BRÜEL & KJÆR

NÆRUM · DENMARK

BRÜEL & KJÆR application notes

Relevant Hi-Fi tests at home

Paper of the 47th
Audio Engineering Society Convention

1974-02-26/29 Copenhagen, Denmark

Hi-Fi Tests

— with 1/3 octave, pink weighted, random noise

By Henning Møller, Brüel & Kjær

ABSTRACT

Measurements on Hi-Fi systems are relevant when they give the same result as listening tests.

The measuring method most suited for this requirement is the 1/3 octave, pink weighted, random noise method used in the actual listening room. Another consideration may be phase response.

When this method is used in the actual room, it gives information about the combination loudspeaker - room.

Measurements on 5 different speakers in 3 different rooms show a considerable room dependence, but good correspondence with listening tests; that is the combination loudspeaker - room must be optimized.

In practice the measurements can be carried out in many ways - the simplest method requires only a test record and a portable sound level meter.

INTRODUCTION

Hi-Fi material for home use today has got a quite reasonable standard.

Nevertheless it is very unusual that the artistic experience of listening to a Hi-Fi set approaches even approximately the experience of being in the concert hall.

The limits for true Hi-Fi systems must be looked for outside the actual system.

It can not be denied that the limitations have partly psychological reasons - just the missing visual sensation will change the experience - but moreover there are actual limitations both preceding and following the Hi-Fi system, namely in the record and in the listening room.

An improvement preceding the system, that is in the grammophone record, is mainly an economic question and in practice the user has no influence on it. Direct recording on your own equipment is in fact the only alternative.

For the consumer, an improvement following the system seems therefore to be the best solution. In this "paper" we shall look more closely at these possibilities.

We will show that the most suitable measuring method for improvement of a Hi-Fi system seems to be the 1/3 octave, pink weighted, random noise method, used in the actual listening room. We will see results from this method from 5 different speakers in 3 different rooms. We will show that the method can be made very simple and inexpensive, so the practical user enjoys using the method, but also that it can be made so professional that it will give the results at once. We shall indicate supporting measuring methods and keep measuring results, and listening results together. We shall see, that there is almost perfect agreement between measured results and subjective listening results. This is something completely fundamental.

Finally we shall see the results of some supporting measurements showing the dependence of microphone and speaker position. We shall see the frequency spectra of the music examples used for the listening tests, and what the reverberation time was in the 3 rooms.

RELEVANT MEASURING METHODS

It is well known, that the output voltage of an electric circuit is very dependent upon the actual loading. For the Hi-Fi set this analog means that the acoustic loading - i.e. the listening room - more or less decides the final result.

Therefore, it is very important that the Hi-Fi system matches the room.

To be relevant, the measurement must be made under the normal acoustic working conditions, and of course measurement must not alter these conditions.

A great many investigations have been made recently to find suitable measuring methods in listening rooms (Ref. 1 and 2).One has especially - as will be heard elsewhere at this convention - in the "loudspeaker investigation" (Ref. 1) tried almost all possible measuring methods to find the most suitable; that is, the one which best agrees with results from listening tests. It was found that the objective measuring method which corresponds best with subjective judgements is the 1/3 octave, pink weighted, random noise method and further that phase response and power characteristics also correspond reasonably well, but less significantly.

Traditionally one has concentrated only on the amplitude response and ignored the phase response. Obviously the reason being the practical problems of making loudspeaker phase response measurements. Techniques using tone bursts or Fast Fourier Transform have provided the only possibilities until now, before Brüel & Kjær produced a phasemeter with a delay line.

The importance of phase measurements will be discussed in detail at the AES convention in California next month. Here we will concentrate on amplitude measurements in normal rooms.

We selected 3 rooms, 5 loudspeakers and 5 people as listeners.

The 3 rooms are shown in Fig. 1, 2, and 3.

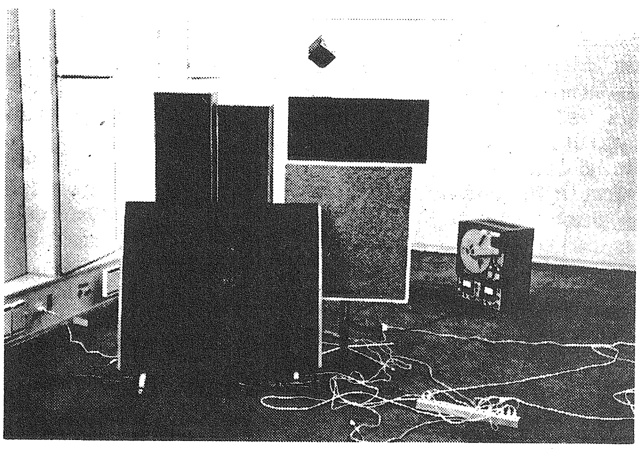


Fig. 1 Room L 1.

Fig. 2 Room L 2.

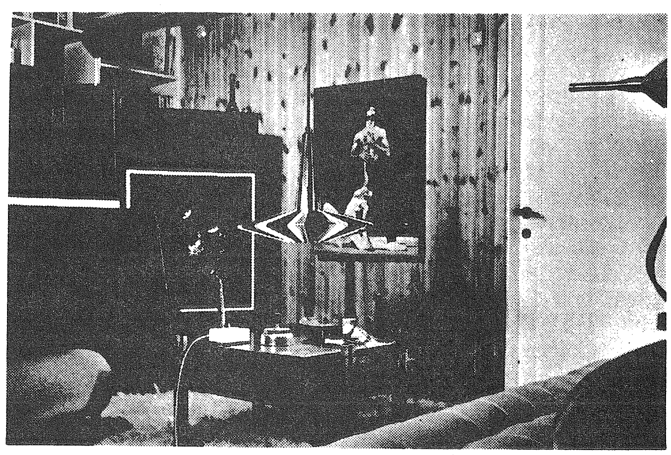


Fig. 3 Room L 3.

The loudspeaker positions used are partly because the minimum distance between the speakers is considered, and partly that the speakers should have acoustic conditions as equal as possible.

Later we shall see that the dependence on loudspeaker positioning was rather limited.

Measurements as well as listening tests were made in all three rooms.

15 CURVES

To give an immediate impression of the results, we will look at the 15 curves which the 5 speakers, in the 3 rooms, gave from measurements with 1/3 octave bandpass noise at the listening place. Later, we will go much closer into how, in fact, these measurements were made.

Fig. 4 shows the 15 measurement curves.

The three vertical columns show, from the left, listening rooms L 1, L 2, and L 3. The five horizontal rows show the five speakers H 1, H 2, H 3, H 4, and H 5.

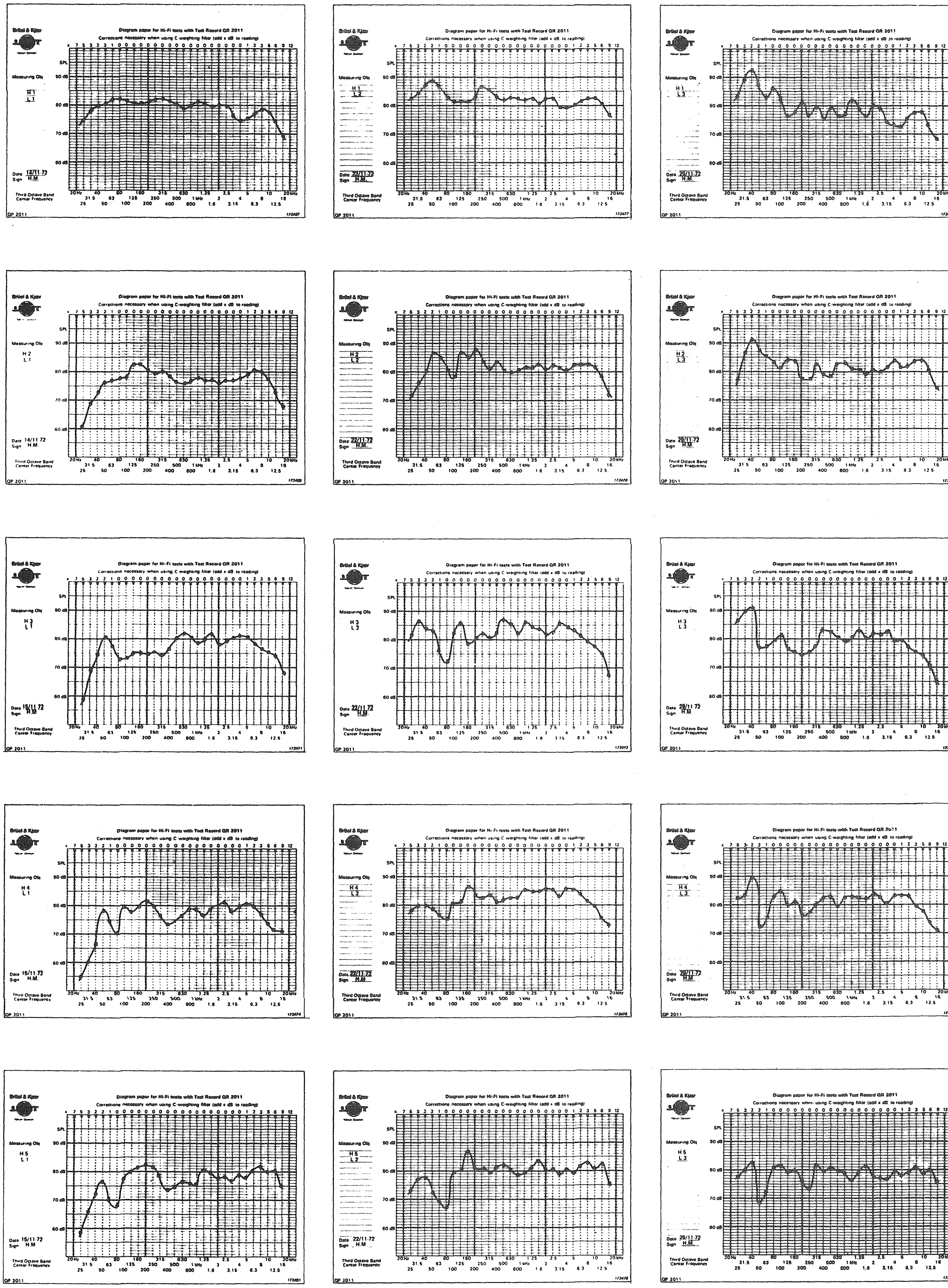
If we look at the three top charts, we see curves for the same speaker, but in three different rooms. There certainly seems to be a big difference. In the large room L 1, the curve is fairly even, in room L 2 it is somewhat worse, and in room L 3 it is very uneven. We see too much bass-lift and too many resonances in room L 3.

EVALUATION OF THE CURVES

We will now look at the five charts in the first vertical column - that is the 5 speakers in room L 1. There is no doubt that the uppermost chart is the best, no. 2 is the next best, and the lowest chart is clearly the worst. Which one of no. 3 or no. 4 is the best, could in the first place be difficult to decide. Examining them more closely - as shown later - we will see that no. 4 is better than no. 3.

This order was later found to be the same as the order of preference indicated by listening tests. The closer evaluation of these curves concerns the following two criteria. The first, that the curves should be as smooth and straight as possible, indicating that all frequencies are reproduced at approximately equal level.

When music is recorded under far-field conditions, it will contain a suitable mixture of direct and reflected sound, and the curve ought to be absolutely flat in that case.



These curves show how the 5 loudspeaker responses differ in the 3 rooms. The measurements were made by the portable and inexpensive method Fig.1a. Note that the 3 in-line curves are for the same loudspeaker tested in different rooms

Fig. 4

If the recording is made as a combination of near-field and far-field information, which is in fact normal, the curve ought to boost a little at low frequencies and roll off a little at high frequencies. A suitably shaped curve is shown in Fig. 5.

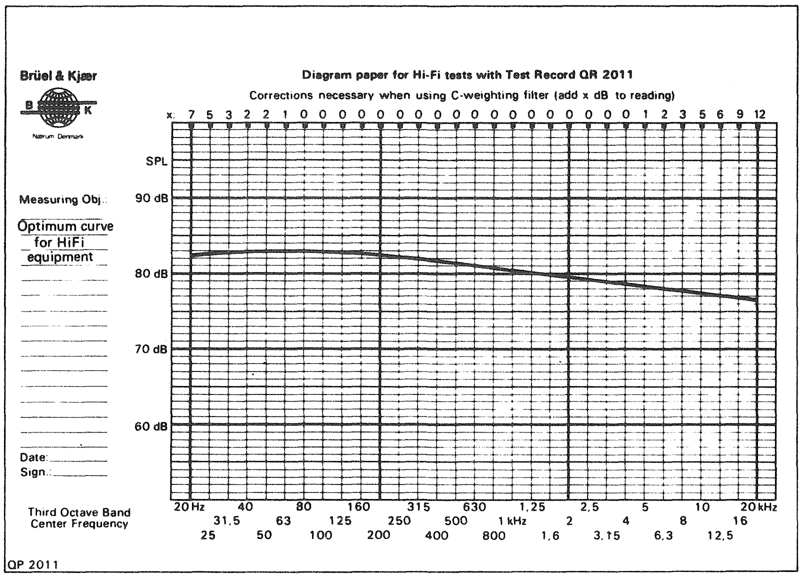


Fig. 5

The curve shows only the necessary tendencies. This curve was derived partly as a result of listening tests and partly by consideration of curves from average concert halls. According to Beranek (Ref. 2) the average concert hall has the same tendency as the curve shown, but at twice the rate. We have chosen only half the rate because most recordings are equally distributed between near-field and far-field recording. Practice has shown that this curve is absolutely reasonable.

The second consideration when evaluating the curves in Fig. 4 is the average frequency content in normal music recording.

As we will see later (Fig. 20) this is typically in the range from 60 Hz to 6 kHz, and therefore this range is given more consideration than the rest of the audible range, when evaluating the curves.

It should be mentioned, that when we made the investigation, we were not really sure of Fig. 5. For instance we cannot call loudspeaker H 5 in room L 3 really bad, just because it did not roll off at high frequencies. If we had done that we would have got even better agreement with the listening tests, as we will see later.

From the above mentioned criteria, our evaluation of the measured curves in Fig. 4 were as follows:

Room L 1:	H1 - H2 - H4 - H3 - H5
Room L 2:	H1 - H2 - H4 - H3 - H5
Room L 3:	H2 - H4 - H5 - H1 - H3

Preference sequence from measurements

- that is in room L 1 we found loudspeaker H 1 the best, H 2 second best, and so on.

LISTENING TESTS

Throughout the listening tests, the loudspeakers were compared two by two. The person listening was asked to choose which of the two loudspeakers, he most wanted to listen to at any given time. All the loudspeakers received "pink" noise and were equalised to the same sound pressure level, by a special box made for the purpose. All the speaker cabinets were covered by a porous cloth, so the cabinets could not be seen during the listening tests.

For each of these two by two comparisons the person listening had to fill out a questionnaire diagram as shown in Fig. 6.

IDENTIFIKATIONSDATA

Forsøgsperson nr........

Lokale nr..............

Sammenligning nr........

EGENTLIGE MÅLINGER

1 Præference I.........

HELHEDSINDTRYK

2 Naturlig............

3 Velafstemt..........

GENERELLE KLANGINDTRYK

4 Nuanceret...........

5 Udflydende..........

6 Præsent.............

7 Fjern...............

8 Åben................

9 Klemt...............

10 Brillant............

11 Mat.................

12 Fyldig..............

13 Tynd................

FREKVENSGANG

14 Jævn................

15 Kraftig bas..........

16 Svag bas............

17 Kraftig mellemtone...

18 Svag mellemtone......

19 Kraftig diskant......

20 Svag diskant.........

FORVRÆNGNING

21 Buldrende...........

22 Hul.................

23 Nasal...............

24 Metallisk...........

25 Spids...............

26 Skinger.............

27 Hård................

28 Blød................

MEST NATURTRO GENGIVELSE AF:

29 Bas.................

30 Mellemtone..........

31 Diskant.............

32 Symfoniorkester......

33 Kammermusik..........

34 Pop.................

35 Præference II........

1 = Højttaler 1 har egenskaben bedst.

2 = Højttaler 2 har egenskaben bedst.

3 = Begge har egenskaben lige godt.

4 = Ingen af højttalerne har egenskaben.

Fig. 6

The questionnaire diagram.

These subjective characteristics in the diagram are almost impossible to translate, and therefore we have let the diagram stay in its original form in Danish.

It is seen, that there are 35 characteristics and for each characteristic the listener was to select which of the two speakers, in his opinion, possessed most of that particular characteristic. The following code was used.

1 = Loudspeaker 1 has most of the characteristic

2 = Loudspeaker 2 has most of the characteristic

3 = Both speakers have the characteristic in equal degree

4 = Neither of the speakers has the characteristic

Using 5 loudspeakers, 5 listeners, 3 rooms and 35 characteristics a total of $(4 + 3 + 2 + 1) \times 5 \times 3 \times 35 = 5250$ comparisons were made.

During all the comparisons, six different short music pieces were used: Wagner Opera, Modern String Quartet, Organ Music from a church, Beat, Jazz, and Popular Music, to ensure that the results are independent of the type of music material used.

RESULTS OF LISTENING TESTS

It would not be reasonable to go into details of the statistical treatment of this material here. Let us simply examine the main result Fig. 7.

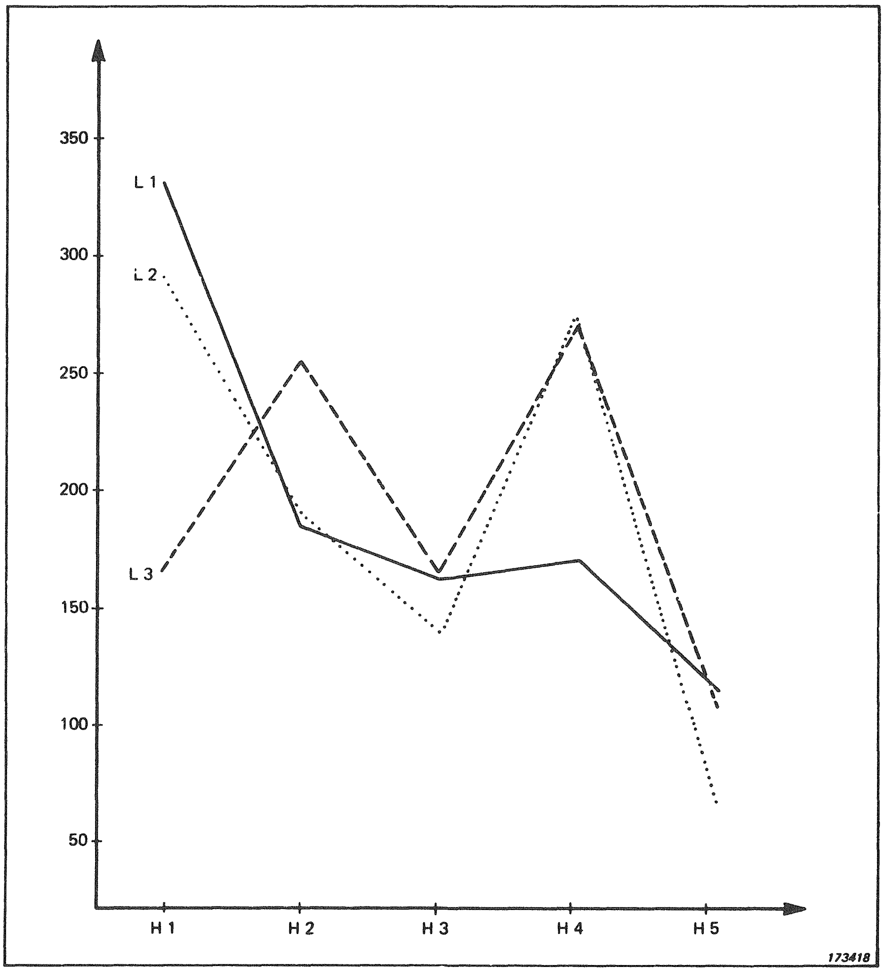


Fig. 7

Curves showing how the subjective quality evaluation of a particular loudspeaker strongly depends on the actual listening room. The y-axes shows the number of times a loudspeaker was preferred. These results are based on answers to positively orientated questions

The three curves show the number of times a given loudspeaker has been generally characterized as being the best one for each room. It is seen, as mentioned in the introduction, that the results are strongly dependent upon the room. The result for loudspeaker H 1 differs for instance 100% from room L 3 to room L 1, and the result for H 4 is much worse in L 1 than in the other rooms.

This is a real problem for a customer who hears these five loudspeakers demonstrated by a dealer, with a demonstration room like room L 1. He decides on loudspeaker H 1, when price is not taken into consideration. But then he finds that his own room is like room L 3, - that is, he should never have selected H 1, but rather H 4 instead.

From the curves in Fig. 6 we can make the following preference sequence:

Room L1: H1 – H2 – H4 – H3 – H5
Room L2: H1 – H4 – H2 – H3 – H5
Room L3: H4 – H2 – H1 – H3 – H5

Preference sequence from listening tests

SIMILARITIES BETWEEN MEASURED RESULTS AND LISTENING RESULTS

If we now compare this preference sequence from the listening tests with the one we got from measurements, we see that the only essential difference in the results is that loudspeaker H 5 in room L 3 was placed as no. 3 from measurements, while from listening tests it was placed as no. 5. As mentioned earlier this is based on the original evaluation - today Fig. 5 would be considered more important and the result would be even better than shown here. The difference between loudspeakers H 2 and H 4 in rooms L 2 and L 3 is so small that it is almost impossible to state a preference, either from listening tests or from measurements.

An example of how the speakers were distributed for the different characteristics is shown in Fig. 8.

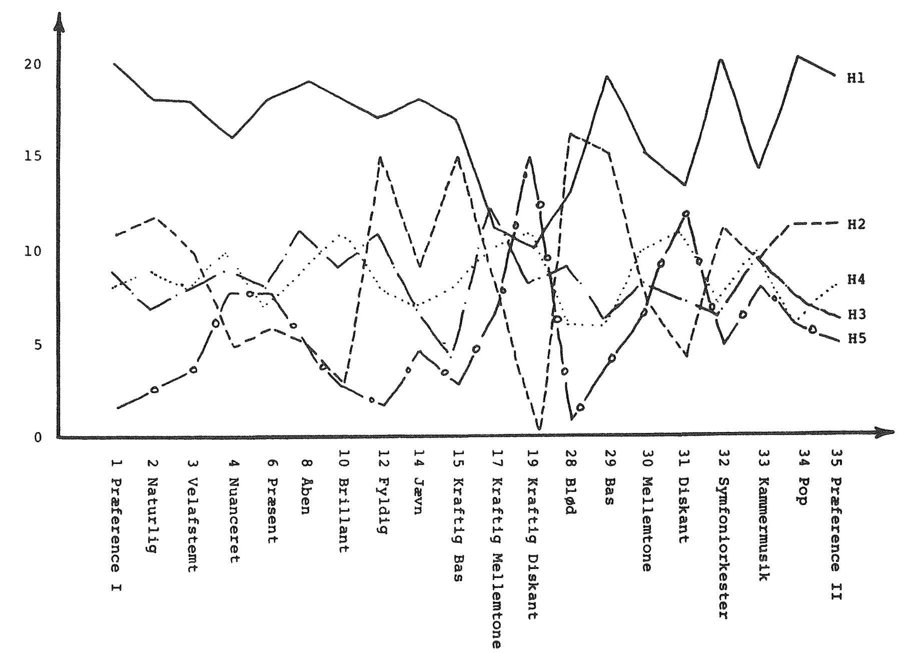


Fig. 8 The number of preferences for the speakers as a function of the characteristics.

The curves show the number of times the different speakers (H 1, ..., H 5) have been characterised as better than the ones they were compared with, as a function of the positively oriented questions in the questionnaire diagram Fig. 6.

This figure (Fig. 8) is valid for room L 1 but corresponding curves were also made for the other rooms, of course, as well as for the negative characteristics. The results from the positive and the negative characteristics were almost the same.

An example of how the listeners were distributed is shown in Fig. 9. There is one curve for each person. It is seen that for each speaker there was one person who voted appreciably different from average. But as it was a different person each time, they can be considered in reasonable agreement. The average voting must be considered to be consistent, they were also all experienced critical listeners.

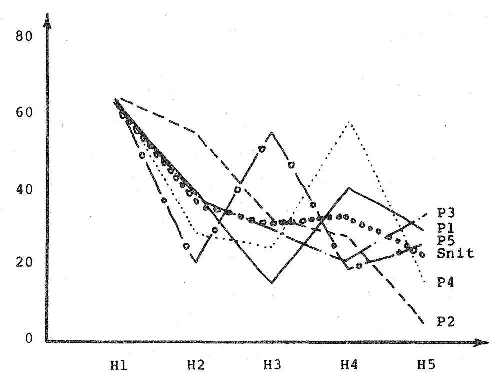
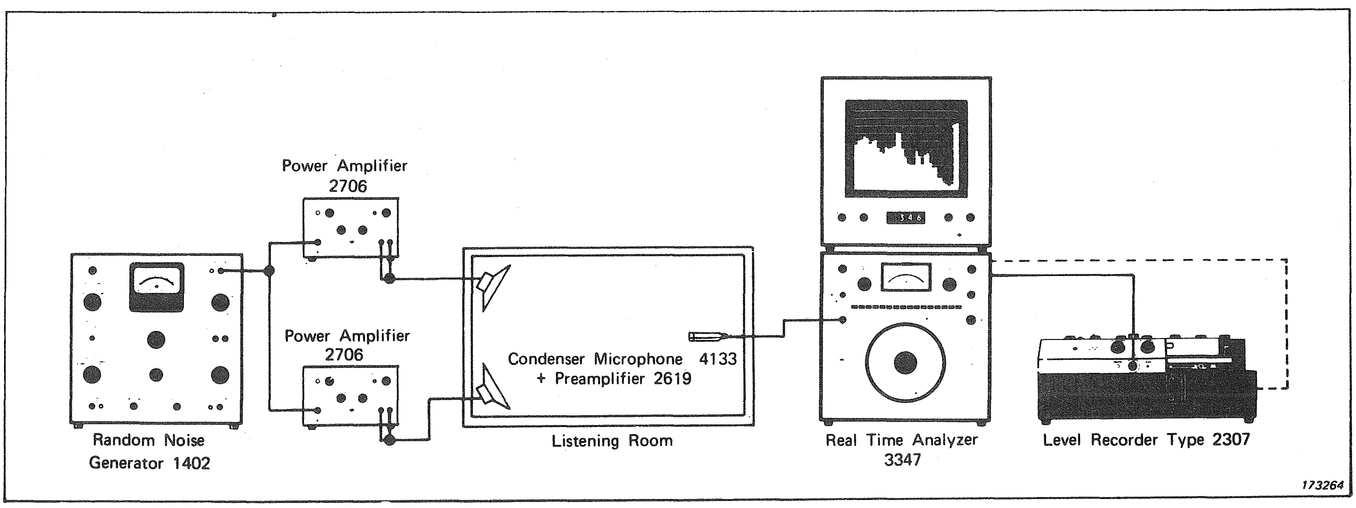


Fig. 9

Example of how the listeners were distributed.

PROFESSIONAL MEASURING METHODS

Until now we have only discussed the measuring method as "the 1/3 octave pink weighted, random noise method". Now let us look closer, to see how the method is used.



Measuring set-up with "pink" noise 20 Hz — 20 kHz. The result is immediately read out on the screen

Fig. 10

We see the professional version Fig. 10. Here the Noise Generator Type 1405/2 plays broad band "pink noise" through the system, and the Real Time Analyzer is used as the measuring instrument. In the Real Time Analyzer all the 1/3 octave filters are connected in parallel, which means, that each 1/3 octave band is measured simultaneously and continuously. This is called "Real Time Analysis".

"Pink noise" looks as shown in Fig. 11 when it is sent directly to the Real Time Analyzer - each column is seen to have about the same height.

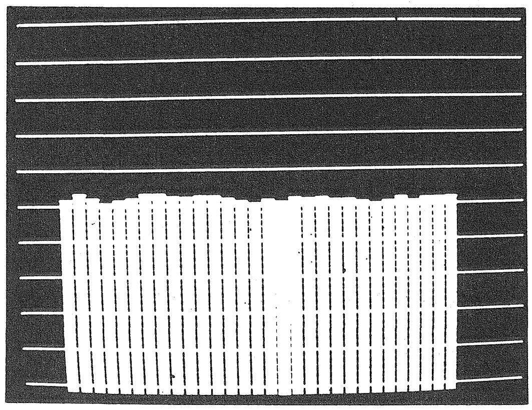


Fig. 11 Spectrum for pink noise.

If we put white noise in, it appears as shown in Fig. 12. White noise contains all frequencies at a constant amplitude. Nevertheless we see a slope of + 3 dB/octave, and that is because we use a logarithmic scale, where the actual bandwidth increases proportionally with frequency - 1/3 octave at low frequencies is just a few Hz, while 1/3 octave at high frequencies covers several thousand Hz. When the voltage for white noise is proportional to the squareroot of the bandwidth, the increase will be only 3 dB/octave.

To get the flat curve, that we wanted on the screen, we had to correct the white noise with a -3 dB/octave filter. This is the signal called "pink noise". Pink noise is nothing but white noise weighted - 3 dB/octave, and it is used together with the usually employed logarithmic scale with constant percentage bandwidth. White noise is used with constant bandwidth.

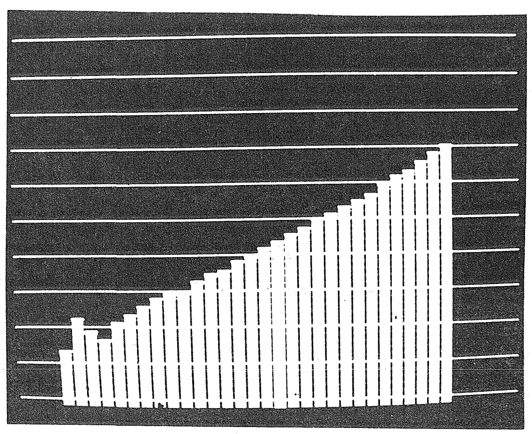
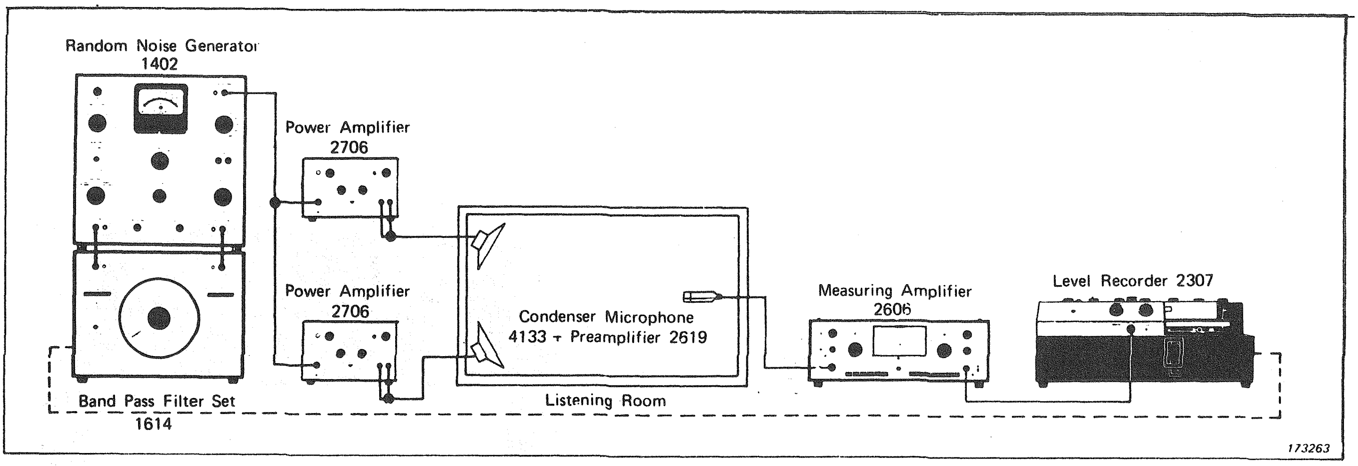


Fig. 12 Spectrum for white noise.

SEMIPROFESSIONAL METHODS

A somewhat cheaper measuring method is shown in Fig. 13. In this case we apply each 1/3 octave bandwidth individually and take a broad band level measurement. The signal looks as shown in Fig. 14.



Set-up for "1/3 octave, pink weighted, random noise method"

Fig. 13

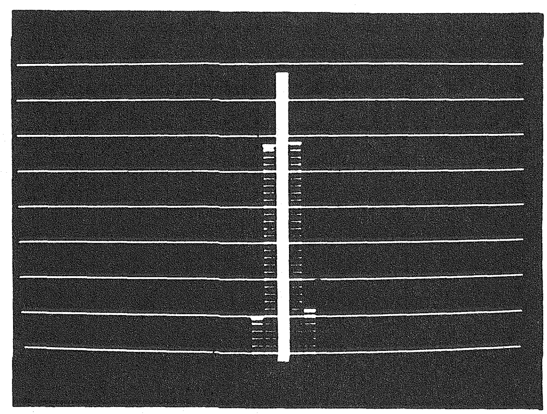


Fig. 14 Spectrum for 1/3 octave.

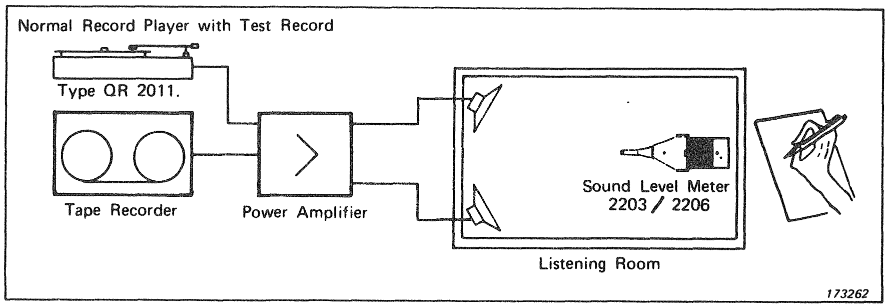
This method is slower than the professional method, because we had to make 30 measurements, one for each 1/3 octave, while in the professional we measured them all at the same time.

In practice this principle can be reversed, that is, we can send out the broad band pink noise, and measure selectively in 1/3 octaves one at a time. In this case we risk burning-out the tweeters because the noise is applied to the speakers for quite a long time.

The optimum signal-to-noise ratio is reached by both sending out and measuring in 1/3 octave bands. In practice there is no difference in the results from these different methods.

THE SIMPLE METHOD

The same principle as used in the semiprofessional method is also used in the simple method.



The portable and inexpensive method. The recording is manual with one point for each 1/3 octave

Fig. 15

We use the Test Record QR 2011 on the generator side, and a Sound Level Meter Type 2206 on the measuring side.

The signals recorded on the test record are the same as those which the Noise Generator and the 1/3 octave filter produced in the semiprofessional method - that is pink noise in 1/3 octave bands.

The microphone and the measuring amplifier are replaced by the Sound Level Meter and the Level Recorder is replaced by the special chart paper QP 2011. The curve is recorded manually.

The only equipment required to make this measurement, is in fact the Test Record QR 2011 and the Sound Level Meter Type 2206. Naturally this simple method is less accurate than the professional methods, but it is an excellent alternative because it is portable and inexpensive, because it does in fact test the whole system from pick-up to loudspeaker - room combination, and because it has the same working conditions as normal music records.

The 15 curves in Fig. 4 are in fact all made using the simple method. The differences between the simple method and the more professional methods are typically ± 1 dB, which, because of the fluctuations normally found in ordinary rooms, can be considered negligible.

In this connection, it seems reasonable to mention that the response in the listening room does not necessarily disclose everything there is to know about the system. Alone, it is not a pure scientific truth.

The picture might be disturbed by the so-called "Non-Minimum Phase Behaviour" (Ref. 3) and possibly by other unknown phenomena. This, however, has normally less influence on the final result, and will therefore not be dealt with further right here. A supporting measurement might be phase.

SUPPLEMENTARY EXPERIMENTS

To support and supply the given results we will now look at the dependence of loudspeaker and microphone positioning. We shall see the relative frequency content of the music examples used. We shall see what the reverberation time was in the three rooms and finally we shall suggest, what one can do, to improve the measured characteristic - and thereby obtain optimum performance.

The dependence of microphone and loudspeaker positioning is not very important in normal rooms. Typical variations fall within the frequency range 50 - 2000 Hz and within ± 5 dB. The remark one normally hears - that the bass increases whenever the speaker comes close to a corner - is not completely true. It is only the upper part of the bass range and the lower part of the mid range which increase. We must admit, however, that subjectively it seems as though the bass increases.

Figs. 16 and 17 show the dependence of positioning for room L 2, and figs. 18 and 19 show it for room L 3.

Fig. 16 The variations in room L 2 for three different microphone positions.

Fig. 17 The variations in room L 2 for three different loudspeaker positions.

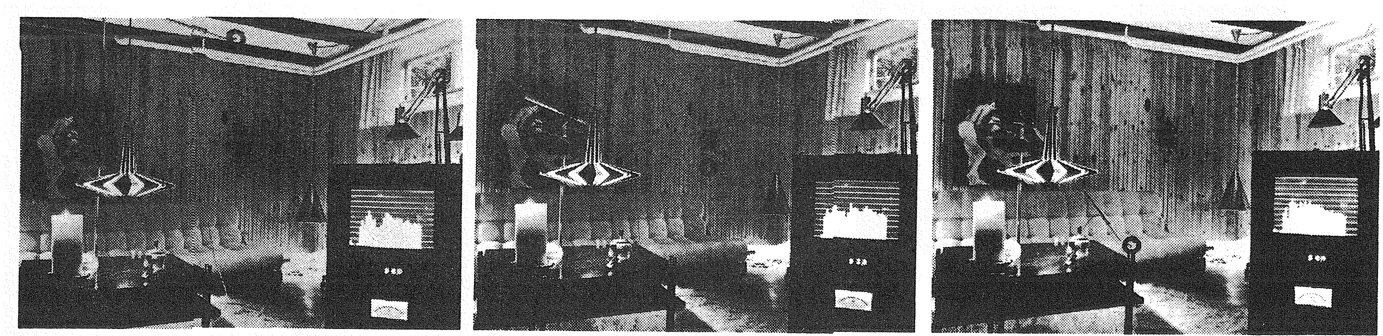


Fig. 18 The variations in room L 3 for three different microphone positions.

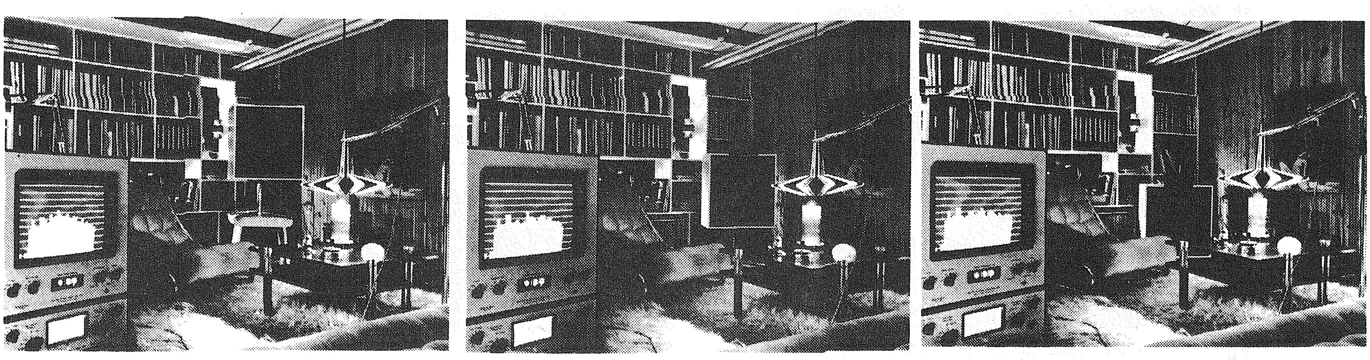
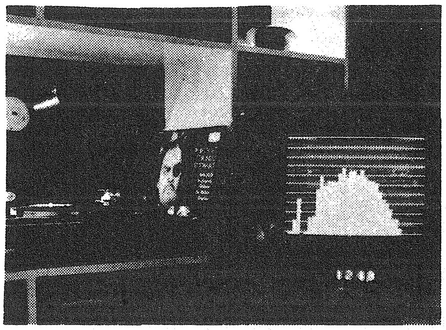


Fig. 19 The variations in room L 3 for three different loudspeaker positions.

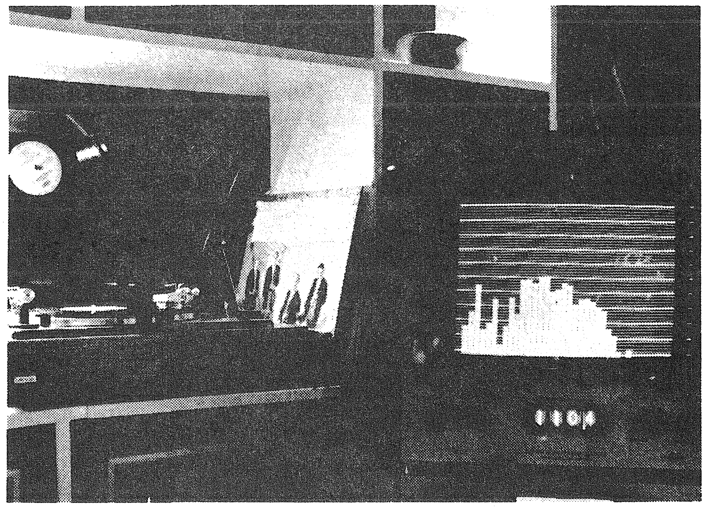
The relative frequency content of the music examples used for the listening tests is shown in fig. 20.

It is seen, for example, that the organ music, M 3, has a quite wide and smooth frequency content.

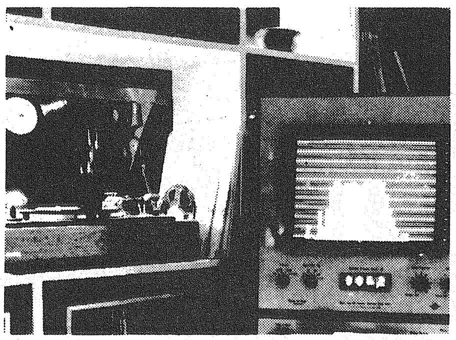
The beat music, M 4, exhibits typical electric-bass around 125 Hz and brass instruments around 1,25 kHz. On the Oscar Peterson recording, M 6, we see the bass around 100 Hz, the piano around 400 Hz and the cymbal around 12,5 kHz.



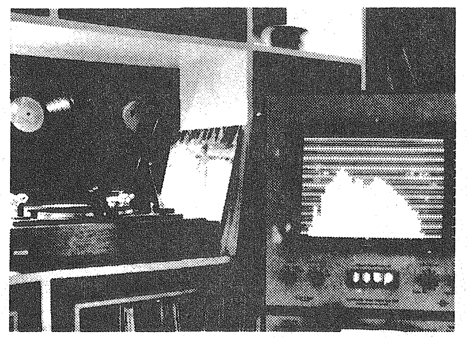
M1 . Wagner: Die Walküre, finale 3. akt. Deutsche Grammophon 135 150.



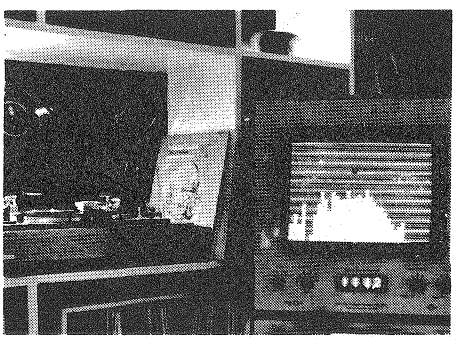
M2 Max Regor: Strygekvartet g-mol, op. 54 nr. 1. 2' sats. Deutsche Grammophon 2530 081.



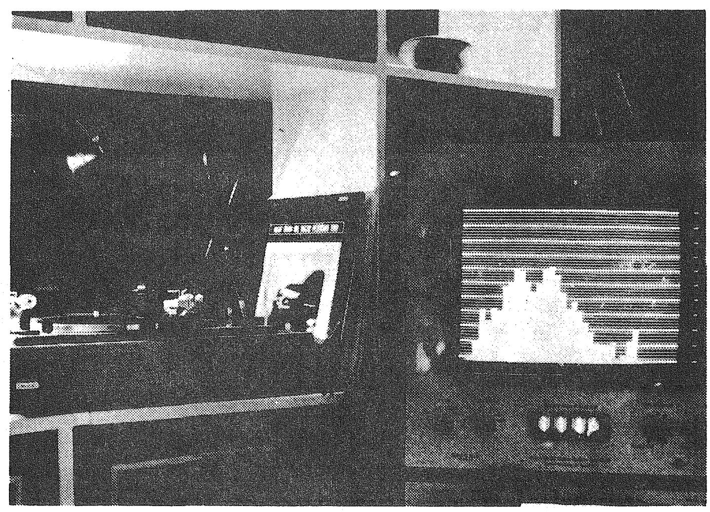
M3 Kirkemusik. Egen optagelse fra Grundtvigskirken.



M4 Spinning Wheel, Shirley Bassey, United Artists UAS 29100.



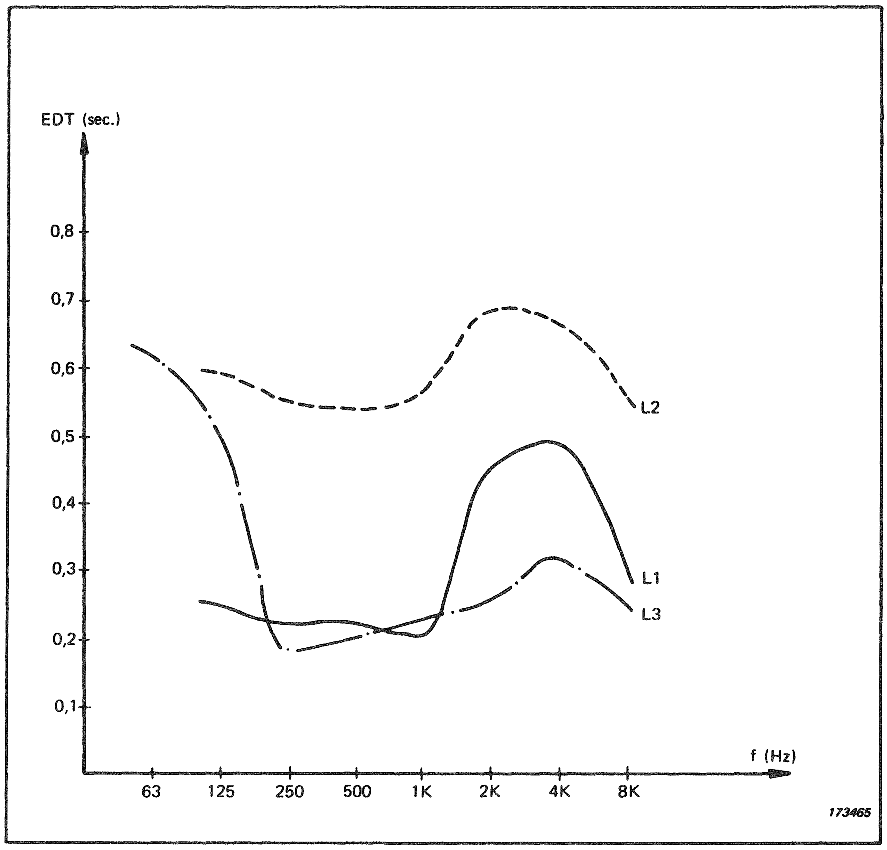
M5 Stan Kenton: Adventure in emotions, part 6, joy. Capitol ST 2424.



M6 Oscar Peterson: Things ain't what they used to be. Verve V 6-8538.

Fig. 20 The spectra of the music examples from the listening tests.

The reverberation time in the three rooms, as a function of frequency, is shown in fig. 21. In fact, it is the so-called "Early Decay Time" (EDT) which is shown, but this is almost the same as the normal reverberation time, the only difference being that consideration is placed on the beginning of the reverberation curve.



Reverberation Time (EDT) versus frequency in the three rooms

Fig. 21

For room L 3, there seems to be good agreement between the long reverberation time at low frequencies and the appreciable bass lift, that we saw from measurements with 1/3 octave noise in this room (compare with fig. 4).

Finally, we shall consider what possibilities exist for correcting the curve measured with the 1/3 octave, pink weighted, random noise method.

One possibility is, of course, to select a Hi-Fi set which, so far as possible, neutralizes the weeknesses of the room. At a dealer who has measured all speakers in his demonstration room, one can at once find the optimal Hi-Fi set.

With a given set one has the possibility of acoustical and electrical corrections. The acoustical correction can be made by changing the reverberation time of the room, that is with furniture, carpets, curtains, wood panels and so on, or by moving the speakers so that resonances and standing waves are avoided as much as possible.

The electrical corrections can be made by various commercially available spectrum shapers, or by building special filters.

Often, a change to the cross-over network will be the easiest. We have a rough example of this on speaker HI where a resonance in the bass system is moved simply by connecting a series resonant circuit directly across the bass speaker terminals. The difference, with and without corrections, is shown in fig. 22.

The result was, as seen, to give better measuring results and, in fact, much better listening results.

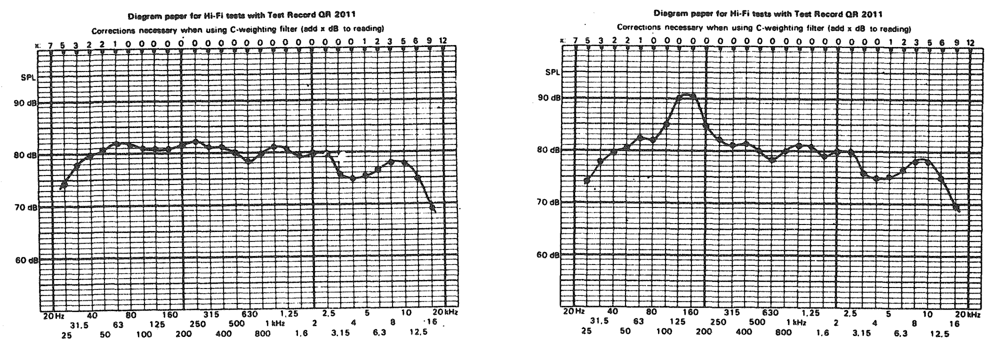


Fig. 22 The amplitude response for loudspeaker H 1 with and without compensation.

CONCLUSION

Since the listening room is an extremely important factor in loudspeaker performance, an objective test method is required, that gives good correlation with subjective listening tests. It is found that pink weighted, random noise in third octave bands, best meets this requirement.

The measurement may be implemented in several ways of various degrees of convenience and expense; (1) Real time third octave analysis. (2) Sequential third octave analysis. (3) Pink noise test record analysis.

The professional methods are relevant in cinemas, theaters, concert halls and especially in recording- and radio studios. The Hi-Fi enthusiast and the small dealer of course require the portable and inexpensive method.

All three methods show excellent agreement with each other, and with subjective tests.

Ref.1. E. Rørbæk Madsen and H. Staffeld
Højttalerundersøgelser
Akademiet for de tekniske videnskaber
Denmark 1972.

Ref.2. Roy F. Allison and Robert Berkovitz
The Sound Field in Home Listening Rooms
Journal of the Audio Engineering Society
July/August 1972, Vol. 20, No.6

Ref.3. Richard C. Heyser
Loudspeaker Phase Characteristics and Time Delay Distortion: Part 1
Journal of the Audio Engineering Society
January 1969, Vol. 17, No. 1.