Karolinska Institutet
Technical Audiology
KTH, 100 44 Stockholm 70

Report TA No. 83
June 1976

DETECTION OF AMPLITUDE DISTORTION BY NORMAL HEARING AND HEARING IMPAIRED SUBJECTS

Alf Gabrielsson, Technical Audiology, Karolinska Institutet, Stockholm
Per-Olov Nyberg, Department of Audiology, Sahlgrenska Sjukhuset, Göteborg
Håkan Sjögren, Technical Audiology, Karolinska Institutet, Stockholm
Lena Svensson, Department of Audiology, Sahlgrenska Sjukhuset, Göteborg

ABSTRACT

Detection of amplitude distortion in excerpts of music and speech and in a sinusoidal stimulus was studied for twelve normal hearing subjects and twelve subjects suffering from sensory-neural hearing loss. Two preliminary experiments were made to find individual "most comfortable levels" and to select appropriate distortion levels, respectively. In the main experiment individual threshold values for quadratic and cubic amplitude distortion were studied by means of the constant method. The threshold values were lowest for the sinusoidal and considerably higher for the music and speech stimuli. The hearing impaired subjects in general had higher threshold values than the normal hearing subjects. There was a big inter-individual variation in threshold values (as well as in "most comfortable levels") within each group. On the basis of the results recommendations for maximum allowed amplitude distortion in hearing aids are suggested.

(Note: This report appears simultaneously as report 77:105 in "Forskningsrapporter från Audiologiska Avdelningen, Sahlgrenska sjukhuset, Göteborg".)

INTRODUCTION

An important goal for construction of amplifier systems, for instance hearing aids, is that the distortions arising in such systems should not affect the perceived sound quality. The distortions should thus be so small that they are not detected by the listener.

There are different types of distortion: linear distortion, amplitude distortion, phase distortion, and transient distortion. This report deals only with amplitude distortion.

Amplitude distortion is described by either harmonic and/or intermodulation distortion. For amplifiers, earphones, magnetic tapes, hearing aids and gramophone records the system transfer function can often be described by a power series of the input signal:

$$\text{output} = \sum_n \text{coeff}_n \cdot (\text{input})^n$$

where "coeff" are distortion coefficients.

The general case, however, is far more complex as distortion is frequency dependent and as feedback normally is used. Furthermore, the nonlinearity of the transfer function may result in what is called center distortion, which means that the system is approximately linear for high signal levels but non-linear for low levels.

It is easily shown that center distortion deteriorates the signal in a way that the sound becomes very unclear and muddy even if the amount of distortion (given as harmonic or intermodulation distortion) is very small.

When discussing the perception of amplitude distortion with complex program material, it is necessary to know the type of nonlinearity. This can be documented either by the transfer function itself at several discrete frequencies, or by the harmonic and/or intermodulation distortion as a function of the output level.

The results in this paper are valid only for power series distortion. Information on the connection between harmonic and intermodulation distortion is given in Appendix A.

The detection of center distortion is excluded, because measurements on all hearing aid types with amplifiers class B-type accepted for the state grant in Sweden 1974 have shown that the amount of center distortion was negligible.

The purpose of the present work is therefore to investigate the detectability of power series amplitude distortion with special regard to its importance for hearing aids. The results should give some information about which maximum amount of amplitude distortion still goes unnoticed by a listener - information that then may be used for the design of amplifier systems like hearing aids.

Detection of amplitude distortion for sinusoidals and for flute and clarinet spectra was studied in an earlier report (Gabrielsson & Sjögren, 1971; 1972). From this work, as well as from many other investigations referred to in that paper, it is evident that the detectability of amplitude distortion may vary very much with the type of stimuli used. In the present investigation, therefore, the stimuli included short excerpts of music and speech and a sinusoidal for comparison with earlier results. Presuming also that the detactability is different for subjects with normal hearing and subjects suffering from hearing loss (the latter ones of special importance here), both these subject categories were included in the investigation.

METHOD

In summary the method included the following steps:
Two preliminary experiments were performed. In one of them, each subject adjusted the sound level of each of the four stimuli to a "most comfortable level". In the other preliminary experiment the detectability for two types of distortion was explored by a simple "up-and-down" method. The results from this experiment were used to select the distortion levels to be used in the main experiment.

In the main experiment the subjects listened monaurally to each stimulus presented two times in immediate succession, one time without distortion and one time with a certain type and level of distortion added. The subject should judge whether there was any difference or not between the two presentations. Threshold values for detection of the distortion were computed individually for each stimulus x distortion type condition.

Stimuli and listening conditions

Four different stimuli were used:

1. Short excerpt (about 5 seconds) from a piece of jazz music performed by Oscar Peterson's trio (piano, bass, drums). Referred to as "Music" in the following.

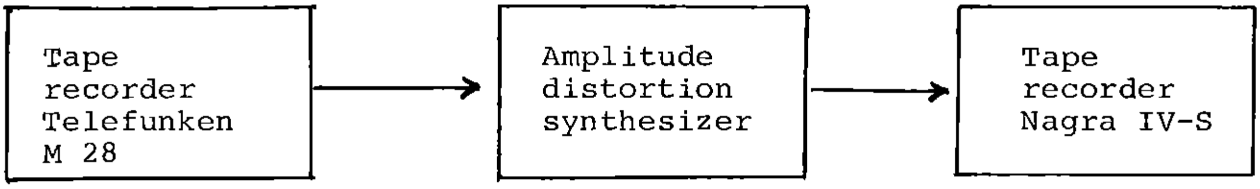
2. Short excerpt (about 5 seconds) from a recording of conversation during a coffee break in the laboratory, some 15 people chatting simultaneously. Referred to as "Chat".

3. Female voice reading a short sentence (about 3 seconds). Referred to as "Voice".

4. Sinusoidal 500 Hz, duration 1.5 seconds, generated by a Madsen OB-70 audiometer. Referred to as "Sinewave".

Submaster tapes, one for each stimulus, were produced from a master tape for the first three stimuli and directly from the audiometer for the sinusoidal. The submaster tapes were used for producing the stimulus tapes in the two preliminary experiments and in the main experiment (see further under Procedure).

The distortion of the stimuli was generated by means of an amplitude distortion synthesizer described in Appendix A. Two types of distortion were used, quadratic and cubic amplitude distortion (called X^2- and X^3-distortion, respectively), see Appendix A. In the preliminary experiment with the "up-and-down" method the distortion was added "on line" to the appropriate parts of the corresponding stimulus tape, that is: the distortion synthesizer was handled manually by the experimenter. For the main experiment, however, the appropriate type and level of distortion was recorded in advance on the stimulus tape by means of the following setup:



The peak levels were controlled by a peak level meter and further control of the recording was made by means of a storage oscilloscope. Further details about the stimulus tape for the main experiment are given under Procedure.

The subject was sitting in an audiometric booth and listened to the stimuli by means of TDH39 headphones, monaurally.

In the preliminary experiment concerning "comfortable listening level" the subject adjusted the sound level of the respective (undistorted) stimuli by means of a stepless potentiometer connected to the experimenter's gain control. The total range of the gain control was 35 dB.

All tape recordings for the main experiment were made at a speed of 38 centimeters/second. The frequency response curve for the tape was kept within ± 1 dB, 40 - 15000 Hz.

Subjects

There were 12 normal hearing subjects (designated as NH subjects in the following), seven males and five females, 23 - 50 years old (seven of them 23 - 30 years). There were also 12 subjects with sensory-neural hearing loss (designated as SN subjects), four males and eight females, 18 - 69 years old (median age 40 years). Their hearing loss is given in Table I. All subjects were paid for their participation.

Procedure

The preliminary experiments as well as the main experiment were run with one subject at a time.

a) "Most comfortable level" (MCL). The stimulus tape for this experiment included 20 presentations in sequence of each of the four stimuli (undistorted). The subject was instructed to "adjust the loudness so as to be comfortable, that is, so that the music, or the speech, or the tone sounds comfortably loud ... the loudness should neither be too soft nor too loud but comfortable." (The word "comfortable" is an approximation to the Swedish word "lagom" which means something that is neither too much, nor too little.) The subject thus listened to the successive presentations of a certain stimulus and adjusted the loudness continuously with his potentiometer until he was satisfied. This was made four times for each of the four stimuli (in total thus 4 x 4 = 16 cases) in a balanced order. Two times the initial sound level was set at a probably too high level by the experimenter, and the other two times it was set at a probably too low level. These initial levels were chosen using information from a crude determination of each subject's "tolerance level" made in connection with the audiometric testing, and were randomized within a range of about 8 - 15 dB. The arithmetic mean of the four adjustments for each stimulus was taken as the subject's MCL for the respective stimulus and was used in the following experiments.

b) Preliminary experiment on detection of distortion. This experiment was designed to give information about which distortion levels should be included in the main experiment. The stimulus tape included at least 10 pairwise presentations of each stimulus with an interval of 1.5 seconds between the two presentations in a pair. The experimenter introduced distortion on one of the presentations (the first or the last) by means of the amplitude distortion synthesizer, and the subject was instructed to judge if there was any difference or not between the two presentations. If the subject answered "Different" the experimenter decreased the amount of distortion occurring within the next pair. If the subject answered "Same" the experimenter increased the amount of distortion in the next pair - and so on until a distortion region was found in which the subject varied between "Same" and "Different" answers, which may be taken as an indication of an approximate threshold value.

This method has been called the "up-and-down" method and is sometimes recommended for a preliminary exploration of threshold values (Guilford, 1954, p. 114). It should be followed up by a more refined method, as for example the constant method which is used in the following main experiment. The "up-and-down" method was applied here in 24 conditions (3 stimuli x 2 types of distortion x 2 different positions of distortion within a pair x 2 different initial levels of distortion, too high or too low), in a randomized order. The "Sinewave" stimulus was omitted since approximate "threshold values"for sinusoidals were available (Gabrielsson & Sjögren, 1971; 1972).

c) Main experiment on detection of distortion.

On the basis of the results from the "up-and-down" method six distortion levels were chosen for inclusion in the main experiment. Expressed as distortion coefficients (see Introduction and Appendix A) they were the following: 0.02, 0.05, 0.15, 0.30, 0.50 and 0.75. They were applied for all stimuli (however, the coefficient 0.75 was omitted for the sinusoidal), for both types of distortion, and for all subjects in order to facilitate comparisons between the results from different stimulus conditions and different subjects.

The stimulus tape included 100 pairwise stimulus presentations, each pair containing one undistorted and one distorted presentation of the respective stimulus condition. However, in eight pairs (4 stimuli x 2 distortion types) both presentations within the pair were undistorted. These eight cases were "blank trials" and were included to get some information about the "decision criterion" used by the subject as emphasized in signal detection theory (Green & Swets, 1966). The 100 pairs on the tape thus arise from the following conditions: (4 stimuli x 2 distortion types x 6 distortion coefficients x 2 positions of distortion within the pair, first or last) + 8 blank trials - 4 (the omitted distortion coefficient at the sinusoidal for two distortion types and two positions) = 100 cases.

These 100 cases appeared on the tape in a randomized order and were subdivided into blocks of ten successive cases each. The order of the ten blocks was randomized independently for each subject and each new test session.

Between the two presentations in a pair there was a silent interval of about 1.5 seconds for "Music" and "Chat", about 1.0 second for "Voice", and about 0.8 seconds for "Sinewave". The pause between successive pairs was about 5 seconds.

The instruction given to the subjects was the following:
"In this experiment we want to investigate how small differences in sound reproduction you can hear. You will listen to short sections with music or speech or a single tone. Each section is presented two times in succession like this (some demonstrations). Your task is to listen very attentively to the two reproductions of a section and compare them. If you hear any difference between the reproductions your write 'Different' (D). If you do not hear any difference between the reproductions you write 'Same' (S). You must write down the answer immediately and then concentrate on the next case which appears within some seconds. After every ten cases there is a short pause of about half a minute".

After the instruction twenty preliminary trials (two blocks from the stimulus tape) were made for practice. Each subject then listened to the whole stimulus tape five times, distributed over three experimental sessions. This means that he made ten judgments per each stimulus condition (stimulus condition = a certain stimulus with a centain type of distortion at a certain distortion level), five for the case that the distortion appeared on the first member of the pair and five for the reverse case. For certain subjects, however, this number of judgments was not enough to get reliable computations for threshold values. These subjects therefore listened another five times to the whole stimulus tape resulting in twenty judgments per stimulus condition (see further details under Results). The sound level of each stimulus presentation was continuously set by the experimenter to correspond to the MCL as determined individually in the preliminary experiment.

The psychophysical method described above corresponds to the constant method (Guilford, 1954). The number of judgments per stimulus condition is smaller than what is often recommended. However, considering the fact that each subject took part in at least eight experimental sessions (including the two preliminary experiments and the main experiment), it was not possible to extend the experiment further. To support the subjects' motivation for their task they were told in connection with the instruction that the results would be of importance for the design of hearing aids, and that it thus was necessary to keep the concentration high even though the task may seem monotonous.

RESULTS

a) "Most comfortable level" (MCL).

The adjusted MCL:s for each subject appear in Table II. Each value in this table is the arithmetic mean of the four adjustments made per stimulus, see Procedure. These values were thus used for the presentation of the stimuli in the later main experiment. (However, some subjects preferred to change one or more of their earlier adjusted levels when they started the main experiment: subjects no. 5, 7, 12, 58, 59, 61, and 62. The values given for these subjects are those actually used in the main experiment, while their originally adjusted levels are given within parentheses.)

In general the MCL:s lie higher for SN subjects than for NH subjects (compare the means over subjects in both groups). However, there is a certain overlap between the groups, and above all there is a big inter-individual variation within both groups, standard deviations are about 8 - 11 dB in the SN group and 7 - 11 dB in the NH group. (It is noted that NH subject 51 has considerably higher MCL:s than the other NH subjects, especially for "Music". This subject was the only one with active experience of sound engineering.)

On the other hand the average intra-individual standard deviation is considerably smaller (Table II, column to the right). It is computed as the average standard deviation of the four adjustments that each subject made for each of the four stimuli (averaging over stimuli), This standard deviation is roughly 2-4 dB for most of the subjects in both groups. The fact that the subjects vary very much in their "comfortable listening level" (big inter-individual variation) but may be rather stable within themselves with regard to this level (smaller intra-individual variation) is quite in line with the results from many earlier investigations on the same topic (see, for instance, Gabrielsson et al., 1974; Gabrielsson & Sjögren, 1976 and other works cited in these papers).

To analyze the data further an analysis of variance was performed on the data from each group using stimuli, subjects and "starting position" (whether the adjustments started from a level above or below the comfortable level, see Procedure) as variables in the analysis. Subjects were treated as a random variable and the analysis conforms to that for a randomized blocks factorial design, mixed model (Kirk, 1968). Besides the obvious individual differences mentioned earlier, there appeared also significant differences as regards the stimuli and the starting position. As seen in Table II the "Sinewave" stimulus was adjusted to a lower level than the first three stimuli, among which the "Voice" was set somewhat higher than the other two (see the group means in Table II). However, there was also a significant interaction between stimuli and subjects, which means that this rank order is not general but varies with different subjects which is easily seen in Table II. As regards the starting position there was a significant difference meaning that a starting position above the comfortable level resulted in a higher adjusted level than when the adjustments started from a position below the comfortable level. Even there, however, there was a significant interaction with subjects. Nearly all subjects in both groups set a higher level when they started from above than when they started from below, but the amount of the difference varied from practically nothing for some subjects up to 6 - 8 dB for some other subjects.

Although not directly related to the problems of this paper, it may be of interest to look at which correlation there is between the size of the hearing loss at different frequencies given in Table I and the sound level adjusted to represent MCL for the different stimuli (Table II). The product moment correlations between these variables appear in Table III. All correlations are positive, which means that there is in general a tendency to adjust to a higher MCL the bigger the hearing loss is - which is, of course, not surprising. However, the correlations are generally low for the frequencies 125, 4000, and 8000 Hz. They are moderately high (0.55 - 0.59) at 250 Hz, and somewhat higher at 500, 1000, and 2000 Hz (roughly 0.60 - 0.80). For all four stimuli the highest correlation occurs at 500 Hz, especially for "Music" and "Chat". Since the "Sinewave" was of the frequency 500 Hz it is natural to expect that its MCL should correlate highest with the hearing loss at 500 Hz. However, its correlation with the surrounding frequencies 250, 1000, and 2000 Hz is not so very much lower than that with 500 Hz.

If the MCL is correlated with the average hearing loss at 500-1000, 500-2000, and 250-2000 Hz (lower part of Table III), the correlations are around 0.80 for "Music", 0.75 for "Chat", 0.68-0.73 for "Voice", and 0.62-0.67 for "Sinewave". A frequency analysis of the "Music", Chat", and "Voice" stimuli would presumably show that most of their spectral contents lies within 250-2000 Hz.

b) <u>Preliminary experiment on detection of distortion</u>

Since the purpose with this preliminary experiment was mainly to give a basis for selection of distortion levels in the main experiment and since the "up-and-down" method is a rather crude method, the results from this experiment are given only in very general terms. For most subjects the "threshold values" were higher for the "Music" and "Chat" stimuli than for the other stimuli and higher for X^3-distortion than for X^2-distortion. However, there were some individual exceptions from these tendencies, and there was in both groups a considerable <u>inter</u>-individual variation with regard to "threshold value" within the same stimulus x distortion type condition. The "threshold values" were in general higher for the SN subjects than for the NH subjects. The effects of different initial levels of the distortion and different positions of the distortion were rather varying with different subjects.

c) Main experiment on detection of distortion

Individual threshold values were computed as that level of distortion which corresponds to 50 per cent "Different" judgements within the respective stimulus condition (linear interpolation process, Guilford 1954). The resulting threshold values expressed in distortion coefficient (the zero integer is omitted) appear in Table IV.

Some subjects made 20 judgements per condition (see Procedure). These subjects are designated with a (+) sign at the subject number. The remaining subjects made 10 judgements per condition.

From the detailed inspection of all individual data it was quite clear that certain threshold values should be regarded with suspicion either because there was a too big number of "false alarms" (that is, reporting a perceived difference at the blank trials) within the respective condition, or because there were obvious "irregularities" in the data (the number of "Different" judgements did not increase with increased distortion level but "reversed" one or more times). These suspect threshold values are marked with a cross (x) sign in Table IV and are excluded from the following presentation.

It is easily seen in Table IV that there is a considerable inter-individual variation in the threshold values within each stimulus condition. Consequently the mean values over subjects (given in the bottom margin for each group) are of very limited value. It is necessary to inspect the data individual for individual to see which relations are more or less general. Furthermore an analysis of variance was performed on the values in Table IV for each of the two groups. The analysis conforms to that of a randomized blocks factorial design (Kirk, 1968) with stimuli and distortion type as fixed variables and subjects as random variable (mixed model, non-additive). From these inspections and analyses the following conclusions are made:

1) There are big differences in threshold values between the different stimuli (the stimulus variable was highly significant in the analysis of variance). The lowest threshold values occur,

as expected, for the "Sinewave". Threshold values for the remaining three stimuli are considerably higher and in most cases highest for "Music" and/or "Chat". It seems that the more "complex" the stimuli are, the more difficult it is to detect the distortion.

2) As regards distortion type (X^2 versus X^3 distortion), the effects vary with stimuli and subjects in rather complex ways (significant interactions between distortion type and stimuli as well as between distortion type and subjects). If threshold values are averaged over subjects (see bottom margin for both groups in Table IV), it seems that threshold values may be higher for X^3 distortion than for X^2 distortion at the "Music" and "Chat" stimuli, while the reverse is true at the "Voice" and "Sinewave" stimuli. However, inspecting the individual threshold values reveals that these relations hold only partially for most subjects.

3) As regards the two subject categories, threshold values are in general higher for the SN subjects than for the NH subjects. However, there is more or less overlapping between the groups for all stimulus conditions (that is, some member(s) of the SN group may have lower threshold values than some member(s) of the NH group within the same stimulus condition).

4) The range of individual threshold values within each stimulus condition is wide in both groups. For the "Music", "Chat", and "Voice" stimuli the lowest individual threshold values are about 0.20 - 0.40 and the highest threshold values bigger than 0.75. In the SN group half or more of the subjects have threshold values higher than 0.75 for "Music" and "Chat", and this also occurs for 1/3 of the subjects in the NH group (at X^3 distortion). For the "Sinewave" stimulus individual threshold values vary between about 0.10 to about 0.40 in the SN group (subject 4 excepted) and between about 0.05 to about 0.25 in the NH group.

5) The percentage of "false alarms" at the blank trials varied inter-individually between 5% (subject 9) and 24% (subject 7) in the SN group with a group mean of 12%. In the NH group the variation was between 5% (subject 60) and 21% (subject 54) with a group mean of 13%. There was a moderate negative correlation (r about -0.60) between the percentage of "false alarms" and the

"mean threshold value" (averaged over stimuli and distortion types) for the subjects. That is, subjects with more "false alarms" tend to have lower threshold values than subjects with fewer "false alarms". However, the correlation is far from perfect, and detailed inspection of the individual data shows that there is a considerable intra-individual variation in the frequency of "false alarms". That is, even subjects with many "false alarms" do not distribute them uniformly over the eight stimulus conditions (4 stimuli x 2 distortion types) but concentrate them to only some of the stimulus conditions, while they may have zero "false alarms" at other stimulus conditions.

It was noted that in the NH group the frequency of "false alarms" was in general higher at the "Music" and "Chat" stimuli than for the other two stimuli. It might be that the difficulty of detecting the distortion at these two programs (and, consequently, the low frequency of "Different" judgements at these programs) "tempted" the subjects to give "Different" judgements at "blank trials" to a higher extent than at the "Voice" and "Sinewave" stimuli, where the distortion was more easily detected.

DISCUSSION

It seems apparent from the present results that detection of amplitude distortion is a function of many physical variables and interactions between them. This finding is in general agreement with the results in several earlier investigations (see Gabrielsson & Sjögren, 1971, 1972, and papers cited there; further Letowski, 1974, and Rakowski, Letowski & Barwicka, 1968). Inter-individual variations in detectability are also often reported in earlier works.

With regard to the spectrum of the stimuli it is usually said that spectra with fewer components are more sensitive to distortion. The fact that the threshold values for the sinusoidal were much lower than for the other stimuli here is in accordance with this statement. The threshold values for "Sinewave" in Table IV are a bit higher than those given in Gabrielsson & Sjögren (1972, Table I). However, the latter ones were obtained for binaural listening, while the present ones refer to monaural listening. The high threshold values for "Music", "Chat", and (to somewhat less degree) "Voice" may be explained with reference to masking effects arising from these complex spectra and to the fact that the ear ignores distortion components of short duration (as found, for instance, in the rapid transient conditions in most music; see Jacobs & Wittman, 1964).

Combining the results from Gabrielsson & Sjögren (1972) and from the present investigation it is found that the lowest threshold values occur for sinusoidals (0.02 - 0.07 in the earlier paper, 0.03 - 0.17 for most NH subjects in the present paper), the highest threshold values for music and speech stimuli (0.40 to more than 0.75 for most NH subjects at "Music", and "Chat", 0.25 - 0.55 at "Voice"), while threshold values for flute and clarinet tones lie between these extremes (0.15 - 0.45 for most subjects, see Table II in the 1972 paper). These results seem to be consistent with the principle that the more "steady-state" the stimulus is and the fewer partials it contains the more sensitive it is for detection of distortion. Conversely, the more "transient" the stimulus is and the more partials it contains, the more difficult it is to detect the distortion.

As regards the relative effects of X^2 and X^3 distortion there is in the present data a suggested interaction with the kind of stimuli (higher threshold values for X^3 distortion at "Music" and "Chat" but higher threshold values for X^2 distortion at "Voice" and "Sinewave"). An interaction between distortion type and spectrum was also found in the earlier report (Gabrielsson & Sjögren, 1972) and could there be explained with reference to different changes in the spectra of the stimuli when X^2 and X^3 distortion was added, respectively. Something similar may be the case here too, but a detailed analysis of this is very difficult to do because of the complexity of the present stimuli. Moreover, the suggested interaction is not general over all subjects.

The practical consequences of the results for design of hearing aids or other amplifier systems are not quite easy to arrive at since the detection of amplitude distortion apparently is very dependent on the type of stimuli and also varies very much with different individuals. It is thus generally necessary to continue the research using new samples of stimuli and subjects. Bearing in mind the limitations of the present study, the following points seem necessary to consider when trying to suggest limits for maximum allowed amplitude distortion.

a) On the whole, threshold values for detection of amplitude distortion in reproductions of "realistic" stimuli like music and speech seem to be rather high. The lowest threshold values occurring for certain NH subjects at the "Music", "Chat", and "Voice" stimili are about 0.05 - 0.30, for SN subjects about 0.20 - 0.35 (Table IV). As seen in Table V, the frequency of these lowest threshold values is small in both groups (about 10%), and the overwhelming majority of threshold values occur at considerably higher values (in the SN group even more than half the threshold values are higher than 0.70). Transformed to per cent harmonic distortion (Appendix A) a threshold value of 0.05 corresponds roughly to 2.5% quadratic distortion and about 1.25% cubic distortion. Threshold values of 0.70 correspond to about 35% quadratic and about 17.5%cubic distortion. Threshold values varying within about the same ranges as here are reported also for other stimuli, for instance by Jacobs & Wittman (1964) and Rakowski, Letowski & Barwicka (1968).

b) The threshold values here were obtained during concentrated listening. The subjects were "set" to observe small differences and to report any difference theycould hear (see the instruction). As in the previous report (1972) these conditions were expected to set the subject's "decision criterion" at a rather low level. The fact that the percentage of "Different" responses at the blank trials was in average as high as 12-13% may be taken as indication that this expectation was fulfilled. During less concentrated listening in "real life" and with no special attitude to observe any distortions, the threshold values might be higher than those reported here. On the other hand, it may be argued that continued listening by a hearing aid over many years will increase the sensitivity to distortion (practice effects leading to lower threshold values were observed in the 1972 report). It is also noted that the present threshold values refer to monaural listening which ought to be taken into consideration at least for NH subjects.

c) The definition of a threshold as the value at which 50% "Different" judgments are obtained is a common one but, of course, arbitrary. Other definitions may be wanted and can be applied to the present data by using Table VI, which presents the average percentage of "Different" judgments occurring at the different distortion coefficients for each of the four stimuli. By means of this table it is possible to compute differently defined threshold values, for instance, defined as corresponding to 75% "Different" judgments or to 25% "Different" judgments etc. It should be noted that the data in Table VI are group data (averaged over subjects), and that there is a considerable <u>inter</u>-individual variation behind these data. The "irregularities" seen at some places in Table VI (that is, the percentage value does not always increase or decrease monotonically as a function of the value of the distortion coefficient) are due to some individual peculiarities and are thus not general.

With regard to all information given in this report a recommendation about maximum allowed amplitude distortion can be made in many different ways. In our opinion the limit should be set at a value for which the frequency of detection would be "low" independently of which listeners, which stimuli, which listening conditions etc may be the case. Referring to the present data (expecially Tables IV and V) we conclude that a power series amplitude distortion corresponding to a distortion coefficient of 0.15 or less for SN subjects (roughly 7.5% quadratic and 3.75% cubic distortion) and 0.05 or less for NH subjects (roughly 2.5% quadratic and 1.25% cubic distortion) would be rarely detected. Detection of distortion in a sinusoidal stimulus may occur at lower distortion levels, but this unrealistic case is excepted from this discussion. However, the 0.05 limit for NH subjects may also be satisfactory as regards the flute and clarinet tones investigated earlier (Gabrielsson & Sjögren, 1972, Table II; the lowest threshold values reported there were about 0.06 - 0.12).

In terms of the information given in Table VI it is seen that the percentage of detection at a distortion coefficient of 0.15 and 0.05 (for SN and NH subjects, respectively) is only slightly higher than that which occurs at the corresponding blank trials (distortion coefficient 0.00).

These values for maximum allowed non-linear distortion seem to be quite possible to obtain for electro-acoustic systems today. The limit for distortion in hearing aids used in Sweden is 10% but wanted to be below 3%. Most hearing aids on the market today have less distortion than 3% at 5 dB below maximum output.

For hi-fi-systems used for the reproduction of music, however, a limit of 1.25% cubic distortion may be difficult to obtain. Tape recorders are often used at a level where the distortion is more than 3%. The loudspeaker components are often non-linear in the base region giving quadratic distortion in the order of 10%.

Telephone systems, finally, sometimes show higher distortion components than what is suggested here as a limit. In this case, however, it must be born in mind that the results reported here are found for systems with flat frequency response and further work has to be done using varying frequency responses.

Finally it is once more emphasized that the results and recommendations given here are based on a limited sample of subjects and stimuli and so continued research is necessary. For certain other technical points not treated here the reader is referred to the discussion in Gabrielsson & Sjögren (1972).

ACKNOWLEDGEMENTS

The authors express their gratitude to Bertil Johansson for valuable discussions. The project was supported by the Swedish Board for Technical Development.

REFERENCES

GABRIELSSON, A. & SJÖGREN, H. (1971). Detection of amplitude distortion in flute and clarinet spectra. Rep. Psychol. Lab. Univ. Uppsala no. 107.

GABRIELSSON, A. & SJÖGREN, H. (1972). Detection of amplitude distortion in flute and clarinet spectra. J.Acoust. Soc. Amer., 52, 471-483.

GABRIELSSON, A., JOHANSSON, B., JOHANSSON, Bodil, LINDBLAD, A.-C., PERSSON, L. (1974). Assessment of comfort and discomfort levels for pure tone, a methodological study. Rep. Technical Audiology, Karolinska Institutet, Stockholm no. 74.

GABRIELSSON, A. & SJÖGREN, H. (1976). Preferred listening levels and perceived sound quality at different sound levels in "high fidelity" sound reproduction. Rep. Technical Audiology, Karolinska Institutet, Stockholm, no. 82.

GREEN, D.M. & SWETS, J.A. (1966). Signal detection theory and psychophysics. New York: Wiley.

GUILFORD, J.P. (1954). Psychometric methods (2nd ed.), New York: McGraw-Hill.

JACOBS, J.E. & WITTMAN, P. (1964). Psychoacoustics, the determining factor in stereo disc distortion. J. Audio Eng. Soc., 12, 115-123.

KIRK, R.E. (1968). Experimental design. Proceduresfor the behavioral sciences. Belmont, California: Brooks/Cole Publishing Company.

LETOWSKI, T. (1974). Difference limen for nonlinear distortion in sine signals and musical sounds. Lab. Musical Acoustics, The Academy of Music. Warszawa, Poland.

RADOWSKI, A., LETOWSKI, T. & BARWICKA, K. (1968). Subjective appraisal of nonlinear distortions in music recordings. Archiwum Akustyki (Warszawa), 3,247-256.

APPENDIX A

Harmonic distortion is measured using one sinusoidal as test signal, intermodulation is measured using two (or more). There are two common ways of measuring intermodulation (see IEC standard 268) giving difference-frequency distortion and modulation distortion. In the first case the two components in the test signal have equal amplitude, for the second case a relation 4:1 is used. Intermodulation distortion is thus a common name for two (may be more in the future) test methods. When the degree of intermodulation distortion is given in data sheets, it is usually meant to be difference-frequency distortion.

The relation between harmonic och difference-frequency distortion can be calculated.

Let the test signal (X) be two sinusoidal tones with equal amplitude

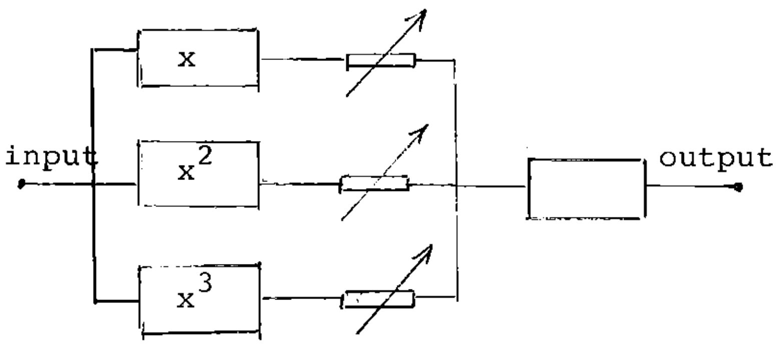
$$X = \sin \omega_1 t + \sin \omega_2 t$$

If then the system transfer function can be described by a power series, the output of the system (Y) is:

$$Y = \sum_{n=1}^{\infty} a_n (\sin \omega_1 t + \sin \omega_2 t)^n$$

Evaluation of the expression is easily done for n = 2, 3, ... The resulting amplitudes and frequencies are given in Table A.1 and A.2. The terms a_n are called distortion coefficients and describe both harmonic and intermodulation distortion. They are in the general case frequency dependent.

The non-linearities were generated in analog quadratic and cubic devices and mixed with the linear signal.



Block diagram of amplitude distortion synthesizer

For the quadratic function a function generator from Philbrick (PSQ-P) was used. The cubic term was designed using the fact that $e^{3\log x}=x^3$. The frequency response of all terms is flat 10 - 40000 Hz, transient distortion is negligible, noise -95 dB(A) and unwanted distortion products below -60 dB. The cubic and quadratic functions work satisfactorily within an input amplitude range of 25 dB.

Table A.1. Harmonic and difference frequency distortion at distortion coefficient a_2.

Frequency	Amplitude	
$2\omega_1$	$0.5\ a_2$	Harmonic distortion
$2\omega_2$	$0.5\ a_2$	Harmonic distortion
$/\omega_2 - \omega_1/$	a_2	Intermod. distortion
$\omega_1 + \omega_2$	a_2	Intermod. distortion

Table A.2. Harmonic and difference frequency distortion at distortion coefficient a_3.

Frequency	Amplitude	
$3\omega_1$	$0.25\ a_3$	Harmonic distortion
$3\omega_2$	$0.25\ a_3$	
$2\omega_1 - \omega_2$	$0.75\ a_3$	Intermodulation distortion
$2\omega_1 + \omega_2$	$0.75\ a_3$	
$2\omega_2 - \omega_1$	$0.75\ a_3$	
$2\omega_2 + \omega_1$	$0.75\ a_3$	

Table I. Hearing loss (dB) at 125 - 8000 Hz for SN subjects

	Frequency							
Subject	125	250	500	1000	2000	4000	8000	Ear
1	40	50	55	55	50	60	75	Right
2	25	40	45	55	55	45	20	Right
3	40	30	25	45	55	35	25	Right
4	35	45	55	80	70	75	65	Left
5	25	20	25	30	40	65	65	Right
6	5	5	15	35	60	50	60	Left
7	40	40	45	55	45	55	70	Left
8	5	20	30	55	60	40	45	Left
9	20	20	30	45	45	40	55	Left
10	55	50	45	45	50	45	55	Right
11	40	50	80	85	85	80	85	Left
12	15	35	55	60	60	55	50	Left

Table II. Adjusted MCL:s, dB(A), slow meter action, for all subjects at four different stimuli. Mean and standard deviation over subjects appear in the bottom margin and the average intra-individual standard deviation to the right. SN = sensory-neural hearing loss, NH = normal hearing. SD = standard deviation.

	Stimuli				
SN subjects	Music	Chat	Voice	Sinewave	Intra-ind.SD
1	85	84	88	79	2.7
2	82	87	91	82	1.9
3	70	72	82	74	3.9
4	86	85	86	74	2.4
5	71(64)	73(63)	73(78)	65(60)	2.5
6	78	81	88	75	4.2
7	86	88(79)	88	80	5.1
8	91	88	90	87	3.8
9	78	81	82	71	4.7
10	96	98	98	92	3.4
11	109	108	106	99	1.7
12	94(87)	90	90	81	3.3
Mean	85.3	86.3	88.5	79.9	
SD	11.2	9.9	8.2	9.4	

NH subjects	Music	Chat	Voice	Sinewave	Intra-ind.SD
51	98	87	93	81	4.8
52	72	74	87	69	3.6
53	68	66	75	70	2.0
54	73	75	78	73	3.3
55	70	65	71	67	2.0
56	76	74	74	70	2.3
57	66	68	63	60	2.9
58	61(52)	63(51)	63(57)	55(60)	2.6
59	71	73(68)	73	65	2.6
60	61	63	73	65	5.3
61	51	48	63(58)	55	3.0
62	66(62)	68(61)	68	65	2.5
Mean	69.4	68.7	73.4	66.3	
SD	11.2	9.4	9.3	7.4	

Table III. Product moment correlations between the size of hearing loss at different frequencies (rows) and MCL for the different stimuli (columns).

Frequency Hz	Stimuli			
	Music	Chat	Voice	Sinewave
125	0.24	0.30	0.29	0.30
250	0.59	0.59	0.55	0.55
500	0.81	0.78	0.69	0.65
1000	0.71	0.66	0.62	0.56
2000	0.67	0.62	0.68	0.60
4000	0.45	0.40	0.25	0.16
8000	0.43	0.40	0.24	0.18
500-1000	0.79	0.75	0.68	0.62
500-2000	0.80	0.76	0.73	0.66
250-2000	0.79	0.76	0.72	0.67

Table IV. Threshold values expressed in distortion coefficient for the detection of quadratic (X^2) and cubic (X^3) amplitude distortion. SN = sensory-neural hearing loss, NH = normal hearing. See further explanations in text.

	Stimuli							
	Music		Chat		Voice		Sinewave	
SN subjects	X^2	X^3	X^2	X^3	X^2	X^3	X^2	X^3
1^+	>.75	>.75	>.75	.75	.69	.60	.21	.19
2	>.75	>.75	>.75	>.75	.75	.63	.15	.18
3^+	>.75	>.75	$<.02^x$	$.43^x$	.50	.40	.33	.40
4^+	>.75	.35	>.75	.50	.23	.26	>.75	.17
5^+	$.04^x$	.50	$.05^x$	.67	>.75	.54	.12	.14
6^+	>.75	>.75	>.75	>.75	$>.75^x$	.47	.25	.15
7	.43	>.75	.34	>.75	.40	.38	$<.02^x$	$.05^x$
8^+	>.75	.40	>.75	>.75	.63	.41	.24	$<.02^x$
9	.67	.75	.67	>.75	.75	.54	.42	.34
10	$<.02^x$	$.02^x$	>.75	>.75	>.75	>.75	.30	.30
11^+	.75	>.75	.63	>.75	.25	.30	.40	$.13^x$
12	.36	>.75	.34	>.75	.36	.33	.18	.23
Mean	>.67	>.66	>.65	>.72	>.55	>.47	.30	.23
NH subjects								
51	.54	.61	.34	.50	.59	.44	.11	.15
52	.56	>.75	.39	.67	>.75	.56	.12	.09
53^+	$<.02^x$	.50	.20	.60	.04	.25	.17	.07
54^+	.30	$.15^x$	.30	>.75	.32	.12	$.05^x$	$.04^x$
55	.67	.46	.50	>.75	.35	.37	.24	.19
56^+	.58	.63	$.20^x$	.75	.32	.24	.07	.04
57	.58	.55	.38	.40	.46	.38	.13	.10
58^+	$.15^x$	$.30^x$	.35	.63	$.04^x$	.20	.10	.04
59	.75	>.75	.75	.75	.25	.41	.10	.03
60	.70	>.75	.50	>.75	.54	.40	.24	.23
61^+	.61	>.75	.38	$>.75^x$	.64	.53	.17	.08
62	.43	.43	$<.02^x$	$<.02^x$	.50	.36	.08	.03
Mean	.57	>.62	.41	>.66	>.43	.36	.14	.10

Table V. Frequency and percentage of threshold values within different regions of the distortion coefficient for SN subjects and NH subjects at "Music", "Chat", and "Voice" stimuli.

	SN subjects			NH subjects		
Dist. coeff.	Frequency	Percentage	Cumulated percentage	Frequency	Percentage	Culmulated percentage
<.10	0	0.0	0.0	1	1.6	1.6
.10 - .19	0	0.0	0.0	1	1.6	3.2
.20 - .24	1	1.5	1.5	3	4.8	8.0
.25 - .29	2	3.1	4.6	2	3.2	11.2
.30 - .34	4	6.2	10.8	5	7.9	19.1
.35 - .39	4	6.2	17.0	8	12.7	31.8
.40 - .44	5	7.7	24.7	6	9.5	41.3
.45 - .49	1	1.5	26.2	2	3.2	44.5
.50 - .54	5	7.7	33.9	8	12.7	57.2
.55 - .59	0	0.0	33.9	6	9.5	66.7
.60 - .64	4	6.2	40.1	6	9.5	76.2
.65 - .69	4	6.2	46.3	2	3.2	79.4
>.70	35	53.7	100.0	13	20.6	100.0

Table VI. Percentage "Different" judgments at different distortion coefficients in average over the members of each subject group.

SN subjects

		Distortion coefficient						
		.75	.50	.30	.15	.05	.02	.00
"MUSIC"	X^2	54	40	30	28	22	26	15
	X^3	45	28	28	21	28	24	11
"CHAT"	X^2	55	43	36	30	30	31	10
	X^3	37	28	29	24	32	30	11
"VOICE"	X^2	65	40	28	20	20	15	7
	X^3	80	55	25	13	11	12	17
"SINEWAVE"	X^2	-	76	58	36	31	22	10
	X^3	-	89	69	44	28	26	13

NH subjects

"MUSIC"	X^2	84	47	35	26	22	23	17
	X^3	63	38	25	26	30	23	15
"CHAT"	X^2	90	70	37	24	37	25	17
	X^3	54	37	31	35	27	27	16
"VOICE"	X^2	92	60	36	23	28	18	13
	X^3	95	81	32	20	19	18	10
"SINEWAVE"	X^2	-	97	89	60	23	18	8
	X^3	-	98	93	72	40	21	7