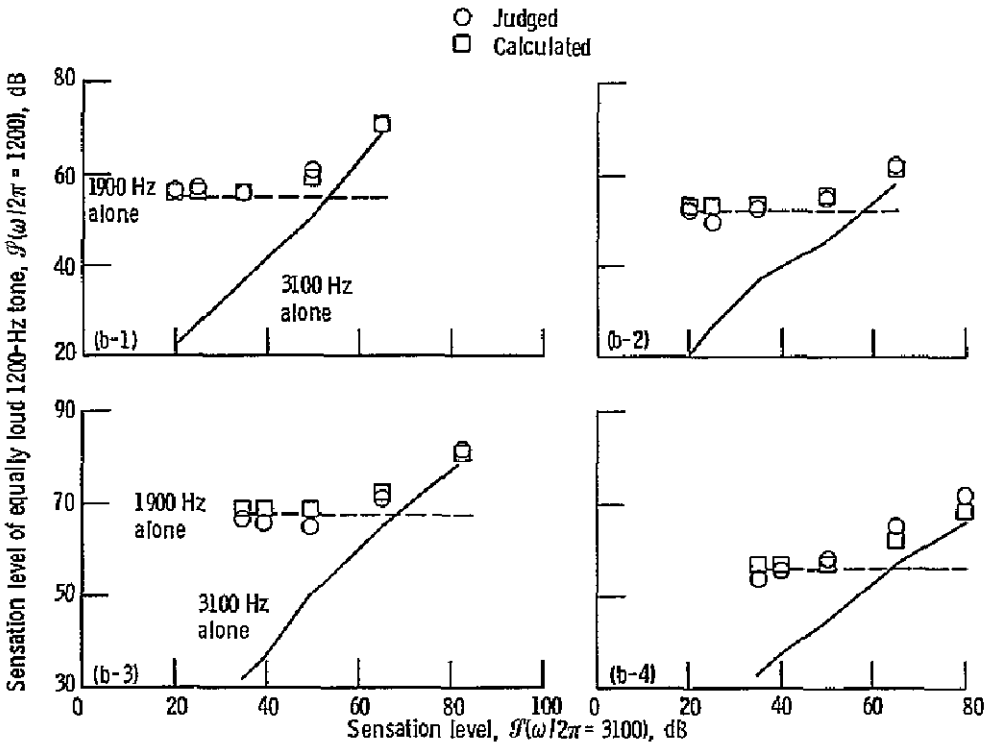
(a-2) Reference 57, figure 7(b); listener W.R.T.

(a-3) Reference 57, figure 8(a); listener A.M.S.

(a-4) Reference 57, figure 8(b); listener W.R.T.

(a) $\omega_1/2\pi$, 400 hertz; $\omega_2/2\pi$, 1000 hertz.

FIGURE 16.—Comparison of judged and calculated sensation levels of unequally loud pairs of tones. (Data from Small and Thurlow (ref. 57).)



(b–1) Reference 57, figure 9(a); listener A.M.S.
(b–2) Reference 57, figure 9(b); listener W.R.T.
(b–3) Reference 57, figure 10(a); listener A.M.S.
(b–4) Reference 57, figure 10(b); listener W.R.T.
(b) $\omega_1/2\pi$, 1900 hertz; $\omega_2/2\pi$, 3100 hertz.

FIGURE 16.—Continued.

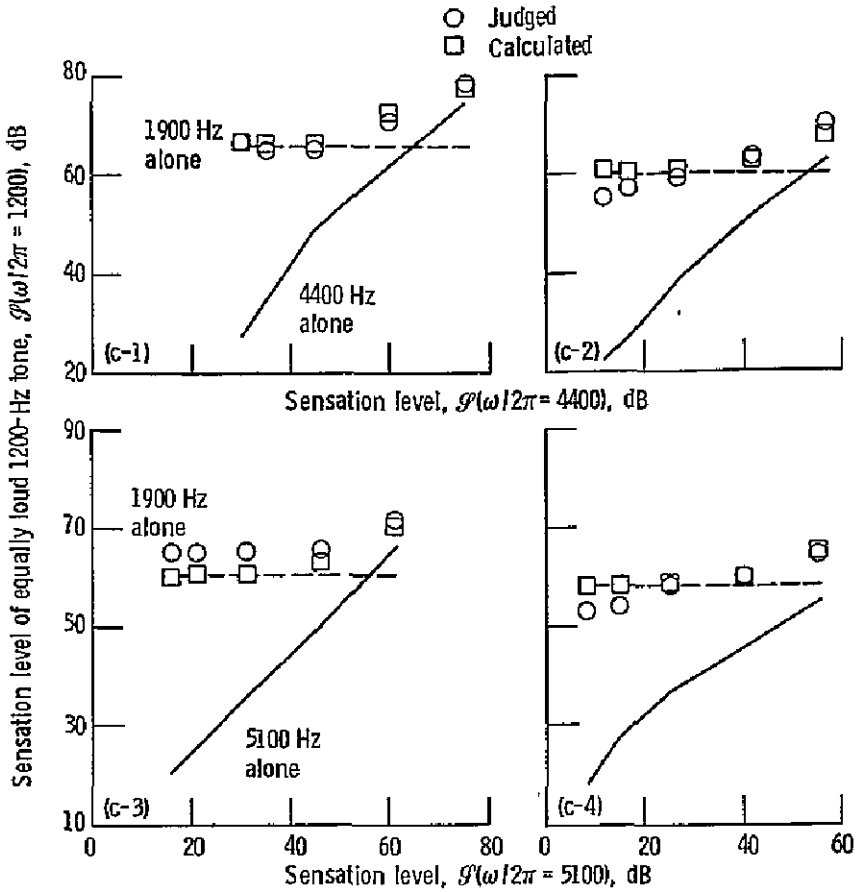
Thus, equation (87) can be used to compute the sound-pressure levels of the adjusted-level tones from the data in figure 16. The difference between the values of $S(\omega_n)$ found from equation (87) and those found from equation (85) using the abscissa values for $\mathcal{S}(\omega_n)$ in figure 16 and published data on $C(\omega_n)$ (ref. 19 or fig. 8) is an indication of the inconsistency of the data (differences as great as 10 dB). Possibly this inconsistency is not unexpected since there is no guarantee that the listeners' hearing was equal in both ears.

If $w(\omega_2)\rightarrow 0$, then, for suprathreshold loudnesses, equation (84) reduces to

$$\mathcal{L}=10^{-2}\left[(1+\frac{1}{2}k)Q(\omega_1)w(\omega_1)\right]^{1/2}$$

in which the quantity in brackets is excessive by an amount $(1/2)kQ(\omega_1)w(\omega_1)$, unless k is set equal to zero. This has been done in figure 16, except where the sound-pressure levels of the two tones are within 1 decibel of each other.

Among all data, except in the instance just cited, the average difference between calculated and judged levels appears to be at least as



(c-1) Reference 57, figure 11(a); listener A.M.S.
(c-2) Reference 57, figure 11(b); listener W.R.T.
(c-3) Reference 57, figure 12(a); listener A.M.S.
(c-4) Reference 57, figure 12(b); listener W.R.T.
(c) $\omega_1/2\pi$, 1900 hertz; $\omega_2/2\pi$, 4400 hertz and $\omega_1/2\pi$, 1900 hertz; $\omega_2/2\pi$, 5100 hertz.

FIGURE 16.—Concluded.

small as the average difference of the observations by the two observers.

D. H. Howes' experiment (1950).—By listening monaurally through an earphone, four trained listeners equated the loudness of pairs of equally intense tones, and two trained listeners equated the loudness of pairs of equally loud tones to that of a 1-kilohertz reference tone (ref. 58). The following equally intense component pairs were presented at sound-pressure levels of 40, 50, 60, 70, 80, and 90 decibels for each component:

Pair	Separation, mels	$\omega_1/2\pi$, Hz	$\omega_2/2\pi$, Hz	k
1	250	1900	(1420, 2450)	0
2	500	1420	2450	2

(The mel is a subjective unit of frequency or pitch. A 1000-hertz tone at a sound-pressure level of 40 decibels is defined as having a pitch of 1000 mels. A 2000-mel tone has twice the pitch.) For pair 1, two values of $\omega_2/2\pi$ are given in the table because these frequencies were alternated on successive presentations of the pair. Loudness judgment data were presented separately for each sound-pressure level and for each tone pair in order to demonstrate the effect of these quantities on loudness. However, since the average of loudness judgments by each of the two listeners differed by factors up to 3, the indicated nonlinear dependence of loudness on sound-pressure level will not be stressed. Thus, the data listed in table II, which contains results for various numbers of tones, will be averaged over all sound-pressure levels.

For 250-mel separation the tones were in adjacent critical bands, whereas for 500-mel separation the tones were separated by two empty critical bands. Since the tone separation was supercritical and the tone frequencies were in the flat response range of the internal auditory system, the predicted loudness of the tone pairs should be given by equation (84) with $k=0$ (250-mel separation), $k=2$ (500-mel separation), and the Q's$=1$. For the equally intense tones at suprathreshold loudness it is simpler to determine the ratio of the loudness of the two tones relative to that of one of the tones and then express the result in terms of the loudness-level difference ΔL (cf. eq. (54)). The predicted and judged loudness-level differences ΔL are compared in table II, which includes both the average and median of the judged level differences. The judged and calculated loudness levels L are compared in figure 17.

For the judgment data the average and median differences ΔL are 6.5 phons for tone pair 1 (250-mel separation) and about 8.2 phons for tone pair 2 (500-mel tone separation), whereas the calculated values are 5.2 phons in the former case and 7.2 phons in the latter case. Thus, the average difference between theory and experiment is about 1 phon.

D. H. Howes also presented data on judged loudness levels of equally intense tones centered at 1000 hertz, and with each tone at a sound-pressure level of 60, 70, 80, or 90 decibels and spaced 250 mels apart. The resultant difference ΔL between the loudness level of the tone complexes and that of the central tone is shown in table III as a function of the number of tones and the sound-pressure level of each component. The loudness levels L are compared in figure 17. In table III the average and median level differences ΔL over all sound-pressure levels are also compared with the theoretical prediction. For two tones, with $k=1$, the disagreement between calculation and judgment is about 1 phon for listener J. M. and is negligible for listener D. H.

TABLE II.—LOUDNESS OF EQUAL-INTENSITY TONE COMPLEXES CENTERED AT 1900 HERTZ

[Data from D. H. Howes (ref. 58).]

$S(\omega_1)=S(\omega_2)=\ldots$, dB	Number of tones											
	2		3		4		5		6		8	11
	Tone separation, mels											
	250	500	250	500	250	500	250	500	250	500	250	250
	$\Delta L \equiv L(\omega_1, \omega_2, \ldots) - L(\omega_m)$, phons [a]											
40	7	8	14	13	18	19	20	21	23	22	26	29
50	8	10	15	15	19	20	23	25	25	23	28	31
60	8	8	13	14	16	19	20	22	23	23	26	28
70	6	8	9	14	14	19	16	20	18	21	23	26
80	5	7	9	12	12	16	14	18	17	19	20	23
90	5	9	8	11	12	15	14	16	16	17	19	21
Average ΔL, phons	6.5	8.3	11.3	13.2	15.2	18	17.8	20.3	20.3	20.8	23.7	26.3
Median ΔL, phons	6.5	8	11	13.5	15	19	18	20.5	20.5	21.5	24.5	27
Calculated ΔL, phons (approx.)	5.2	7.2	9.9	12.0	12.9	15	15.2	16.6	16.6	18.8	19.4	22.4

[a] For equally intense tones.

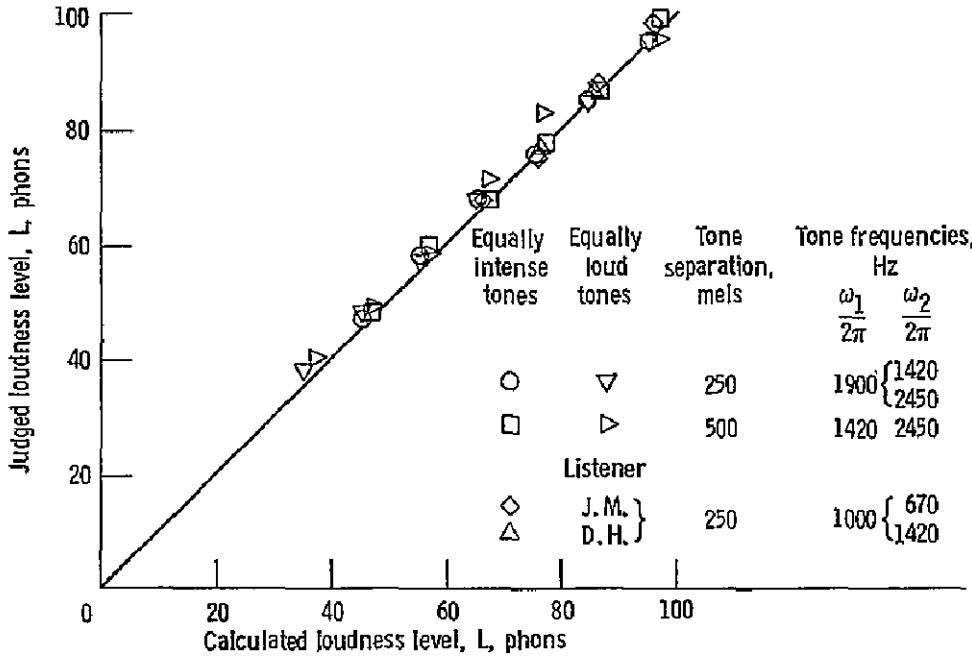


FIGURE 17.—Comparison of judged and calculated loudness levels for two tones. (Data from D. H. Howes (ref. 58).)

Finally, similar loudness judgment data are presented in table IV for complexes of equally loud tones (rather than equally intense tones) centered at 1900 hertz and spaced 250, and then 500, mels apart at the same sound-pressure levels as before. The test data in table IV and figure 17 are presented in reference 58 as loudnesses determined from an experimentally judged loudness function of questionable validity (cf. ref. 10). Thus, Howes' loudness data have been reconverted to loudness levels using his table II in order to compute the loudness-level differences ΔL and their averages and medians, which are also listed in table IV herein. For two tones, with $k=0$ (250-mel spacing) or $k=2$ (500-mel spacing), the average difference between these results and theory is about 1.5 phons for tone pair 1 (250-mel tone separation) and about 2 phons for tone pair 2 (500-mel separation).

Scharf's experiment (1959).—By binaural listening through earphones, eight to twelve listeners equated the loudness of a pair of equally loud tones to that of a reference tone held at a fixed loudness (ref. 43). By this procedure the loudness of each pair as a function of tone spacing was determined with constant loudness of each tone as a parameter. All judgments were made in the vicinity of the loudness threshold. The 1500-hertz comparison tone coincided with the geometric-mean frequency of the two tones.

The judged loudness level of each tone and the equal-loudness curves of Robinson and Dadson (ref. 20 or fig. 8) modified (dashed

TABLE III.—LOUDNESS OF EQUAL-INTENSITY TONE COMPLEXES CENTERED AT 1000 HERTZ WITH 250-MEL SPACING BETWEEN COMPONENTS

[Data from D. H. Howes (ref. 58).]

$S(\omega_1)=S(\omega_2)=\ldots$, dB	Listener											
	J. M.						D. H.					
	Number of tones											
	2	3	4	5	6	7	2	3	4	5	6	7
	$\Delta L \equiv L(\omega_1, \omega_2, \ldots) - L(\omega_m)$, phons [a]											
60	8	15	18	24	25	29	8	13	17	21	20	24
70	5	12	16	19	21	24	7	10	14	18	20	22
80	8	10	16	18	18	20	6	9	15	13	16	17
90	8	11	12	14	17	18	5	8	8	11	13	15
Average ΔL, phons	7.2	12	15.5	18.8	20.2	22.8	6.5	10	13.5	15.8	17.2	19.5
Median, ΔL, phons	8	11.5	16	18.5	19.5	22	6.5	9.5	14.5	15.5	18	19.5
Calculated ΔL, phons (approx.)	6.3	10.5	13.4	15.5	17.2	18.6	6.3	10.5	13.4	15.5	17.2	18.6

[a] For equally intense tones.

TABLE IV.—LOUDNESS OF EQUALLY LOUD TONE COMPLEXES CENTERED AT 1900 HERTZ

[Data from D. H. Howes (ref. 58).]

$S(\omega_1)=S(\omega_2)=\ldots$, dB	Number of tones											
	2		3		4		5		6		8	11
	Tone separation, mels											
	250	500	250	500	250	500	250	500	250	500	250	250
	$\Delta L \equiv L(\omega_1, \omega_2, \ldots) - L(\omega_m)$, phons [a]											
30	8.5	10.5	12.0	11.0	14.5	16.5	18.0	20.2	20.0	----	23.0	----
40	8.5	9.0	12.5	14.5	15.5	18.5	19.0	19.5	22.5	----	23.0	----
50	7.0	8.5	13.0	14.0	17.0	20.0	20.0	22.0	23.0	28.5	30.0	33.0
60	8.0	12.0	17.0	15.0	18.0	20.0	21.0	25.0	24.5	27.5	28.5	34.0
70	5.5	13.0	13.5	15.5	16.0	20.5	20.0	27.5	23.0	29.5	28.0	35.0
80	4.5	7.5	9.5	11.5	12.0	14.5	13.5	20.5	16.0	21.0	20.5	27.0
90	4.5	5.5	8.5	8.5	11.0	11.0	11.5	8.5	13.0	18.0	17.5	24.0
Average ΔL, phons	6.6	9.4	12.3	12.9	14.9	17.3	17.6	20.5	20.3	24.9	24.3	30.6
Median ΔL, phons	7.0	9.0	12.5	14.0	15.5	18.5	19.0	20.5	22.5	27.5	23.0	33.0
Calculated ΔL, phons (approx.)	5.2	7.2	9.9	12.0	12.9	15.0	15.2	16.6	16.6	18.8	19.4	22.4

[a] For equally loud tones.

portion in fig. 8) for earphone listening were used to determine the sound-pressure level of each tone at the eardrum. These values permitted the overall loudness level to be computed. The predicted overall loudness level of the tone pairs as a function of tone separation with sensation level (not shown) of the comparison tone as parameter are shown in figure 18. For each parameter value the results would fall on a horizontal line if the measurements and calculation procedure were ideal. In the figure the horizontal lines are nominal loudness level values of the pair. The dashed line at the bottom of the figure indicates the interquartile range of the original data, that is, the phon range containing the second and third quartiles of the data. It is evident that nearly all of the calculated loudness levels agree with the nominal loudness levels to within a difference less than the interquartile range.

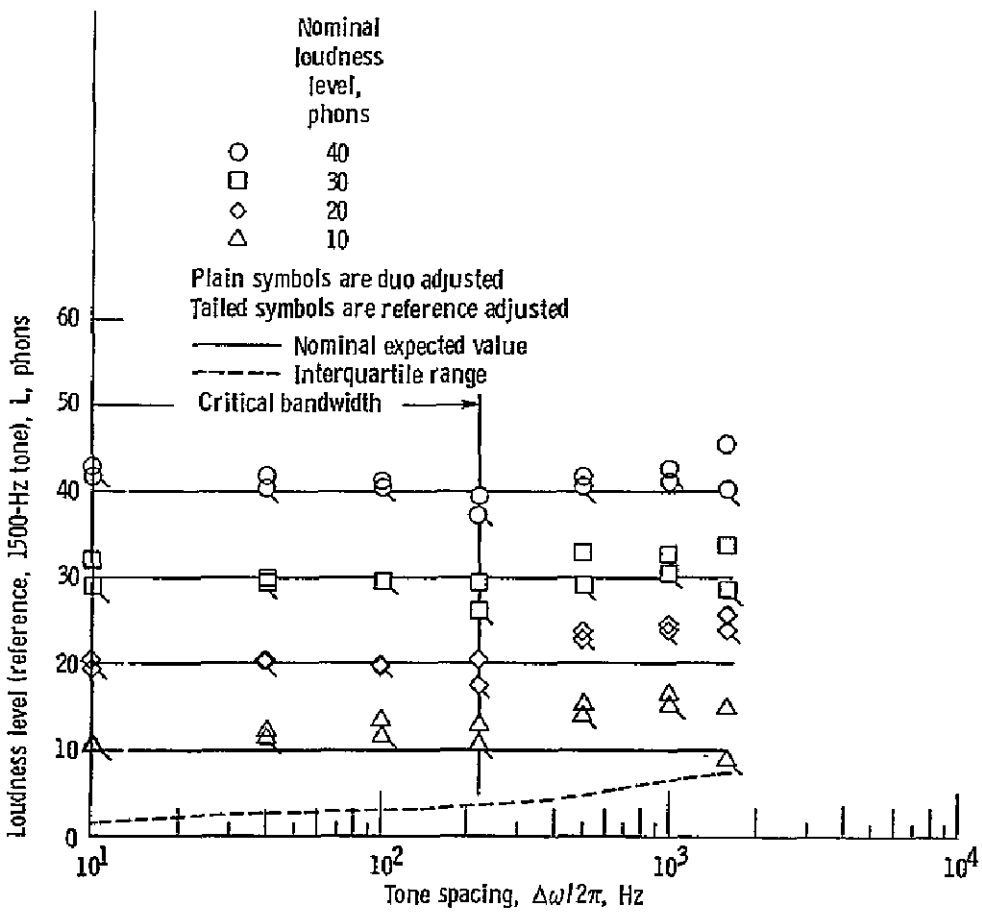


FIGURE 18.—Calculated loudness level of equally loud tone duo near threshold as function of tone spacing. Arithmetic-mean frequency, $\omega_{am}/2\pi$, 1500 hertz. (Data from Scharf (ref. 43).)

Three Tones

The sound consists of three discrete frequencies, so that

$$p(t)=A_1\cos(\omega_1 t+\phi_1)+A_2\cos(\omega_2 t+\phi_2)+A_3\cos(\omega_3 t+\phi_3)$$

For three tones spaced subcritically, equation (33) becomes

$$\begin{aligned}\mathcal{L}=10^{-2}\{&Q_r(\omega_1)w(\omega_1)+Q_r(\omega_2)w(\omega_2)+Q_r(\omega_3)w(\omega_3)\\&+2\mathcal{J}_r(\omega_1,\omega_2)|\mathcal{N}(\omega_1,\omega_2)|w(\omega_1,\omega_2)\\&\times\operatorname{sinc}[(\omega_2-\omega_1)T/2\pi]\cos[(\omega_2-\omega_1)(t-T/2)+\phi_2-\phi_1]\\&+2\mathcal{J}_r(\omega_1,\omega_3)|\mathcal{N}(\omega_1,\omega_3)|w(\omega_1,\omega_3)\\&\times\operatorname{sinc}[(\omega_3-\omega_1)T/2\pi]\cos[(\omega_3-\omega_1)(t-T/2)+\phi_3-\phi_1]\\&+2\mathcal{J}_r(\omega_2,\omega_3)|\mathcal{N}(\omega_2,\omega_3)|w(\omega_2,\omega_3)\\&\times\operatorname{sinc}[(\omega_3-\omega_2)T/2\pi]\cos[(\omega_3-\omega_2)(t-T/2)+\phi_3-\phi_2]\}^{1/2}\\&-10^{-2}\mathcal{R}_{qt}^{1/2}\end{aligned}\tag{88}$$

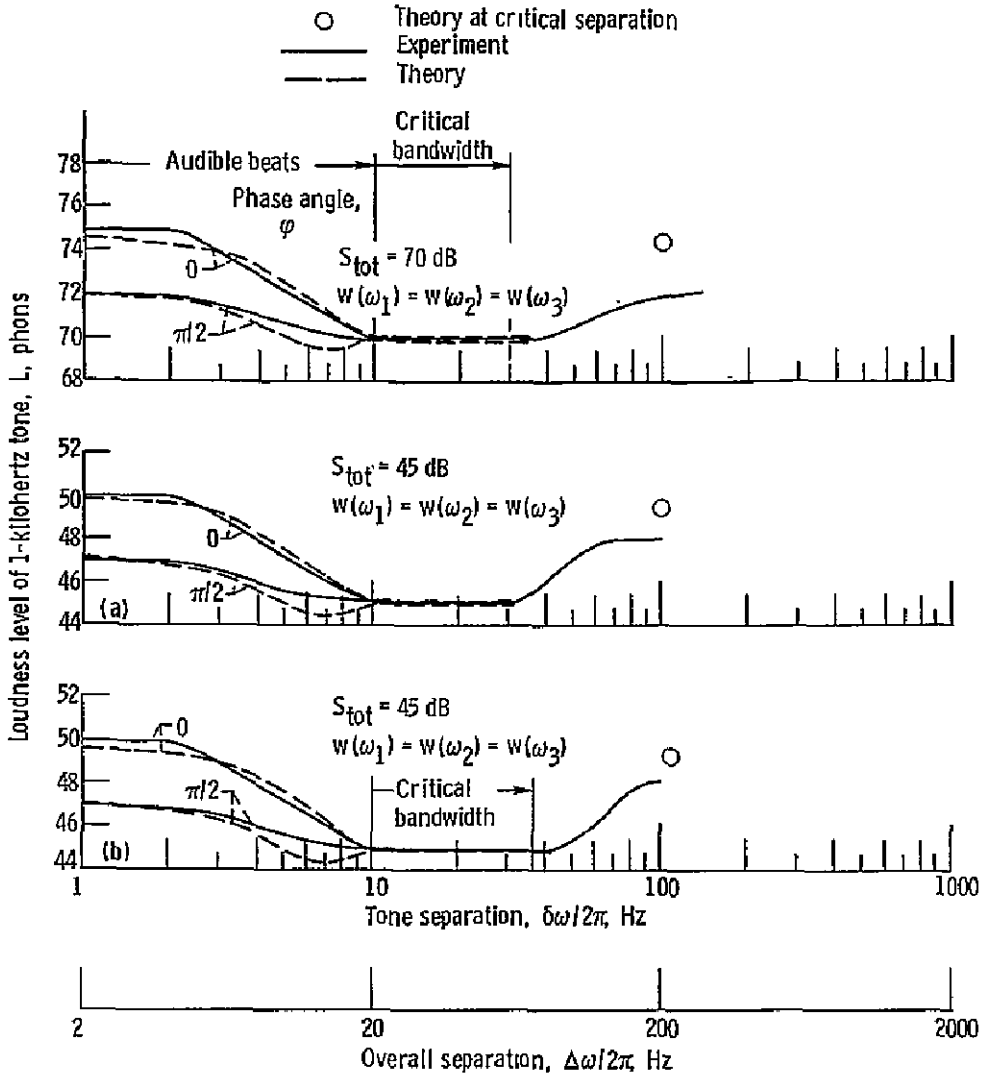
If the frequencies of three tones are not too low or too high, or if they are closely spaced, then all the Q_r's in this equation may be replaced by $Q_r(\omega_2)$, say, in analogy with the case of two tones.

For three supercritically spaced tones, equation (40) becomes

$$\begin{aligned}\mathcal{L}=10^{-2}\Big\{&\frac{1}{2}(3+k)[Q_r(\omega_1)w(\omega_1)+Q_r(\omega_2)w(\omega_2)+Q_r(\omega_3)w(\omega_3)]\\&+(4/\pi)[\beta_{12}^2\mathcal{J}_r(\omega_1,\omega_2)|\mathcal{N}(\omega_1,\omega_2)|w(\omega_1,\omega_2)\\&+\mathcal{J}_r(\omega_1,\omega_3)|\mathcal{N}(\omega_1,\omega_3)|w(\omega_1,\omega_3)\\&+\beta_{23}^2\mathcal{J}_r(\omega_2,\omega_3)|\mathcal{N}(\omega_2\ \omega_3)|w(\omega_2,\omega_3)]\Big\}^{1/2}\\&-10^{-2}\Big[\sum_{k=1}^{3}(\mathcal{R}_{qt})_k\Big]^{1/2}\end{aligned}\tag{89}$$

where the subscripts now denote the tone number, and k denotes the total number of empty critical bands between the three tones.

Bauch's experiment (1956).—Bauch and Zwicker equated the suprathreshold loudness of a triad of equally spaced tones, heard monaurally through an earphone, with that of the central tone alone (ref. 37). The results are displayed in figure 19 as loudness level of the triad when held at a fixed sound-pressure level as a function of the tone separation, which was varied from narrow (1 Hz) to supercritical (up to 2 kHz). However, the maximum overall separation was not sufficient to achieve peak loudness. Loudness levels were judged for center frequencies $\omega_2/2\pi=400$, 500, 1000, and 4000 hertz, a variety of fixed overall sound-pressure levels ranging from 30 to 70 decibels, and for three ratios, 1/2, 1, 2, of the double amplitude of the outer frequencies to the amplitude of the center frequency. In addition, results were obtained for two phases of the center frequency which differed by $\pi/2$.



(a) Center frequency, 400 hertz.
(b) Center frequency, 500 hertz.

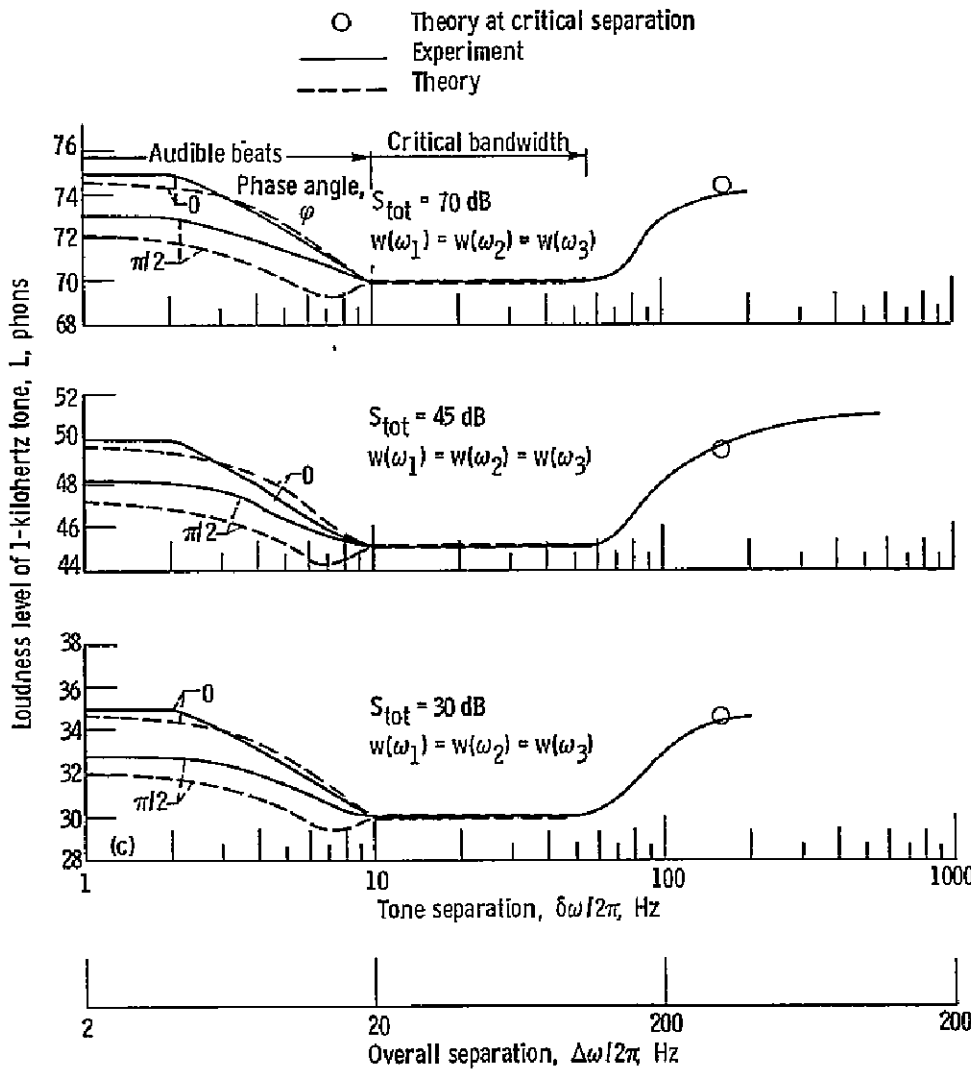
FIGURE 19.—Loudness level of three tones of equal intensity as function of tone separation. (Data from Bauch (ref. 37).)

Since all tone frequencies corresponded to the flat response range of the auditory system, all the Q's in the appropriate loudness equations (88) and (89) can be set equal to unity. The selected tone-amplitude ratios and corresponding acoustic power ratios are as follows:

$A_1/A_2 = A_3/A_2$	1/4	1/2	1
$w(\omega_1)/w(\omega_2) = w(\omega_3)/w(\omega_2)$	1/16	1/4	1
$w_{tot}/w(\omega_2)$	9/8	3/2	3

where

$$w_{tot} = w(\omega_1) + w(\omega_2) + w(\omega_3)$$



(c) Center frequency, 1 kilohertz.
FIGURE 19.—Concluded.

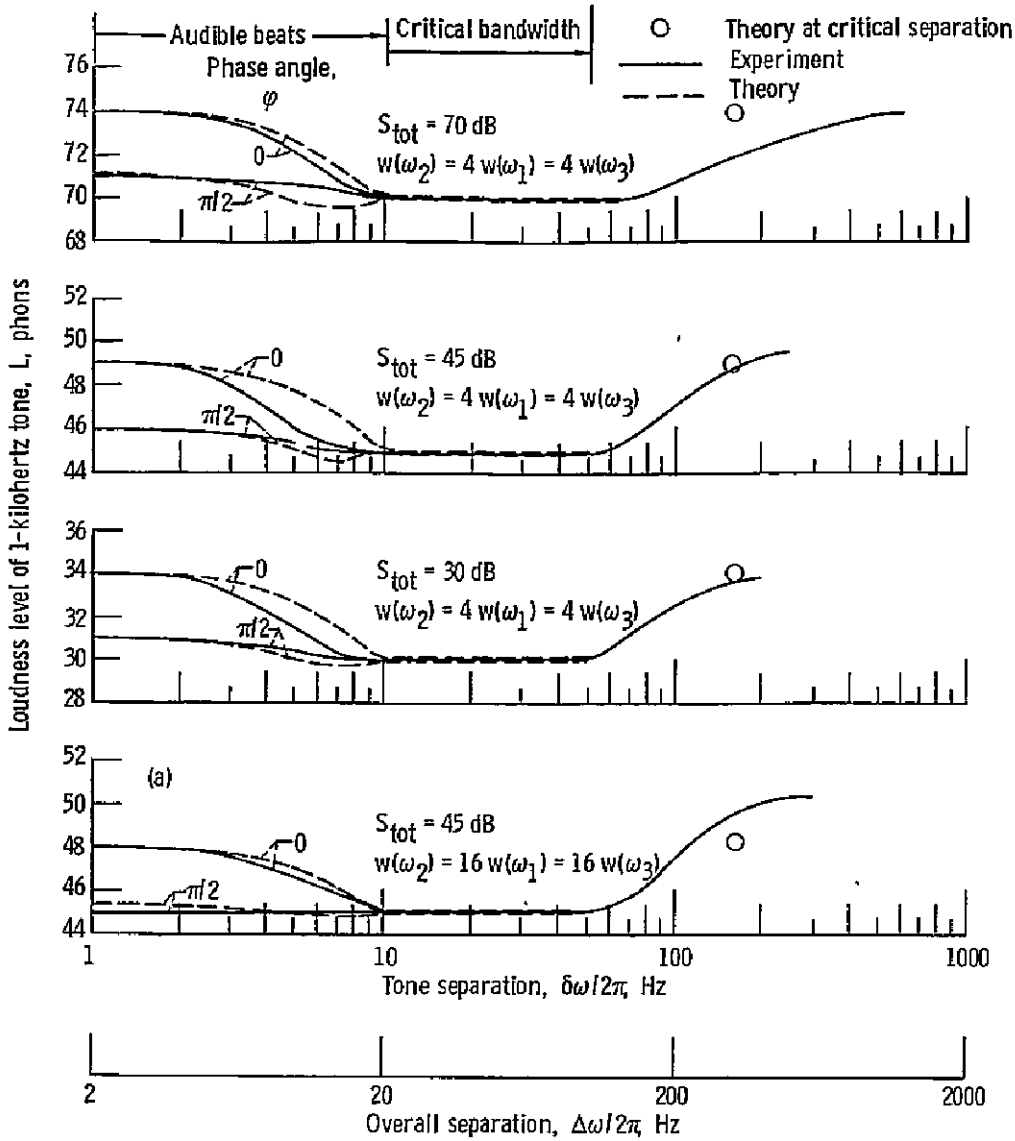
From equation (88) it can be shown that, for a symmetric spectrum of three suprathreshold tones uniformly spaced over a subcritical frequency range,

$$\begin{aligned}\mathcal{L}/\mathcal{L}_{cr} = (1 + (2/w_{tot})\{&2w(\omega_1,\omega_2)\,\mathrm{sinc}[(\omega_2-\omega_1)T/2\pi] \\ &\times\cos[(\omega_2-\omega_1)(t-T/2)+(\phi_1+\phi_3)/2] \\ &\times\cos[(2\phi_2-\phi_1-\phi_3)/2]+w(\omega_1)\,\mathrm{sinc}[2(\omega_2-\omega_1)T/2\pi] \\ &\times\cos[2(\omega_2-\omega_1)(t-T/2)+\phi_3-\phi_1]\})^{1/2}\end{aligned}$$

where

$$\mathcal{L}_{cr} = 10^{-2} w_{tot}^{1/2}$$

is the loudness of the triad when its total range is slightly less than a critical bandwidth. The loudness $\mathcal{L}$ fluctuates temporally at the rotational frequencies $\omega_2-\omega_1$ and $2(\omega_2-\omega_1)$. The maximum loudness is obtained if $\phi_1=\phi_2=\phi_3=0$. For this value of the phase angles the loudness still fluctuates. However, a listener tends to classify the sub-



(a) Center frequency, 1 kilohertz.

FIGURE 20.—Loudness level of three tones of unequal intensity as function of tone separation. (Data from Bauch (ref. 37).)

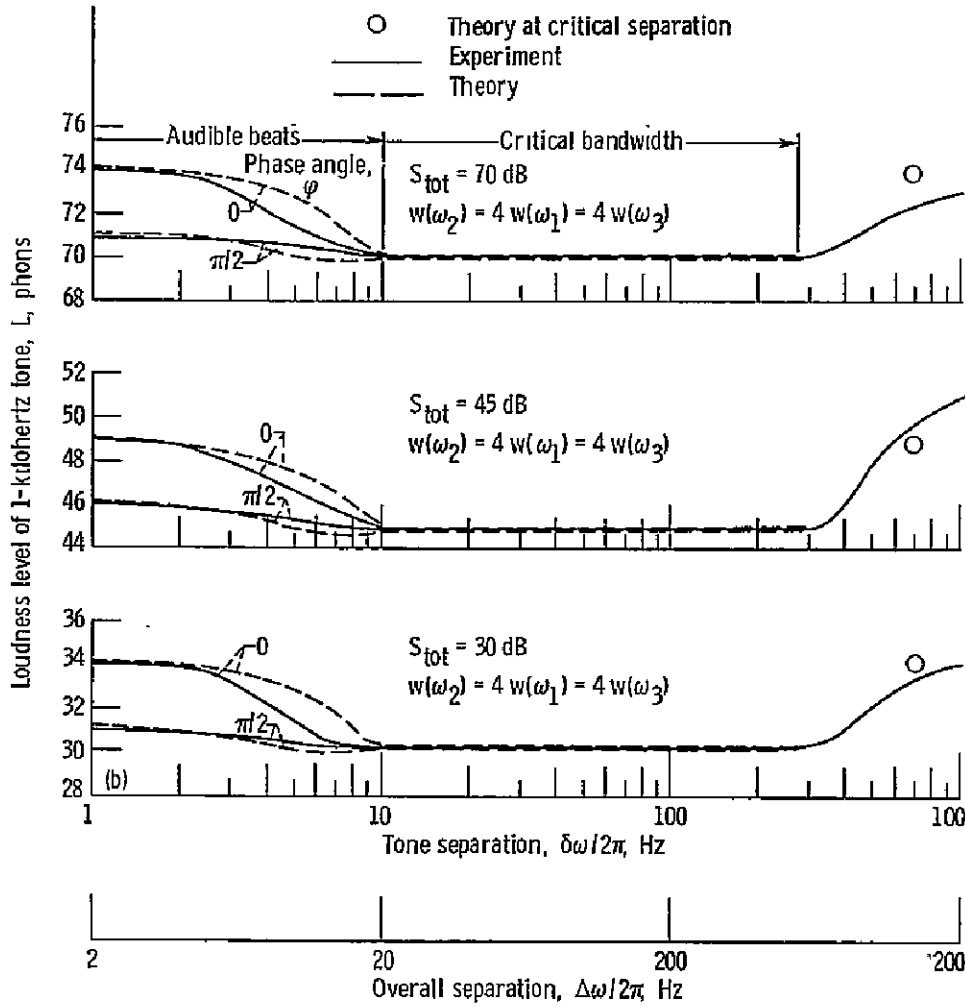
jective intensity of the sound according to its maximum loudness, which, relative to $\mathcal{L}_{cr}$, is given by

$$(\mathcal{L}/\mathcal{L}_{cr})_{max} = (1+(2/w_{tot})\{2w(\omega_1,\omega_2)\operatorname{sinc}[(\omega_2-\omega_1)T/2\pi] + w(\omega_1)\operatorname{sinc}[2(\omega_2-\omega_1)T/2\pi]\})^{1/2}$$

If the phase of the center frequency is shifted by $\pi/2$, so that $\phi_1=\phi_3=0$, $\phi_2=\pi/2$, then

$$(\mathcal{L}/\mathcal{L}_{cr})_{max} = \{1+2[w(\omega_1)/w_{tot}]\operatorname{sinc}[2(\omega_2-\omega_1)T/2\pi]\}^{1/2}$$

If the adjacent tones are spaced supercritically in adjacent critical bands, then, since the Q's$=1$, and $w(\omega_1)=w(\omega_3)$, equation (89) reduces to



(b) Center frequency, 4 kilohertz.
FIGURE 20.—Concluded.

$$\mathcal{L}/\mathcal{L}_{cr} = \{(3/2) + (4/\pi w_{tot})\,[w(\omega_1) + 2w(\omega_1,\omega_2)]\}^{1/2}$$

For lesser spacings the adjacent tones are spaced subcritically whereas the outer tones may still be spaced supercritically. Hence the method for category AB should be appropriate. However, because the detailed functional dependence of β_{mn} on tone spacing is unknown, in the present instance, where the value of an individual β_{mn} may be highly significant, the loudness in the interval between supercritical overall spacing and critical individual spacing of the tones will not be evaluated. However, calculated levels for critical individual spacing can be computed. When more tones are involved, the value of an individual β_{mn} is not so crucial, since only a minimum number of frequency components is spaced slightly supercritically, in which case β_{mn} is most sensitive to spacing.

The loudness levels given by the preceding formulas are compared in figures 19 (equally intense tones) and 20 (unequally intense tones)

with the corresponding loudness judgment data for the various center frequencies and sound pressures. The theoretical loudness ratios have been converted to loudness-level differences ΔL and thence to loudness levels L by means of equations (54) and (55), respectively. In all instances the predicted loudness levels in the subcritical tone-separation range are within 2 phons of the judged levels. Moreover, the effect of the phase of the central tone is properly accounted for quantitatively. The small downward digression of the theoretical curves for $\phi=\pi/2$ in the subcritical range is hardly significant. For critically separated tones the predicted loudness levels are within 2.5 phons of the judged values in all instances.

D. H. Howes' experiment (1950).—Loudness levels of the following triads were judged as in the preceding two-tone experiment (ref. 58):

Triad	Spacing, mels	$\omega_1/2\pi$, Hz	$\omega_2/2\pi$, Hz	$\omega_3/2\pi$, Hz	k
1	250	1420	1900	2450	1
2	500	1000	1900	3120	5

The tones were equally intense and spaced supercritically. Overall loudness judgments were made with each tone at the sound-pressure levels previously mentioned. The data and their comparison with predictions are presented in table II and figure 21.

The loudness of each triad is given by equation (89). For each tone, $Q\approx 1$ for practical purposes. For triad 1, $k=1$. For triad 2, $k=5$. Hence, the ratio of the loudness of triad 1 to that of the center tone is

$$\mathcal{L}(\omega_1,\omega_2,\omega_3)/\mathcal{L}(\omega_2)=3.13$$

The corresponding loudness-level difference is

$$\Delta L=9.9 \text{ phons}$$

For triad 2 the corresponding results are

$$\mathcal{L}(\omega_1,\omega_2,\omega_3)/\mathcal{L}(\omega_2)=3.98$$

$$\Delta L=12.0 \text{ phons}$$

These results are listed in table II. The average experimental level differences were $\Delta L=11.3$ phons for triad 1, and $\Delta L=13.2$ phons for triad 2. These values disagree with the predictions by less than 1.5 phons.

For a triad of equally intense tones centered at 1000 hertz and spaced 250 mels apart, the average difference between calculated and judged loudness levels is about 1 phon for listener J. M. and less than 1 phon for listener D. H., as found from table III.

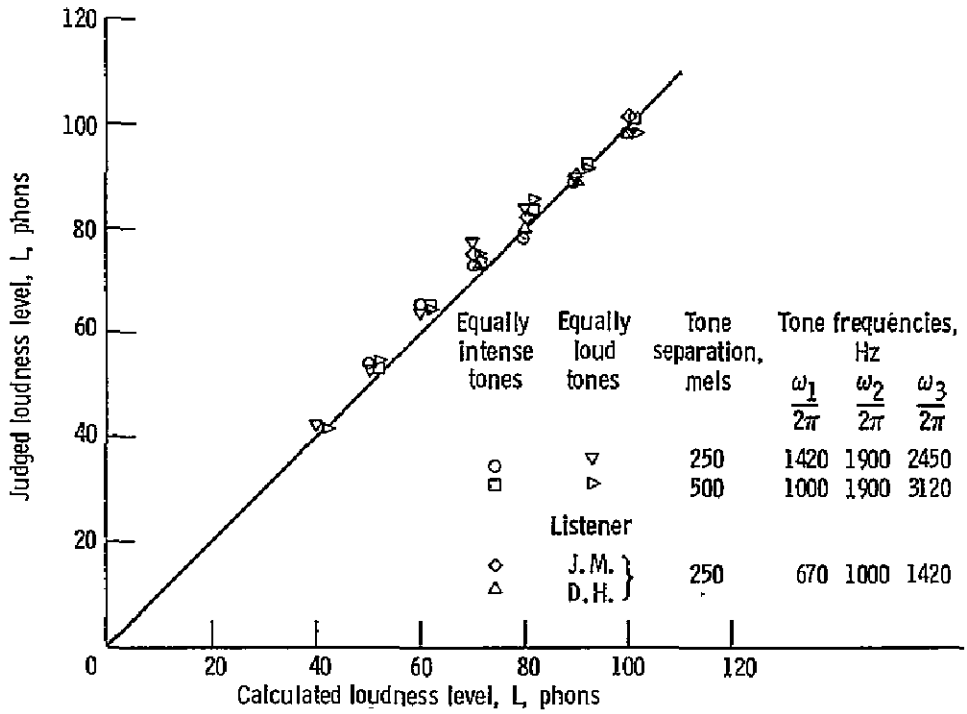


FIGURE 21.—Comparison of judged and calculated loudness levels for three tones. (Data from D. H. Howes (ref. 58).)

If equation (73) is used, for three *equally loud* tones centered at 1900 hertz and spaced 250 mels (and then 500 mels) apart, the average difference between calculated and judged loudness levels is found to be about 2.5 phons for 250 mel spacing and about 1.5 phons for 500 mel spacing (see table IV).

Howes also reported on an unequal loudness, three-tone experiment in which the loudness of three tones of unequal intensity was equated to that of three equally intense reference tones. Both triads were centered at 1900 hertz and spaced 250 mels, and then 500 mels, apart. The loudness level of the central component of one triad was set at 60 decibels, which equaled the loudness level of each component of the reference triad. The low frequency component of the former triad was set at a fixed attenuation below the level of the central component. Then the level of a high-frequency component was adjusted until the loudness of the triad equaled that of the reference triad.

Since $Q \approx 1$ for all frequencies concerned, and since $\mathcal{L}(\omega_1,\omega_2,\omega_3) = \mathcal{L}[\mathcal{L}(\omega_1) = \mathcal{L}(\omega_2) = \mathcal{L}(\omega_3)]$, where the left side refers to the trio of unequally loud tones and the right side refers to the reference triad of equally loud tones, it can be shown using equation (89) that

$$w(\omega_3)/w(\omega_2)=\frac{16}{(3+k)^2\pi^2}\left\{1+\left[\frac{w(\omega_1)}{w(\omega_2)}\right]^{1/2}-\frac{\pi}{4}\left\{18+\frac{72}{\pi}+\frac{16}{\pi^2}\right.\right.$$
$$+\left(12+\frac{24}{\pi}\right)k+2k^2-\left(\frac{24}{\pi}-\frac{32}{\pi^2}+\frac{8}{\pi}k\right)\left[\frac{w(\omega_1)}{w(\omega_2)}\right]^{1/2}$$
$$\left.\left.-\left(9-\frac{16}{\pi^2}+6k+k^2\right)\frac{w(\omega_1)}{w(\omega_2)}\right\}^{1/2}\right\}^2$$

where ω_3 is the high frequency, and ω_2 is the center frequency. When expressed in decibels, this ratio yields a sound-pressure-level difference which is added to the 60-decibel level of the center frequency to obtain the predicted sound-pressure level $S(\omega_3)$ of the high-frequency tone.

In figure 22 the measured sound-pressure level of the high-frequency component is shown as a function of the attenuation, in decibels, of the low-frequency component. In the theoretical calculations, $k=1$ for 250-mel spacing and $k=5$ for 500-mel spacing. The two theoretical curves differ by less than 1 decibel. Although the curves indicated by judgment data and by calculations are generally similar in shape, the experimental sound-pressure level of the high-frequency tone for equal loudness exceeds the theoretical value by up to 6 decibels. Since the sound-pressure level $S(\omega_3)$ increases by only a small amount, say 3 decibels, as the low frequency attenuation is changed from −10 to −40 decibels, it is possible that the disagreement between theory and experiment is primarily due to experimental error. It is difficult to imagine how a reduction of power in the low-

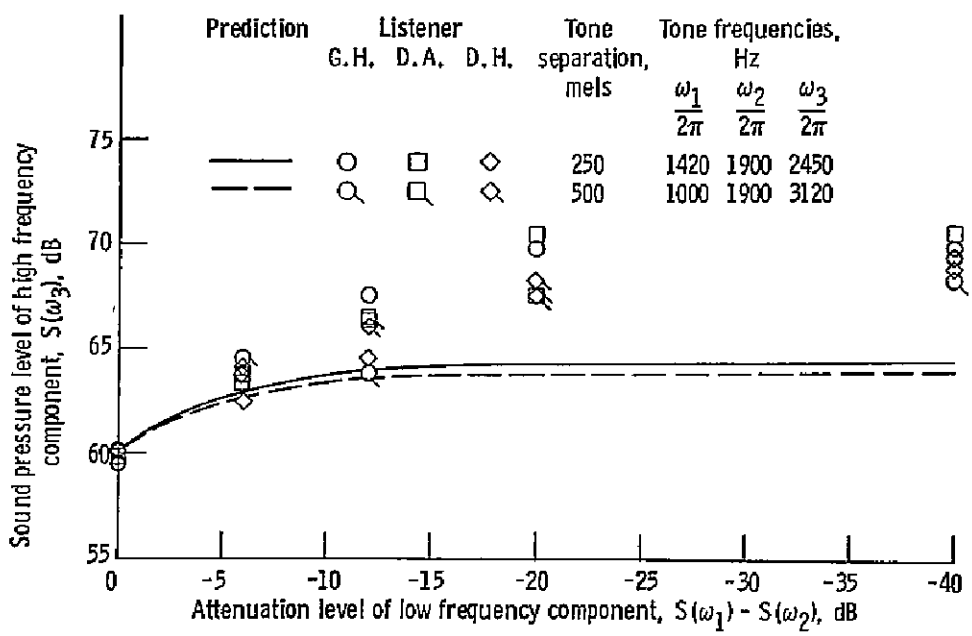


FIGURE 22.—Sound-pressure levels of components of tone triad with loudness equaling that of similar triad of 60-decibel tones. Tone frequency $\omega_2/2\pi$, 1900 hertz; $S(\omega_2)$, 60 decibels. (Data from D. H. Howes (ref. 58).)

frequency component from 10^{-1} to 10^{-4} that of the center frequency would require doubling the power of the (most intense) high-frequency tone, especially since the relative power in the low-frequency tone is so low to start with.

Scharf's experiment (1962).—Nine trained listeners equated the loudness of a triad of equally spaced tones to that of the center frequency (2000 Hz) alone (ref. 59). The intensities of the outer frequency tones were varied relative to that of the center tone, but the overall sound pressure of the triad, or of the comparison tone, was held constant. Judgments were made for overall separations of 280, 1100, and 2200 hertz and a variety of sound-pressure levels. The first overall separation is approximately critical, and the other two separations are supercritical. Hence, equation (88) should predict the loudness if $|\omega_3-\omega_1|/2\pi=280$ hertz, and equation (89) should be appropriate if $|\omega_3-\omega_1|/2\pi=1100$ or 2200 hertz.

For the subcritically separated triad with a symmetrical spectrum and the intensity of the outer tones less than, or equal to, that of the central tone, it was found from tests at overall sound-pressure levels of 20, 50, and 80 decibels that the loudness is independent of the intensity of the outer tones relative to that of the center tone. This result is confirmed by equation (88) since $\sum_{n=1}^{3} Q(\omega_n)w(\omega_n)=\text{constant}$ for any given fixed overall sound pressure.

For the corresponding supercritically spaced triads, the following conditions applied:

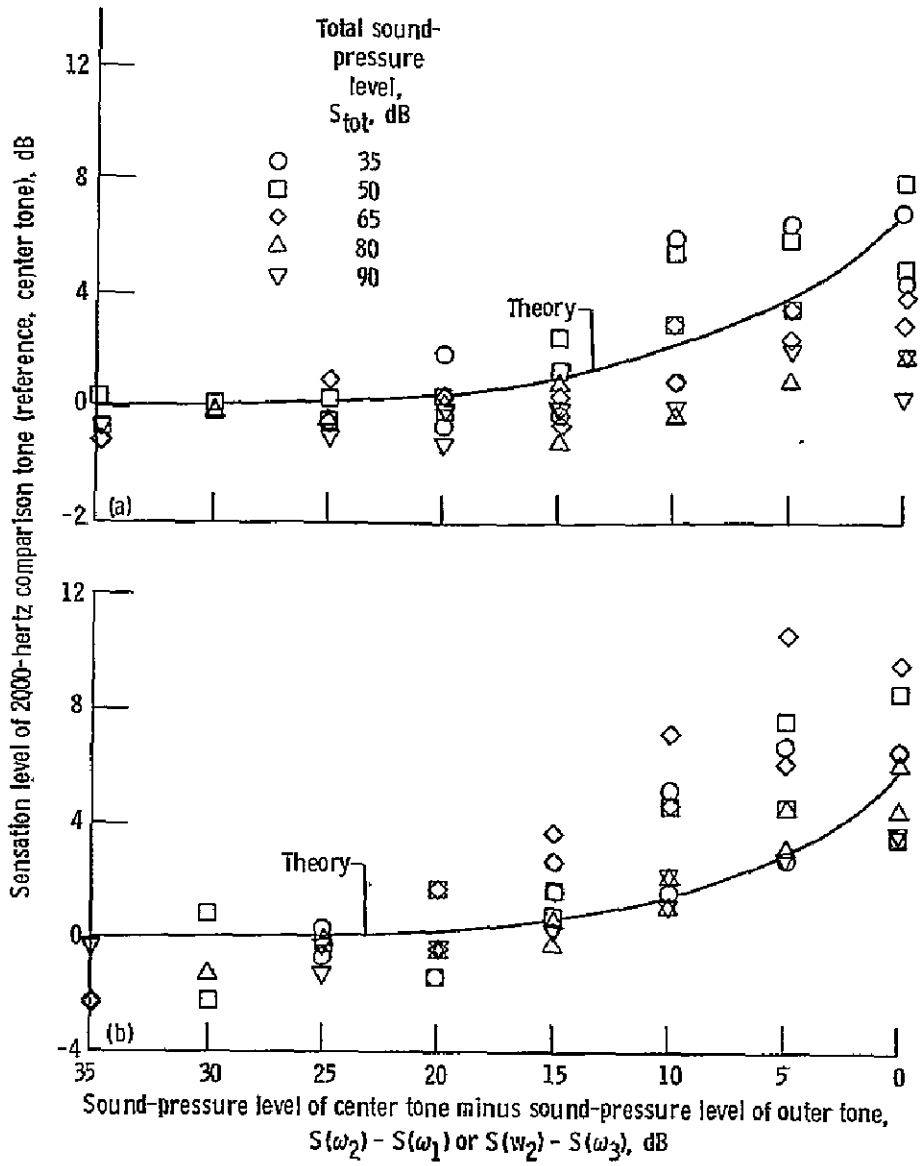
Triad	$\omega_1/2\pi$, Hz	$\omega_2/2\pi$, Hz	$\omega_3/2\pi$, Hz	k
1	1450	2000	2550	1
2	900	2000	3100	5

For practical purposes, $Q\approx 1$ for each tone. The tone frequencies are nearly the same as those selected by D. H. Howes (ref. 58). As shown in figure 23 the overall loudness rose as the intensities of the outer tones approached that of the center tone.

It follows from equation (89) that the loudness of the supercritically spaced, symmetric triad relative to that of the triad when the center tone is strongly dominant is given by

$$\mathcal{L}(\omega_1,\omega_2,\omega_3)/\mathcal{L}(0,\omega_2,0)$$
$$=\left\{\frac{1}{3}(3+k)+\frac{16}{3\pi}\left[\frac{w(\omega_1)}{w(\omega_2)}\right]^{1/2}+\left[\frac{2}{3}(3+k)+\frac{8}{3\pi}\right]\frac{w(\omega_1)}{w(\omega_2)}\right\}^{1/2}$$

since $w(\omega_3)=w(\omega_1)$. The curves given by this specialization of equation (89) are compared in figure 23 with the preceding data. The theoretical curves have been shifted so that they coincide with the



(a) Triad 1; tone spacing $\delta\omega/2\pi$, 550 hertz; overall tone spacing, $\Delta\omega/2\pi$, 1100 hertz.

(b) Triad 2; tone spacing $\delta\omega/2\pi$, 1100 hertz; overall tone spacing, $\Delta\omega/2\pi$, 2200 hertz.

FIGURE 23.—Loudness of tone triad with symmetric spectrum centered at 2000 hertz as function of attenuation of outer tones. (Data from Scharf (ref. 59).)

nominal data curve given in reference 59 when the central tone intensity is dominant.

The agreement between theory and experiment shown in figure 23 appears only fair because of the expanded scale and the wide scatter of the data. This scatter results primarily from a systematic tendency for the loudness to rise more for the central range of the sound-pressure level S_{tot}, in agreement with other data (ref. 58 and tables II

to IV). This effect is contained only implicitly in the β_{mn}. The functional dependence of this effect cannot yet be explicitly predicted from the theory and, more importantly, is not present in data from all sources (e.g., ref. 13).

In addition to loudness judgments of triads describing symmetric spectra, tests were also made of the same triads describing spectra with constant slope on a semi-logarithmic plot.

It follows from equation (89) that the loudness of a triad of supercritically spaced tones with a constant sloped spectrum is given, relative to the loudness for infinite slope (that of the outer tone), by

$$\mathcal{L}(\omega_1,\omega_2,\omega_3)/\mathcal{L}(\omega_1,0,0)=\left\{\frac{1}{3}(3+k)\left[1+\frac{w(\omega_2)}{w(\omega_1)}+\frac{w(\omega_3)}{w(\omega_1)}\right]\right.$$
$$\left.+\frac{8}{3\pi}\left[\left(\frac{w(\omega_2)}{w(\omega_1)}\right)^{1/2}+\left(\frac{w(\omega_3)}{w(\omega_1)}\right)^{1/2}+\left(\frac{w(\omega_2)}{w(\omega_1)}\frac{w(\omega_3)}{w(\omega_1)}\right)^{1/2}\right]\right\}^{1/2}$$

for negative slope and by

$$\mathcal{L}(\omega_1,\omega_2,\omega_3)/\mathcal{L}(0,0,\omega_3)=\left\{\frac{1}{3}(3+k)\left[\frac{w(\omega_1)}{w(\omega_3)}+\frac{w(\omega_2)}{w(\omega_3)}+1\right]\right.$$
$$\left.+\frac{8}{3\pi}\left[\left(\frac{w(\omega_1)}{w(\omega_3)}\frac{w(\omega_2)}{w(\omega_3)}\right)^{1/2}+\left(\frac{w(\omega_1)}{w(\omega_3)}\right)^{1/2}+\left(\frac{w(\omega_2)}{w(\omega_3)}\right)^{1/2}\right]\right\}^{1/2}$$

for positive slope.

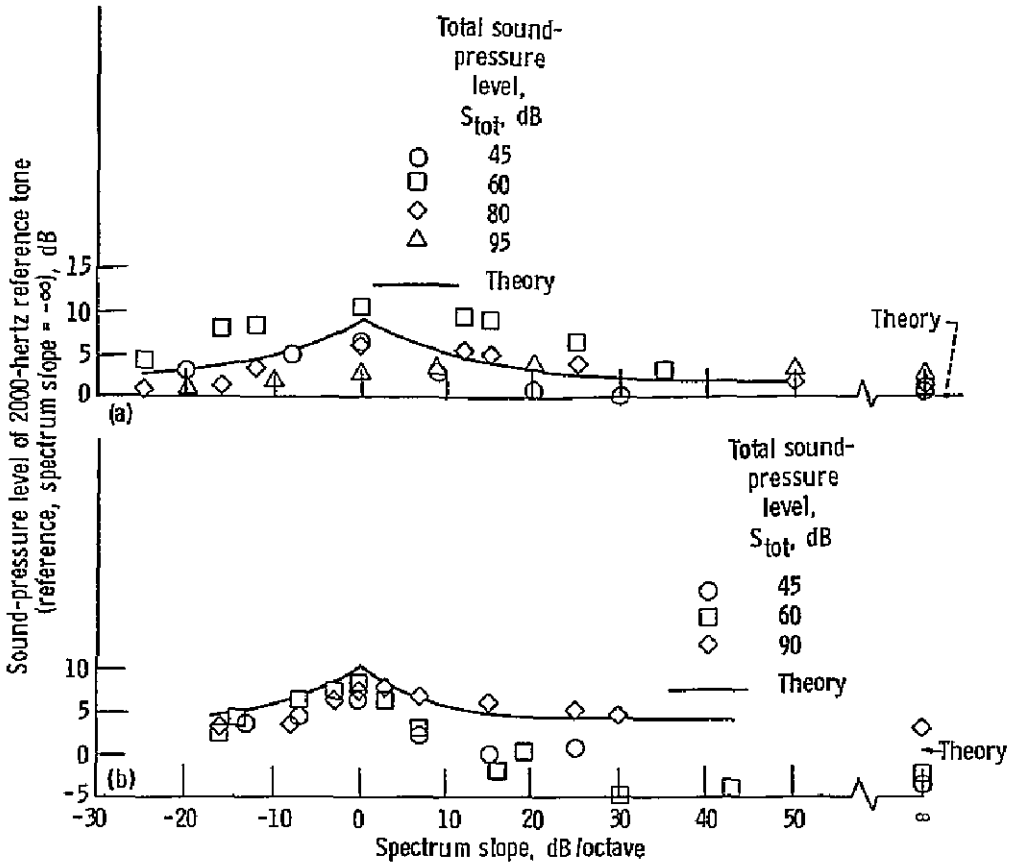
The experimental and theoretical results are compared in figure 24, in which the reference sound-pressure level is set to zero when the spectrum slope is minus infinity. Considering the scatter of the data the judged and calculated results are in fairly good agreement. There is no evidence that the loudness of the triad depends on the algebraic sign of the spectrum slope.

The reported interquartile range of sound-pressure level variability averaged over all observers was 4 decibels for the experiments on symmetric spectra and 6.5 decibels for the experiments on sloped spectra. In the latter case the overall variability was sometimes as great as 20 decibels.

Four Tones

The explicit expressions given for the loudness of two- and three-tone complexes should be sufficient to show how the loudness equations expand as the number of tones is increased.

D. H. Howes' experiment (1950).—Howes' studies (ref. 58) of four or more tones were similar to those already described. The following quartets were studied:



(a) Triad 1; overall tone spacing, $\Delta\omega/2\pi$, 550 hertz.
(b) Triad 2; overall tone spacing, $\Delta\omega/2\pi$, 1100 hertz.

FIGURE 24.—Loudness of tone triad centered at 2000 hertz as function of its spectrum slope. (Data from Scharf (ref. 59).)

Quartet	Separation, mels	$\omega_1/2\pi$, Hz	$\omega_2/2\pi$, Hz	$\omega_3/2\pi$, Hz	$\omega_4/2\pi$, Hz	k
1	250	(1000) 1420	1420 1900	1900 2450	2450 (3120)	2
2	500	670	1420	2450	4000	8

(Loudness judgments for 250-mel tone separation were made by alternating the first and second row of tones with the reference tone. Thus, the frequencies 1000 and 3120 Hz were alternated on successive presentations.) Quartets of equally intense tones centered at 1000 hertz and spaced 250 mels apart were also studied.

For the tone quartet with 250-mel spacing,

$$\mathcal{L}(\omega_1,\omega_2,\omega_3,\omega_4)/\mathcal{L}(\omega_j)=4.42$$

which corresponds to

$$\Delta\mathcal{L}=12.9 \text{ phons}$$

For 500-mel spacing the corresponding results are

$$\mathcal{L}(\omega_1,\omega_2,\omega_3,\omega_4)/\mathcal{L}(\omega_j)=5.62$$

$$\Delta L=15.0 \text{ phons}$$

In table II the predictions are compared with the judged loudness levels of the preceding tone quartets. For quartet 1 the difference between calculations and judgments is approximately 2 phons, whereas for quartet 2 the difference is approximately 3.5 phons.

For equally intense tones centered at 1000 hertz and spaced 250 mels apart, the disagreement between calculations and judgments is, as shown in table III, approximately 2.5 phons for listener J. M. and 0.5 phons for listener D. H.

For equally loud tones centered at 1900 hertz (quartets 1 and 2 previously tabulated) the disagreement between calculations and judgments shown in table IV is about 2 phons for 250-mel spacing and about 2.5 phons for 500-mel spacing.

In all cases the judged and calculated loudness levels are compared graphically in figure 25.

Zwicker, Flottorp, and Stevens' experiment (1957).—The loudness of quartets of uniformly spaced, equally intense tones was judged by comparison with a reference tone at the center frequency (ref. 38). Judgments were by 6 to 22, apparently naive, subjects utilizing binaural listening through earphones. Determinations of loudness were made as a function of tone spacing about the center frequencies 500, 1000, and 2000 hertz and for a fixed overall sound-pressure level of 57.5 decibels. An additional similar set of tests was performed using a tone quartet centered at 1000 hertz and at three other sound-pressure levels.

The tone spacing covered both subcritical and supercritical ranges. For the subcritical range the loudness of the quartet is given by equation (33), whereas for the supercritical range, equation (40) is appropriate.

No information was given as to the phase of the tones at $t=0$. It is evident from equation (33) and from the results for two and three tones that phase relations have an important effect on loudness if adjacent tones are separated by less than 10 hertz. Since a listener tends to characterize a fluctuating sound by its maximum loudness, it is to be expected that the four-tone data shown in figures 26 and 27 should tend to rise for overall tone spacing less than 30 hertz. Some rise appears in some of the data. However, if the rise is actually a systematic one it is certainly minimal, so that the initial phases of the tones may have been essentially random. The loudness is otherwise effectively independent of tone spacing for subcritical overall spacing, in agreement with equation (33). For supercritical overall spacing the loudness rises, as in the case of three tones. The calculated

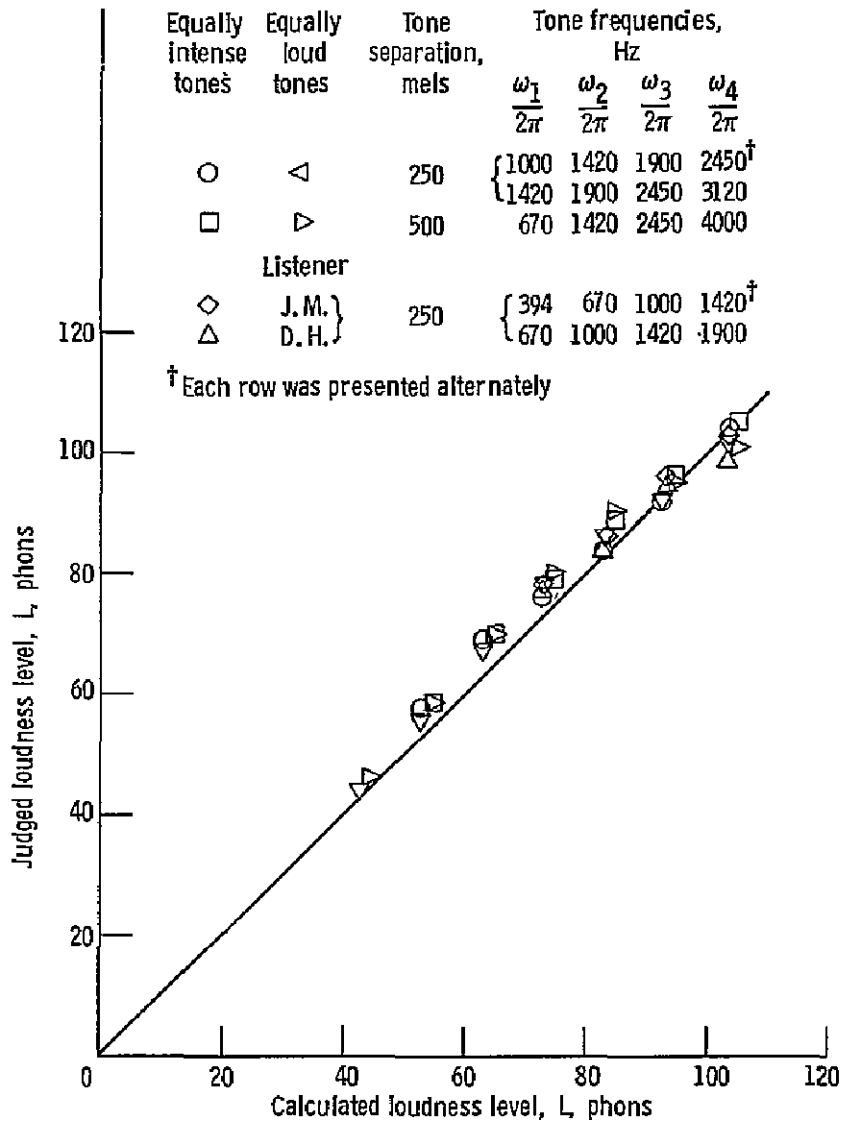
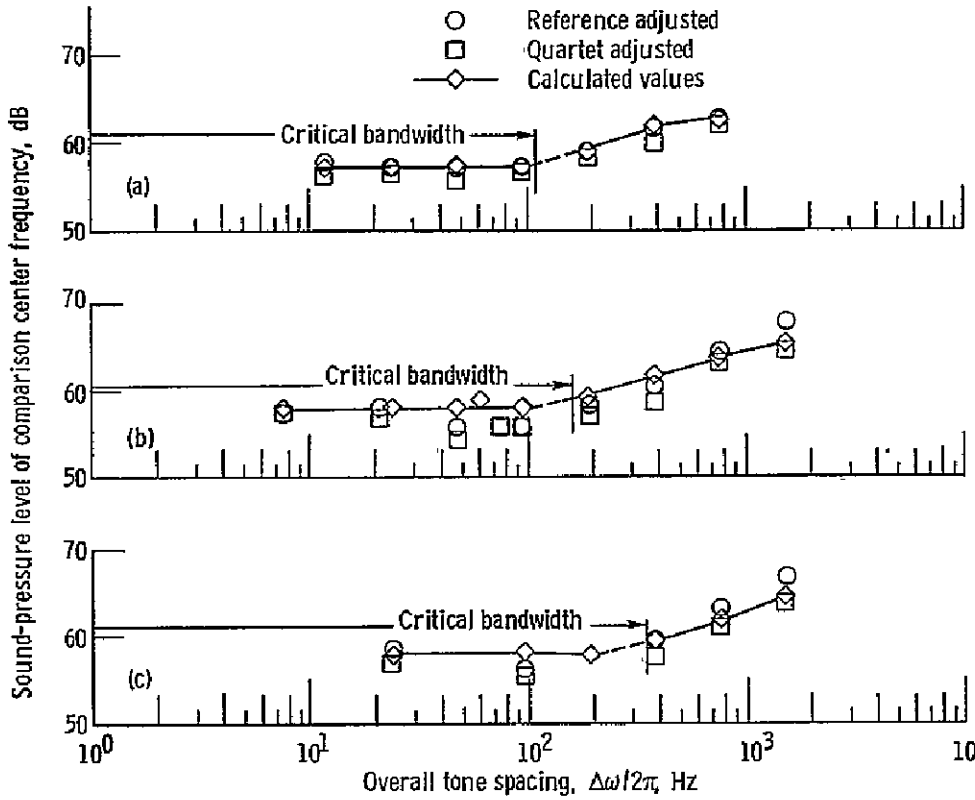


FIGURE 25.—Comparison of judged and calculated loudness levels for four tones. (Data from D. H. Howes (ref. 58).)

and judged loudness levels are generally in good agreement. The rise is between 5 and 10 phons for the maximum spacing studied. These results may be compared with those of D. H. Howes (ref. 58).

Some of the present results are for center frequencies of 1000 and 2000 hertz with a maximum overall tone spacing of 1536 hertz (512 Hz between tones). D. H. Howes' results included those for 1000- and 1900-hertz center frequencies with an overall spacing up to 1230 hertz in the former case and averaging 1575 hertz in the latter case. Thus, the given conditions in the two experiments are comparable. The loudness level increases, ΔL, of the quartets relative to the loudness level of one of the tones are quite similar in the two experiments, as shown in table V in which the level increases can be compared with each other as well as with the calculated values. Among all



(a) Geometric-mean frequency, $\omega_r/2\pi$, 500 hertz.
(b) Geometric-mean frequency, $\omega_r/2\pi$, 1000 hertz.
(c) Geometric-mean frequency, $\omega_r/2\pi$, 2000 hertz.

FIGURE 26.—Loudness level of equally intense tone quartet as function of uniform tone spacing showing effect of arithmetic-mean frequency $\omega_{am}/2\pi$. (Data from Zwicker, Flottorp, and Stevens (ref. 38).)

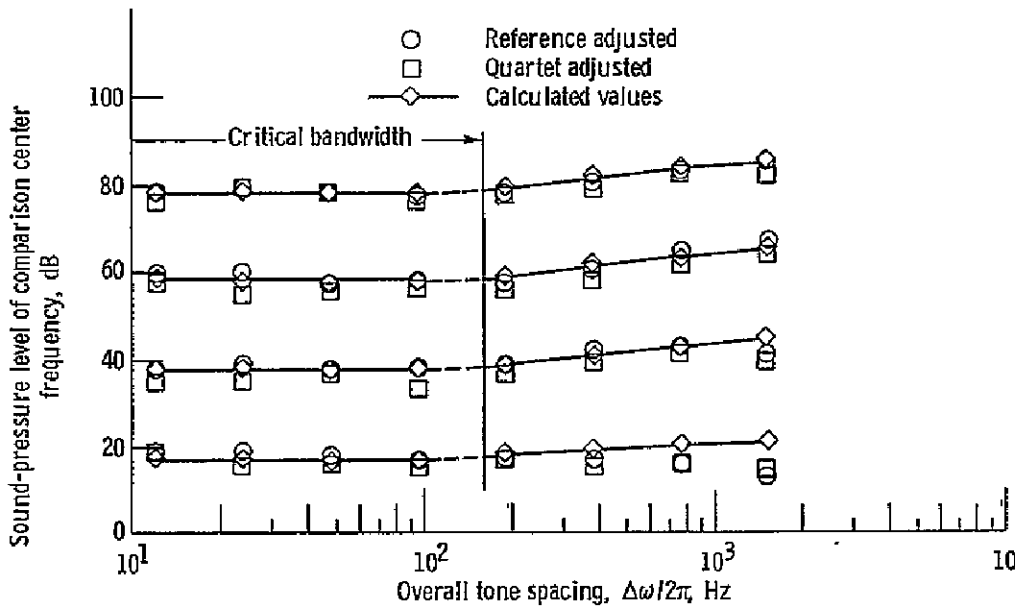


FIGURE 27.—Loudness level of equally intense tone quartet as function of uniform tone spacing showing effect of sound-pressure level. Geometric-mean frequency, $\omega_r/2\pi$, 1000 hertz. (Data from Zwicker, Flottorp, and Stevens (ref. 38).)

TABLE V.—COMPARISON OF FOUR-TONE LOUDNESS DATA OF D. H. HOWES (REF. 58) AND OF ZWICKER, FLOTTORP, AND STEVENS (REF. 38) WITH THEORY

Author	Reference	Center frequency, Hz	Overall tone spacing, Hz	Total sound-pressure level, S_{tot} dB	Experimental ΔL, phons	Calculated ΔL, phons
D. H. Howes	58	1900	1575	56	15.5	12.9
Zwicker, Flottorp, and Stevens	38	2000	1536	57.5	13.5	12.7
D. H. Howes	58	1000	1230	76	15.2	13.4
Zwicker, Flottorp, and Stevens	38	1000	1536	77.5	14.2	14.0

appropriate comparisons the maximum disagreement is about 2.5 phons.

Zwicker, Flottorp, and Stevens (ref. 38) also reported on an experiment by Scharf who found that the overall loudness level of a quartet of subcritically spaced (150 Hz overall spacing), equal-intensity tones is a linear function of the loudness level of a component, that is, the two loudness levels differ by a constant. This is in agreement with the present theory, as shown in figure 28.

The loudness of irregularly spaced quartets of tones was also determined relative to that of a tone approximating the center frequency

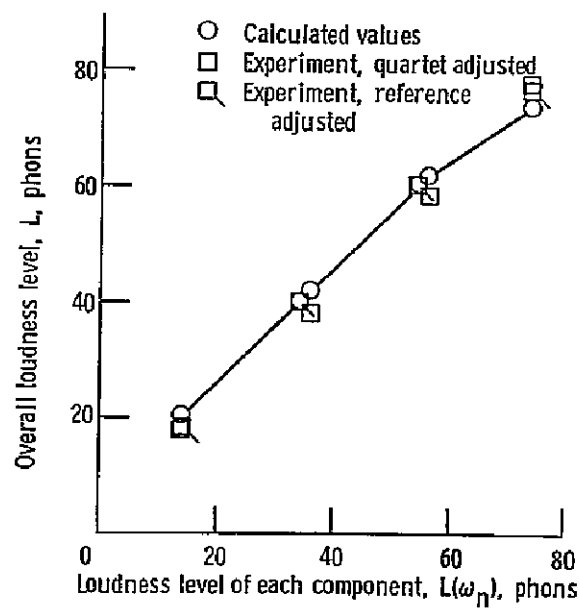
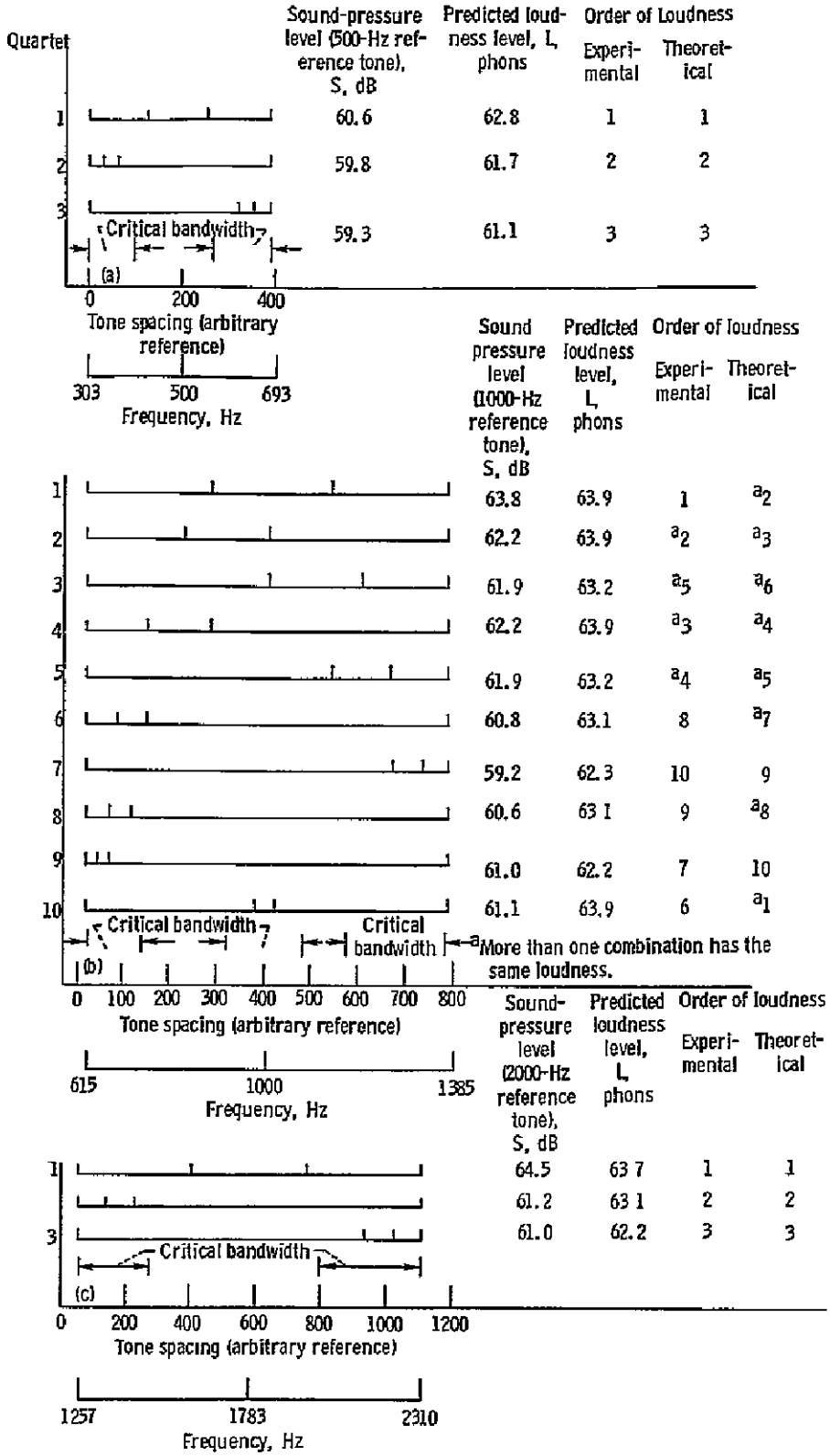


FIGURE 28.—Calculated loudness level of quartet of equally intense tones with subcritical overall spacing as function of loudness level of one tone. Geometric-mean frequency, $\omega_r/2\pi$, 1000 hertz; overall tone spacing, $\Delta\omega/2\pi$, 150 hertz. (Data from Zwicker, Flottorp, and Stevens (ref. 38).)

of the quartet. Three basic quartets were studied, namely, one centered at 500 hertz with an overall spacing of 390 hertz, one centered at 1000 hertz with an overall spacing of 770 hertz, and one centered at 1783 hertz with an overall spacing of 1053 hertz. In all cases the overall sound-pressure level of the quartet was 57.5 decibels. The results are shown in figure 29. For the ten irregularities tested about the 1000-hertz center frequency, the total calculated and experimental ranges of loudness variability were only 1.7 and 4.6 phons, respectively. In the other tests the ranges were the same or less. The order of loudness of the various irregularities are listed in figure 29. The judged and calculated orders of the quartets centered at 1000 hertz are in fair agreement, considering the restricted range of loudnesses. Perfect agreement of order is obtained for the other two quartets. Over all tests the predicted loudness levels are, on the average, about 1.5 phons different than the test levels.

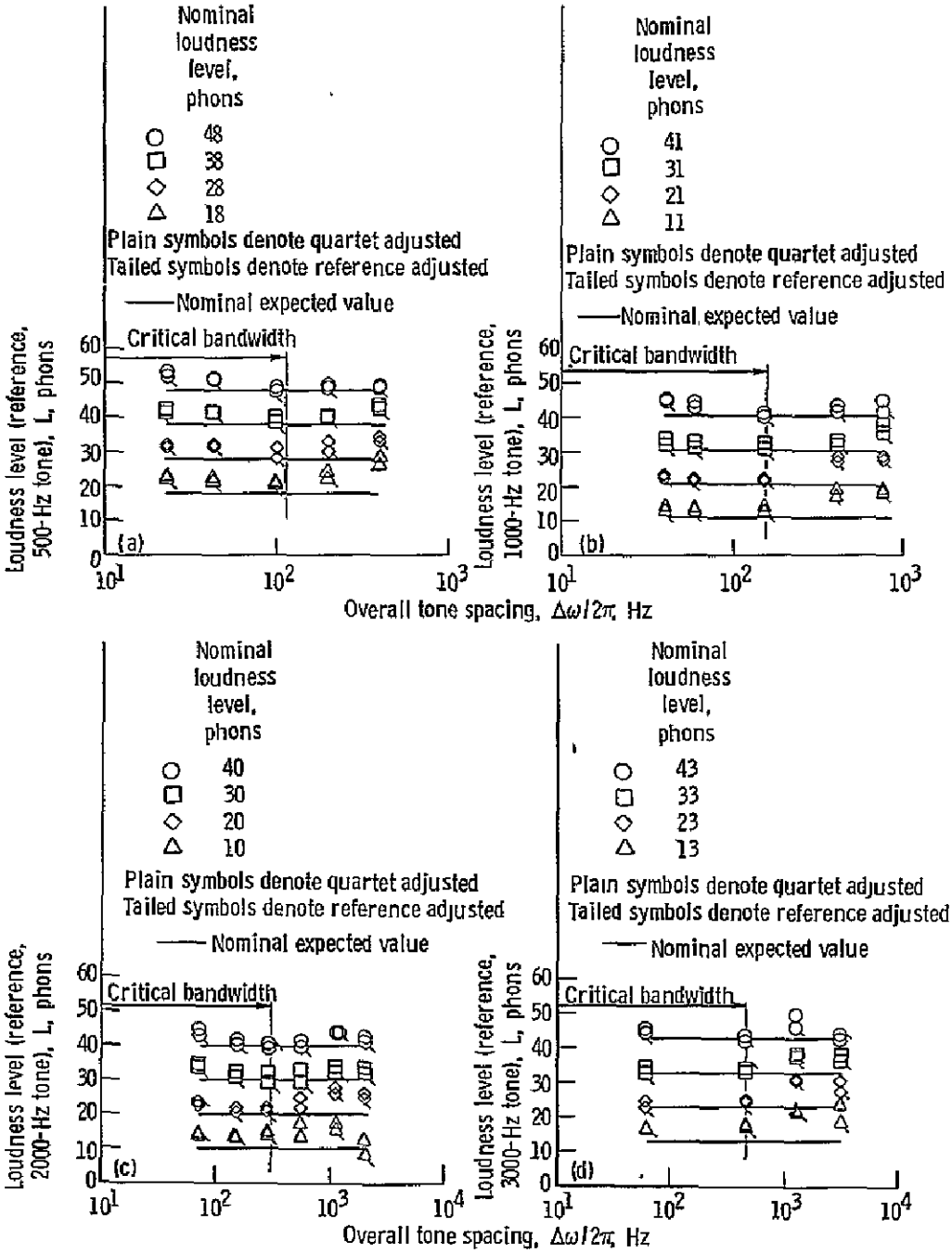
Zwicker, Flottorp, and Stevens concluded that maximum loudness is achieved for uniform spacing of the tones. This is in essential agreement with the present theory according to which, for equal intensity tones, the power from each tone is summed if the tones are subcritically spaced overall and each pair is more than 10 hertz apart. Otherwise, if the overall spacing is supercritical the power interaction terms in equation (40) indicate that a significant increase in loudness can be expected.

Scharf's experiment (1959).—Eight to twelve listeners listening binaurally through earphones equated the loudness of a quartet of equally loud tones near the loudness threshold to that of a reference tone held at a fixed loudness (ref. 43). The loudness of each tone as a function of tone spacing was determined with constant loudness of the combination as a parameter. The tones were apparently spaced in constant geometrical ratio of frequencies, and the reference-tone frequency was at the geometric-mean frequency of the quartet. Quartets with mean frequencies of 500, 1000, 2000, and 3000 hertz were studied. As shown in figure 30, most of the calculated loudness levels agree with the corresponding nominal levels (horizontal lines) to within the expected interquartile range obtained using tone pairs (cf. fig. 18). One cause of the discrepancy between theory and experiment is the fact that the tones were specified according to their loudness rather than sound-pressure level. The sound-pressure levels were determined by interpolation of the Robinson-Dadson, equal-loudness curves (dashed curves in fig. 8) with the external transmittance excluded. Any error in the interpolation would tend to be magnified because, near threshold, the loudness is determined by the difference of nearly equal quantities. Thus, in figure 30 the deviation from the nominal loudness level tends to be greatest near threshold.



(a) Arithmetic-mean frequency, $\omega_{am}/2\pi$, 500 hertz.
(b) Arithmetic-mean frequency, $\omega_{am}/2\pi$, 1000 hertz.
(c) Arithmetic-mean frequency, $\omega_{am}/2\pi$, 1783 hertz.

FIGURE 29.—Loudness of irregularly spaced quartets of equal intensity tones. (Data from Zwicker, Flottorp, and Stevens (ref. 38).)



(a) Arithmetic-mean frequency, $\omega_{am}/2\pi$, 500 hertz.

(b) Arithmetic-mean frequency, $\omega_{am}/2\pi$, 1000 hertz.

(c) Arithmetic-mean frequency, $\omega_{am}/2\pi$, 2000 hertz.

(d) Arithmetic-mean frequency, $\omega_{am}/2\pi$, 3000 hertz.

FIGURE 30.—Calculated loudness level of equally loud tone quartet near threshold as function of uniform tone spacing. (Data from Scharf (ref. 43).)

Five Tones

D. H. Howes' experiment (1950).—The following quintets were studied with monaural earphone listening by four trained subjects

listening to equally intense tones or two subjects listening to equally loud tones (ref. 58):

Quintet	Separation, mels	$\omega_1/2\pi$, Hz	$\omega_2/2\pi$, Hz	$\omega_3/2\pi$, Hz	$\omega_4/2\pi$, Hz	$\omega_5/2\pi$, Hz
1	250	1000	1420	1900	2450	3120
2	500	394	1000	1900	3120	5100

Similar measurements were made for a quintet of equally intense tones centered at 1000 hertz and spaced 250 mels apart.

The judged and calculated loudness levels for equally intense tones are shown in tables II and III. The results for equally loud tones are presented in table IV. All results are displayed graphically in figure 31. For equally intense tones centered at 1900 hertz the average disagreement between test and theory is about 2.5 phons for

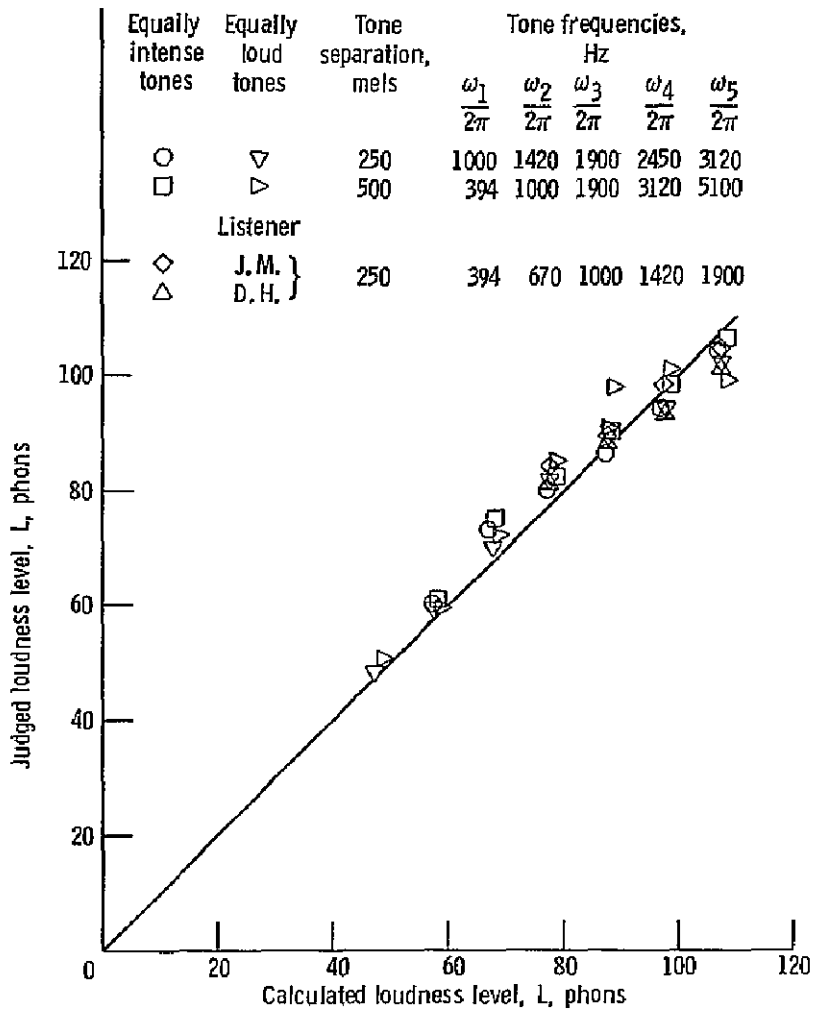


FIGURE 31.—Comparison of judged and calculated loudness levels for five tones. (Data from D. H. Howes (ref. 58).)

TABLE VI.—DATA FROM QUIETZSCH'S EXPERIMENTS (REF. 60)

(a) Noise sources and their loudness levels

Noise	Identification of sound or sound source	Subjective loudness level, phons
1	Murmuring, after passing through a wall.	74.5
2	Busing truck in stand.	98
3	Ventilator noise at exit of a 5-m long channel.	88
4	Murmuring (spectrum not changed).	60.5
5	Large compressed air hammer in room.	103
6	Murmuring, after passing through wall.	53.5
7	Passing freight train at 12 m distance.	91.5
8	Music (spectrum not changed).	60
9	Lohmann auxiliary motorcycle in stand, measured 7 m to right rear.	92.5
10	Aircraft motor noise inside test stand (level decreased by about 44 dB).	101
11	1. Electrical machine.	105
12	Four-cylinder motorcycle (BMW) in stand, measured at 5 m distance in open, nonbuiltup area.	84 5
13	Music, after passing through wall.	53
14	Portable drill, drilling aluminum (closed room).	82
15	Four-cylinder motorcycle (BMW) (see noise 12).	103
16	Noises, after passing through wall.	49
17	Noise from milling machine, taken at 2 m distance.	83
18	2. Electrical machine.	105.5
19	Shaking of screws in metal box.	76
20	3. Electrical machine (motor-generator).	104
21	Noise at exit of aircraft motor test stand, at 1.5 m distance toward the side.	92
22	Impulses with noise (impulse rate 30 s^{-1}).	78
23	Eight tone mixture.	80.5
24	Noise, with uniformly decreasing spectrum.	61
25	White noise (up to 10 kHz).	74.5
26	Dropping of water on roofing felt with stone underlayment.	79.5
27	Screw machine with felt pad on wooden table.	82
*28	Hissing of compressed air escaping from nozzle.	82
29	1. Automobile horn (fanfare type).	86
30	Impulse noise (pulse rate 2250 s^{-1}).	69
31	Hammer blows on iron plate.	73.5
32	Tone mixture of 20 harmonic components with statistical phase relations to each other.	71
33	2. Automobile horn.	86
34	Noise, with steadily increasing spectrum.	66.5
35	White noise, after passing through wall.	62.5
*36	Buzzing noise.	87
*37	Electric bell.	87.5

* Noises 28, 36, and 37 are original sounds (not produced by loudspeakers).

TABLE VI.—Continued.

(b) Octave-band sound-pressure levels of 37 noises

Geometric-mean frequency, $\omega_\nu/2\pi$, Hz	Lower cut-off frequency, Hz	Upper cutoff frequency, Hz
71	50	100
141	100	200
283	200	400
566	400	800
1131	800	1 600
2262	1600	3 200
4525	3200	6 400
9051	6400	12 800

Noise	Total sound-pressure level, S_{tot}, dB	Geometric-mean frequency, $\omega_\nu/2\pi$, Hz							
		71	141	283	566	1131	2262	4525	9051
		Sound-pressure level, S, dB (in octave)							
1	75.	55	71.5	70.5	67	58.5	50	<27	----
2	98.5	97	87.5	85.5	86.5	86.5	83.5	71	54
3	83	70	69	77.5	77	74	74	64.5	50
4	55	31	43	49	52	45	40	26	----
5	97.5	87.5	89.5	91	90	87	87.3	84	81.8
6	53	33	49.5	48.5	45	36.5	28	<15	----
7	86.5	82.5	81.5	75	73.5	74	74 5	65.5	53
8	53.5	28.5	38	48	51.5	44	36.5	21	----
9	83	67	70.5	70.5	73	75.5	79.5	73	63 5
10	92.5	81.5	85	81	84	83.5	84 5	82.5	77
11	95	69	78	85.5	90.5	89.5	88	86	80.5
12	77	73	71.5	71.5	67.5	63.5	55	49.5	44
13	49	29.5	39.5	45.5	45	37.5	26	<16	----
14	69.5	50.5	43.5	46	55.5	67	63.5	61 5	57.5
15	98	87.5	95.5	94	86	82	71	71	69
16	48.5	37	45.5	45	36.3	28.5	26	<18	----
17	70	50	48.5	59.5	61.5	64.5	66	53.5	45 5
18	95	75.5	79	90	92	86.5	87.5	79	74.5
19	61.5	43.5	41	37	56.5	55	56	50.5	49.5
20	94	83	90	89.5	86 5	84	75.5	79.5	73.5
21	94	94	78	61	58	56.5	57	58	56.5
22	62	47.8	54.5	51	51.5	53	55.5	55.2	47
23	64	56	53.5	55	55	55.5	54	54	53
24	48.5	29	41	44 5	43	41	38.5	29	17
25	57	37.5	40	44.5	46.5	47.5	50.5	51.5	51.5
26	62.5	53	50.5	46	49	51	55	58	51
27	65.5	60	55.5	52.5	51	59.5	57	55	53
28	75	35	34.5	36	39.3	41	52	66	74
29	70.5	----	----	48	56.8	68.2	67.3	50.8	40

TABLE VI.—Continued.

(b) Continued

Noise	Total sound-pressure level, S_{tot}, dB	Geometric-mean frequency, $\omega_r/2\pi$, Hz							
		71	141	283	566	1131	2262	4525	9051
		Sound-pressure level, S, dB (in octave)							
30	55.5	----	----	----	----	----	53 2	50 5	46
31	54.5	41.5	44.5	34.5	40.5	48.5	49.5	49	40
32	53	----	39	41.5	43.5	47.5	49.5	33	----
33	70	----	----	53.5	46.5	47.5	70	58.5	48.5
34	52.5	25	31	29.5	30.5	35	41	45.5	50
35	47	41	41	40	38	34.5	36	36	34
36	66.5	32.5	40	48	51	64	56	49	58.5
37	75.5	32.5	32.5	40.5	40	41	53	60	75

(c) Sound-pressure levels of each component of noises 23, 29, 30, 32, and 33

Number of the component	Frequency, Hz	Sound-pressure level, dB
Noise 23		
1	71	56
2	142	53.5
3	283	55
4	566	55
5	1 132	55.5
6	2 264	54
7	4 528	54
8	9 056	53
Noise 29		
1	410	56.5
2	820	64
3	1 230	66
4	1 640	66
5	2 050	60
6	2 460	55
7	2 870	44
8	3 280	45
9	3 690	38
10	4 100	44
11	4 510	43
12	4 920	42
13	5 330	40
14	5 740	36
15	6 150	34
16	6 560	31
17	6 970	33
18	7 380	28
19	7 790	31
20	8 200	33
21	8 610	31
22	9 020	29
Noise 30		
1	2 250	53.2
2	4 500	50.5
3	6 750	45
4	9 000	39.5
5	11 250	34
Noise 32		
1	127	39
2	254	38
3	381	39.5
4	508	38
5	635	38.5
6	762	40.5
7	889	39
8	1 016	39
9	1 143	40.5
10	1 270	40.5
11	1 397	40
12	1 524	40
13	1 651	39
14	1 778	39.5

TABLE VI.—Concluded.

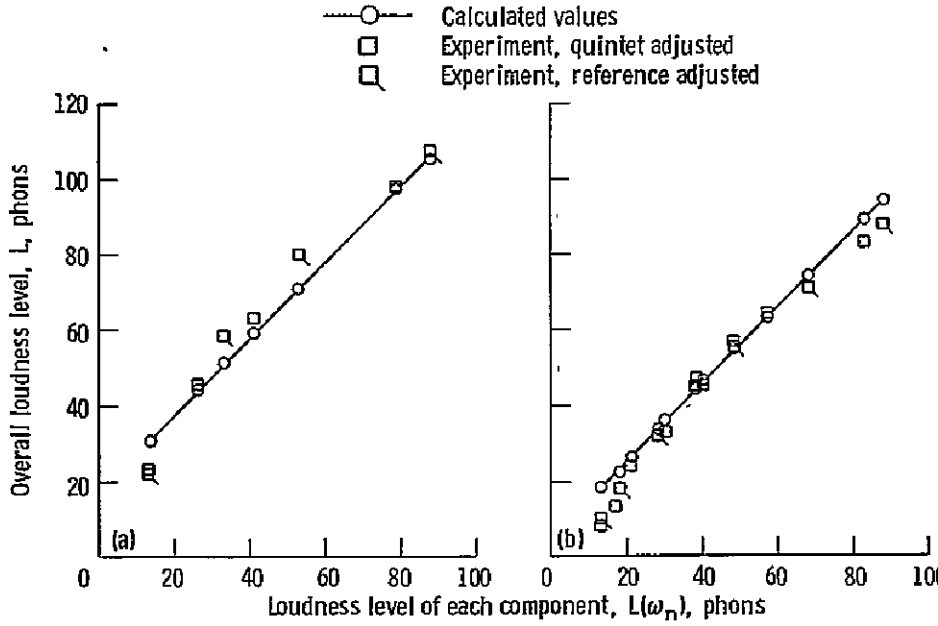
(c) Concluded

Number of the component	Frequency, Hz	Sound-pressure level, dB	Number of the component	Frequency, Hz	Sound-pressure level, dB
15	1 905	39.5	11	3 520	56
16	2 032	38	12	3 840	43
17	2 159	39.5	13	4 160	41
18	2 286	37.5	14	4 480	34
19	2 413	38.5	15	4 800	33
20	2 540	38	16	5 120	----
			17	5 440	40
Noise 33			18	5 760	35
1	320	53	19	6 080	36
2	640	46.5	20	6 400	38
3	960	42	21	6 720	----
4	1 280	31	22	7 040	40
5	1 600	44	23	7 360	43.5
6	1 920	50	24	7 680	44.5
7	2 240	56	25	8 000	39
8	2 560	69	26	8 320	34
9	2 880	65	27	8 640	33
10	3 200	52	28	8 960	33

250-mel spacing and less than 4 phons for 500-mel spacing. For equally intense tones spaced 250 mels apart and centered at 1000 hertz the difference between theory and experiment is about 3 phons for listener J. M. and is negligible for listener D. H. For equally loud tones the average disagreement is about 3 phons for 250-mel spacing and 4 phons for 500-mel spacing.

Zwicker, Flottorp, and Stevens' experiment (1957).—Zwicker, Flottorp, and Stevens reported on data obtained by Scharf on the loudness levels of quintets of (assumed) uniformly spaced, equally intense tones as a function of the level of one component for earphone listening (ref. 38). In one case the overall spacing was 5000 hertz with a geometric-mean frequency of 1800 hertz, whereas the other case involved an overall spacing of 1800 hertz with a geometric-mean frequency of 1250 hertz. The judged and calculated results are compared in figure 32. Except at the lowest levels (<20 phons), the maximum absolute difference between the calculated and judged loudness levels of the quintets is less than 9 phons in the former case and 6 phons in the latter case.

Quietzsch's experiment (1955).—At least 20 listeners with normal hearing compared the loudness of a tone quintet with that of a 1-kilohertz tone under conditions of free-field, binaural listening in an anechoic chamber (ref. 60). The unequally intense tones consisted



(a) Geometric-mean frequency, $\omega_r/2\pi$, 1800 hertz; overall tone spacing, $\Delta\omega/2\pi$, 5000 hertz.

(b) Geometric-mean frequency, $\omega_r/2\pi$, 1250 hertz; overall tone spacing, $\Delta\omega/2\pi$, 1800 hertz.

FIGURE 32.—Calculated loudness level of equally intense tone quintets with supercritical overall spacing as function of loudness level of one tone. (Data from Zwicker, Flottorp, and Stevens (ref. 38).)

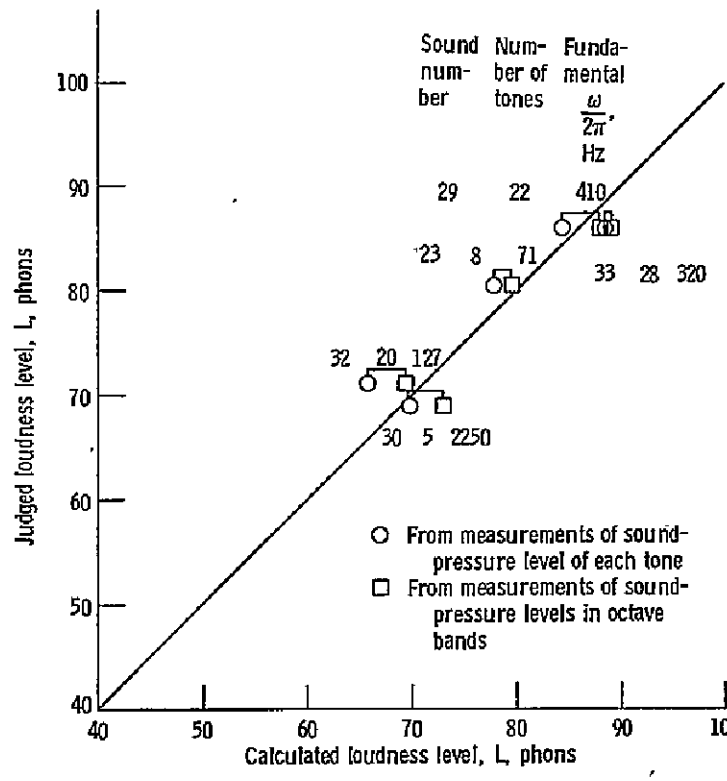


FIGURE 33.—Judged loudness level of multiple tones as function of calculated loudness level. (Data from Quietzsch (ref. 60).)

of an impulsive sound with a fundamental pulse rate of 2250 per second plus its first four harmonics. The frequencies and corresponding sound-pressure levels are listed in table VI(c). Actually, the sound-pressure levels were measured in two ways, namely by measuring the sound-pressure level of each tone and also by measuring the sound-pressure level of the quintet in octave bands. Of course, the former method is more appropriate, but the latter method may be more common, especially if it is not realized that the spectrum is discrete. As shown in figure 33, the disagreement between the judged and calculated loudness levels is less than 4 phons in both cases. As in other cases to follow, the levels calculated from octave-band data are greater (by as much as 4 phons) than those calculated from narrow-band data.

Six Tones

D. H. Howes' experiment (1950).—The loudness of the following sextets of equally intense (four trained listeners), and then equally loud (two trained listeners), tones centered at 1900 hertz was determined, as before, by monaural earphone listening (ref. 58):

Sextet	Separation, mels	$\omega_1/2\pi$, Hz	$\omega_2/2\pi$, Hz	$\omega_3/2\pi$, Hz	$\omega_4/2\pi$, Hz	$\omega_5/2\pi$, Hz	$\omega_6/2\pi$, Hz
1	250	[a](670)	1000	1420	1900	2450	3120
		1000	1420	1900	2450	3120	[a](4000)
2	500	610	670	1420	2450	4000	6600

[a] Frequencies in parentheses were alternated.

Judgments were also made for a sextet of equally intense tones centered at 1000 hertz and spaced 250 mels apart.

The various experimental results are listed in tables II, III, and IV and shown graphically in figure 34. For the tones centered at 1900 hertz the disagreement between the calculated and judged results is, for the equally intense tones, less than 4 phons for 250-mel separation and about 2 phons for 500-mel separation. In the case of equally intense tones centered at 1000 hertz, the disagreement is about 2.5 phons for listener J. M. and less than 0.5 phon for listener D. H. For the equally loud tones the disagreement is about 5 phons for 250-mel separation and about 7 phons for 500-mel separation.

Seven Tones

D. H. Howes' experiment (1950).—Only a septet of equally intense tones spaced in 250-mel increments about a 1000-hertz center frequency was considered (ref. 58). Experimental results and theo-

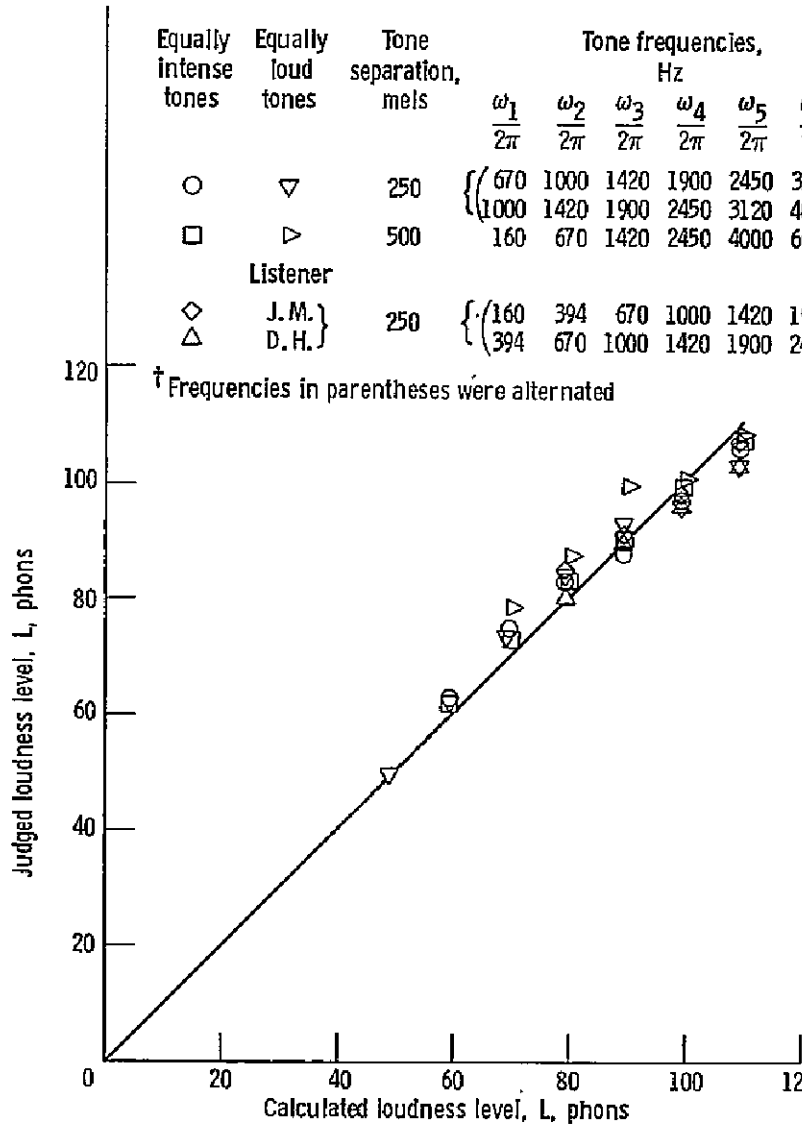


FIGURE 34.—Comparison of judged and calculated loudness levels for six tones. (Data from D. H. Howes (ref. 58).)

retical predictions for these tones are compared in table III and figure 35. For listener J. M. the average disagreement between theory and experiment is about 4 phons, whereas the corresponding difference is 1 phon for listener D. H.

Eight Tones

D. H. Howes' experiment (1950).—Only an octet of tones spaced 250 mels apart about a 1900-hertz center frequency was considered (ref. 58). The specific frequencies comprising the octet were (394),

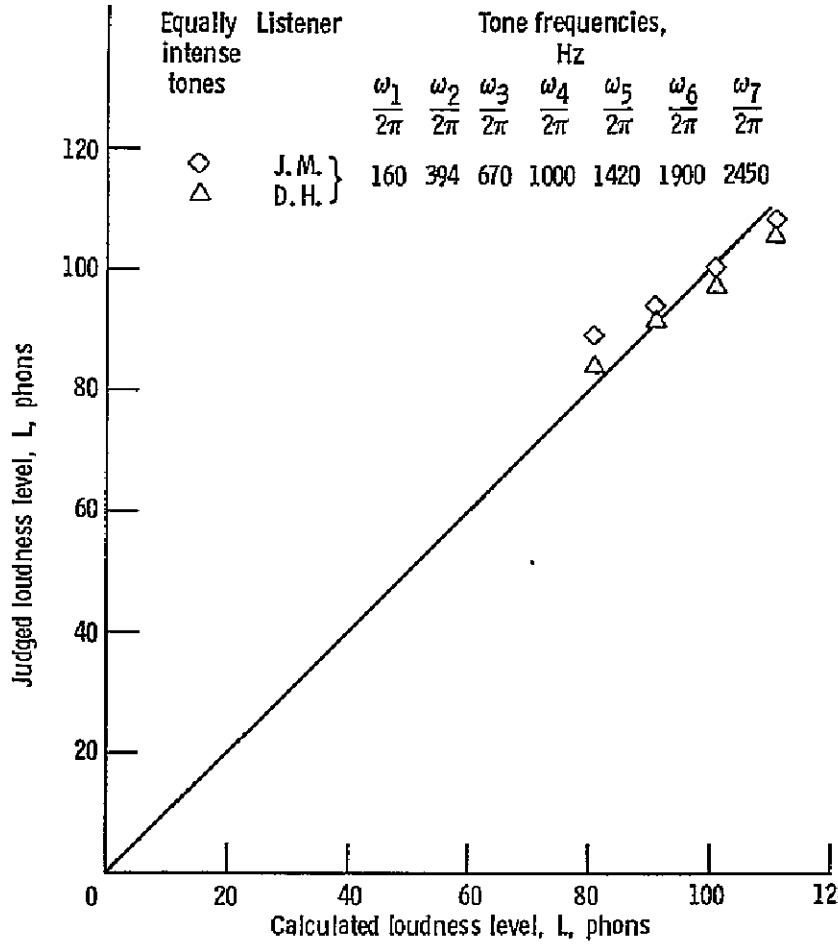


FIGURE 35.—Comparison of judged and calculated loudness levels for seven tones. (Data from D. H. Howes (ref. 58).)

670, 1000, 1420, 1900, 2450, 3120, 4000, and (5100) hertz, in which the frequencies in parentheses were alternated.

Judgment results and corresponding calculations are compared in tables II and IV and in figure 36 for octets of equally intense and equally loud tones, respectively. For both complexes the average difference between calculated and judged loudness levels is about 4.5 phons.

Quietzsch's experiment (1955).—As in the five-tone test, the loudness level of an octet of tones was judged under conditions of binaural listening in an anechoic chamber (ref. 60). The equally intense tones consisted of a 71-hertz fundamental plus seven harmonics $m=2^n$, where $n=1,2,3,\ldots$. The frequency and precise sound-pressure level (ave. $S=54.5$ dB) of each tone are listed in table VI. The tone levels were measured separately. Also, the level of the octet was measured in octave bands. As shown in figure 33, the judged and calculated (narrow band) loudness levels differ by about 3 phons. For the octave band measurements the disagreement is about 1 phon.

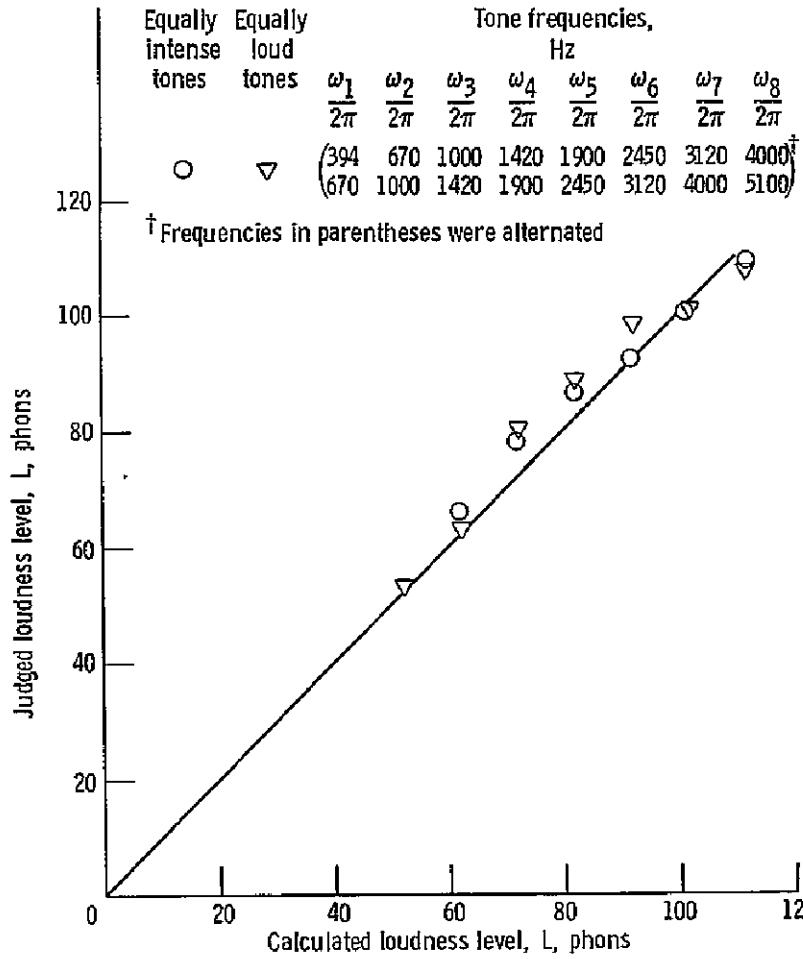


FIGURE 36.—Comparison of judged and calculated loudness levels for eight tones. Tone separation, 250 mels. (Data from D. H. Howes (ref. 58).)

For 16 arbitrary noises and 24 listeners, the standard deviation of the judged loudness levels was 2.5 phons, whereas the maximum deviation from the mean for any individual value of loudness level was 9.5 phons. Thus, the calculated value compares favorably with the test results.

Ten Tones

Fletcher and Munson's experiment (1933).—Eleven subjects listening binaurally through earphones judged loudness levels of ten tones by comparison with a 1-kilohertz tone (ref. 13). Data were reported for equally loud tones and for unequally loud tones with even numbered tones attenuated 5 decibels, and then 10 decibels, below the odd

numbered tones, where the lowest frequency tone is designated as number 1. The following tone combinations were studied:

Decuplet	Overall frequency range, $\Delta\omega/2\pi$, Hz	Tone spacing, $\delta\omega/2\pi$, Hz
Equally loud tones		
1	50 to 500	50
2	1400 to 1895	55
3	100 to 1000	100
4	3000 to 3900	100
5	1100 to 3170	230
6	260 to 2600	260
7	530 to 5300	530
L of even-numbered tones 5 dB less than L of odd-numbered tones		
8	1725 to 2220	55
L of even-numbered tones 10 dB less than L of odd-numbered tones		
9	1725 to 2220	55

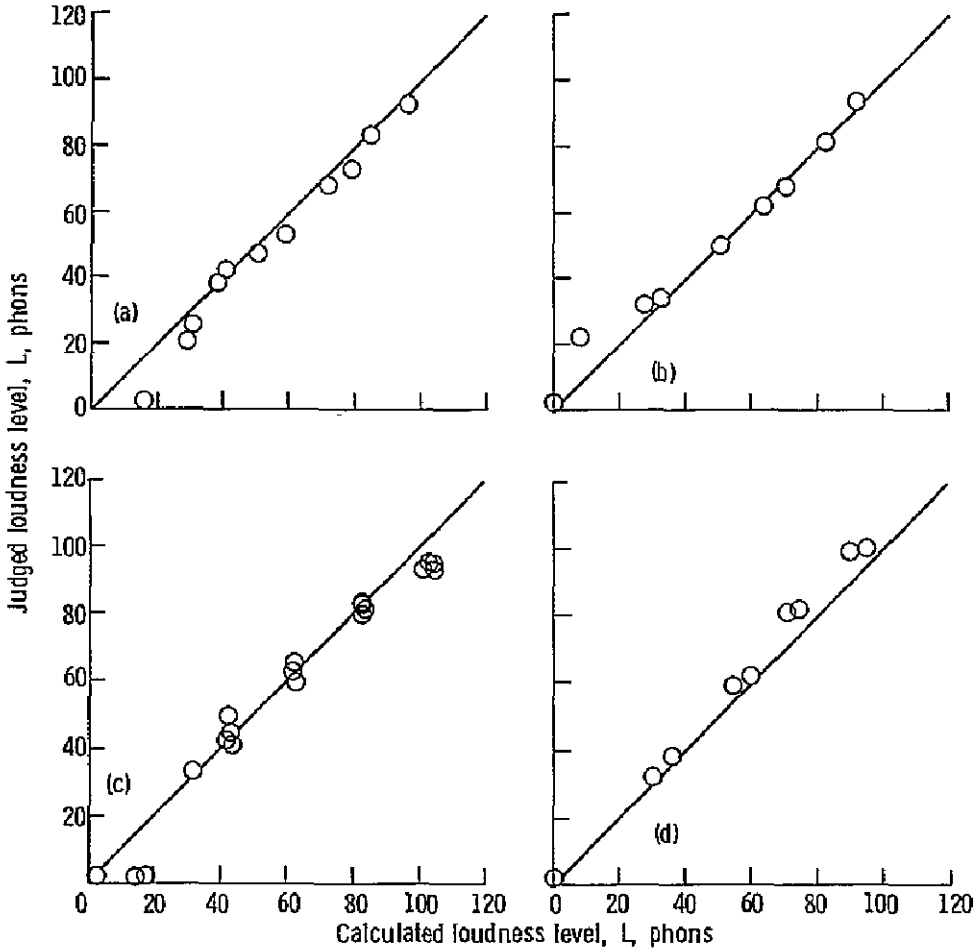
The judged loudness levels are shown graphically in figure 37 for comparison with the theoretical predictions.

In all examples the overall frequency range extended over at least two critical bands. Frequently two or more tones are included in one critical band. The only empty critical bands occurred with decuplet 6, for which $k=1$, and 7 for which $k=4$.

In order to predict the loudness of subcritically spaced tones the phase angles of the tones should be known at some precise time. Fletcher's experiment included subcritically spaced tones; but the phase angles were not determined. In the absence of these data and because of the large number of tones comprising the complex, it is sufficient to assume that the phase angles were uniformly distributed in the interval 0 to 2π radians. In equation (33) the expectation of $w(\omega_m,\omega_n)$ for subcritically spaced tones is then zero if $\omega_m \neq \omega_n$, but remains unity if $\omega_m=\omega_n$ (so that $\phi_m=\phi_n$). Hence, equation (33) reduces to

$$\mathcal{L}=10^{-2}\left[\sum_{n=1}^{N}\mathcal{Q}_r(\omega_n)w(\omega_n)\right]^{1/2}-10^{-2}(\mathcal{R}_{qt})^{1/2}$$

if the phase angles of the tones are uniformly distributed.



(a) $\omega_1/2\pi$, 50 hertz; $\omega_{10}/2\pi$, 500 hertz; $\delta\omega/2\pi$, 50 hertz.

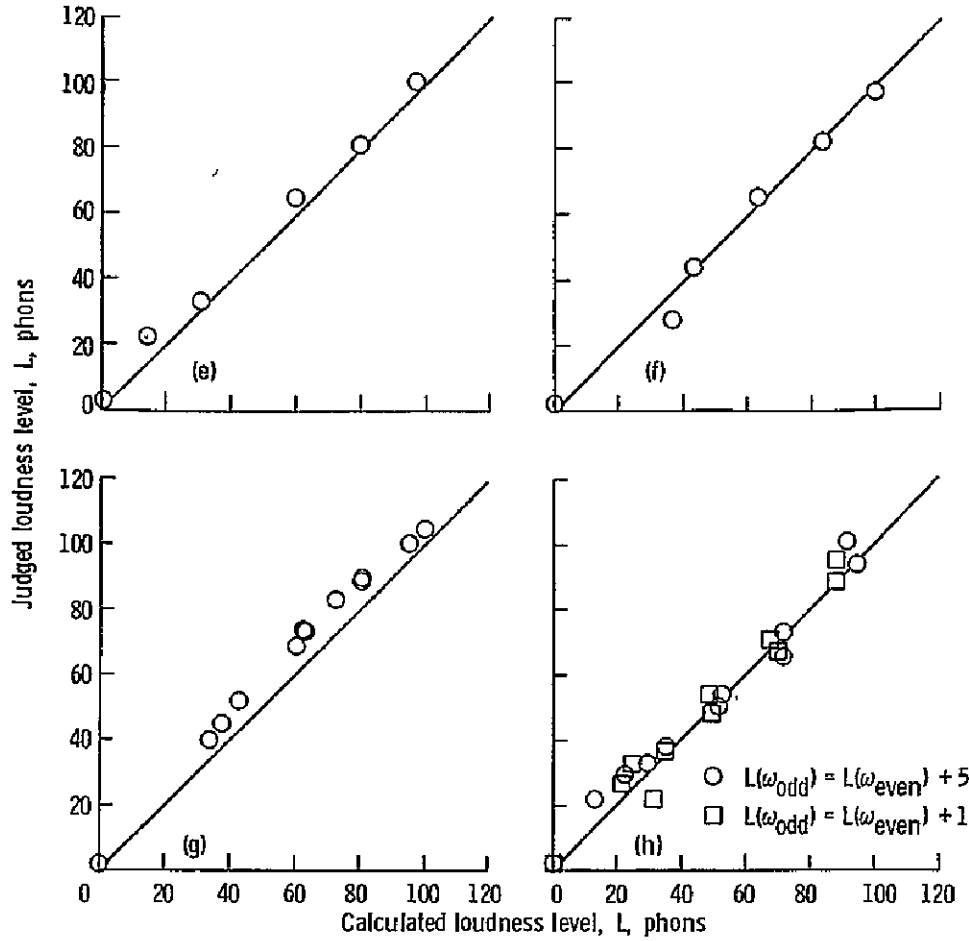
(b) $\omega_1/2\pi$, 1400 hertz; $\omega_{10}/2\pi$, 1895 hertz; $\delta\omega/2\pi$, 55 hertz.

(c) $\omega_1/2\pi$, 100 hertz; $\omega_{10}/2\pi$, 1000 hertz; $\delta\omega/2\pi$, 100 hertz.

(d) $\omega_1/2\pi$, 3000 hertz; $\omega_{10}/2\pi$, 3900 hertz; $\delta\omega/2\pi$, 100 hertz.

FIGURE 37.—Comparison of judged and calculated loudness levels for ten tones. Tones equally loud, except in (b). (Data from Fletcher and Munson (ref. 13).)

The probable error of the experimental overall loudness levels was claimed to be approximately ±2 decibels. However, as is evident from the data in reference 13, for decuplet 8, the variation of judged loudness for a given selected loudness was as great as 7 decibels between different tests involving the same set of tones. For suprathreshold, overall-loudness levels the median loudness-level increase ΔL was always within 1 decibel of the average increase. The average absolute difference between the calculated and judged loudness levels varies from about 1 to 6 phons for the various complexes. The difference between calculated and judged levels does not appear to be systematic and appears to be approximately the same magnitude as that



(e) $\omega_1/2\pi$, 1100 hertz; $\omega_{10}/2\pi$, 3170 hertz; $\delta\omega/2\pi$, 230 hertz.

(f) $\omega_1/2\pi$, 260 hertz; $\omega_{10}/2\pi$, 2600 hertz; $\delta\omega/2\pi$, 260 hertz.

(g) $\omega_1/2\pi$, 530 hertz; $\omega_{10}/2\pi$, 5300 hertz; $\delta\omega/2\pi$, 530 hertz.

(h) $\omega_1/2\pi$, 1725 hertz; $\omega_{10}/2\pi$, 2220 hertz; $\delta\omega/2\pi$, 55 hertz.

FIGURE 37.—Concluded.

obtained using Fletcher and Munson's empirical formula, which was derived to fit these specific data.

Eleven Tones

D. H. Howes' experiment (1950).—Eleven tones spaced 250 mels apart about a 1900-hertz center frequency were considered (ref. 58). The specific frequencies comprising this complex were 160, 394, 670, 1000, 1420, 1900, 2450, 3120, 4000, 5100, and 6600 hertz.

The experimental and calculated loudness levels are compared in tables II and IV and figure 38. For the complex of equally intense tones the average difference between theory and experiment is about 4 phons whereas for the equally loud tone complex the average difference is about 9 phons.

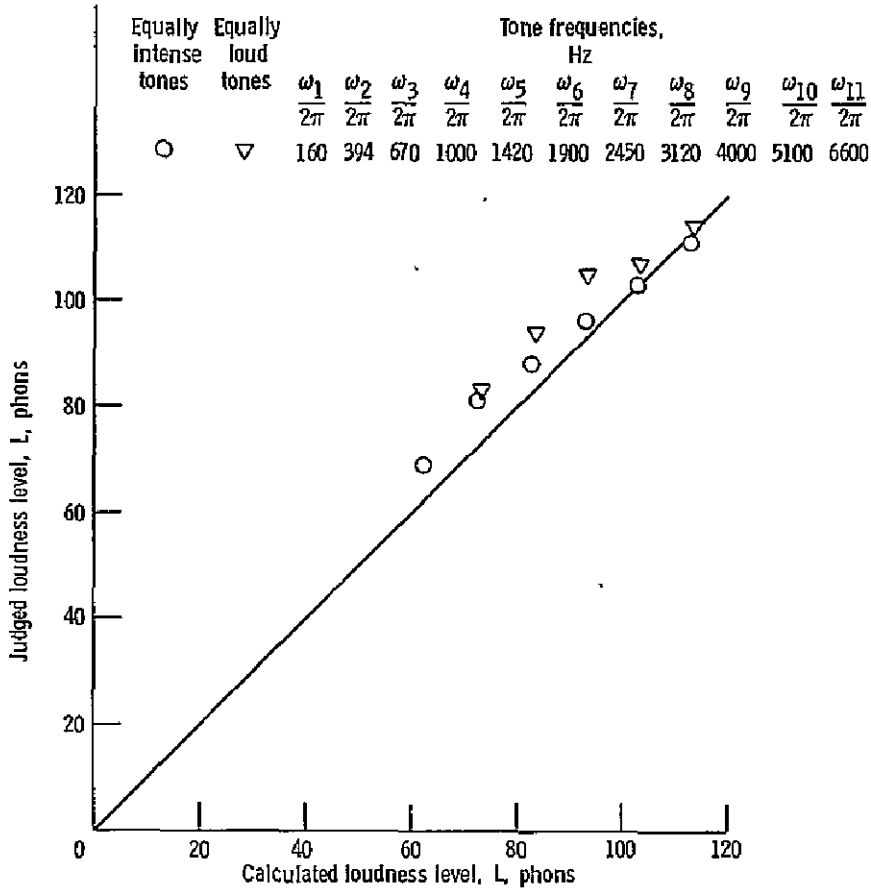
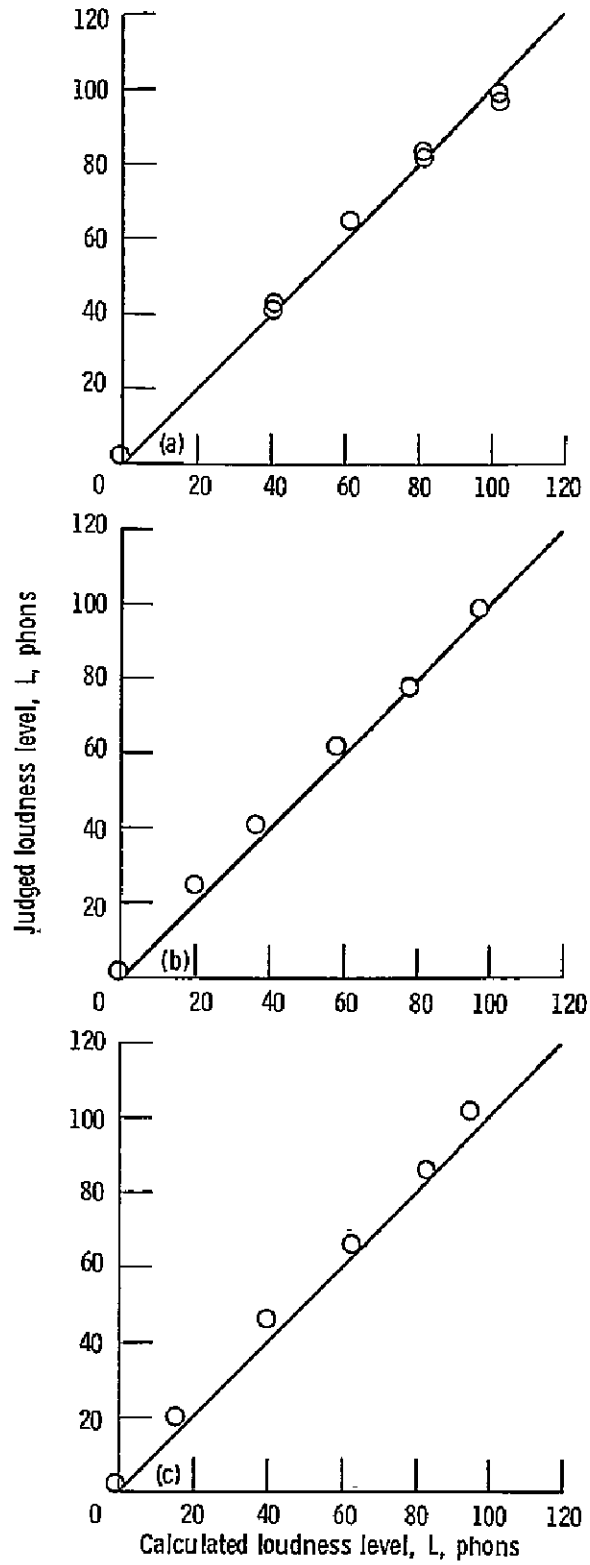


FIGURE 38.—Comparison of judged and calculated loudness levels for eleven tones. Tone separation, 250 mels. (Data from D. H. Howes (ref. 58).)

Fletcher and Munson's experiment (1933).—The loudness of the following eleven-tone complexes was judged by eleven subjects binaurally through earphones (ref. 13):

Undecuplet	Overall frequency range, $\Delta\omega/2\pi$, Hz	Tone spacing, $\delta\omega/2\pi$, Hz
Equally loud tones		
1	1000 to 2000	100
2	1150 to 2270	112
3	1120 to 4520	340
L of even-numbered tones 5 dB less than L of odd-numbered tones		
4	57 to 627	57
5	3420 to 4020	60
L of even-numbered tones 10 dB less than L of odd-numbered tones		
6	57 to 627	57
7	3420 to 4020	60



(a) $\omega_1/2\pi$, 1000 hertz; $\omega_{II}/2\pi$, 2000 hertz; $\delta\omega/2\pi$, 100 hertz

(b) $\omega_1/2\pi$, 1150 hertz; $\omega_{II}/2\pi$, 2270 hertz; $\delta\omega/2\pi$, 112 hertz.

(c) $\omega_1/2\pi$, 1120 hertz; $\omega_{II}/2\pi$, 4520 hertz; $\delta\omega/2\pi$, 340 hertz.

FIGURE 39.—Comparison of judged and calculated loudness levels for eleven equally loud tones. (Data from Fletcher and Munson (ref. 13).)

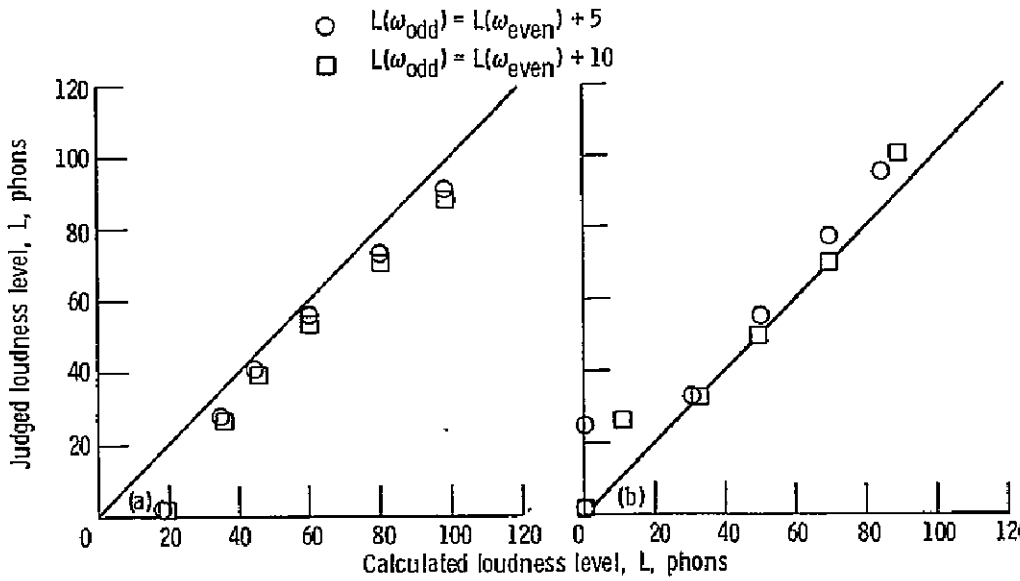
The experimental results and theoretical predictions are compared graphically in figures 39 and 40. There were no empty critical bands.

As in the case of the 10 tone data, the probable error of the experimental overall loudness levels was claimed to be approximately ±2 decibels. For suprathreshold overall-loudness levels the median loudness-level increase ΔL was within 1 phon of the average increase, except with undecuplet 7 where the median is about 3.5 phons less than the average value. The divergence between the median and average is accounted for by one test result. Undecuplet 7 happens to be one for which the overall bandwidth is subcritical. For the tone complex in this case, and also with undecuplet 5, the loudness level should be given approximately by the sound-pressure level corresponding to summation of the relative powers over all tones. Where the agreement between theory and experiment is poor, there is an implied inconsistency between the data of Fletcher and Munson and those of Zwicker and Feldtkeller (ref. 27), which were used to establish critical bandwidths.

For suprathreshold tests the average absolute difference between the calculated and judged loudness levels varies from about 2 to 9 phons. The difference between theoretical and experimental results does not appear to be systematic.

Twenty Tones or More

Quietzsch's experiment (1955).—Under conditions previously cited, the loudness levels of 20, 22, and 28 tones were judged (ref. 60). The



(a) $\omega_1/2\pi$, 57 hertz; $\omega_{11}/2\pi$, 627 hertz; $\delta\omega/2\pi$, 57 hertz.

(b) $\omega_1/2\pi$, 3420 hertz; $\omega_{11}/2\pi$, 4020 hertz; $\delta\omega/2\pi$, 60 hertz.

FIGURE 40.—Comparison of judged and calculated loudness levels for eleven tones. (Data from Fletcher and Munson (ref. 13).)

20 tones consisted of a 127-hertz fundamental and its first 19 harmonics, all at the same sound-pressure level (approx. 39 dB) and in random phase. The 22 tones, generated by an automobile horn, had a 410-hertz fundamental plus the first 21 harmonics, all at different sound-pressure levels listed in table VI. The 28 tones, generated by another type of automobile horn, consisted of a 320-hertz fundamental plus its first 27 harmonics, all at different sound-pressure levels listed in table VI(c). The tone sound-pressure levels were measured separately and in octave bands.

For the 20-tone stimulus, the disagreement between the judged and calculated loudness levels shown in figure 33 is 5.5 phons. The difference for the complex measured in octave bands is about 1.5 phons. For the 22-tone stimulus, the corresponding differences are 1.5 phons and 2 phons, respectively. For the 28-tone stimulus, the disagreements are 2.5 phons for the tone measurements and 3 phons for the octave measurements. As before, these results compare favorably with the standard deviation of 2.5 phons among 16 arbitrary noises and a maximum deviation from the mean of 9.5 phons for any individual value.

Loudness as Function of Number of Tones Contained Within Fixed Frequency Band

Scharf's experiment (1959).—Scharf studied the loudness of tone complexes as a function of the number of tones with the overall bandwidth, center frequency, and overall sound pressure held constant (ref. 61). All tones were equally intense and located in the flat frequency response range of the auditory system. The loudness of the complexes was determined by comparison with a 1500-hertz tone using eight subjects listening binaurally through earphones. Complexes of 2, 3, 4, and 8 uniformly spaced tones centered at 1500 hertz were studied for overall bandwidths of either 175 hertz (subcritical) or 1600 hertz (supercritical) at nominal sound-pressure levels of 25, 50, 75, and 90 decibels.

It was found that, for subcritical bandwidth, the sound-pressure levels of the complex and comparison tone were equal for equal loudness, independent of the number of tones in the complex. For supercritical bandwidth the sound-pressure level of the comparison tone exceeded (by up to 11 dB, depending on the selected nominal sound-pressure level) that of the complex when the loudnesses were equal. Again, the result was independent of the number of tones in the complex.

Similar results were obtained for random noise. Specifically, the loudness of a pair and a quartet of equally intense tones with an overall spacing of 3400 hertz centered at 2600 hertz essentially

C-2

equalled that of a white noise with similar bandwidth and center frequency when the overall sound pressure of the three complexes was the same.

The preceding results for subcritical and supercritical complexes are consistent with equations (33) and (40) (also see category AB for subcritically spaced tones spread over a supercritical band), respectively, which indicate that loudness depends on input power in each critical band and not on the number of tones producing that power. The condition of constant input power in each critical band was satisfied in the experiment because the tones were equally intense and uniformly spaced over an interval of uniform frequency response, and the overall sound-pressure level was held constant.

LOUDNESS OF SPECIFIC SOUNDS WITH CONTINUOUS SPECTRA (COMPARISON OF MODEL WITH EXPERIMENT)

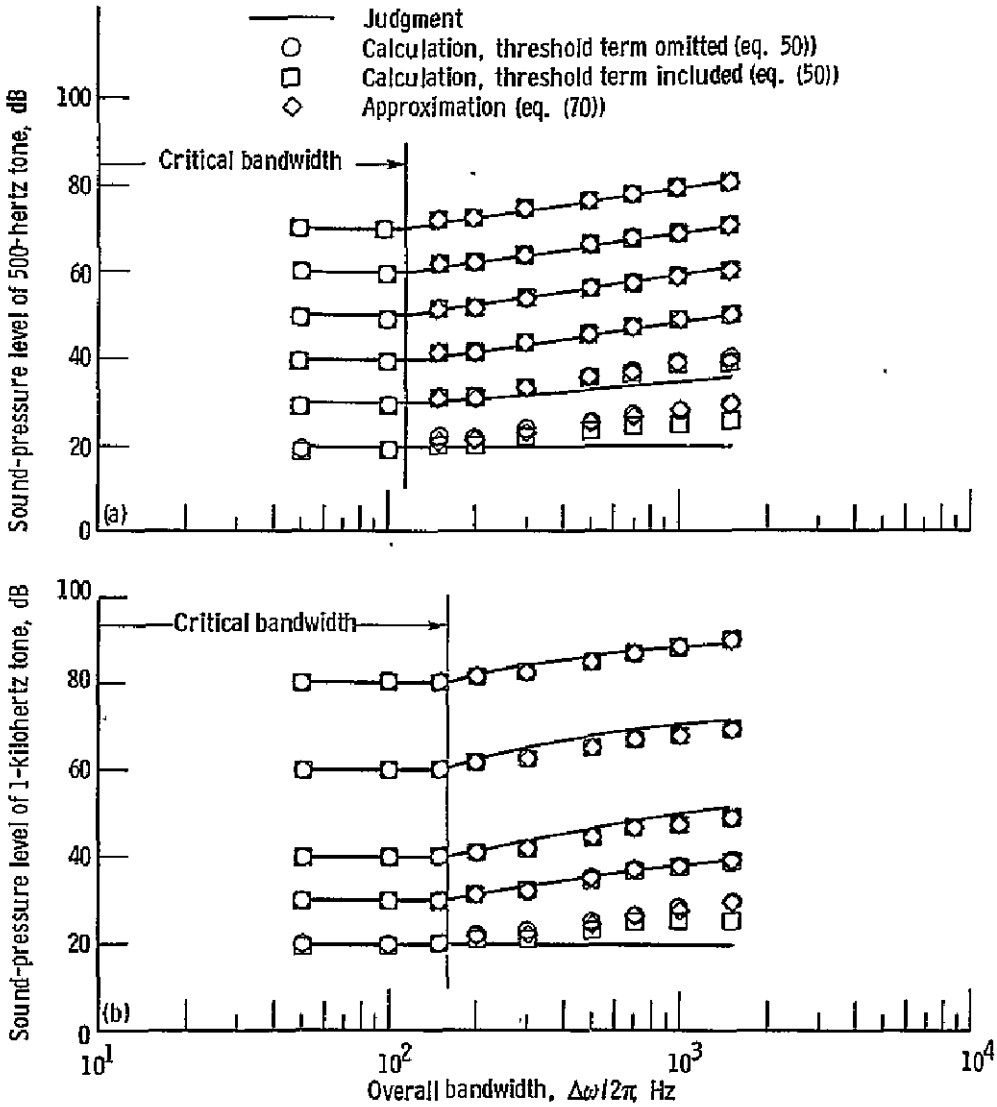
In this section the loudness of specific sounds having continuous spectra is calculated and compared with psychoacoustic data reported by various investigators. Sounds with continuous spectra are here presumed to include no tones. In fact, however, tones may be present but are not evident either because the bandwidth of the spectrum analyzer was too broad to resolve them or because their spectral densities were less than that of the continuous spectrum.

Effect of Bandwidth

Zwicker and Feldtkeller's experiment (1955).—The loudness level of band-limited white noise was judged by comparison with a 1-kilohertz tone as a function of bandwidth with the overall sound-pressure level held constant (ref. 27). Judgments were made by several (possibly as many as seven) listeners utilizing monaural earphone listening. Judgments were made for several sound-pressure levels and band geometric-mean frequencies of 500, 1000, 2000, and 4000 hertz. Both subcritical and supercritical bandwidths were considered. The judged and calculated (using eq. (50)) loudness levels are compared as a function of bandwidth in figure 41. Among all the data the maximum increase in loudness level for supercritical bandwidth exceeded the subcritical value by about 14 phons, although the sound-pressure level was held constant. However, near the loudness threshold, judged loudness tended to be independent of the noise bandwidth. The agreement between the judged and calculated loudness levels is generally excellent. The difference between judged and calculated levels exceeds 5 phons in only one instance. Levels calculated with and without the inclusion of the threshold term show

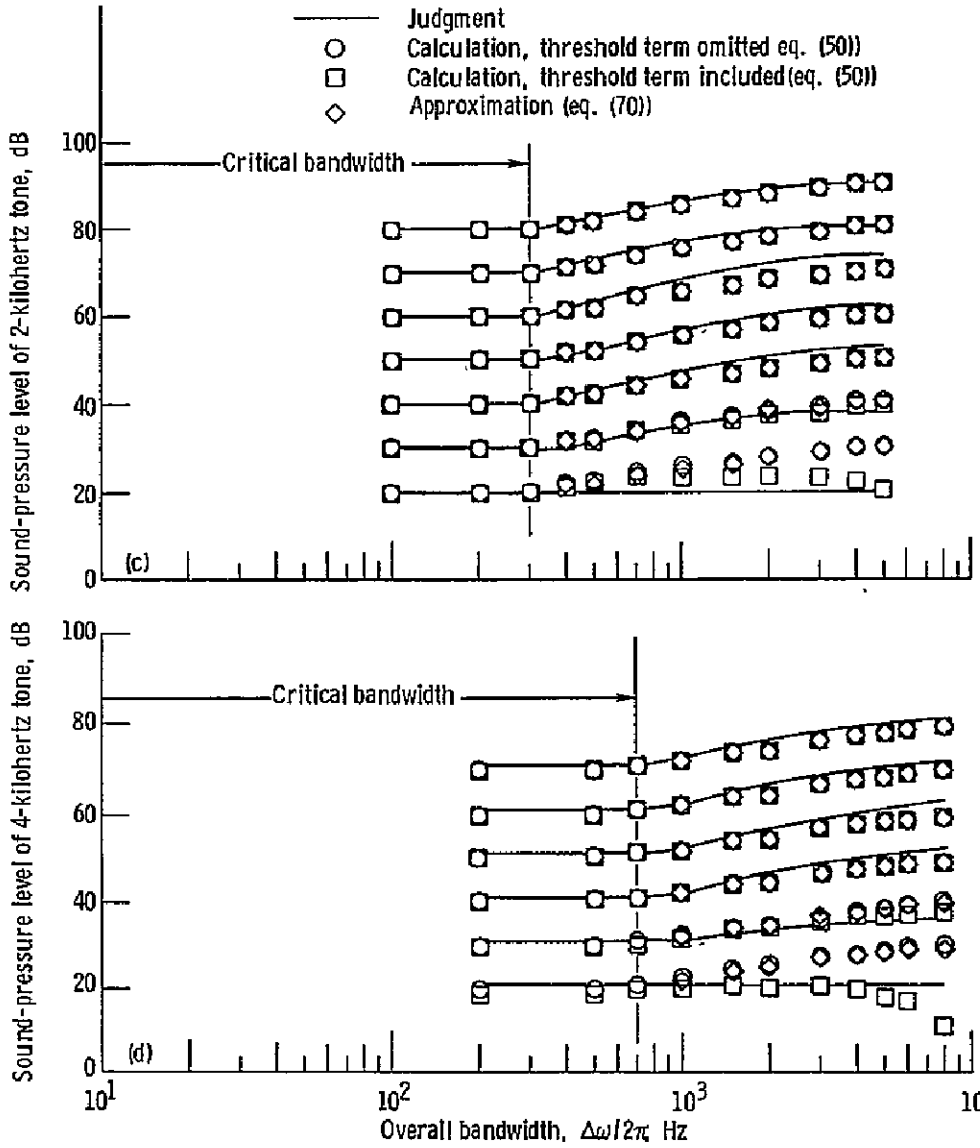
that the threshold term is significant only for loudness levels less than 40 phons and nearly always results in improved agreement with the judged loudness level. Levels calculated using the approximation equation (70) are in good agreement with those calculated from equation (50), except near the loudness threshold when equation (50) clearly yields better agreement with judgment data. However, even near the loudness threshold, equation (70) yields slightly better agreement with judgment data than does equation (50) with the threshold term omitted.

Niese's experiment (1961).—In comparing the loudness of band-limited white noise by monaural and binaural listening through earphones, loudness judgments as a function of bandwidth were recorded



(a) Noise geometric-mean frequency $\omega_r/2\pi$, 500 hertz.
(b) Noise geometric-mean frequency $\omega_r/2\pi$, 1000 hertz.

FIGURE 41.—Loudness level of band-limited, white noise as function of bandwidth. (Data from Zwicker and Feldtkeller (ref. 27).)



(c) Noise geometric-mean frequency $\omega_r/2\pi$, 2000 hertz.
(d) Noise geometric-mean frequency $\omega_r/2\pi$, 4000 hertz.
FIGURE 41.—Concluded.

for both listening modes at a fixed overall sound-pressure level of 60 decibels (ref. 62). Disregarding the judged loudness difference (which was not conclusive) between the two modes, the judged and calculated loudness levels are compared in figure 42 as a function of the noise bandwidth. The maximum disagreement is 5 phons.

Zwicker, Flottorp, and Stevens' experiment (1957).—The experiment of Zwicker and Feldtkeller on loudness determination as a function of bandwidth at constant overall sound pressure was repeated for bands of noise with geometric-mean frequencies of 440, 1420, and 5200 hertz for binaural earphone listening by 12 listeners (ref. 38). However, loudness comparisons were made with a

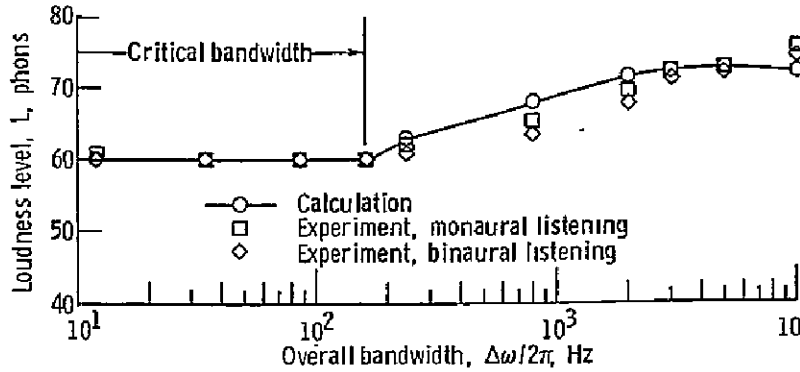
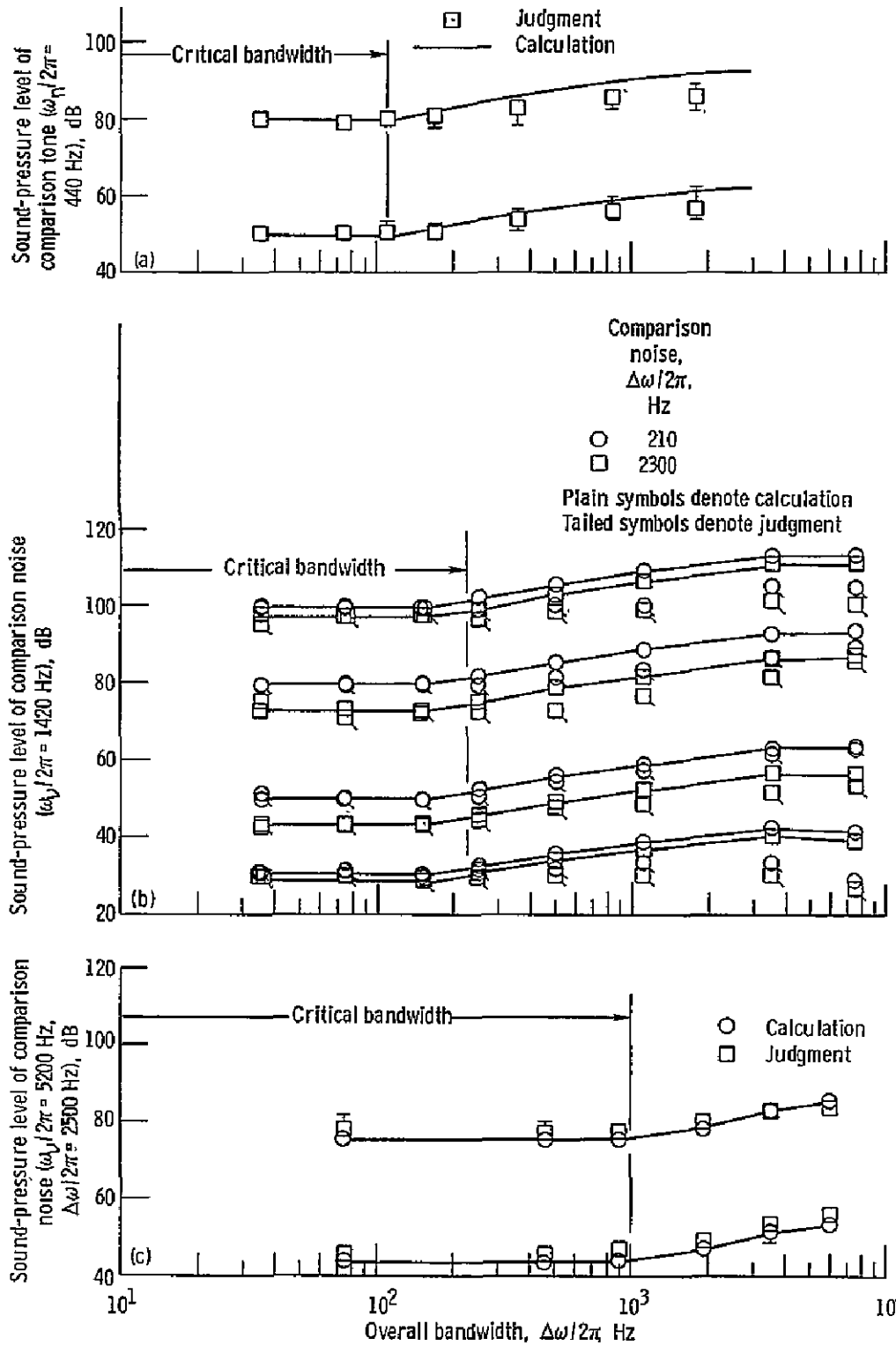


FIGURE 42.—Loudness level of band-limited, white noise at constant overall sound pressure as function of bandwidth. Geometric-mean frequency, $\omega_r/2\pi$, 1000 hertz. (Data from Niese (ref. 62).)

440-hertz tone or bands of noise rather than with a 1-kilohertz tone. Thus the results shown in figure 43 indicate the variation of loudness level as a function of bandwidth but not necessarily the absolute loudness levels. Although the variation of loudness level as a function of bandwidth is qualitatively similar to that found by Zwicker and Feldtkeller, the agreement between judgments and predictions for supercritical bandwidths is not quite as good as it was in that former experiment. The disagreement between theory and experiment is greatest at the maximum bandwidth and for the highest and lowest sound pressures. The judged levels often tend to be less than the predicted values whereas in the case of Zwicker and Feldtkeller's results there is no consistent systematic difference. This difference is especially important when comparing Zwicker and Feldtkeller's results for geometric-mean frequencies of 1000 and 2000 hertz with those of Zwicker, Flottorp, and Stevens for 1420 hertz. Clearly the results of these two experiments are sufficiently inconsistent, say by up to 4 or 5 phons, that the disagreement between theory and experiment cannot be regarded as systematic, except possibly at very low and very high sound pressures.

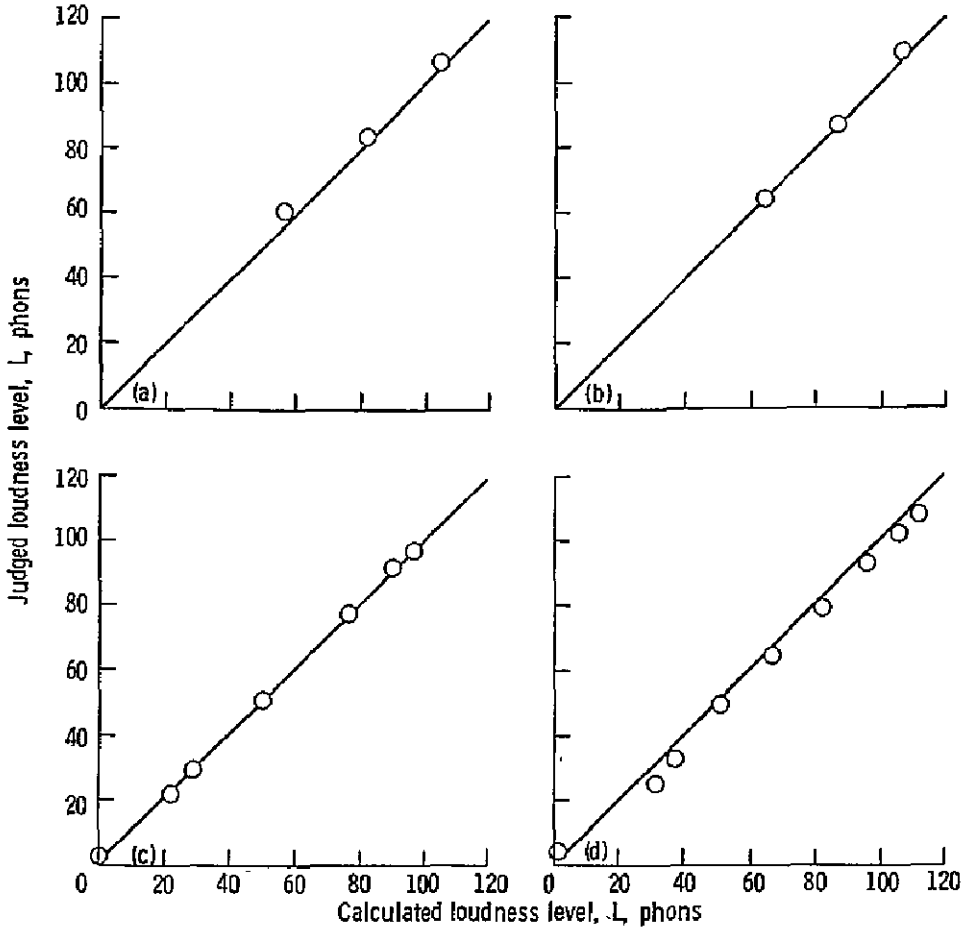
Octave Bands

Robinson and Whittle's experiment (1964).—The loudness of octave bands of noise was judged by free-field exposure to plane waves incident from the front in an anechoic chamber (ref. 63). The test data were modified by the authors to correspond to diffuse waves. The number of listeners varied from 20 to 30 with normal hearing. Judgments were made as a function of sound-pressure level for eight



(a) Geometric-mean frequency, $\omega_\nu/2\pi$, 440 hertz.
(b) Geometric-mean frequency, $\omega_\nu/2\pi$, 1420 hertz.
(c) Geometric-mean frequency, $\omega_\nu/2\pi$, 5200 hertz.

FIGURE 43.—Loudness level of band-limited, white noise as function of bandwidth. (Data from Zwicker, Flottorp, and Stevens (ref. 38).)

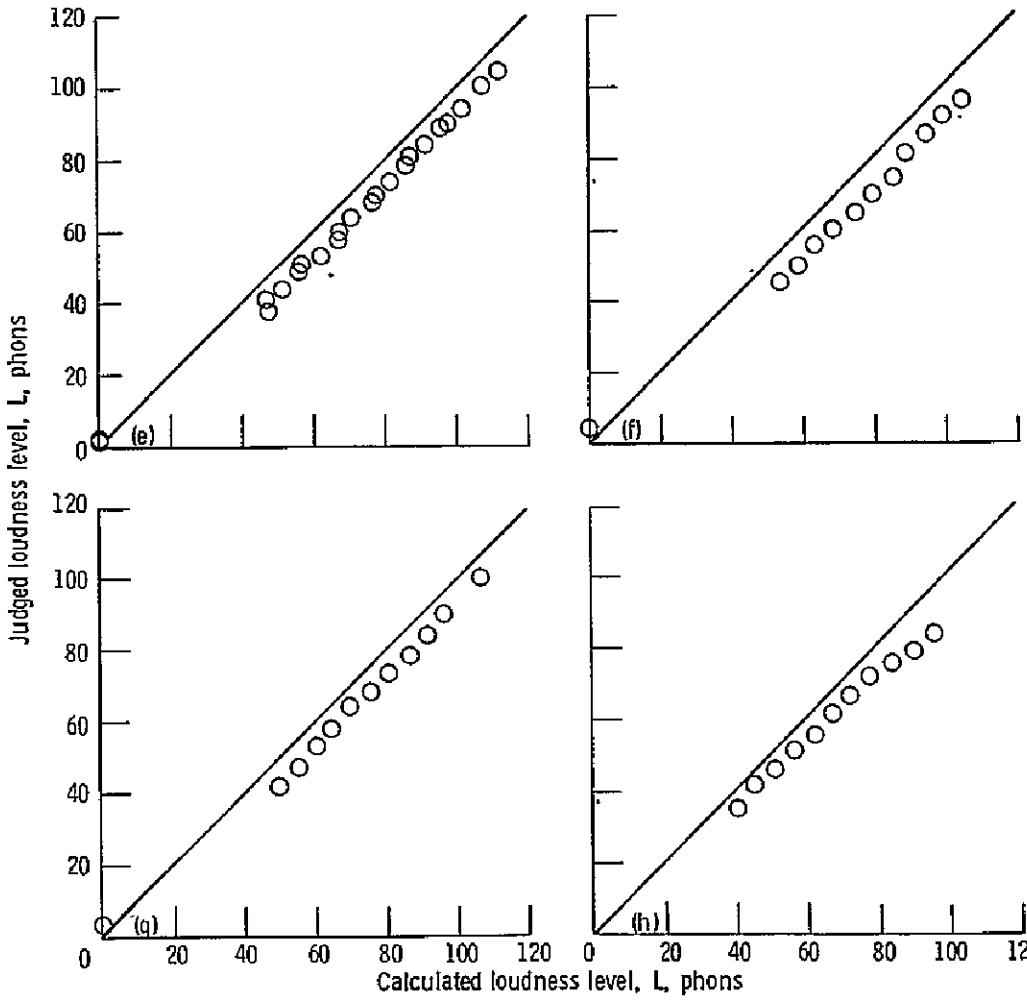


(a) Geometric-mean frequency, $\omega_\nu/2\pi$, 58 hertz.
(b) Geometric-mean frequency, $\omega_\nu/2\pi$, 111 hertz.
(c) Geometric-mean frequency, $\omega_\nu/2\pi$, 274 hertz.
(d) Geometric-mean frequency, $\omega_\nu/2\pi$, 564 hertz.

FIGURE 44.—Comparison of judged and calculated loudness levels for octave bands of noise. (Data from Robinson and Whittle (ref. 63).)

octave bands with geometric-mean frequencies ranging from 58 to 8000 hertz.

The judged and calculated loudness levels are compared in figure 44. For the octaves centered at 58, 111, 274 (the value in the original text appears to be a misprint) and 3750 hertz the agreement between judged and calculated levels is excellent with differences no more than 2 or 3 phons. For the other octaves there is a systematic disagreement, as great as 9 phons for the octave centered at 1950 hertz. Among other experiments a similar disagreement does not appear. Thus, it seems that the disagreement may result from a systematic experimental error. In particular, for the octave centered at 940 hertz the sound-pressure level is only about 1 decibel less than that associated with an equally loud 940 hertz tone. Since the correspond-



(e) Geometric-mean frequency, $\omega_\nu/2\pi$, 940 hertz.

(f) Geometric-mean frequency, $\omega_\nu/2\pi$, 1950 hertz.

(g) Geometric-mean frequency, $\omega_\nu/2\pi$, 3750 hertz.

(h) Geometric-mean frequency, $\omega_\nu/2\pi$, 8000 hertz.

FIGURE 44.—Concluded.

ing octave bandwidth is 665 hertz, whereas the critical bandwidth is only 160 hertz, a considerably greater difference in sound-pressure levels for equal loudness should be expected. This further reinforces the suspicion that the observed loudness-level differences are attributable to experimental error.

Cremer, Plenge, and Schwarzl's experiment (1959).—A total of 41 mostly naive listeners with normal hearing determined the loudness of octave bands of noise by comparison with a one-third-octave noise centered at 1 kilohertz (ref. 64). Listening was in the free field in a reverberant chamber. Thus, the sounds were diffuse. Seven noises, each one octave in bandwidth, the lowest with a lower cutoff frequency of 100 hertz and the highest with an upper cutoff frequency of 12 800 hertz, were studied. Loudness judgments were made

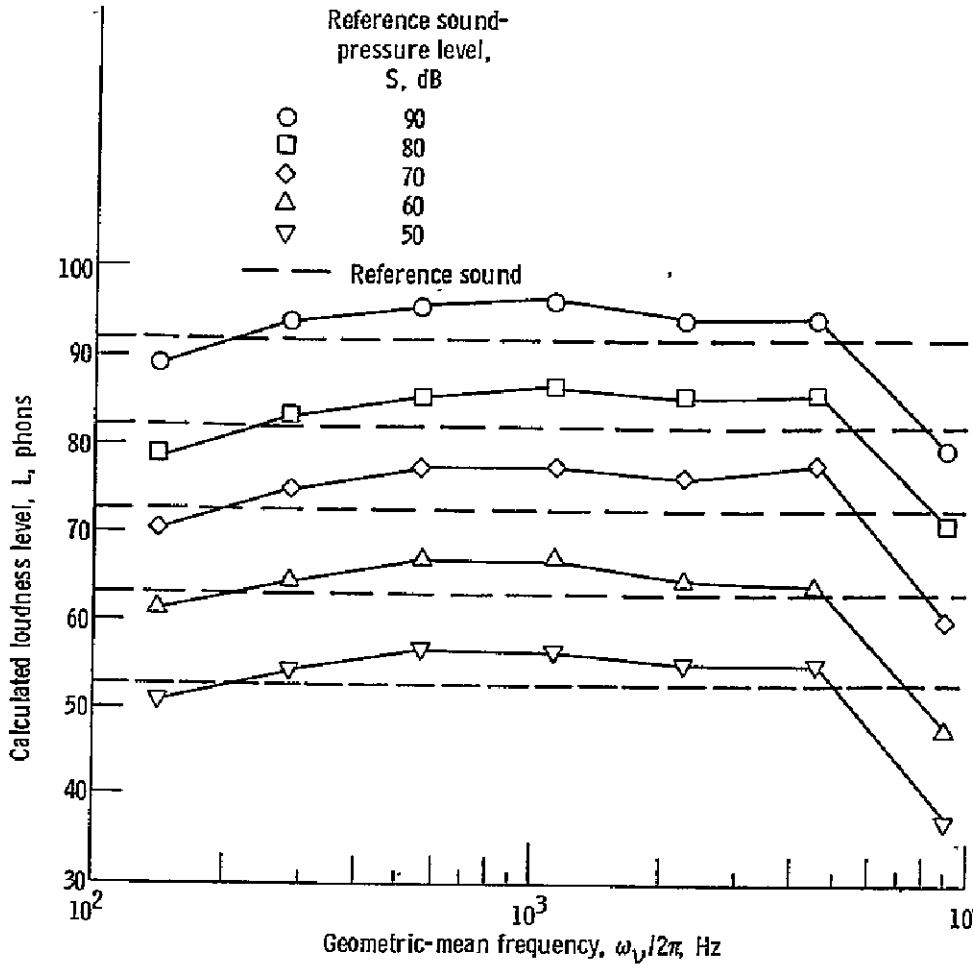


FIGURE 45.—Calculated loudness level of judged equally loud octave bands of noise as function of geometric-mean frequency of octave. (Data from Cremer, Plenge, and Schwarzl (ref. 64).)

and sound-pressure levels recorded for one-third-octave reference noise levels from 50 to 90 decibels in 10 decibel steps.

The calculated loudness level of each octave noise at each indicated overall-sound-pressure level is shown in figure 45. For each given reference level the corresponding calculated loudness levels of the octave noises should be the same for each center frequency since they were adjusted for equal judged loudness. The calculated loudness levels of the one-third-octave reference sounds are also shown in figure 45. (Actually, these values were estimated from equivalent calculations performed in connection with Zwicker and Feldtkeller's data shown in fig. 41.) The authors estimated a standard deviation of about 3 decibels among the loudness judgments. With the exception of the calculations for the highest octave, all calculated loudness levels corresponding to a given loudness lie within a maximum total range of 8 phons with a maximum disagreement with the reference loudness level of 5 phons. All calculations for the highest octave yield

a low level, as much as 15.5 phons less than the reference level. Whether this disagreement results from experimental or theoretical errors is not completely clear, although the fact that the empirical external transmittance level is approximately 2 decibels less than the measured value at 9 kilohertz (fig. 9) accounts for part of the discrepancy.

G. Jahn's experiment (1959).—The loudness of noises with octave bandwidths was equated to that of a one-third-octave band of noise having a geometric-mean frequency of 1 kilohertz (ref. 65). Judgments were performed in the free field in a reverberant chamber, so that the sound was diffuse. Judgments were made for fixed intensity of the one-third-octave reference noise corresponding to measured loudness levels of 40 and 55 phons. For each loudness level, comparisons with octave noises having geometric-mean frequencies ranging from 100 to 12 800 hertz were made. In the former case, 10 listeners were used; whereas in the latter case, six listeners were used.

Preliminary tests indicated that, for equal loudness, the sound-pressure level of the reference one-third-octave noise should be 2 decibels less than that of the 1-kilohertz tone. Also, for equal loudness at 1 kilohertz, the sound-pressure level for diffuse waves was found to be 3 decibels less than for plane waves from the front. (The sign

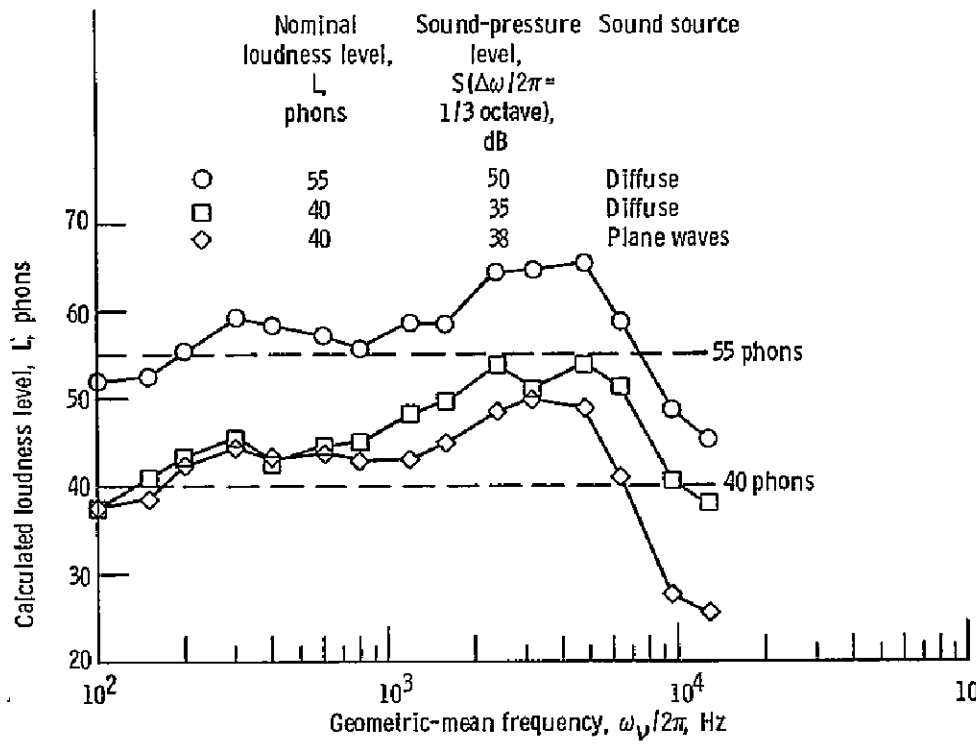


FIGURE 46.—Calculated loudness level of judged equally loud octave bands of noise as function of geometric-mean frequency of octave. (Data from G. Jahn (ref. 65).)

of this difference is in opposition to that reported in ref. 46 and shown in fig. 9 herein.) These results were used to establish the loudness levels of the reference noises. The measured sound-pressure levels of the octave noises judged equally loud were used to calculate the loudness levels of the octave noises. For a given reference level the calculated loudness levels for all octaves should be the same. However, as shown in figure 46, this is not the case, except in the interval 200 to 2000 hertz. G. Jahn's data cover the same frequency range as those of Cremer, Plenge, and Schwarzl (ref. 64), which do yield a relatively constant loudness level, except at 9 kilohertz. Hence, the two sets of data are inconsistent, except near 10 kilohertz, where a similar falloff in calculated loudness level is obtained. However, both of these results tend to disagree with those of Robinson and Whittle at 8 kilohertz, where the calculated loudness level is greater, not less, than the judged level.

If Robinson and Whittle's value (−2.5 dB at 1000 Hz) of the external-transmittance-level difference between diffuse and plane waves replaces the value (3 dB) assumed by G. Jahn, then the nominal levels (40 and 55 phons) shown in figure 46 must be raised by 5.5 phons. Consequently, the judged octave levels are in much better agreement with the nominal values over a much greater frequency range.

Stevens' experiment (1956).—Twelve listeners equated the loudness of octave bands of noise to a variety of reference, band-limited noises (ref. 5). Sometimes the octave noise was adjusted, and other times the reference noise was adjusted. Judgments were made in a diffuse sound field. The fixed sound was set at a nominal loudness level of 72 phons. Octave cutoff frequencies ranged from less than 75 hertz to 9600 hertz.

As shown in figure 47 the dependence of loudness on center frequency appears similar to that obtained by G. Jahn, and the effect of the choice of fixed sound is negligible in comparison with the frequency dependence. The similarity of the two frequency dependences is probably fortuitous since the maximum predicted loudness occurs at 1700-hertz center frequency according to Stevens' data and in the interval 3 to 5 kilohertz according to G. Jahn's data. Stevens' data exhibit good reproducibility but poor agreement with predicted loudness levels. Consideration of other data reported by Stevens in the same report suggests that this poor agreement might be expected. For example, semi-interquartile ranges up to 7 decibels and standard deviations up to 9 decibels were reported. Thus, the scatter of the data was relatively large. Also, it is reported that, for equal loudness, the sound-pressure levels of an octave band of noise with geometric-mean frequency at 850 hertz (cutoff frequencies, 600 and 1200 Hz) and a 1000-hertz tone differ by only 1 decibel.

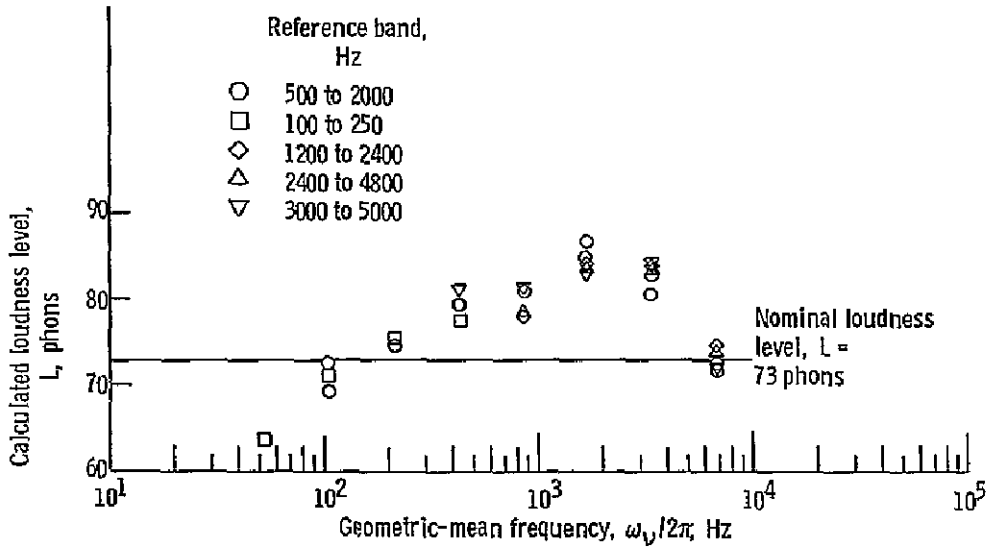


FIGURE 47.—Calculated loudness level of judged equally loud octave bands of noise as function of geometric-mean frequency of octave and for various reference noises. (Data from Stevens (ref. 5).)

The preponderance of other data and the present theory indicate that the difference should be considerably greater, say 3 to 7 decibels.

Various Bands

Zwicker's experiment (1966).—The loudness of three different noises was judged by 12 naive listeners with normal hearing by free-field listening in an anechoic chamber (ref. 66). The three sounds consisted of a one-octave, two-octave, and broadband noise which were equated in loudness to a 1-kilohertz tone. The spectra were measured with a one-third-octave, audio-spectrum analyzer. The octave-band noise at an overall sound-pressure level of 80 decibels was centered at 250 hertz. The 2-octave-band noise at an overall sound-pressure level of 70 decibels was centered at 1 kilohertz. The broadband noise, imposed at an overall sound-pressure level of 70 decibels, peaked near 200 hertz, and its spectral density possessed a 2-decibel-per-octave slope at higher frequencies.

The calculated and predicted loudness levels are compared in figure 48. In comparison with the calculated loudness levels, the judged values were low, although in at least one of the examples the mean variation of the judgments included the calculated value. (For the 2-octave noise the published tabulated and graphical loudness levels disagree. The former is 77 phons whereas the latter is 78.5 phons. The tabulated value is shown in fig. 48 herein.) The result that the judged loudness of the one-octave and two-octave noises is low relative to the calculated values is inconsistent with the corresponding

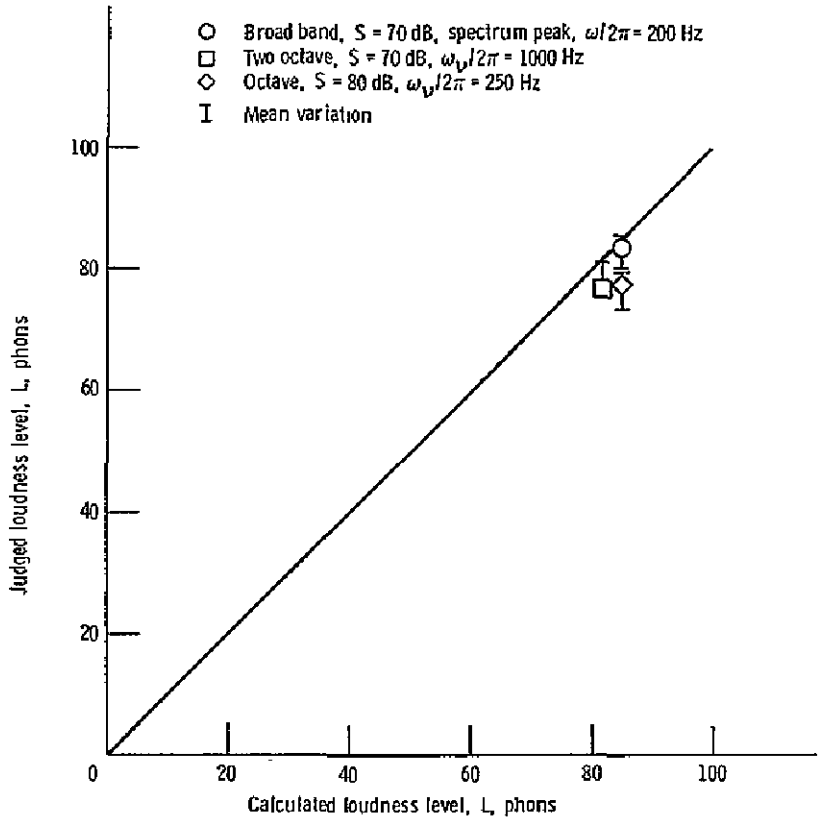


FIGURE 48.—Comparison of judged and calculated loudness levels for three noises. (Data from E. Zwicker (ref. 66).)

results of Zwicker and Feldtkeller and of Robinson and Whittle, where the judged values tend to be in good agreement with the predicted values, as shown in figures 41(a) and 44(c), respectively. However, the present results tend to agree with those of Zwicker, Flottorp, and Stevens (fig. 43(a)), where the corresponding judged values are also lower than the predicted values.

Sums of Contiguous Bands

Stevens' experiment (1956).—Naive listeners (probably 12 in number) equated the loudness of bands of noise (usually octaves) to that of a comparison noise previously calibrated subjectively against a 1-kilohertz tone so as to determine loudness levels (ref. 5). The loudness of sums of contiguous groups of these noise bands was also judged in order to determine how the loudness of the sum of bands of noise is related explicitly to the loudness of each band, not to the acoustic power in each band. However, from the loudness given for each band the corresponding loudness level can be read from a table (table III) given by Stevens. Next, the sound-pressure level at the eardrum can be approximated from the relation $S_e(\omega)=L(\omega)-N(\omega)$, which applies precisely to a tone. Finally, using the approxi-

mate value of S_e determined for each band, the loudness of a sum of contiguous noise bands may be calculated and the result compared with the judged value. This comparison is shown in figure 49. Except for the lowest (75 to 300 Hz) and highest (1420 to 10 000 Hz) bands of noise, the data generally lie within 3 phons of a line displaced about 8 phons from the ideal line, for which judged and calculated levels are equal. Since the displacement is independent of sound pressure and frequency, the most likely cause of the shift is experimental calibration error. For the lowest and highest bands an additional frequency calibration error may have existed. The low frequency problem is mentioned in the text, and serious nonuniformity of the earphone response above 2 kilohertz is evident in figure 2 of reference 5. Finally, median level values were reported rather than the higher levels corresponding to the average of power measurements.

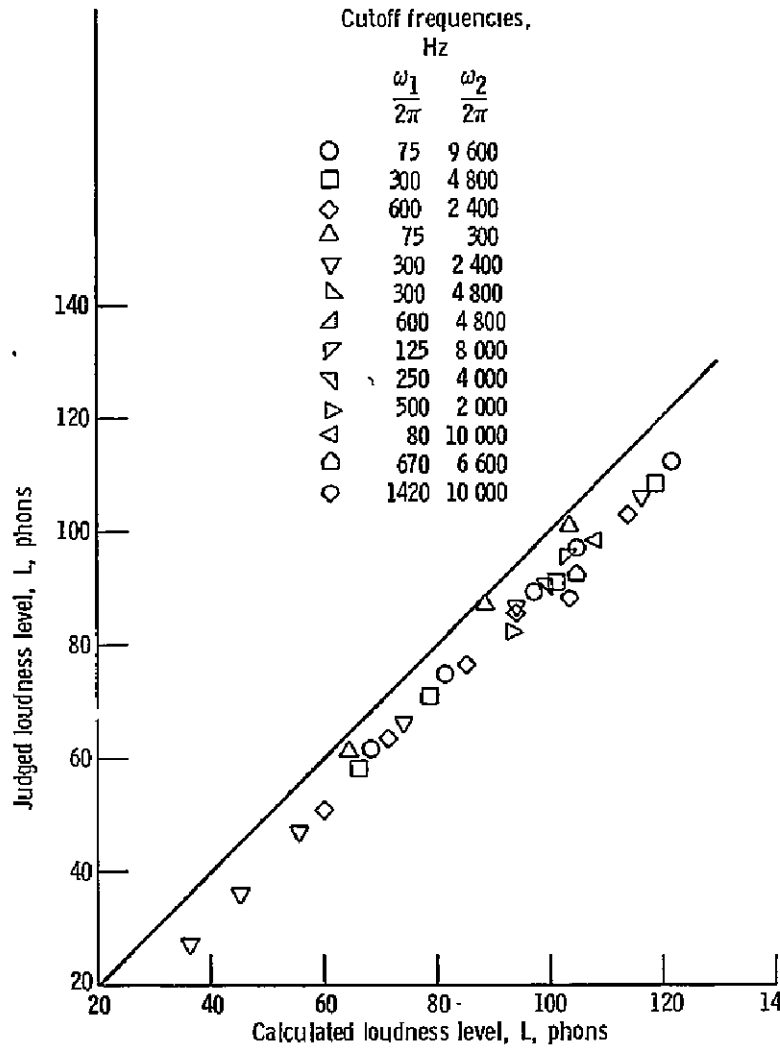


FIGURE 49.—Judged loudness level of contiguous bands of noise as function of calculated loudness level. (Data from Stevens (ref. 5, table I).)

Sums of Noncontiguous Bands

Stevens' experiment (1956).—An experiment similar to that for contiguous bands of noise was performed using noncontiguous bands (usually octaves) of noise (ref. 5, table III). The loudness of the bands and of the sum of noncontiguous bands was judged by groups of 20 or 12 listeners with earphones.

The comparison between the judged and calculated loudness levels of the sum of noncontiguous bands shown in figure 50 is very similar to that for contiguous bands in that the resultant data generally lie within 3 phons of a line axially displaced about 9 phons from the ideal line. As before, the displacement is attributed to experimental calibration error.

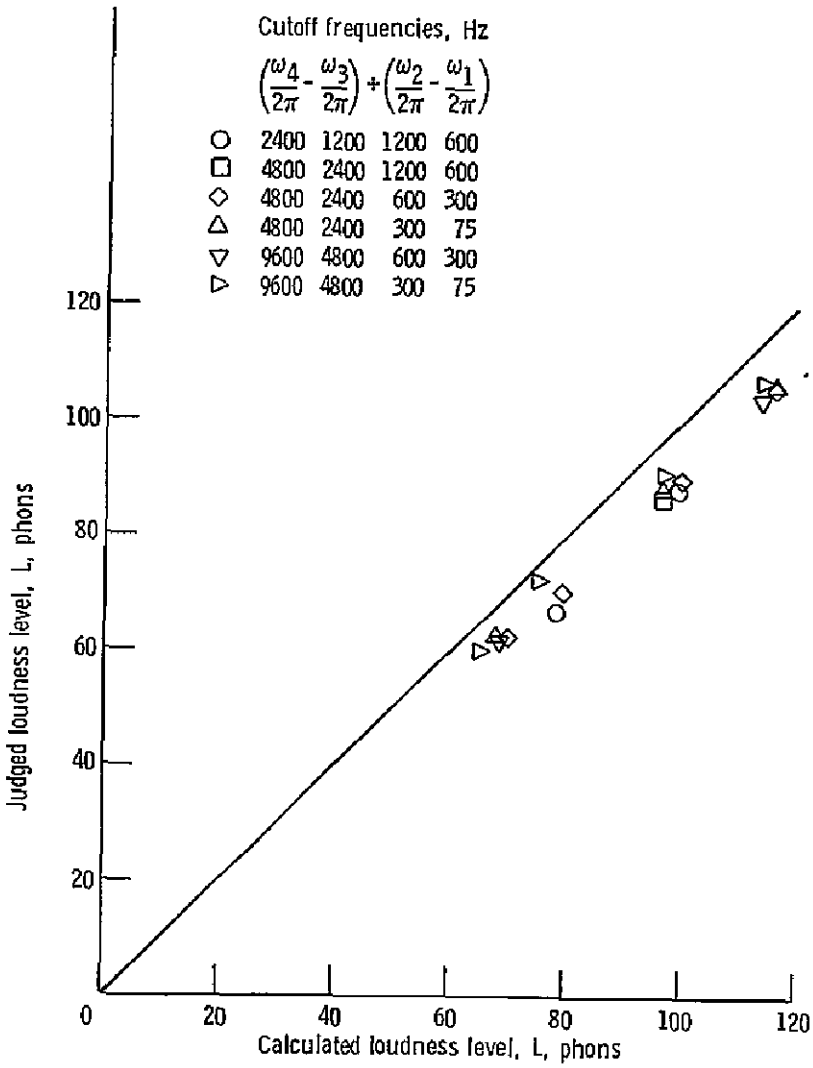


FIGURE 50.—Comparison between judged and calculated loudness levels of separated bands of noise. (Data from Stevens (ref. 5, table III).)

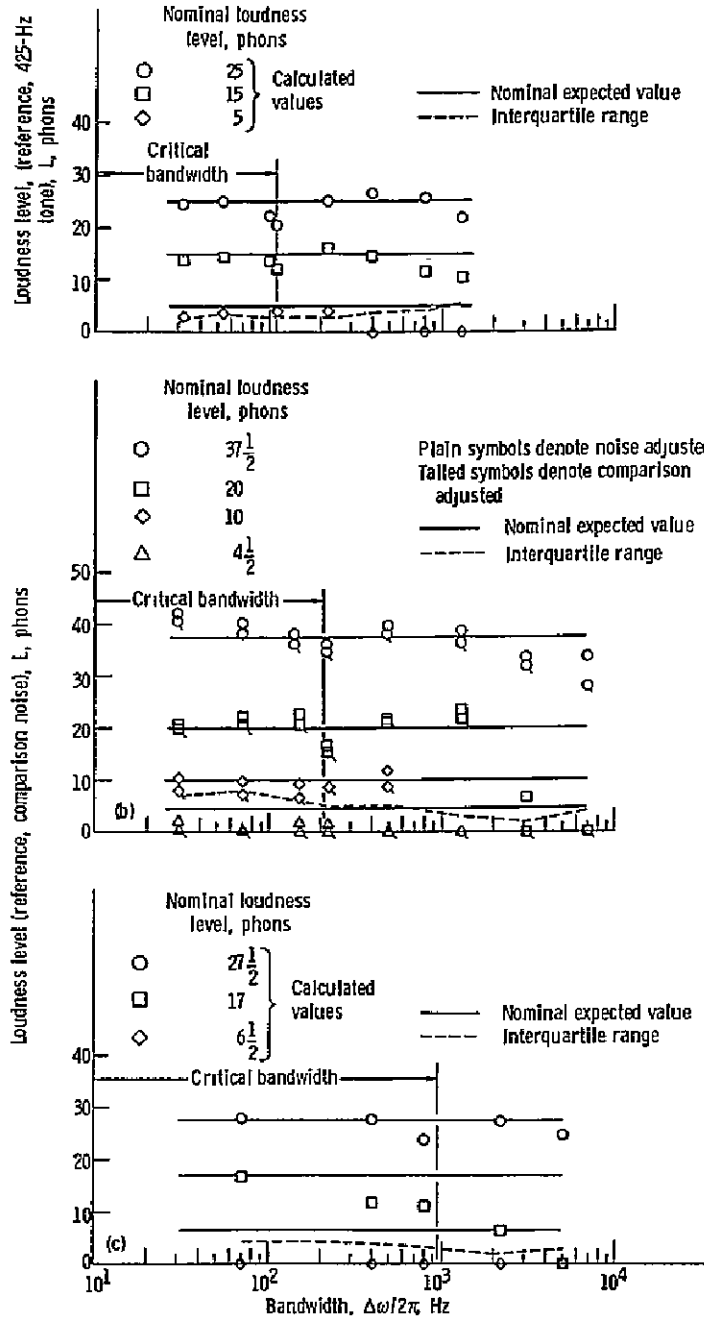
Effect of Bandwidth Near Loudness Threshold

Scharf's experiment (1959).—In addition to studies of the loudness of equally loud tone complexes as a function of tone spacing at constant overall loudness near the loudness threshold, similar studies were made of white noise near the loudness threshold (ref. 43). Bands of noise with geometric-mean frequencies of 425, 1370, and 5050 hertz were studied as a function of bandwidth. The noise centered at 425 hertz was compared with a 425-hertz tone, that centered at 1370 hertz with a band of noise 2340 hertz wide, and that at 5050 hertz with a band of noise 1730 hertz wide. The center frequencies of the comparison noises were not mentioned.

In the case of the tones, the loudness level of each of the tones was given relative to that of the reference tone rather than a 1-kilohertz tone. If the internal transmittance level at the reference frequency differed significantly from that at 1 kilohertz, then the given data values would have to be shifted by the amount of the reference-level difference in order to compute proper values of the loudness level. The case of the four-tone data centered at 3000 hertz and nearest the loudness threshold was the only one for which the correction might have been required. However, the correction was not included in the calculations.

In the case of the white noise a correction is clearly indicated for the data centered at 425 hertz (correction $= \pm 3$ dB) and nearest the loudness threshold as well as for all data centered at 5050 hertz. The correction was included in the loudness-level calculations for the noise centered at 425 hertz. The correction could not be incorporated in the calculations for the noise centered at 5050 hertz because the comparison sound was a broadband noise, and its center frequency was not specified.

The calculated and nominal (horizontal lines) loudness levels are compared in figure 51, wherein the dashed lines signify the interquartile ranges. Except for two points at the lowest level, the loudness levels calculated for the noise centered at 425 hertz agree with the nominal values to within a difference less than the interquartile range. For the other two noises the agreement is not as good, especially at the broader bandwidths. It is likely that this disagreement results because the levels used in the calculations are relative to the comparison noise rather than a 1-kilohertz tone. With the latter reference the indicated sound-pressure levels would be greater. Near threshold the transmittance at the outer limits of the noise bands would, hence, be increased. Therefore, the calculated loudness levels for the broader bandwidths would be increased relative to the calculated levels for the narrower bandwidths, as desired.



(a) Comparison, 425-hertz tone; geometric-mean frequency, $\omega_r/2\pi$, 425 hertz.
(b) Comparison, noise with 2340-hertz bandwidth; geometric-mean frequency, $\omega_r/2\pi$, 1370 hertz.
(c) Comparison, noise with 1730-hertz bandwidth; geometric-mean frequency, $\omega_r/2\pi$, 5050 hertz.

FIGURE 51.—Calculated loudness level of band of white noise near threshold as function of bandwidth. (Data from Scharf (ref. 43).)

White Noise

Feldtkeller's experiment (1956).—In connection with a comparison of loudness judgment data for repetitive impulses and a 1-kilohertz

tone, some data for white noise in the frequency range from 0 to 20 000 hertz were also displayed (ref. 67), apparently for earphone listening. The sound pressure level of the white noise as a function of that for an equally loud, 1-kilohertz tone was shown, so that the calculated loudness level of the noise as a function of that of the tone could be compared. The resultant comparison is shown in figure 52. The greatest disagreement between the judged and calculated loudness levels is 8 phons, which occurs in the middle of the studied sound-pressure level range.

Pollack's experiment (1951).—The loudness of white noise in an interval somewhat less than 60 to 5800 hertz was judged in comparison with a 1-kilohertz tone by an unspecified number of listeners using earphones (ref. 68). No information on the precise shape of the spectrum was given other than the fact that the earphone output was flat within ±2 decibels over the range 60 to 5700 hertz. The judged and calculated loudness levels are compared in figure 53 for stimulus sound-pressure levels ranging from 10 to 110 decibels. The agreement for the highest sound-pressure levels is not especially good, with level differences up to 9 phons and the calculated levels on the high side. However, in the middle range from 50 to 80 phons

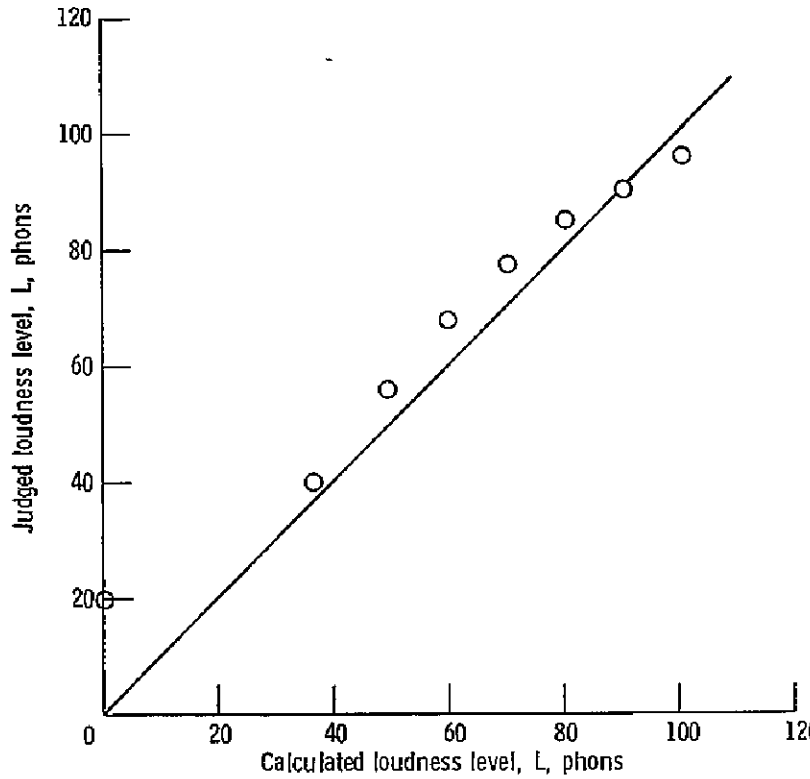


FIGURE 52.—Comparison of judged and calculated loudness levels for white noise in frequency interval of 0 to 20 000 hertz. (Data from R. Feldtkeller (ref. 67).)

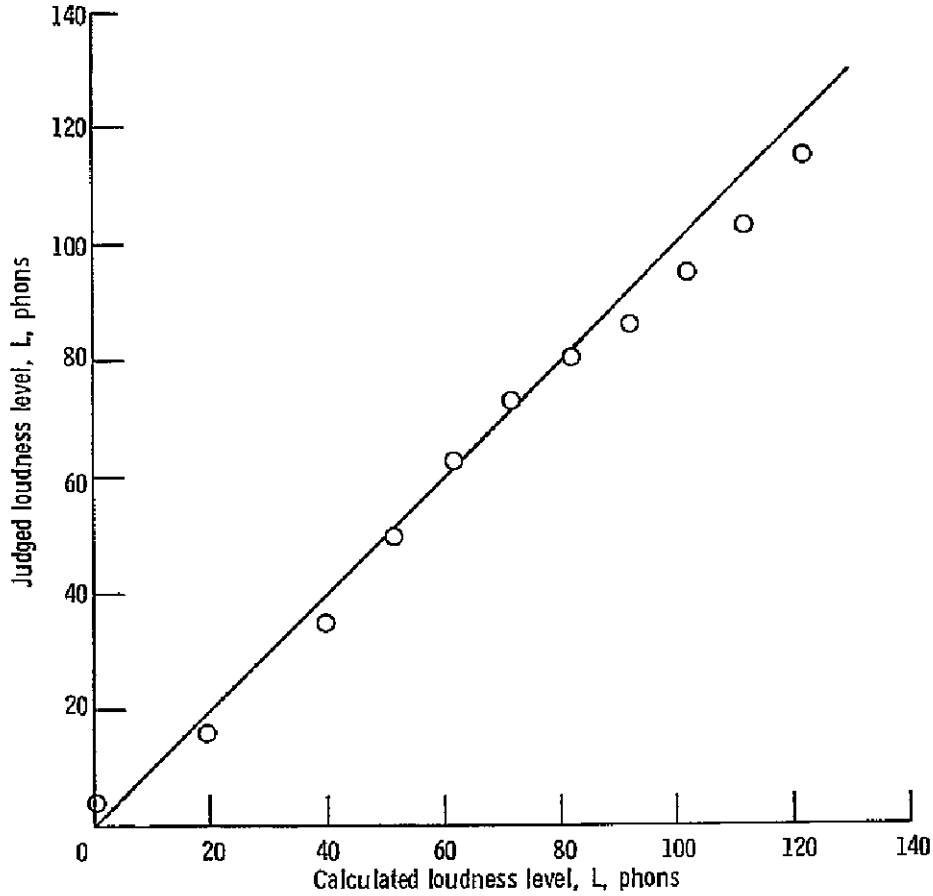


FIGURE 53.—Comparison of judged and calculated loudness levels of broadband (60- to 5800-Hz) white noise. (Data from Pollack (ref. 68).)

the agreement is good, with level differences generally 3 phons, or less.

Miller's experiment (1947).—The loudness of white noise in an interval from 150 to 7000 hertz was judged in comparison with a 1-kilohertz tone by 12 listeners wearing earphones (ref. 69). The spectrum of the noise was flat to within ±5 decibels. The judged and calculated loudness levels are compared in figure 54 wherein it has been assumed that the judged loudness levels were based on indicated sensation levels of the reference tone. Otherwise judged and calculated levels would differ by 10 phons, the claimed difference between the sound-pressure level and sensation level of the noise. Also, the difference is almost the same as that (8 phons) found from Stevens' data (see fig. 49) taken in the same laboratory using what appears to be similar equipment calibrated in a similar way. The precise explanation for what appears to be a calibration error is not evident. However, the error may be accounted for by the fact that the loudness threshold for the 1-kilohertz tone may have been well above the assumed reference pressure $p_{ref}=20$ micropascals.

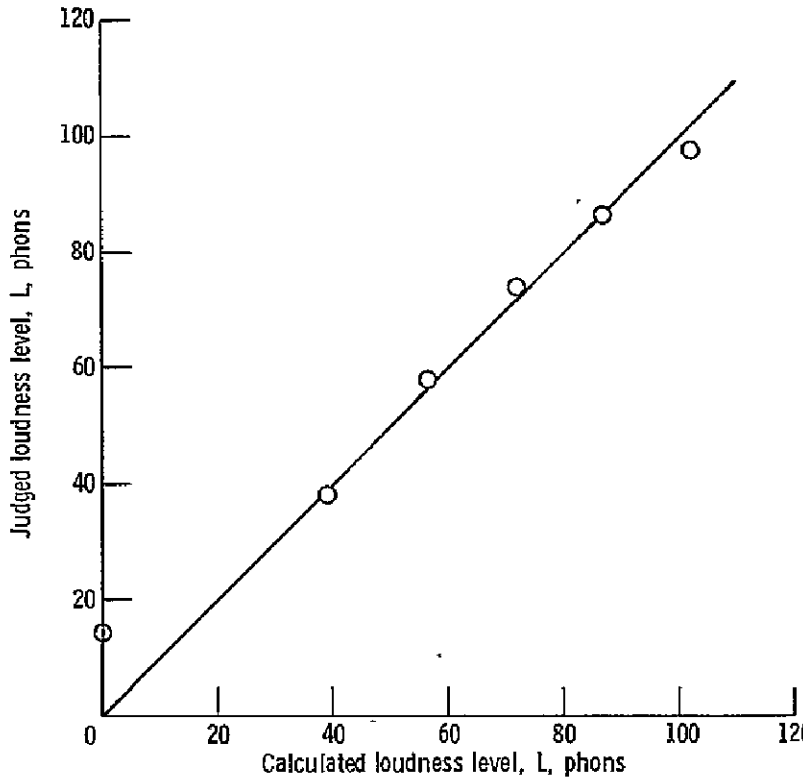


FIGURE 54.—Comparison of judged and calculated loudness levels for white noise in frequency interval 150 to 7000 hertz. (Data from Miller (ref. 69).)

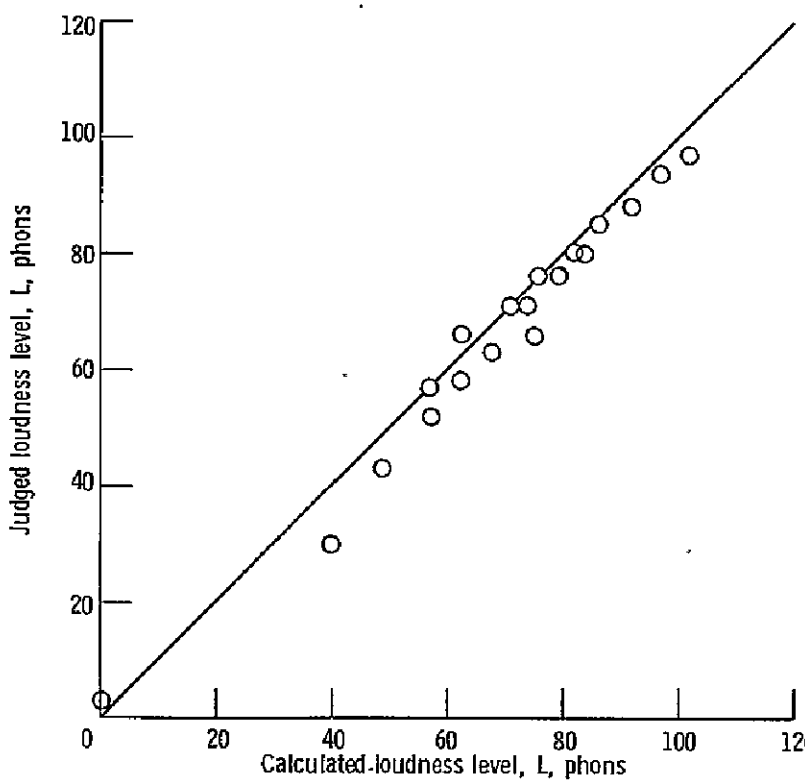


FIGURE 55.—Comparison of judged and calculated loudness levels for white noise in frequency interval 40 to 15 000 hertz. (Data from Brittain (ref. 70).)

Brittain's experiment (1939).—The loudness of white noise in an interval from 40 to 15 000 hertz was judged in comparison with a 1-kilohertz tone by 10 naive listeners in the free field (ref. 70). The spectrum was flat to within ±10 decibels, based on the response of the sound production system. The judged and calculated loudness levels are compared in figure 55 for noise sound-pressure levels ranging from 19 to 84 decibels. The disagreement between the judged and calculated loudness levels is 10 phons or less in all samples and usually less than 4 phons in the intermediate range of loudness.

Niese and Köhler's experiment (1958).—Thirty-five listeners equated the loudness of bands of noise to that of the geometric-mean frequency of the band (ref. 71). In the first experiment, loudness was studied as a function of bandwidth with the acoustic power increasing in proportion to the bandwidth. Free-field listening was used. The judged and calculated results are shown in figure 56 in which the maximum bandwidth barely exceeds the critical bandwidths in any instance. The agreement of the judgments and calculations is good.

The second experiment was similar, except that the overall sound pressure was held constant. This experiment was to determine whether amplitude-modulated and frequency-modulated noise would result in a different dependence on bandwidth. The results for free-field listening are shown in figure 57 and for earphone listening in figure 58. No difference is evident. However, for supercritical bandwidths the loudness rise reported by others is not observed.

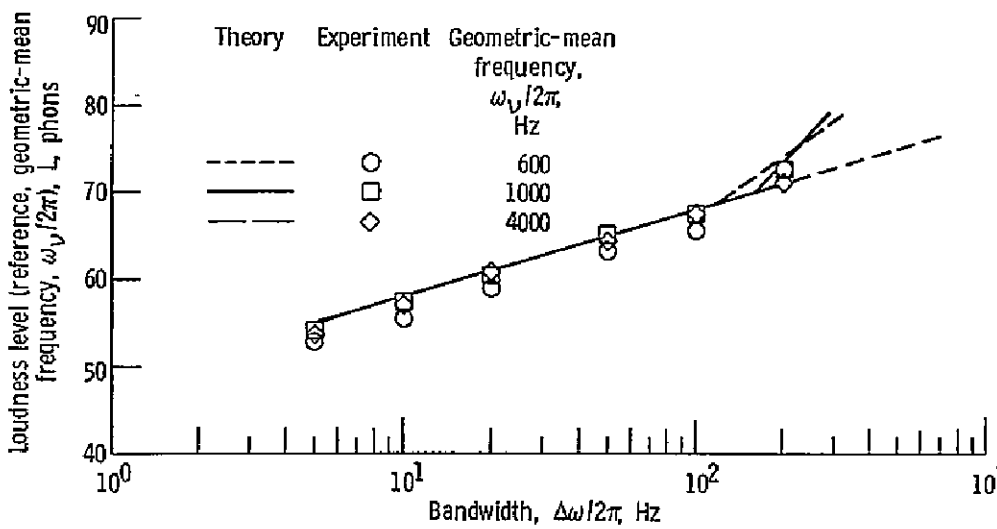
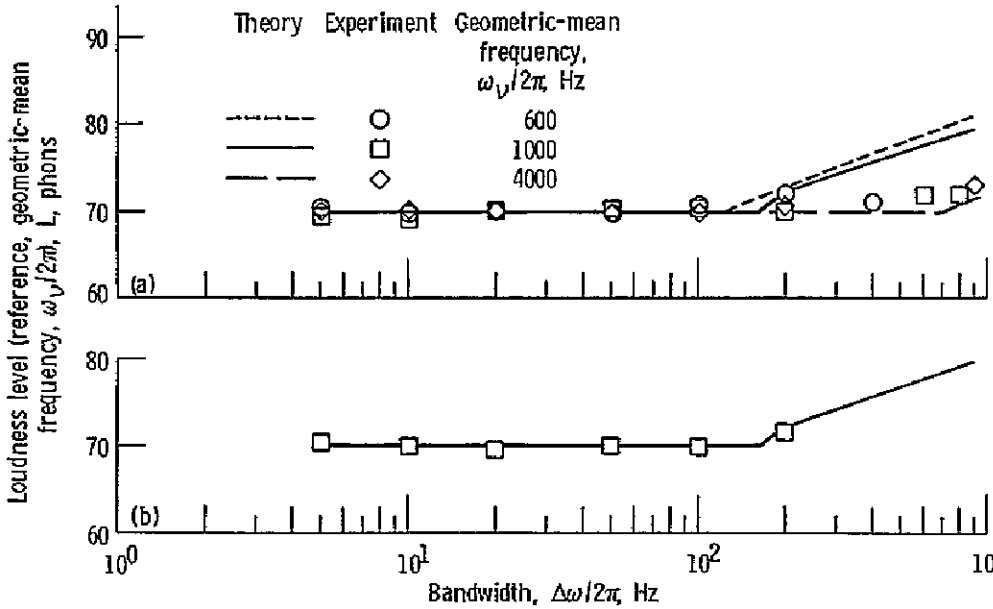
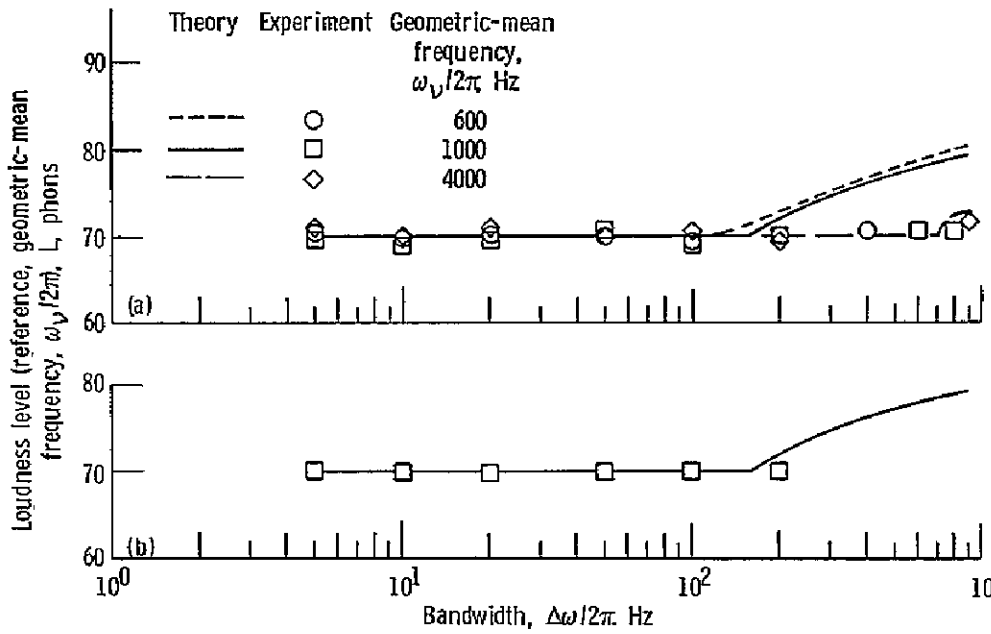


FIGURE 56.—Comparison of judged and calculated loudness levels as function of bandwidth for various mean frequencies if acoustic power is doubled for each doubling of bandwidth. Free-field listening. (Data from Niese and Köhler (ref. 71).)



(a) Amplitude modulated noise.
(b) Frequency modulated noise.

FIGURE 57.—Comparison of judged and calculated loudness levels as function of bandwidth for various mean frequencies at constant overall sound pressure. Free-field listening. (Data from Niese and Köhler (ref. 71).)



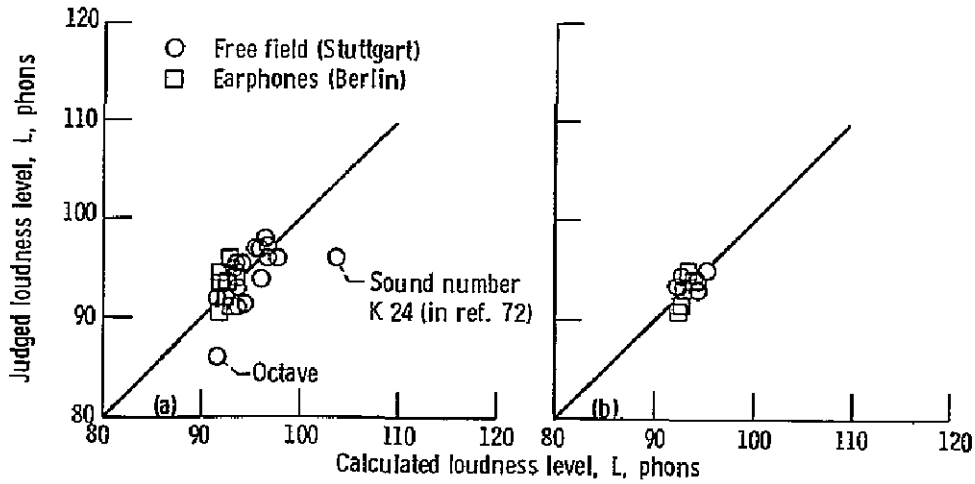
(a) Amplitude modulated noise.
(b) Frequency modulated noise.

FIGURE 58.—Comparison of judged and calculated loudness levels as function of bandwidth for various mean frequencies at constant overall sound pressure. Earphone listening (Data from Niese and Köhler (ref. 71).)

Broadband Noises

Lübcke, Mittag, and Port's experiments (1964).—These experiments (ref. 72) and the following experiment by M. Jahn (ref. 73) are especially important because the loudness of the same magnetic-tape recordings of machinery and other noises was judged independently by three different groups at three different facilities by means of free-field and earphone listening. In Berlin, 10 subjects listened through earphones. In Stuttgart, 12 subjects with normal hearing made judgments under free-field conditions in an anechoic chamber. The imposed sound-pressure level of all sounds was about 80 decibels. The sound-pressure-level spectrum was reported for one-third-octave bandwidths. Loudness levels were determined by equating the loudness of the test sound to that of a critical band of noise centered at 1 kilohertz, rather than to a 1-kilohertz tone. The latter two sounds were assumed to be equally loud at the same sound pressure.

The judged and calculated loudness levels are compared in table VII and in figure 59(a). The disagreement between the two results is less than 3 phons for nearly all of the examples. Except for two of the noises, the maximum disagreement is less than 3.5 phons. Of the exceptions, the spectral data for the noise designated K24 may be erroneous since the tabulated data do not agree with the spectrum shown in the original report. (The original published data were selected from 89 noises.) The difference shown for the octave noise centered at 1 kilohertz is 5.5 phons, similar to the 6 phon difference obtained by Robinson and Whittle for a 940-hertz center frequency (fig. 44(e)). However, this agreement between experiments may be



(a) Noises. (b) Noises containing intense tones.

FIGURE 59.—Judged loudness level as function of calculated loudness level of various engine noises (Data from Lübcke, Mittag, and Port (ref. 72).)

TABLE VII.—LOUDNESS OF ENGINE NOISES

[Data from Lübcke, Mittag, and Port (ref. 72).]

Noise designation	Experimental L, phons	Calculated L, phons	Loudness level difference, ΔL, phons	Experimental L, phons	Calculated L, phons	Loudness level difference, ΔL, phons
	Stuttgart, free field			Berlin, earphones		
P7	97	95.7	1.3	92	93.1	—1.1
M4	93	93.7	—.7	92	92.3	—.3
M18	92	92 3	—.3	92.5	92.1	.4
D3	96	97.8	—1.8	91.5	92.8	—1.3
D8	91	93.7	—2.7	90 5	91.8	—1.3
D14	91.5	94.5	—3 0	91	93.2	—2.2
K1	98	96.4	1.6	94	92.9	1.1
K10	95.5	94.3	1.2	93.5	92.6	.9
K11	95	93.1	1.9	93.5	91.7	1.8
K19	97	96.8	.2	94	93.4	.6
K21	97	96.0	1.0	----	----	----
K24	96	103.8	—7.8	----	----	----
J5	95.5	93.9	1.6	----	----	----
J12	96	93.9	2.1	----	----	----
Octave	86	91.6	—5.6	----	----	----
USN	96	96.8	—.8	96	92.8	3.2
Absolute average ΔL, phons			2.1	Absolute average ΔL, phons		1.3

fortuitous since in other experiments a different result was obtained.

The average difference between theory and experiment is 2.1 phons for free-field listening and 1.3 phons for earphone listening.

The loudness levels of all the sounds studied were recalculated using the approximation equation (70). The levels given by this equation usually were higher than those given by the "exact" equation (50). The maximum difference in loudness level indicated by the two equations was 1 phon.

M. Jahn's experiment (1965).—Twenty-eight listeners, some trained, some untrained, all with normal hearing, judged the loudness (ref. 73) of some of the same recorded engine noises studied by Lübcke, Mittag, and Port (ref. 72). The judgments were made in an anechoic chamber by equating the loudness of the engine noise to that of a supercritical, 200-hertz bandwidth noise centered geometrically at 1 kilohertz. In order to obtain judged loudness levels relative to a 1-kilohertz tone the measured levels were reduced by 2 phons. The consequent comparison between the judged and calculated loudness levels is as appears in figure 60. The maximum scatter of the results about an average is less than 2 phons; however, the average disagreement between the judgments and calculations is about 6.5 phons, possibly due to a calibration error.

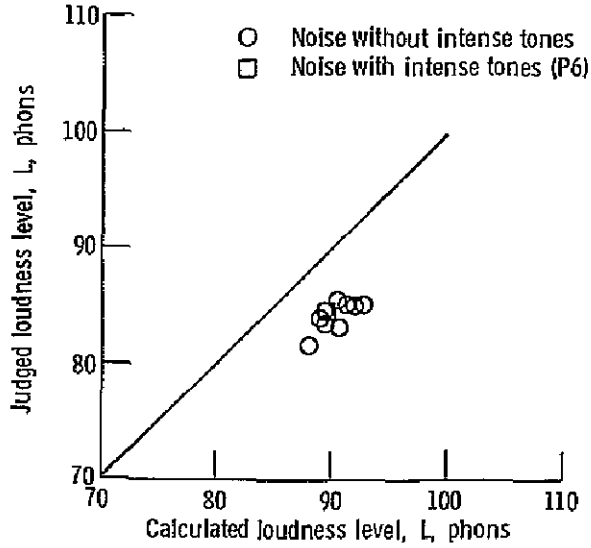


FIGURE 60.—Judged loudness level as function of calculated loudness level of various engine noises. (Data from M. Jahn (ref. 73).)

The calculated results in figure 60 were from data obtained by sound-pressure level measurements in one-third-octave bands. Calculations using data obtained from octave-band measurements yielded loudness levels from about 0.5 to 2 phons higher, which illustrates the effect of insufficient spectral resolution. A similar lack of resolution of data measured in one-third-octaves might account for some of the residual difference between theory and experiment.

Warren's experiment (1973).—A total of 600 naive listeners judged the sound-pressure level reduction necessary to reduce the loudness of broadband white noise by one-half (ref. 74). Listening was through earphones or in the free field in a semi-reverberant room. Careful efforts were made to eliminate experimental biases. Over the sound-pressure-level range from 45 decibels through 90 decibels, halving the loudness was found to correspond to a sound-pressure level reduction of 6 decibels. This implies that loudness is proportional to the square root of the imposed acoustic power, in complete agreement with the presently proposed theory over the frequency range for which the auditory transmittance is uniform, say from about 200 to 4000 hertz for listening through earphones.

Quietzsch's experiment (1955).—The loudness of 32 noises with continuous spectra was judged by at least 20 listeners in the free field (ref. 60). Some tests (22 noises) were performed in an anechoic chamber, whereas others (10 noises) were performed in a lecture hall. In the anechoic chamber the judgments were made relative to a 1-kilohertz tone whereas in the lecture hall the comparisons with the 1-kilohertz reference tone were made indirectly by comparison with a one-third-octave bandwidth noise with cutoff frequencies at 900

and 1120 hertz. Most of the sounds listed in table VI were generated from magnetic tape recordings using loudspeakers, but three sounds (noises 28, 36, and 37) were generated by the original sources. The spectra of the noises were measured in octave bands. The resultant calculated loudness levels are compared in figure 61 with the judged values. The maximum scatter of the data about the ideal line is 11 phons with a systematic shift of about 3 phons, which can be ascribed to the fact that the spectra were measured in octave bands and to averaging error.

Corliss and Winzer's experiment (1965).—From magnetic-tape recordings, nine listeners equated the loudness of a tapping machine and of three women with steel-tipped shoe heels acting on various concrete-backed, hard floor surfaces to that of a narrowband noise centered at 1000 hertz and a warble tone produced by a 955- and a 1005-hertz tone (ref. 75). The loudness judgments were performed in the free field in a reverberant chamber. Because of standing waves

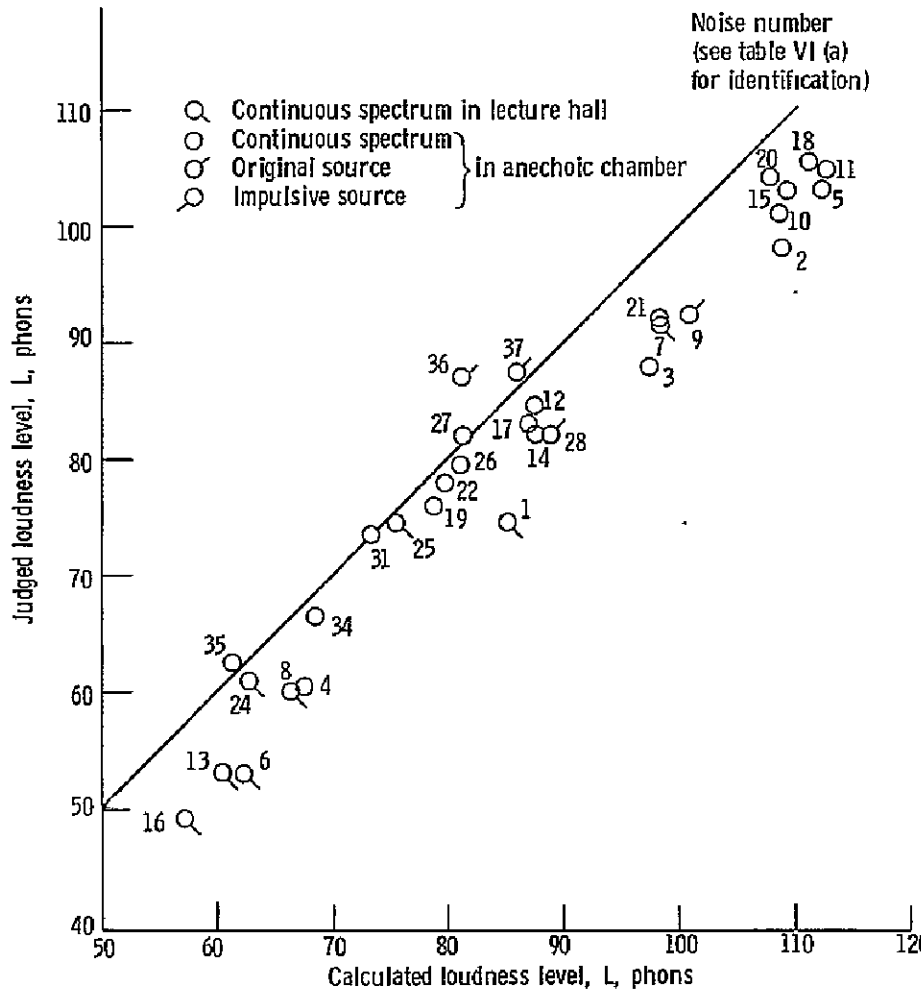


FIGURE 61.—Judged loudness level as function of calculated loudness level of various sounds. (Data from Quietzsch (ref. 60).)

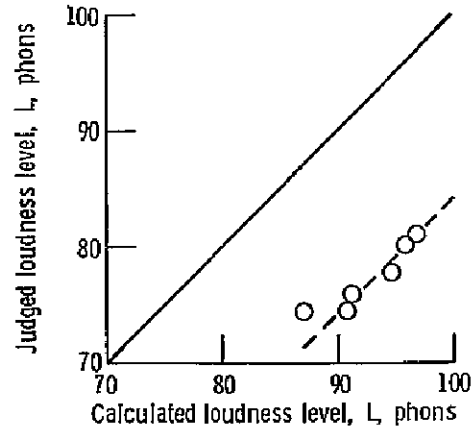


FIGURE 62.—Judged loudness level as function of calculated loudness level. (Data from Corliss and Winzer (ref. 75).)

in the reverberant chamber the two reference sounds were compared in loudness to a 1-kilohertz tone by listening through earphones in order to determine the loudness levels of the tapping sounds.

The judged and calculated levels are compared in figure 62. Both the judged and calculated levels occur essentially in the same order of relative loudness and are all grouped within 3 phons of a line paralleling the ideal line. However, the data are shifted approximately 16 phons from the ideal line. This shift may result partly from the fact that the sounds were impulsive rather than continuous. However, it was claimed that the impulsiveness was smoothed out by the room reverberance. More likely sources of error are calibration errors, especially in using earphones, and the use of a warble-tone reference, because of the influence of its unsteadiness on subjective judgments.

Robinson and Bowsher's experiment (1961).—Sound-pressure levels corresponding to equal judged loudness were determined for five helicopters from magnetic tape recordings listened to by a total of 570 listeners, ten at a time, in a "moderately" reverberant room (ref. 76). Since the loudness of the noises was not compared with that of

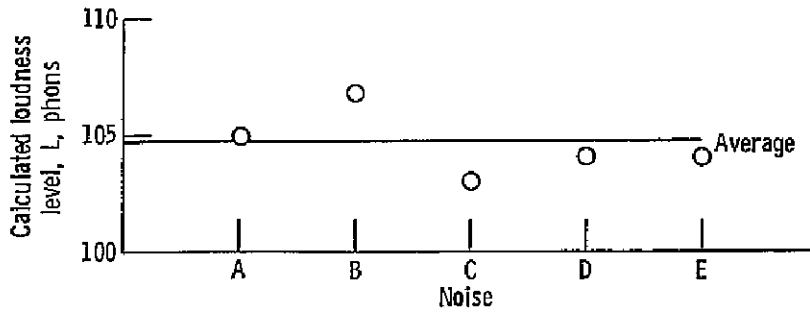


FIGURE 63.—Calculated loudness level for five helicopter noises judged equally loud. (Data from Robinson and Bowsher (ref. 76).)

a 1-kilohertz tone, absolute loudness levels could not be determined experimentally. From the measured one-third-octave, sound-pressure-level spectra the loudness levels were computed, with the results shown in figure 63. The calculated average sound-pressure level was 104.8 phons with a maximum spread of about 4 phons among the individual values.

Copeland et al. experiment (1960).—Recordings of five sounds of aircraft, or simulated aircraft, were played for 1578 listeners (up to 20 at a time) (ref. 77). The sounds were imposed in diffuse form by means of multiple loudspeakers in an auditorium. Listeners judged the louder of two of the sounds imposed at fixed sound-pressure levels. From these data the sound-pressure levels at which the sounds would apparently be equally loud could be estimated. No comparison with a reference tone was made which would permit absolute loudness levels to be determined. However, from the estimated sound-pressure levels and corresponding spectra measured in octaves, loudness levels could be calculated. The results are shown in figure 64. Neglecting sound "E," which contains introduced pure-tone components, the range of calculated loudness levels of the other four sounds was 4 phons. Ideally this range should, of course, be zero phons.

LOUDNESS OF SPECIFIC SOUNDS WITH MIXED SPECTRA

In some psychoacoustic studies it is common to treat equally intense, discrete tones and noises as if they produced quite different sensation levels obeying different empirical rules. In the present loudness theory this differentiation does not occur. Rather, response to both tones and noise appears as a natural consequence of speciali-

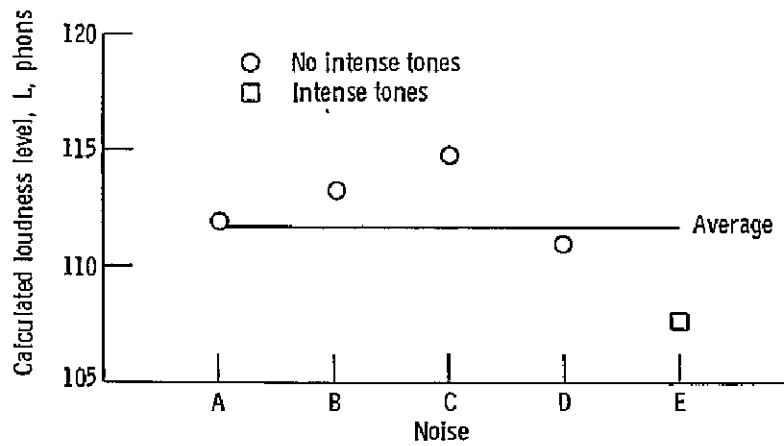


FIGURE 64.—Calculated loudness levels of aircraft noises subjectively judged equally loud. (Data from Copeland, et al. (ref. 77).)

zation of a general theory which implicitly incorporates both types of sound. The next few experiments illustrate the success of this approach when applied to noises containing intense pure-tone components.

Lübcke, Mittag, and Port's Experiment (1964)

Among the noises studied, five contained strong pure-tone components (ref. 72). The judged and calculated loudness levels of these sounds are compared in table VIII and in figure 59(b). The maximum difference between these two levels for both free field and earphone listening is 1.8 phons. The average difference is 1.0 phon for free field listening and 1.4 phons for earphone listening.

M. Jahn's Experiment (1965)

Of the noises studied by Lübcke, Mittag, and Port (ref. 72) and then by M. Jahn (ref. 73), only one (noise P6) contained strong pure-tone components. For this sound the difference between the judged and calculated loudness level for free field listening was 4.5 phons, as shown in figure 60. It should be recalled, however, that most of this difference may result from a calibration error.

Quietzsch's Experiment (1955)

Some of the noises previously discussed would appear from the description of the source (table VI) and the exhibited spectra (ref. 60) to contain strong pure-tone contributions. Among these noises might be noises number 31 and 37. The difference between the judged and calculated loudness levels for these noises is 0 and 1.5 phons, respectively, as shown in figure 61.

Copeland et al. Experiment (1960)

Two of the five sounds (D and E) evaluated by free-field listening (ref. 77) were identical except that sound E had a 3-kilohertz square wave added, which, however, hardly contributed to measured octave levels. Nevertheless, the sound-pressure level of sound E was judged to be 3.3 decibels less than that of sound D for equal loudness. Thus, for judged equal loudness the calculated loudness level of sound D was 3.3 phons greater than that of sound E, as shown in figure 64. Although this difference is not great, the total spread among the five sounds is 7.3 phons, which is considerably greater than that in the better controlled laboratory experiments of Lübcke, Mittag, and Port and of M. Jahn.

TABLE VIII.—LOUDNESS OF ENGINE NOISES CONTAINING STRONG PURE-TONE COMPONENTS

[Data from Lübcke, Mittag, and Port (ref. 72).]

Noise designation	Experimental L, phons	Calculated L, phons	Loudness level difference, ΔL, phons	Experimental L, phons	Calculated L, phons	Loudness level difference, ΔL, phons
	Stuttgart, free field			Berlin, earphones		
P6	95	95.3	−0.3	91.5	92.7	−1.2
M7	93.5	92.4	.1.1	91	92.5	−1.5
G1	94	94.2	−.2	94	93.0	1.0
G4	93	94.6	−1.6	95	93.3	1.7
J11	94.5	92.7	1.8	----	----	----
Absolute average ΔL, phons			1.0	Absolute average ΔL, phons		1.4

RELATION BETWEEN PRESENT THEORY AND A-WEIGHTED, SOUND-PRESSURE LEVEL

For many years the A-weighted, sound-pressure level has been generally regarded as the best, though unsatisfactory, physical measure of the subjective intensity level of a sound (e.g., ref. 78). As various alternative loudness prediction procedures are proposed and then fail, use of the A-weighted level is revived, especially because of its simplicity and the fact that sound-level meters provide an A-weighted reading. The A-weighted level is supposed to be in accordance with inversion of equal-loudness contours near 40 phons determined by modifying Fletcher and Munson's equal-loudness curves (ref. 13) to account for a diffuse source (ref. 53, p. 22). This weighting is derived from measurements using earphones. It does not take into account the dependence of the internal filter on sound pressure. When this factor and standard equal-loudness curves (ref. 20) were adopted, improved agreement with loudness judgments resulted (ref. 79). Although the calculation procedure in reference 79 appeared to yield loudness levels in good agreement with a limited set of judgment data it was subsequently found to disagree with a much larger body of judgment data included in this report, just as A-weighted sound-pressure level, especially with the cited improvements, sometimes represents correctly the relative loudness of certain limited groups of similar sounds, but nevertheless completely fails to predict correctly the absolute loudness levels of steady sounds in general.

It has been demonstrated that the theory proposed in this report does predict absolute values of loudness levels in accordance with judged values. Moreover, the present theory can be used to show when the A-weighted sound-pressure level provides useful information on loudness levels. Specifically, according to the present theory

two suprathreshold sounds with the same spectrum shape possess a loudness ratio equal to the square root of the ratio of their spectral densities. This applies for spectra with subcritical or supercritical bandwidths since all values of w or W in all of the loudness equations are multiplied by the same constant in going from one sound to the other. The same relation also applies to the A-weighted, effective sound pressure when it is assumed to represent loudness. Thus, neglecting the weak effect of sound pressure on transmittance, the A-weighted sound-pressure levels will otherwise correctly yield the loudness level difference between two sounds which have the same spectrum shape. Moreover, if, in addition, the spectrum bandwidths are subcritical and beats are not significant, the A-weighted, sound-pressure levels equal the true loudness levels to the extent that the A-weighting corresponds to the true weighting of the auditory system. However, in general, the present theory indicates that the A-weighted sound-pressure level is not simply related to loudness level and may fall more than 15 decibels (phons) below the loudness level (cf. fig. 41).

DISCUSSION AND REMARKS

It is important to recognize that the present loudness calculations are based on an operational theory of the auditory system rather than on an empirical formula selected to fit some psychoacoustic data, as is often the case. The empirical aspects of the theory have been minimized so as to yield a precise loudness prediction minimally influenced by empirical functions. The theory predicts loudnesses in fairly good agreement with a large body of, but not all, loudness judgment data. However, where the predictions and judgments fail to agree, the judgment data usually conflict with other data which do agree with the theory. None of the data were taken by the author. Therefore, the theory is not biased in an attempt to fit the author's own data, nor is it designed to fit a specific limited set of data from one source. Rather, in an effort to gain credibility the theory has been designed to account most simply for the maximum amount of physiological and psychoacoustic data while obeying physical principles. Of course, inconsistency of some of the data limits this possibility.

The generally good agreement between theory and experiment (level differences less than 3 to 5 phons, say) implies that biases imposed by the variety of psychoacoustic judgment techniques employed by the experimentalists and temporal fluctuations of the auditory system are not significant enough to preclude repeatability of results among various experiments. This is not to say that these effects are insignificant. They may, in fact, account for a large part of the variability of the data as well as much of the disagreement between

theory and experiment. Since the calculations yield loudness levels relative to a fixed reference, namely a 1-kilohertz tone, calibration differences among different experiments should be exposed. A calibration error should appear as a constant difference between theory and experiment among all tests in a given experiment. Potential large calibration errors are indicated in a few experiments (cf. figs. 49, 50, 54, 60, and 62).

The method of averaging raw judgment data may introduce a bias in the results. Mean-squared pressures should be averaged rather than the decibel equivalents. In a typical experiment the former average was found to yield a level more than 3 decibels greater than the latter average (ref. 52). This error alone might account for the residual difference between predicted and judged loudness levels in some experiments. Unfortunately the raw judgment data is normally not included with published results, so that the appropriate averages cannot now be computed.

Although the periodic phenomena (for example, phase locking of neural discharges and periodicity pitch) incorporated in the present theory are generally not observed for frequencies greater than about 5 kilohertz, at higher frequencies discharge periodicities nevertheless exist. These, in fact all discharge periodicities, may be associated with pitch rather than frequency (ref. 22). Hence, there is apparently no reason to expect that the basic mechanisms for loudness production are any different at high stimulus frequencies.

A more refined analysis than the present one might involve transmittance of signals at all frequencies in all channels with different filtering of each frequency in each channel. Unfortunately the complexity of the filtering process for multiple tone stimuli in individual neurons (ref. 42) indicates that this approach might require a much more complicated analysis than the present one, but with little, if any, better agreement between calculated and judged loudness because of the variability of loudness judgments.

Channels with associated critical bands containing frequencies included in the imposed sound spectrum are most strongly stimulated. As a result, the broader the bandwidth of a sound with a fixed overall sound pressure, the louder the sound will tend to be, at least until its spectrum extends into a frequency range of reduced transmittance. The broader bandwidth sound is louder because more channels are strongly stimulated. Since the acoustic frequency range includes 24 channels, the loudness level may be increased by 14 phons, or even more, simply by broadening the sound bandwidth. The larger increases would be expected if the spectrum is broadened into a frequency range of increased transmittance. An additional increase of up to 3 phons may result from tonal interaction.

The most effective loudness reduction for any sound should be obtained by attenuating the sound at frequencies associated with maximum output power W_q.

It should be possible to construct on the basis of the present theory a loudness meter which would perform the necessary calculations electronically and indicate the loudness, or loudness level in real time.

It seems quite possible that the present theory may be readily applicable for predicting the loudness of impulsive sounds as well as steady sounds. The question of whether critical bands are significant for impulsive sounds as well as for steady sounds is undecided. From some studies it is concluded that critical bands develop as a function of exposure time (ref. 80) whereas from other studies it is concluded that critical bands are not time dependent (ref. 81). In either case the present theory might be valid without major modifications.

SUMMARY OF RESULTS

From a study of the relation between any steady sound stimulus and its judged intensity, or loudness, the following results and conclusions emerged:

1. Loudness is a measure of the electric power (defined over the auditory integration time$\approx$0.1 s) transmitted by the auditory nerve toward the central nervous system in response to an acoustic stimulus.

2. The auditory system is a squaring device in that the electric power transmitted by the auditory nerve is a function of the square root of the input sound power. The squaring results from transformation of the acoustic waveform from mechanical to electrical form.

3. Sound stimuli which excite the same auditory neurons interact. Stimuli which do not excite the same neurons do not interact. The interaction affects loudness significantly.

4. The maximum number of audible harmonics which may interact by exciting the same neurons is about 5. This number is about the same as the observed maximum number of harmonics (5 to 7) which may be heard in response to a multitone stimulus.

5. Audible beats resulting from closely spaced tones (beats of imperfect unisons) or from tones which are nearly harmonics of one another (beats of mistuned consonances) are a consequence of the squaring process and of the finiteness of the auditory integration time. The finite integration time leads to time-dependent, cross-power contributions from the interaction of the tones. The loudness of the beats tends to decline as the tone separation is increased because the integration time approaches and then exceeds the period of the beats. The predicted amplitude of beats as a function of frequency difference and phase difference is in good agreement with experimental data.

6. The nerve fibers comprising the auditory nerve may be classified, according to their best, or characteristic, frequencies into transmission channels having bandwidths equal to critical bandwidths for loudness. Individual fiber bandwidths generally exceed associated critical bandwidths. Thus, any signal at a given frequency may be transmitted in several channels in addition to the one characteristic channel which includes the signal frequency within its associated critical band. The interaction of signal frequencies associated with different characteristic channels may stimulate in each transmission channel a cross-power contribution corresponding to secondary tones which increase the overall loudness of the sound.

7. The critical bandwidth for loudness tends to equal the frequency of onset of the secondary tone which, in turn, equals the characteristic lowest response frequency of neurons with best frequencies lying in the critical band. The regime of secondary tones is limited to the frequency range in which neuron bandwidths clearly exceed associated critical bandwidths.

8. Sounds with a vanishing spectrum in intermediate critical bands stimulate neurons characterizing these "empty" bands as well as neurons characterizing critical bands corresponding to stimulus frequencies. Stimulation of neurons associated with the empty bands increases the loudness.

9. Tones which are individually subliminal might, in combination, stimulate a loudness response if they lie within the same critical band.

10. Minimum loudness for a given overall sound pressure occurs for sounds with a spectral bandwidth equal to a critical bandwidth for loudness, or slightly less.

11. Except near the loudness threshold, for a fixed overall sound pressure the loudness of a sound with supercritical bandwidth tends to a maximum as its bandwidth is increased and then declines as the bandwidth is increased further. The sound is louder for broader bandwidth because more channels are strongly stimulated. Since the acoustic frequency range includes 24 channels, the loudness level may be increased by 14 phons, or even more, by broadening the bandwidth. An additional increase of up to 3 phons may result from tonal interaction. The decline for very broad bandwidth occurs because the sound then extends into a frequency range of reduced system transmittance.

12. Using published data from a large number of experiments involving sounds with discrete spectra (tones), continuous spectra (noise), and mixed spectra (noise plus tones), generally good agreement between judged and calculated loudness levels was found. Where considerable disagreement was encountered for a given experiment, results of another similar experiment usually indicated good agreement between predicted and judged loudness levels.

13. Von Békésy's loudness judgment tests of two equally intense tones of the same frequency show that a 6 decibel reduction in the sound-pressure level of a tone corresponds to halving its loudness, in agreement with several other studies as well as the present theory.

14. A change of n decibels in the overall sound-pressure level of a suprathreshold sound without changing its spectrum shape results in a change of approximately n phons in its loudness level.

15. The A-weighted, sound-pressure-level difference approximately indicates the loudness level difference for sounds with the same spectrum. The A-weighted, sound-pressure level equals the loudness level for subcritical bandwidth sounds (without beats) in spectral regions where the signal is not attenuated relative to the system transmittance at 1 kilohertz. However, in general, A-weighted sound-pressure level is not simply related to loudness level, and the two levels may differ considerably, say by more than 15 decibels.

APPENDIX A
SYMBOLS

A	sound-pressure amplitude
A	$\left[1-\frac{J_1(2r\omega/c)}{2r\omega/c}\right]^2 \sin^2(\omega l/c)$ appendix F
a	complex amplitude, $(A/2)e^{i\phi}$
B	$\left[\frac{2}{(2r\omega/c)^2}K_1(2r\omega/c)\sin(\omega l/c)-\cos(\omega l/c)\right]^2$, appendix F
C	sound-pressure level at loudness threshold, dB
c	function given by eq. (65a)
c	speed of sound in air, appendix F
d	function given by eq. (65b)
$E_{m,n}$	$e^{-i[(\omega_m-\omega_n)t+\phi_m-\phi_n]}$
$H(\omega)$	frequency response of auditory system
$\mathcal{H}(\omega)$	frequency response of internal auditory system
$\mathcal{H}_e(\omega)$	frequency response of external auditory system
$h(t)$	impulse response of auditory system
i	$\sqrt{1}$
$J_k(\omega_n)$	frequency response of channel k to rotational frequency ω_n
J_1	first-order Bessel function of first kind, appendix F
K	dimensional transmission factor
K_1	modified Bessel function, appendix F
$\mathcal{K}$	dimensionless transmission factor
$\mathcal{K}_1$	dimensional constant, $0.5\times10^{-3}\ \mu\text{Pa}^{-1}$
k	total number of empty critical bands
k,n	integers
L	loudness level, phons
ΔL	loudness level difference, eqs. (54) and (55)
$\mathcal{L}$	loudness, sones
$\mathcal{L}_{cr}$	loudness of sound spanning a critical band, sones
l	length of auditory meatus

M	total number of transmission channels spanning sound spectrum
N	total number of tone stimuli in sound
N	internal transmittance level
$\mathcal{N}(\omega_n)$	internal transmittance
$\mathcal{N}(\omega_m,\omega_n)$	internal cross-transmittance
$P(\omega)$	sound-pressure spectral density, s
$\mathcal{P}$	normalized, "effective" sound-pressure defined over auditory integration time, eq. (2)
p	instantaneous sound pressure
p_{ref}	reference effective sound pressure, 20 μPa
$Q(\omega_m)$	total transmittance of auditory system
$Q(\omega_m,\omega_n)$	total cross-transmittance of auditory system
q	"output" of auditory system
r	total number of neural discharges over all neurons in 0.1 s
r	radius of auditory meatus, appendix F
$\mathcal{R}_q(t)$	normalized autocorrelation function of output in a transmission channel
$\mathcal{R}_\varepsilon$	real part
S	sound-pressure level, dB
S_A	A-weighted sound-pressure level, dB(A)
$\mathcal{S}$	sensation level, dB
$\operatorname{sinc}\xi$	$\operatorname{sinc}(\pi\xi)/\pi\xi$, eq. (16)
T	auditory integration time, 0.1 s
T	external transmittance level
$\mathcal{T}(\omega_n)$	external transmittance
$\mathcal{T}(\omega_m,\omega_n)$	external cross-transmittance
t	time
t_c	characteristic time denoting a particular cycle of oscillation within auditory integration period
$t_{\max}$	$-\dfrac{\phi_m-\phi_k-\theta_{mk}}{\left(\dfrac{\epsilon}{\gamma}-1\right)\omega_k}$
Δt	$2\pi/(\omega_m-\omega_k)$
$w(\omega_m)$	dimensionless power of tone
$w(\omega_m,\omega_n)$	dimensionless cross-power of two tones
$W_q(\omega)$	"output" power spectral density, s
$w_q(\omega_m)$	dimensionless "output" power of tone
w_{tot}	dimensionless total power of tones
z	complex quantity
α	function given by eq. (63)
β	function given by eq. (64)

β_{mn}	intertone coefficient for supercritically spaced tones
γ,ϵ,σ	positive integers
Δ,δ	difference
$\delta(\omega)$	Dirac delta function
δ_{mn}	Kronecker delta
η	positive number, $0\leq\eta<1$
θ_{mn}	$\tan^{-1}[\mathcal{N}_i(\omega_m,\omega_n)/\mathcal{N}_r(\omega_m,\omega_n)]$
κ	constant in eq. (4) for loudness, 10^{-2}
κ'	constant in fundamental eq. (3) for loudness
λ	positive number
ν	$\omega''-\omega$
ξ	$\nu T/2\pi$
τ	delay time
Φ	$\phi_m-\phi_k-\theta_{mk}$
ϕ	phase angle
ψ	$\tan^{-1}(\beta/\alpha)$
$\Omega(\omega_\alpha,\Delta\omega_\alpha)\equiv\Omega(\omega_\alpha)$	dimensionless acoustic power in subcritical frequency band $(\Delta\omega/2\pi)_\alpha$ centered at frequency $\omega_\alpha/2\pi$
Ω_{cr}	dimensionless acoustic power in critical band
$\Omega_{cr}(\omega_\mu,\omega_\kappa)$	dimensionless acoustic cross-power between critical bands, eq. (51)
Ω_q	dimensionless "output" power in subcritical frequency band
ω	rotational frequency
$\omega_a/2\pi$	lower cutoff frequency of ideal, spectrum-analyzer filter having subcritical bandwidth
$\omega_b/2\pi$	upper cutoff frequency of ideal, spectrum-analyzer filter having subcritical bandwidth
$\omega_c/2\pi$	center frequency of internal auditory system
$\omega_l/2\pi$	lower cutoff frequency of internal auditory system
$\omega_u/2\pi$	upper cutoff frequency of internal auditory system
$\omega_{am}/2\pi$	arithmetic-mean frequency
$\omega_\nu/2\pi$	geometric-mean frequency
$\Delta\omega_{mk}$	$\gamma\omega_m-\epsilon\omega_k$
$\square$	rectangular function

Subscripts:

cr	in critical bandwidth
e	at eardrum
i	imaginary part
k	integer denoting channel or tone in channel

$L=1$	refers to condition when loudness equals one sone
m,n	integer denoting tone
max	maximum
r	real part
t	loudness threshold
α	integer denoting subcritical frequency band
μ,κ	integer denoting critical band
ν,ρ	integer denoting tone

Superscript:

$*$	complex conjugate

APPENDIX B
SPECTRAL REPRESENTATION OF OUTPUT POWER-SPECTRAL-DENSITY W_q

The output power-spectral-density

$$W_q(\omega,t,T) = (1/Tp_{ref}^2)\int_{-\infty}^{\infty} e^{-i\omega\tau}d\tau \int_{t-T}^{t} q(t')q(t'+\tau)dt' \quad (12)$$

may be expressed as a function of frequency rather than of time by applying the time convolution theorem (ref. 82) to the convolution of q given by equation (5). Thus,

$$\begin{aligned} q(t') &= \int_{-\infty}^{\infty} h(t'')p(t'-t'')dt'' \qquad (5)\\ &= (1/2\pi)\int_{-\infty}^{\infty} H(\omega'')P(\omega'')e^{i\omega''t'}d\omega''\\ &= (1/2\pi)\int_{-\infty}^{\infty} H^*(\omega'')P^*(\omega'')e^{-i\omega''t'}d\omega'' \end{aligned}$$

where

$$H(-\omega) = H^*(\omega) \qquad \text{and} \qquad P(-\omega) = P^*(\omega)$$

because $h(t)$ and $p(t)$ are real. If the theorem is also applied to $q(t'+\tau)$, then

$$\begin{aligned} W_q(\omega,t,T) &= (1/Tp_{ref}^2)\int_{-\infty}^{\infty} e^{-i\omega\tau}d\tau \int_{t-T}^{t} dt'[(1/2\pi)\\ &\quad \int_{-\infty}^{\infty} H^*(\omega'')P^*(\omega'')e^{-i\omega''t'}d\omega'' \ (1/2\pi)\\ &\quad \int_{-\infty}^{\infty} H(\omega')P(\omega')e^{i\omega'(t'+\tau)}d\omega']\\ &= (1/2\pi Tp_{ref}^2)H(\omega)P(\omega)\int_{-\infty}^{t} e^{-i(\omega''-\omega)t'}dt'\\ &\quad \int_{t-T}^{t} e^{-i(\omega''-\omega)t'}dt' \end{aligned}$$

because

$$\int_{-\infty}^{\infty} e^{-i(\omega-\omega')\tau}d\tau = 2\pi\delta(\omega-\omega')$$

However,

$$\int_{t-T}^{t} e^{-i(\omega''-\omega)t'}dt' = \frac{\sin[(\omega''-\omega)T/2]}{(\omega''-\omega)/2}e^{-i(\omega''-\omega)(t-T/2)}$$

so that

$$\begin{aligned} W_q(\omega,t,T) &= (1/2\pi p_{ref}^2)H(\omega)P(\omega)\\ &\quad \int_{-\infty}^{\infty} H^*(\omega'')P^*(\omega'')\frac{\sin[(\omega''-\omega)T/2]}{(\omega''-\omega)T/2}e^{-i(\omega''-\omega)(t-T/2)}d\omega'' \end{aligned} \quad (13)$$

Assume that $H(\omega'')$ and $P(\omega'')$ are constant in the range of ω'' for which the integrand in this equation contributes appreciably to the integral. This range is approximately

$$2|\omega''-\omega|T/2<2\pi$$

or

$$2|\omega''-\omega|/2\pi<2/T\approx 20 \text{ hertz}$$

since

$$\sin(\pi\xi)/\pi\xi \equiv \operatorname{sinc}\xi$$

is small unless $|\xi|<1$. Let $\nu\equiv\omega''-\omega$. Then,

$$\begin{aligned} W_q(\omega,t,T) &= (1/2\pi p_{ref}^2)\,|H(\omega)|^2|P(\omega)|^2 \int_{-\infty}^{\infty} \operatorname{sinc}(\nu T/2\pi)\, e^{-i\nu(t-T/2)}\, d\nu \\ &= (1/Tp_{ref}^2)\,|H(\omega)|^2|P(\omega)|^2 \int_{-\infty}^{\infty} \operatorname{sinc}\xi\, e^{-2\pi i\xi[(t/T)-1/2]}\, d\xi \end{aligned}$$

where

$$\xi\equiv\nu T/2\pi$$

Finally, by virtue of the integration shown in reference 83,

$$W_q(\omega,t,T)=\square\left(\frac{t}{T}-\frac{1}{2}\right)|H(\omega)|^2|P(\omega)|^2/Tp_{ref}^2$$

where

$$\square\left(\frac{t}{T}-\frac{1}{2}\right)\equiv\begin{cases} 1 & \text{if } \left|\frac{t}{T}-\frac{1}{2}\right|<\frac{1}{2} \quad \text{or} \quad 0<t<T \\ 0 & \text{if } \left|\frac{t}{T}-\frac{1}{2}\right|>\frac{1}{2} \quad \text{or} \quad 0>t>T \end{cases}$$

is the rectangular function.

APPENDIX C
LOCATION OF CRITICAL BANDS

Since the loci of the mean frequencies and cutoff frequencies of the channels, or critical bands, depend on the sound spectrum, a method for choosing these loci must be established. Probably the most realistic procedure is to locate the geometric-mean frequency of a critical band so that the maximum possible spectral power falls within one critical band. Then, the loci of all other critical bands can be determined relative to this reference band using bandwidths found from figure 3, or 7. A simpler alternative is to locate the geometric-mean frequency of the reference band at the frequency of the maximum spectral density of the sound. An even simpler alternative is to let the upper cutoff frequency of the uppermost critical band coincide with the upper cutoff frequency of the measured spectrum and then to determine the loci of all other critical bands relative to this reference band. This last procedure avoids the problem of choosing the reference band when the maximum spectral density occurs at several nonadjacent frequencies and was used, where appropriate, in all calculations reported herein.

APPENDIX D
CONCERNING BEATS OF MISTUNED CONSONANCES

Beats of mistuned consonances occur if

$$\gamma\omega_m \approx \epsilon\omega_k$$

where γ and ϵ are positive integers, at least one of which is not unity. Let

$$\omega_m = (\epsilon/\gamma)\,\omega_k + (\Delta\omega_{mk}/\gamma)$$

where $\Delta\omega_{mk} << \omega_k$. Then,

$$\omega_m - \omega_k = [(\epsilon/\gamma) - 1]\,\omega_k + (\Delta\omega_{mk}/\gamma) \tag{D1}$$

Consider the integral of the cosine in equation (37). When equation (D1) is introduced in this integral, it becomes

$$\begin{aligned}(1/T)\ \int_{t-T}^{t} \cos\left\{\left[\left(\frac{\epsilon}{\gamma}-1\right)\omega_k + \frac{\Delta\omega_{mk}}{\gamma}\right]t' + \Phi\right\} dt' \\ = (1/T)\ \int_{t-T}^{t}\left\{\cos\left[\left(\frac{\epsilon}{\gamma}-1\right)\omega_k t' + \Phi\right]\cos\left(\frac{\Delta\omega_{mk}}{\gamma}t'\right)\right. \\ \left. -\sin\left[\left(\frac{\epsilon}{\gamma}-1\right)\omega_k t' + \Phi\right]\sin\left(\frac{\Delta\omega_{mk}}{\gamma}t'\right)\right\} dt'\end{aligned}$$

where, for brevity,

$$\Phi \equiv \phi_m - \phi_k - \theta_{mk}$$

Over each cycle of $(\epsilon/\gamma - 1)\,\omega_k$ only one-half cycle includes discharges. Also, since $(\epsilon/\gamma - 1)\,\omega_k >> \Delta\omega_{mk}/\gamma$, the terms involving $\Delta\omega_{mk}$ are approximately constant over the cycle. Therefore, over each cycle of $(\epsilon/\gamma - 1)\,\omega_k$, with period Δt, the contribution to the preceding integral becomes

$$\begin{aligned}(1/\Delta t)\int_{t_{max}-\Delta t/4}^{t_{max}+\Delta t/4}\left\{\cos\left[\left(\frac{\epsilon}{\gamma}-1\right)\omega_k t' + \Phi\right]\cos\left(\frac{\Delta\omega_{mk}}{\gamma}t'\right)\right. \\ \left. +\cos\left[\left(\frac{\epsilon}{\gamma}-1\right)\omega_k t' + \Phi + \frac{\pi}{2}\right]\sin\left(\frac{\Delta\omega_{mk}}{\gamma}t'\right)\right\} dt' \\ = (1/2)\,\mathrm{sinc}\left[\left(\frac{\epsilon}{\gamma}-1\right)\omega_k \Delta t/4\pi\right]\cos\left(\frac{\Delta\omega_{mk}}{\gamma}t_c\right)\end{aligned}$$

where

$$t_{max} \equiv -\frac{\phi_m - \phi_k - \theta_{mk}}{\left(\frac{\epsilon}{\gamma} - 1\right)\omega_k}$$

and t_c is a characteristic time denoting the particular cycle. Since

$$\Delta t \equiv \frac{2\pi}{\left(\frac{\epsilon}{\gamma} - 1\right)\omega_k}$$

the integrated result becomes

$$(1/\pi)\cos\left(\frac{\Delta\omega_{mk}}{\gamma} t_c\right)$$

This is valid for each cycle of $(\epsilon/\gamma - 1)\omega_k$. To be consistent with other terms in equation (37), which are treated as full-wave rectified, this result should be multiplied by 2. The result is, then,

$$(2/\pi)\cos\left(\frac{\Delta\omega_{mk}}{\gamma} t_c\right)$$

In time T there are $\sigma + \eta$ cycles of $(\epsilon/\gamma - 1)\omega_k$, where σ is an integer, $0 \leq \eta < 1$, and $\sigma >> \eta$. Therefore, the summation of the preceding expression over the entire interval $t - T \leq t' < t$ yields

$$\begin{aligned}
&\frac{2}{\pi(\sigma+\eta)\Delta t_c} \sum \cos\left(\frac{\Delta\omega_{mk}}{\gamma} t_c\right) \Delta t_c \\
&\approx \frac{2}{\pi(\sigma+\eta)\Delta t_c} \int_{t-T}^{t} \cos\left(\frac{\Delta\omega_{mk}}{\gamma} t_c\right) dt_c \\
&= (2/\pi)\,\mathrm{sinc}\left(\frac{\Delta\omega_{mk}}{2\pi\gamma} T\right) \cos\left[\frac{\Delta\omega_{mk}}{\gamma}(t - T/2)\right]
\end{aligned}$$

where the sum may be approximated by the integral because $\Delta t_c << T$.

APPENDIX E
PLANCHEREL'S THEOREM FOR LIMITED INTEGRATION TIME

Define the running-interval Fourier transform

$$P(\omega,t,T)=\int_{t-T}^{t} p(t')e^{-i\omega t'}\,dt'$$

Hence,

$$P(\omega,t,T)P^*(\omega,t,T)=\int_{t-T}^{t} p(t')e^{-i\omega t'}\,dt'\cdot\int_{t-T}^{t} p^*(t'')e^{i\omega t''}dt''$$

so that

$$\begin{aligned}\int_{-\infty}^{\infty}|P(\omega,t,T)|^2d\omega&=\int_{t-T}^{t} p(t')\,dt'\int_{t-T}^{t} p^*(t'')\,dt''\cdot\int_{-\infty}^{\infty} e^{i\omega(t''-t')}d\omega\\&=2\pi\int_{t-T}^{t} p(t')\,dt'\int_{t-T}^{t} p^*(t'')\,\delta(t'-t'')\,dt''\\&=2\pi\int_{t-T}^{t} p^2(t')\,dt'\end{aligned}$$

because $p(t)$ is real. Therefore, by virtue of equation (43),

$$(1/2\pi)\int_{-\infty}^{\infty} W(\omega,t,T)\,d\omega=(1/Tp_{ref}^2)\int_{t-T}^{t} p^2(t')\,dt' \qquad (46)$$

where

$$|P(\omega,t,T)|^2=\square\left(\frac{t}{T}-\frac{1}{2}\right)|P(\omega)|^2$$

APPENDIX F
EXTERNAL TRANSMITTANCE FUNCTION

The external ear resembles a closed, flanged tube. The transmittance function for a closed tube is known (ref. 84). Unfortunately, the theoretical result does not agree with experiment unless the former is modified. If the theoretical formula is squared, then

$$\mathcal{T}(\omega)=(A^2+B^2)^{-1} \tag{F1}$$

where

$$A\equiv\left[1-\frac{J_1(2r\omega/c)}{2r\omega/c}\right]^2\sin^2(\omega l/c)$$

$$B\equiv\left[\frac{2}{(2r\omega/c)^2}K_1(2r\omega/c)\sin(\omega l/c)-\cos(\omega l/c)\right]^2$$

J_1 is the first order Bessel function of the first kind, and K_1 is a modified Bessel function defined by

$$K_1(\xi)=\frac{2}{\pi}\sum_{n=1}^{\infty}(-)^{n+1}\frac{\xi^{2n+1}}{(2n+1)\left[(2n-1)!!\right]^2}$$

where

$$(2n-1)!!=(2n-1)(2n-3)\cdots5\cdot3\cdot1$$

Also, r is the radius, l is the length of the tube (auditory meatus), and c is the speed of sound in air. The empirical and experimental functions $T(\omega)$ are compared in figure 9. The calculated results were obtained by assuming that $2r/c=2.035\times10^{-5}$ second, and $l/c=4.542\times10^{-5}$ second.

For free-field listening the real part of the external cross-transmittance $\mathcal{T}(\omega_m,\omega_n)$ occurs in those terms of the loudness equation which involve spectrum interactions. To evaluate $\mathcal{T}_r(\omega_m,\omega_n)$, with A and B defined as before, let

$$A\equiv A_m,\ B\equiv B_m \text{ if } \omega=\omega_m$$

$$A\equiv A_n,\ B\equiv B_n \text{ if } \omega=\omega_n$$

Let

$$\mathcal{H}_e^*(\omega_m) = (A_m - iB_m)^{-1}$$

$$\mathcal{H}_e(\omega_n) = (A_n + iB_n)^{-1}$$

Then,

$$\mathcal{J}_r(\omega_m, \omega_n) = \frac{A_m A_n + B_m B_n}{(A_m^2 + B_m^2) \quad (A_n^2 + B_n^2)}$$

which reduces to equation (F1) if $m = n$.

REFERENCES

1. SCHARF, BERTRAM: Loudness and Noisiness—Same or Different? Inter-Noise 74. John C. Snowdon, ed., Inst. Noise Control Eng., 1974, pp. 559–564.
2. ANGEVINE, OLIVER L.: Individual Differences in the Annoyance of Noise. Sound and Vibration, vol. 9, no. 11, Nov. 1975, pp. 40–42.
3. NIESE, H.: Beitrag zur Relation Zwischen Lautstärke und Lästigkeit von Geräuschen. Acustica, vol. 15, no. 2, 1965, pp. 236–243.
4. BERANEK, L. L.; ET AL.: Calculation and Measurement of the Loudness of Sounds. J. Acoust. Soc. Am., vol. 23, no. 3, May 1951, pp. 261–269.
5. STEVENS, S. S.: Calculation of the Loudness of Complex Noise. J. Acoust. Soc. Am., vol. 28, no. 5, Sept. 1956, pp. 807–832.
6. STEVENS, S. S.: 1971 Rayleigh Gold Medal Address: Calculating the Perceived Level of Light and Sound. J. Sound Vib., vol. 23, no. 3, Aug. 1972, pp. 297–306.
7. U.S.A. Standard Procedure for Computation of Loudness of Noise. USAS S3.4–1968, United States of America Standards Inst., Inc., Mar. 26, 1968.
8. FLANAGAN, JAMES L.: Speech Analysis Synthesis and Perception. Ch. 4, Academic Press, Inc., 1965, pp. 75–118.
9. SIEBERT, WILLIAM M.: Stimulus Transformations in the Peripheral Auditory System. Recognizing Patterns. Studies in Living and Automatic Systems. Paul A. Kolers and Murray Eden, eds. Mass Inst. Tech. Press, Cambridge, Mass., 1968, Chap. 4, pp. 104–133.
10. HOWES, W. L.: Loudness Function Derived from Data on Electrical Discharge Rates in Auditory Nerve Fibers. Acustica, vol. 30, no. 5, 1974, pp. 247–259.
11. KATSUKI, Y.; SUGA, N.; AND KANNO, Y.: Neural Mechanism of the Peripheral and Central Auditory System in Monkeys. J. Acoust. Soc. Am., vol. 34, no. 8, pt. 2, Sept. 1962, pp. 1396–1410.
12. ROSE, JERZY E.; ET AL.: Some Effects of Stimulus Intensity on Response of Auditory Nerve Fibers in the Squirrel Monkey. J. Neurophys., vol. 34, no. 4, 1971, pp. 685–699.
13. FLETCHER, HARVEY; AND MUNSON, W. A.: Loudness, Its Definition, Measurement and Calculation. J. Acoust. Soc. Am., vol. 5, no. 2, Oct. 1933, pp. 82–108
14. STEVENS, S. S.; AND DAVIS, H.: Psychophysiological Acoustics: Pitch and Loudness. J. Acoust. Soc. Am., vol. 8, no. 1, July 1936, pp. 1–13.
15. WEVER, ERNEST GLEN: Theory of Hearing. Dover Publ., Inc., 1970.
16. DERBYSHIRE, A. J.; AND DAVIS, H.: The Action Potentials of the Auditory Nerve. Am. J. Physiology, vol. 113, 1935, pp. 476–504.
17. WEVER, ERNEST GLEN; AND BRAY, C. W.: The Nature of Acoustic Response: The Relation Between Sound Frequency and Frequency of Impulse in the Auditory Nerve. J. Exptl. Psych., vol. 13, no. 5, Oct. 1930, pp. 373–387.

18. Ruggero, Mario A.: Response to Noise of Auditory Nerve Fibers in the Squirrel Monkey. J. Neurophys., vol 36, 1973, pp. 569–587.
19. Robinson, D. W.; and Dadson, R. W.: A Re-Determination of the Equal-Loudness Relations for Pure Tones. Brit. J. Appl. Phys., vol. 7, no. 5, May 1956, pp. 166–181.
20. Normal Equal-Loudness Contours for Pure Tones and Normal Threshold of Hearing Under Free Field Listening Conditions. ISO/R 226, International Organization for Standardization. Dec. 1961.
21. Howes, Walton L.: Relations Among Pure-Tone Sound Stimuli, Neural Activity, and the Loudness Sensation. NASA TN D-6889, 1972.
22. Howes, Walton L.: Neural Coding of High-Frequency Tones. NASA TM X-3374, 1976.
23. Rose, Jerzy E.; et al.: Phase-Locked Response to Low-Frequency Tones in Single Auditory Nerve Fibers of the Squirrel Monkey. J. Neurophys, vol. 30, 1967, pp. 769–793.
24. Brugge, John F.; et al.: Time Structure of Discharges in Single Auditory Nerve Fibers of the Squirrel Monkey in Response to Complex Periodic Sounds. J. Neurophys., vol. 32, no. 3, 1969, pp. 386–401.
25. Wever, Ernest G.: Beats and Related Phenomena Resulting from the Simultaneous Sounding of Two Tones—I. Psych. Rev., vol. 36, 1929, pp. 402–418.
26. Wever, Ernest G.: Beats and Related Phenomena Resulting from the Simultaneous Sounding of Two Tones—II. Psych. Rev., vol. 36, 1929, pp. 512–523.
27. Zwicker, E.; and Feldtkeller, R.: Über die Lautstärke von Gleichförmigen Geräuschen. Acustica, vol. 5, no. 6, 1955, pp. 303–316.
28. Zwicker, E.: Subdivision of the Audible Frequency Range into Critical Bands (Frequenzgruppen). J. Acoust. Soc. Am., vol. 33, no. 2, Feb. 1961, p. 248.
29. Plomp, R.: The Ear as a Frequency Analyzer. J. Acoust. Soc. Am., vol. 36, no. 9, Sept. 1964, pp. 1628–1636.
30. Rose, Jerzy E.; et al.: Some Possible Neural Correlates of Combination Tones. J. Neurophys, vol. 32, no. 3, 1969. pp. 402–423.
31. Schouten, J. F.: The Perception of Pitch. Phillips Tech. Rev., vol. 5, no. 10, Oct. 1940, pp. 286–294.
32. Stevens, Stanley S.; and Davis, Hallowell: Hearing, Its Psychology and Physiology. John Wiley & Sons, Inc., 1938.
33. Kinsler, Lawrence E.; and Frey, Austin R.: Fundamentals of Acoustics. John Wiley & Sons, Inc., 1950.
34. Kiang, Nelson Yuan-Sheng: Discharge Patterns of Single Fibers in Cat's Auditory Nerve. Res Mono. 35, Mass. Inst. Tech. Press, 1965, p. 111f.
35. Von Békésy, Georg: Experiments in Hearing. McGraw-Hill Book Co., Inc., 1960, p. 235.
36. Scharf, Bertram: Critical Bands. Foundations of Modern Auditory Theory, Vol. I., Ch. 5, Jerry V. Tobias, ed., Academic Press 1970, pp. 157–202.
37. Bauch H.: Die Bedeutung der Frequenzgruppe für die Lautheit von Klangen. Acustica, vol. 6, no. 1, 1956, pp. 40–45.,
38 Zwicker, E ; Flottorp, G.; and Stevens, S. S.: Critical Band Width in Loudness Summation. J. Acoust. Soc. Am., vol. 29, no 5, May 1957, pp. 548–557.
39. Small, Arnold M.: Periodicity Pitch. Foundations of Modern Auditory Theory, Vol. 1, Ch. 1, Jerry V. Tobias, ed., Academic Press, 1970, pp 1–54.

40. SCHOUTEN, J. F.: RITSMA, R. J.; AND LOPES CARDOZO, B.: Pitch of the Residue. J. Acoust. Soc. Am., vol. 34, no. 8, pt. 2, Sept. 1962, pp. 1418–1424.
41. WEVER, ERNEST GLEN; AND LAWRENCE, MERLE: Physiological Acoustics. Princeton Univ. Press, 1954, pp. 130–144.
42. ROSE, JERZY E.; ET AL.: Observations on Phase Sensitive Neurons of Anteroventral Cochlear Nucleus of the Cat: Nonlinearity of Cochlear Output. J. Neurophys., vol. 37, no. 1, Jan. 1974, pp. 218–253.
43. SCHARF, BERTRAM: Critical Bands and the Loudness of Complex Sounds Near Threshold. J. Acoust. Soc. Am., vol. 31, no. 3, Mar. 1959, pp. 365–370.
44. PLOMP, R.; AND MIMPEN, A. M.: The Ear as a Frequency Analyzer. II. J. Acoust. Soc. Am., vol 43, no. 4, 1968, pp. 764–767.
45. WIENER, FRANCIS M.; AND ROSS, DOUGLAS A.: The Pressure Distribution in the Auditory Canal in a Progressive Sound Field. J. Acoust Soc. Am., vol. 18, no. 2, Oct. 1946, pp. 401–408
46. ROBINSON, D. W.; WHITTLE, L. S.; AND BOWSHER, J. M.: The Loudness of Diffuse Sound Fields. Acustica, vol. 11, no. 6, 1961, pp. 397–404.
47. ANGELO, ERNEST J., JR.: Electronic Circuits. 2nd ed., McGraw–Hill Book Co., Inc., 1964, p. 503.
48. JOLLEY, LEONARD B. W.: Summation of Series. Dover Publications, Inc., 1961, p. 4.
49. VON BÉKÉSY, GEORG: Sensory Inhibition. Princeton Univ. Press, 1967, p. 116.
50. ZWICKER, E.: Lautstärkeberechnungsverfahren im Vergleich. Acustica, vol. 17, no. 5, 1966, pp. 278–284.
51. WARREN, RICHARD M.: Elimination of Biases in Loudness Judgments for Tones. J. Acoust. Soc. Am., vol. 48, no. 6, pt. 2, Dec. 1970, pp. 1397–1403.
52. JONES, JESS H.: Perceived Sound and the Frequency Response Characteristics of the Human Auditory System. "Progress of NASA Research Relating to Noise Alleviation of Large Subsonic Jet Aircraft" NASA SP–189, 1968.
53. PETERSON, ARNOLD P. G.; AND GROSS, ERVIN E., JR.: Handbook of Noise Measurement. 7th ed. General Radio Co., 1972.
54. STEVENS, S. S.: Procedure for Calculating Loudness: Mark VI. J. Acoust. Soc. Am., vol. 33, no. 11, Nov. 1961, pp. 1577–1585.
55. STEVENS, S. S.: Perceived Level of Noise by Mark VII and Decibels (E). J. Acoust. Soc. Am., vol. 51, no. 2, pt. 2, 1972, pp. 575–601.
56. HOWES, WALTON L.: Measuring Loudness. Phys. Today, vol. 28, no. 8, Aug. 1975, p. 13.
57. SMALL, ARNOLD M., JR.; AND THURLOW, WILLARD R.: Loudness Relations in Two–Component Tones. J. Acoust. Soc. Am., vol. 26, no. 3, May 1954, pp. 381–388.
58. HOWES, DAVIS H.: The Loudness of Multicomponent Tones. Am. J. Psych., vol. 63, no. 1, Jan. 1950, pp. 1–30.
59. SCHARF, BERTRAM: Loudness Summation and Spectrum Shape. J. Acoust. Soc. Am., vol. 34, no. 2, Feb. 1962, pp. 228–233.
60. QUIETZSCH, G.: Objektive und Subjektive Lautstärkemessungen. Acustica, vol. 5, no. 1, 1955, pp. 49–66.
61. SCHARF, BERTRAM: Loudness of Complex Sounds as a Function of the Number of Components. J. Acoust. Soc. Am., vol. 31, no. 6, June 1959, pp. 783–785
62. NIESE, H.: Die Lautstärkebildung bei binauralem Hören komplexer Geräusche Hochfrequenztechn. und Elektroakustik, vol 70, no. 4, 1961, pp. 132–141.

63. Robinson, D. W.; and Whittle, L. S.: The Loudness of Octave Bands of Noise. Acustica, vol. 14, no. 1, 1964, pp. 24–35.
64. Cremer, L.; Plenge, G.; and Schwarzl, D.: Kurven Gleicher Lautstärke mit Octav-Gefiltertem Rauschen im Diffusen Schallfeld. Acustica, vol. 9, no. 2, 1959, pp. 65–75.
65. Jahn, G.: Kurven Gleicher Lautstärke für Octavbandpassrauschen. Hochfrequenztechn. und Elektroakust., vol. 67, no. 6, 1959, pp. 187–189.
66. Zwicker, E.: Lautstärkeberechnungsverfahren im Vergleich. Acustica, vol. 17, no. 5, 1966, pp. 278–284.
67. Feldtkeller, R.: Die Kurven konstanter Lautstärke für Dauertöne und für einzelne Druckimpulse. Frequenz, vol. 10, no. 11, 1956, pp. 356–358.
68. Pollack, Irwin: On the Measurement of the Loudness of White Noise. J. Acoust. Soc. Am., vol. 23, no. 6, Nov. 1951, pp. 654–657.
69. Miller, George A.: Sensitivity to Changes in the Intensity of White Noise and Its Relation to Masking and Loudness. J. Acoust. Soc. Am., vol. 19, no. 4, July 1947, pp. 609–619
70. Brittain, F. H.: The Loudness of Continuous Spectrum Noise and Its Application to Loudness Measurements. J. Acoust. Soc. Am., vol. 11, July 1939, pp. 113–117.
71. Niese, H.; and Köhler, J.: Subjektive Messungen über den Einfluss der Spitzenhaltigkeit von Bandpassrauschen auf die Empfindung der Lautstärke. Hochfrequenztechn. und Elektroakustik, vol. 66, no. 5, 1958, pp. 150–160.
72. Lübcke, E.; Mittag, G.; and Port, E.: Subjektive und Objektive Bewertung von Maschinengeräuschen. Acustica, vol. 14, no. 2, 1964, pp. 105–114.
73. Jahn, M.: Subjektive und Objektive Bewertung von Maschinengeräuschen. Acustica, vol. 16, no. 3, 1965–66, pp. 175–186.
74. Warren, Richard M.: Quantification of Loudness. Am. J. Psych., vol. 86, no. 4, 1973, pp. 807–825.
75. Corliss, Edith L. R.; and Winzer, George E : Study of Methods for Estimating Loudness. J. Acoust. Soc. Am., vol. 38, no. 3, 1965, pp. 424–428.
76. Robinson, D. W.; and Bowsher, J. M.: A Subjective Experiment with Helicopter Noises. J. Roy. Aero Soc., vol. 65, no. 609, Sept. 1961, pp. 635–637.
77. Copeland, W. C. T.; et al.: A Controlled Experiment on the Subjective Effects of Jet Engine Noise. J. Roy. Aero. Soc., vol. 64, no. 589, Jan. 1960, pp. 33–36.
78. Young, Robert W.; and Peterson, Arnold: On Estimating Noisiness of Aircraft Sounds. J. Acoust. Soc. Am., vol. 45, no. 4, 1969, pp. 834–838.
79. Howes, W. L.: Loudness Determined by Power Summation. Acustica, vol. 25, no. 5, 1971, pp. 343–349.
80. Scholl, H.: Das dynamische Verhalten des Gehörs bei der Unterteilung des Schallspektrums in Frequenzgruppen. Acustica, vol. 12, no. 1, 1962, pp. 101–107.
81. Scharf, B.: Loudness and Frequency Selectivity at Short Durations. Frequency Analysis and Periodicity Detection in Hearing. R. Plomp and G. F. Smoorenburg, eds., A. W. Sijtoff, Leiden, The Netherlands, 1970, pp. 455–462.
82. Papoulis, Athanasios: The Fourier Integral and Its Applications. McGraw–Hill Book Co., Inc., 1962, p. 26.
83 Bracewell, Ron: The Fourier Transform and Its Applications. McGraw–Hill Book Co., Inc., 1965, p. 129.
84 Morse, Phillip M.: Vibration and Sound. 2nd ed., McGraw–Hill Book Co., Inc., 1948, p. 259.

85. RITSMA, R. J.: Existence Region of the Tonal Residue. I. J. Acoust. Soc. Am., vol. 34, no 9, Sept. 1962, pp. 1224–1229.
86. FLETCHER, HARVEY: Speech and Hearing in Communication. 2nd ed, D. van Nostrand Co., Inc., 1953, p. 271.

1 Report No NASA RP-1001	2 Government Accession No	3 Recipient's Catalog No
4. Title and Subtitle OVERALL LOUDNESS OF STEADY SOUNDS ACCORDING TO THEORY AND EXPERIMENT		5 Report Date October 1977 6 Performing Organization Code
7. Author(s) Walton L. Howes		8 Performing Organization Report No E-8342 10 Work Unit No
9 Performing Organization Name and Address National Aeronautics and Space Administration Lewis Research Center Cleveland, Ohio 44135		11 Contract or Grant No
12 Sponsoring Agency Name and Address National Aeronautics and Space Administration Washington, D C. 20546		13 Type of Report and Period Covered Reference Publication 14 Sponsoring Agency Code
15 Supplementary Notes		

16 Abstract

A mathematical theory for calculating the loudness of any steady sound from information on its spectrum is constructed from physical principles and psychological and physiological information on mammalian auditory systems. The theory involves filtering, channeling, squaring, half-wave rectification, and time averaging of the signal. The theory accounts for critical bands for loudness, audibility of sounds consisting of subliminal components, audible beats, periodicity pitch, and pitch of the residue These and other psychoacoustic phenomena are explained in terms of electrical activity in the peripheral nervous system. Simple approximations for loudness are derived from the more exact formulas. Loudness predictions are compared with a wide variety of published loudness judgment data with considerable success.

17. Key Words (Suggested by Author(s)) Loudness loudness level, noise threshold, threshold (perception), subliminal stimuli, tone, Psychoacoustics · auditory perception, auditory stimuli, hearing, sound perception; Psychophysiology. ear, auditory nerve, bioelectricity, neurons, central nervous system, brain, peripheral nervous system, electrophysiology, Acoustics: audiofrequencies, noise (sound), noise intensity, sound intensity, noise measurement; sound pressure, Fourier analysis Fourier transformation, transfer functions, bandwidth, beat frequencies, harmonics, Fourier series, power spectra, autocorrelation, broadband, frequency distribution, frequencies, frequency ranges, periodic functions.		18 Distribution Statement Unclassified - unlimited STAR Category 53	
19 Security Classif (of this report) Unclassified	20 Security Classif (of this page) Unclassified	21 No of Pages	22 Price*

* For sale by the National Technical Information Service, Springfield Virginia 22161

☆ U.S GOVERNMENT PRINTING OFFICE 1979 O—246-540