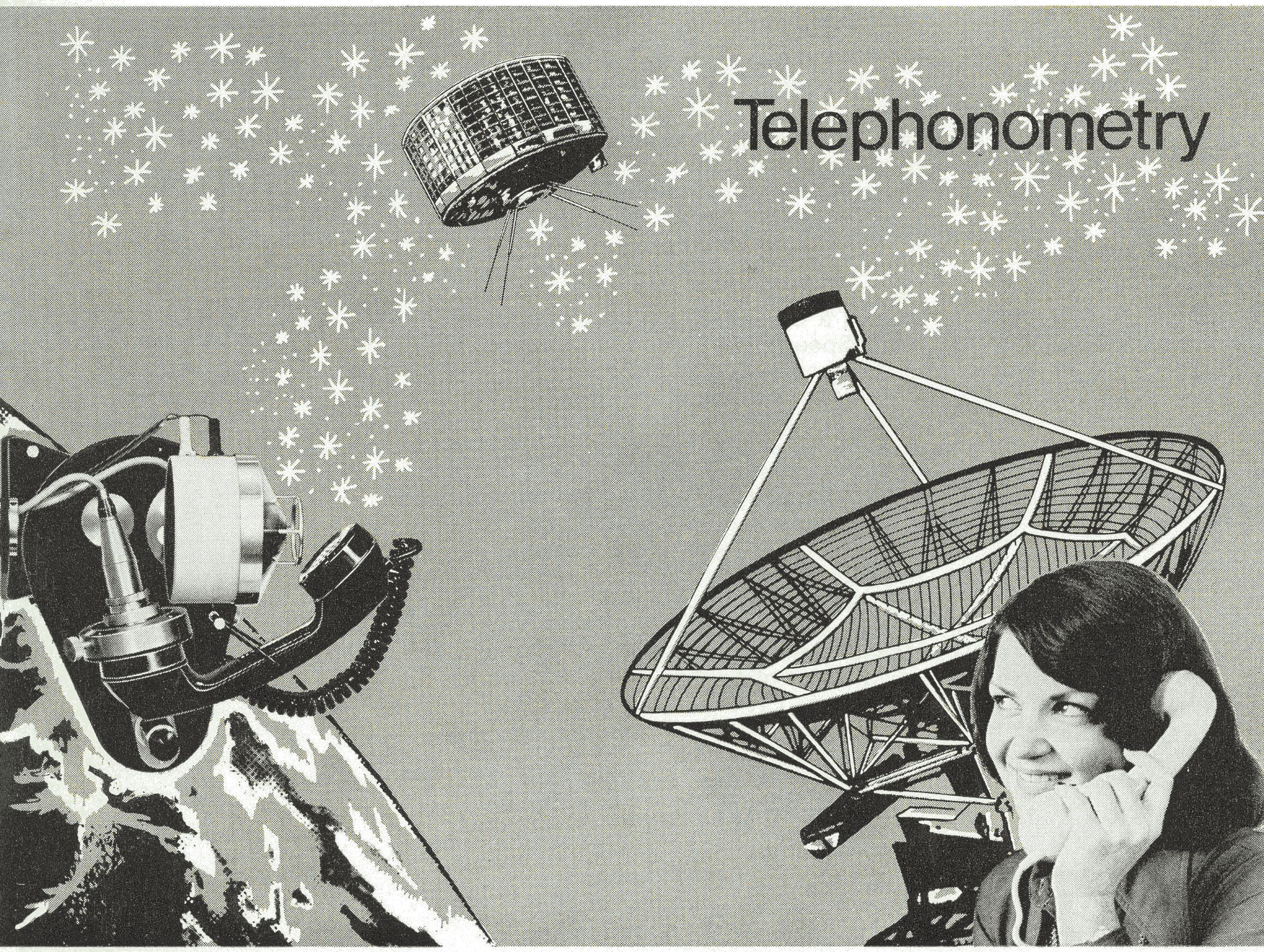
No. 1 1975

issued quarterly

Technical Review

To Advance Techniques in Acoustical, Electrical and Mechanical Measurement



Brüel & Kjær

PREVIOUSLY ISSUED NUMBERS OF BRÜEL & KJÆR TECHNICAL REVIEW

4-1974 Underwater Impulse Measurements.
A Comparison of ISO and OSHA Noise Dose Measurements.
Sound Radiation from Loudspeaker System with the Symmetry of the Platonic Solids.

3-1974 Acoustical Investigation of an Impact Drill.
Measurement of the Dynamic Mass of the Hand-arm System.

2-1974 On Signal/Noise Ratio of Tape Recorders.
On the Operating Performance of the Tape Recorder Type 7003 in a Vibrating Environment.

1-1974 Measurements of averaging times of Level Recorders Types 2305 and 2307.
A simple Equipment for direct Measurement of Reverberation Time using Level Recorder Type 2305.
Influence of Sunbeams striking the Diaphragms of Measuring Microphones.

4-1973 Laboratory tests of the Dynamic Performance of a Turbocharger Rotor-Bearing System.
Measurements on the Resonance Frequencies of a Turbocharger Rotor.

3-1973 Sources of Error in Noise Dose Measurements.
Infrasonic Measurements.
Determination of Resonance Frequencies of Blades and Disc of a Compressor Impeller.

2-1973 High Speed Narrow Band Analysis using the Digital Event Recorder Type 7502.
Calibration Problems in Bone Vibration with reference to IEC R 373 and ANSI S3. 13-1972.
An Investigation of the Near Field Screen Efficiency for Noise Attenuation.

1-1973 Calibration of Hydrophones.
The Measurement of Reverberation Characteristics.
Adaptation of Frequency Analyzer Type 2107 to Automated 1/12 Octave Spectrum Analysis in Musical Acoustics.
Bekesy Audiometry with Standard Equipment.

4-1972 Measurement of Elastic Modulus and Loss Factor of Asphalt.
The Digital Event Recorder Type 7502.
Determination of the Radii of Nodal Circles on a Circular Metal Plate.
New Protractor for Reverberation Time Measurements.

(Continued on cover page 3)

TECHNICAL REVIEW

No. 1 — 1975

Contents

Telephonometry Symposium

General Introduction
On the 12th and 13th November 1974 a special Telephonometry Symposium was held at Copenhagen. Hosts for the symposium were KTAS (Copenhagen Telephone Company) and JTAS (Jutland Telephone Company), and the initiative for the meeting was mainly due to efforts of Mr. P. V. Arlev and S. A Jäger of JTAS.

The symposium consisted of lectures and discussions related to developments in the field of telephonometry, and participants from England, Denmark, Finland, Norway and Sweden contributed to the success of the meeting.

Some of the lectures presented at the symposium are printed in this issue of the Brüel & Kjær Technical Review. Due to space limitations it has not been possible to include all the contributions here. Most of the remaining papers are, however, available in the form of a special publication which can be obtained from B & K upon request.

Major contributions were made by Mr. N. Gleiss (Sweden), Dr. D. L. Richards (U.K.), Mr. R. B. Archbold (U.K.) and Mr. G. J. Barnes (U. K.).

Mr. Gleiss' paper is printed in full below, and so is the contribution presented by Mr. Archbold.

Dr. Richards, whose work is concerned with all aspects of speech transmission, discussed the relationship between subjective and objective standards of transmission performance for telephony. He showed the importance of being able to correlate subjective and objective measurements, and how this could be achieved with the aid of an intermediate reference system (IRS). One of Dr. Richards' major achievements has been the unification of the three separate developments of assessing speech transmission, namely by calculation, by direct subjective measurements, and by objective measurements. He has identified and quantified many of the factors affecting speech transmission, so that their separate effects can be assessed.

In this presentation Dr. Richards demonstrated a computer program, based on his theoretical model for calculation of opinion scores for telephone connections, and showed how expensive and time consuming subjective assessment scores can be replaced by a calculation method with fairly good agreement between the two methods.

His presentation was supported by a number of documents:

1. Calculation of Opinion Scores for Telephone Connections Proceedings, The Institution of Electrical Engineers, Volume 121, Electronics, No. 5, May 1974).
2. Two Post Office Internal Research Memoranda.
3. New Definitions for Loudness Ratings. (Electronic Letters, 19th September 1974, Vol.10, No. 19).
4. Choise of Parameters for Calculating Loudness Ratings of Telephone Speech Paths (Electronic Letters, 31st May 1973, Vol.9 No. 11).
5. New Definitions for Loudness Ratings. (Proceedings, The Institution of Electrical Engineers, Volume 119, Electronics, No.10, October 1972).

Mr. Barnes suggested provision objective equipment for the measurement of loudness ratings based on Dr. Richards' work, and using the B & K 3352, Electroacoustic Telephone Transmission Measuring System, supporting his presentation by two research reports:

1. The Design and Use of Instrumentation for the Determination of Loudness Ratings.
2. Electrical Calibration of Instrumentation for the Determination of Loudness Ratings without the Use of Physical IRS.

15 papers were presented at the symposium with a good balance between subjective and objective measurements. Of these papers 7 are published in this issue of the B & K Technical Review. The remaining papers are listed below, and those for which a full text exists, are available in the form of a B & K Application Note: Telephonometry Symposium 1974.

List of Lectures not published in TR1-1975:

O. LARSSON (Sweden):	What should be required of telephone measuring equipment? Experiences with production control.

H. AHLBERG (Sweden):	Telephone Measuring Equipment.
I. JÄNTTI and T. TUISKU (Finland):	Review of current telephonometric measurements in Finland.
T. ULSETH (Norway):	Measuring methods and experiences from objective measurements of reference equivalents in Norway.
E. LAUKLI (Norway):	Acoustic feedback in telephone sets.
I. SALAMA (Finland):	Factors determining the reproducibility and accuracy requirements of telephonometric measurements. Review of some practical experiences.

Problems in Telephone Measurements

by

Norman Gleiss

Methods of measurements

Telephone networks are constructed primarily for speech transmission. Measurements on a telephone connection or its components are, therefore mostly directed to assessing only such properties as have any influence on the transmitted speech quality.

Such measurements may be carried out either by electroacoustic or psychoacoustic methods. The former involve purely physical, "objective" methods, while the latter may consist of different kinds of conversation and listening tests, or "subjective" methods. These two radically different categories of measuring methods bring up a number of important problems and questions, all of which cannot yet be answered.

For instance, is it possible that the "subjective" assessment of a characteristic of a telephone transmission link is sufficiently "objective" to be used as a base for sales contracts between manufacturer and consumer? The choice between different transmission systems or parts of it may have farreaching consequences. Would it not be better to rely upon well-known physical measuring methods of high precision than to use the more or less diverging judgements from a number of persons?

Evidently, listening tests and electroacoustic measurements are necessary, but often under different conditions. For receivers, for instance, it is possible to set up a specification founded on empirical data and to prescribe the shape of the frequency response curve, measured on a certain artificial ear, as well as the level. At both production and acceptance inspection, physical measurements can therefore be used as a simple and rapid method for checking if a sample meets the specification. In this case a listening test would be neither justified nor practical.

For type tests, the situation is different. As long as a receiver type having a known response curve is manufactured, a check of the electroacoustical efficiency by a fairly simple physical method will suffice. If on the other hand a new design shall be tested that results in a different response curve it will probably be necessary to carry out a listening test, even if the acoustic efficiency is exactly the same as before. It may then appear that speech is judged to sound unnatural over this receiver because of a less suitable shape of the response curve. If the difference between two designs is small in this respect, the tests, even if they are systematic and well planned, have to be very extensive to make certain the difference perceived.

Thus it may be said that physical measurements are preferable but require knowledge in advance of the performance characteristics desired. At best the optimum values of a few measurable parameters are known which in combination define the transmission quality aimed at. More often there exist minimum requirements on one parameter (e. g. efficiency) whereas it is not possible to describe quantitatively the influence of other factors (e. g. distortion). Then it will be necessary to complete the test by a "subjective" measurement.

Standard methods

There are a number of methods for the assessment of the transmission quality of a telephone connection which have been accepted as standard methods by the telecommunications administrations of different countries. Some of them are furthermore recommended by the CCITT*. All these assessment methods are subjective methods, although it has been tried to replace them by objective methods in some cases. Such objective methods may consist of a physical measurement rendering an integrated final value or of a calculation founded on physical data.

Three more or less standardized psychoacoustic methods will be briefly described here.

a) Reference equivalent

This is a measure of how the loudness of speech on a telephone connection is affected by attenuation and bandwidth restriction or, to put it more generally, by amplitude/frequency distortion. The measurement implies a comparison by ear of speech transmitted alternately over the system to be tested and over a reference system (the NOSFER) hav-

* The International Telegraph and Telephone Consultative Committee

ing well-defined transmission properties. The talker continuously repeats a test phrase at constant speech level. Normally a trained group of three persons participates, which makes six talker-listener combinations.

In the reference system there is an attenuator which the listener has to adjust until he judges the loudness to be equal over both systems. The attenuation value chosen tells how much worse the test system is compared to the reference system. If the test system is better than the reference system the reference equivalent becomes negative. Usually the reference equivalent is assessed for sending and receiving separately (Fig.1).

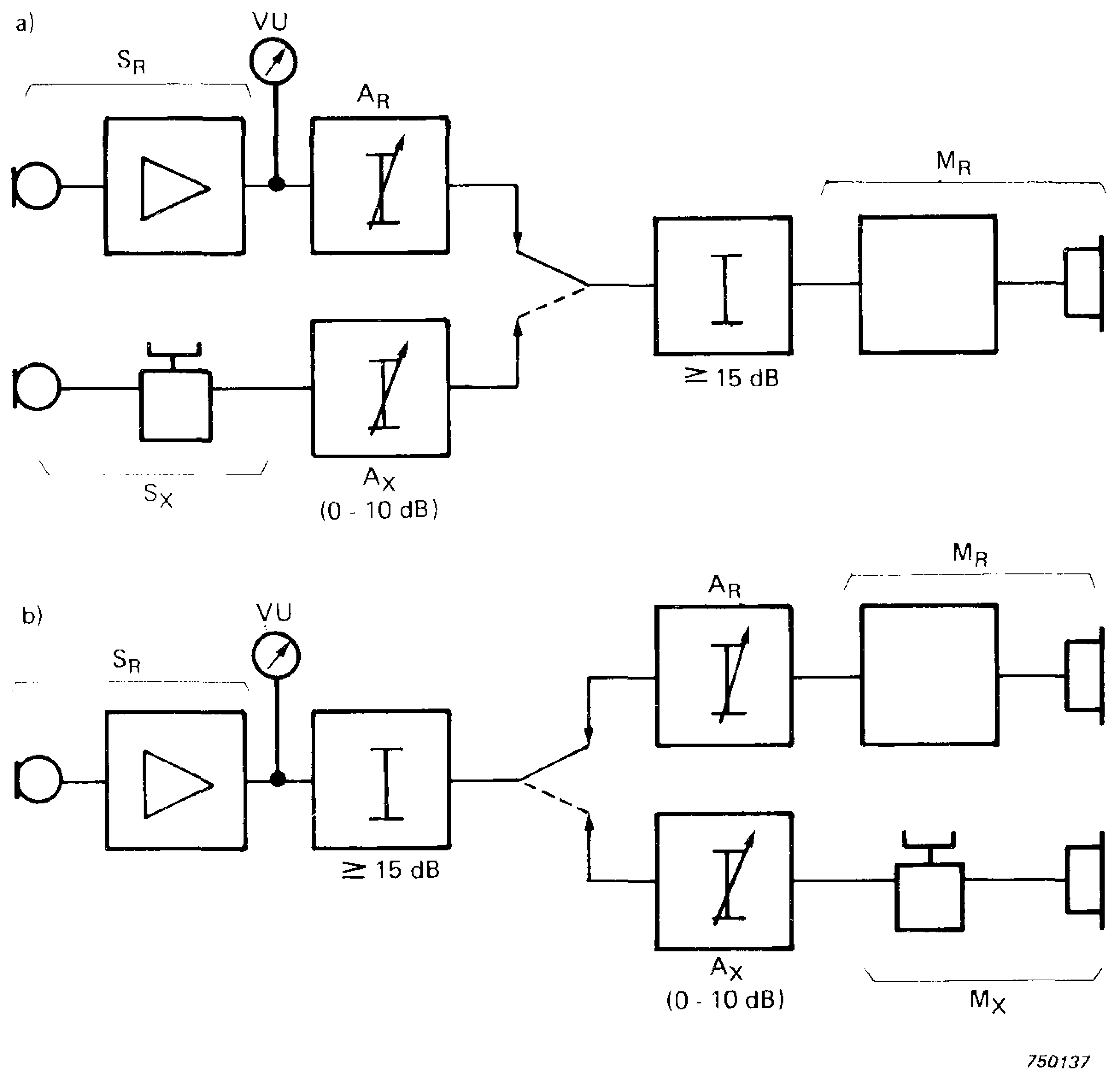


Fig. 1. Block diagram for reference equivalent measurements: a) Sending b) Receiving. Index R denotes the reference system and X the test system

The reference equivalent is one of the oldest measures of telephone quality and is still of great importance in the planning of telephone networks. The method of comparing the arbitrary frequency response of a test system with the practically frequency-independent response of the reference system implies an averaging over the response of the test system and makes possible a comparison between systems which differ very much in the speech level transmitted.

However, the subjective comparison becomes more difficult the greater the difference is between the frequency response of the two systems, because of the change in the timbre of the received speech. Another problem is the deficient additivity of the reference equivalents of separate parts of a link, if there are differences in bandwidth between the parts or compared to the reference system. This deficiency may be reduced by the introduction of an intermediate reference system with nearly the same frequency response characteristics as the telephone systems under test. This is a problem presently considered by a Study Group of the CCITT.

The assessment of reference equivalents is presupposed to be carried out in the absence of circuit noise and room noise. Since it is quite possible that the rank order between transmission systems of different frequency response might be changed by the presence of noise, the validity of the reference equivalent is limited to the undisturbed case.

b) Articulation tests

The intelligibility of speech sounds or words as a measure of the quality of a telephone system has been used for about as long a time as the reference equivalent. Distinction is made betwen the intelligibility of sounds, syllables, words, and sentences. The articulation score is defined as the ratio between the elements correctly understood and the total number of verbal elements transmitted, i. e. vowels and consonants, syllables, words, or complete sentences.

In order to obtain a reliable articulation score for a system it is necessary to carry out an extensive series of measurements. Since the articulation to a high degree depends on the listener's habituation to the task, training series are required, making the tests very time-consuming.

For many years the CCITT used a special method called AEN, where the attenuation in both the test system and the reference system (ARAEN) was adjusted to yield 80% articulation and the level difference then existing was used as a quality measure. However, the method is no longer in use.

c) Opinion tests
Opinion tests may be two-directional (conversation tests) or one-directional (listening tests). In a laboratory conversation test observers converse in pairs over the telephone connection under test and are then asked to score the transmission quality on a five-grade scale (0-4), where each scale value is verbally defined, e. g. ranging from "Bad" to "Excellent". The procedure is generally repeated for different values of a test variable affecting transmission quality, such as attenuation or noise level. As a result, mean opinion scores are obtained, which can be combined into curves (Fig.2). In listening-only tests the speech material is usually presented from taperecordings.

Opinion tests can also be used for the assessment of factors other than overall quality, such as listening effort or loudness.

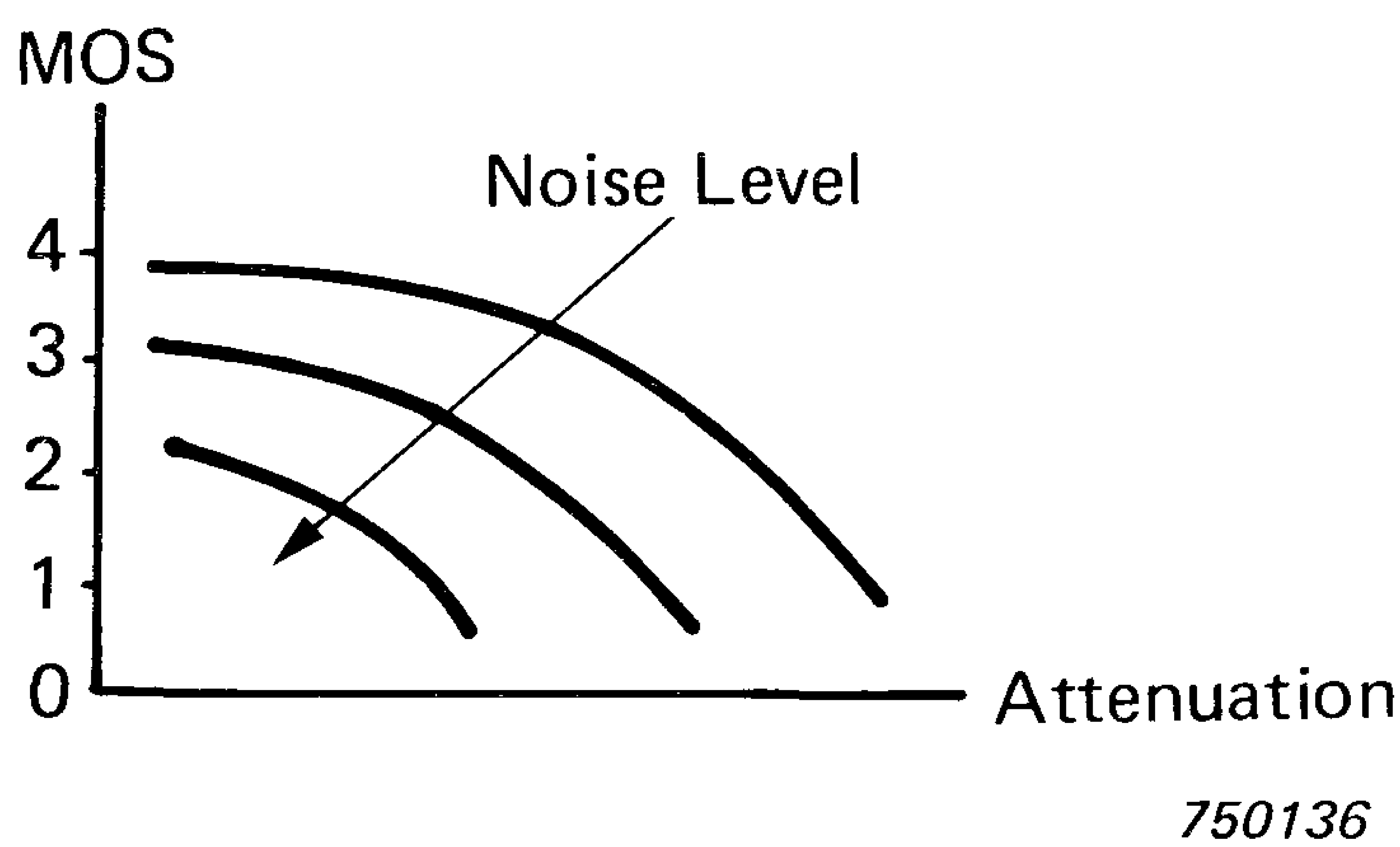


Fig.2. Judgement of telephone transmission quality. The Mean Opinion Score depends on both attenuation and noise level

The uncertainty of subjective measurements
All kinds of measurements may be subject to random and systematic errors. The random errors are supposed to depend on stochastic factors and are therefore expected to be normally distributed around a mean value which for repeated measurements approach zero. The mean of the observations will then approach the true value, if there is no systematic error present. From the observations the standard deviation s of the random error and $s/\sqrt{n}$ of the mean value can be calculated, and from that a confidence interval around the mean is formed. The true value of the quantity measured lies in this interval with a certain, known probability. The interval can be made arbitrarily narrow by increasing the number of measurements, n.

Systematic errors of physical measurements are as far as possible eliminated by the use of recognized measuring methods, by a careful calibration of instruments, and by instructions about the correct procedure. Often a measuring method can be checked by measuring a known quantity.

For subjective measurements the situation is different. Since the purpose is to measure a sensation, there is no possible way of knowning the "true" value. Still, systematic errors can be compensated by the use of well-established methods of experimental psychology and by a careful control of irrelevant factors that may affect the results. At best, one then obtains for a certain subject a distribution of observations where the dispersion consists only of casual fluctuations, the **intraindividual** variation, around his individual mean value. The variation need not be large for a trained observer and is finally given by his sensorial sensitivity. For sound intensity, as an example, the just noticeable difference is of the order of 1 dB, which is the same as the error for electroacoustic measurements.

In telephony, however, there is less interest in the opinion of one single individual than in an average value that is representative for all subscribers in a country. It is necessary to take account of the **interindividual** variation, which is caused by the fact that different persons' sensation of the same stimulus may be consistently different from each other. This variation is often much larger than the intraindividual variation. A laboratory investigation with a small group of observers may therefore involve a sampling error, if the group is not representative of the relevant population. Experience has shown that at least 20 subjects should participate to yield sufficiently valid results. In this respect objections can be raised against the reference equivalent measurements in their present form.

Objective measurements

From what has been said, it may appear that subjective measurements are influenced by a great number of unknown factors, that the dispersion between individuals is large and generalization therefore doubtful, and that the measurements are laborious and time-consuming. Knowing at the same time that objective measurements render very reliable and repeatable results even by quite simple methods, it is easy to conclude that objective methods always are to be preferred.

It is not difficult to meet such reasoning. The truth is that in objective telephone measurements one often very precisely measures something else than what one wants to know.

In psychological test theory, it is usual to distinguish between reliability and validity. By reliability is meant how well a test measures what it does measure, while the validity tells how well a test measures what it should measure. It can safely be stated that subjective methods in telephony have a lower reliability but a higher validity than the objective methods. Since validity is dependent on reliability (but not the reverse) it is important to devote sufficient care to methodological problems in subjective telephone measurements.

Naturally it is desirable to use objective methods as far as possible. In a stage of development, when new situations and problems are encountered, subjective measurements are obviously necessary. On the other hand, for checks and comparative measurements on existing transmission systems, whose properties in principle are known, it is quite possible to use physical test methods also for the assessment of such measures as the reference equivalent. A condition is that the objective method can be validated against the subjective method. This is done by measuring the same system by both methods. If the correlation between the results is sufficiently high, it is generally possible to calibrate the objective measuring apparatus to an acceptable degree of agreement, at least within a limited range of operation.

The problem of validation exists even for relatively simple electroacoustic measurements such as recording the frequency response of a telephone receiver. If the response curve is to give any information about the receiver performance on the listener's ear, the volume coupling the receiver to the measuring microphone should approximate the acoustic impedance of the average human ear. A few years ago the IEC has standardized such an artificial ear for audiometric measurements, which later was adopted by the CCITT for telephonometric measurements after a series of checks comparing telephone receivers on artificial ear models and on real ears. It may be necessary to improve the correlation between these two kinds of data further by introducing an acoustic leak to simulate high-level listening conditions.

The electroacoustic measurement of reference equivalents by a composite apparatus such as OREM involves problems of another kind than those concerned with the design of artificial ears and voices. The apparatus has to simulate the loudness perception process of real listeners, which is not yet sufficiently known to solve the problem for all possible received speech spectra. However, it seems that the perceptual simulation can be done empirically with reasonable accuracy, while it still remains to construct artificial mouths and ears which work for all shapes of telephone handsets. Furthermore, the validation procedure is difficult because of inherent deficiencies in the subjective measurement of reference equivalent, which lead to inconsistent results.

Concerning articulation tests there exist several acknowledged methods which permit the calculation of articulation scores from frequency response curves and noise spectra (1, 2, 3). However, there is no instrumental implementation available yet, nor is there any apparatus for measuring overall speech transmission quality. This raises the question which measures of quality are useful at all.

Transmission quality factors
Much information has been collected on the dependence of speech transmission quality, e. c. as expressed by Mean Opinion Scores, upon system parameters such as attenuation, bandwidth, or noise level. Among other things the data show that there is an optimum received speech level for each noise level value (Fig.3). Regarded in this way,

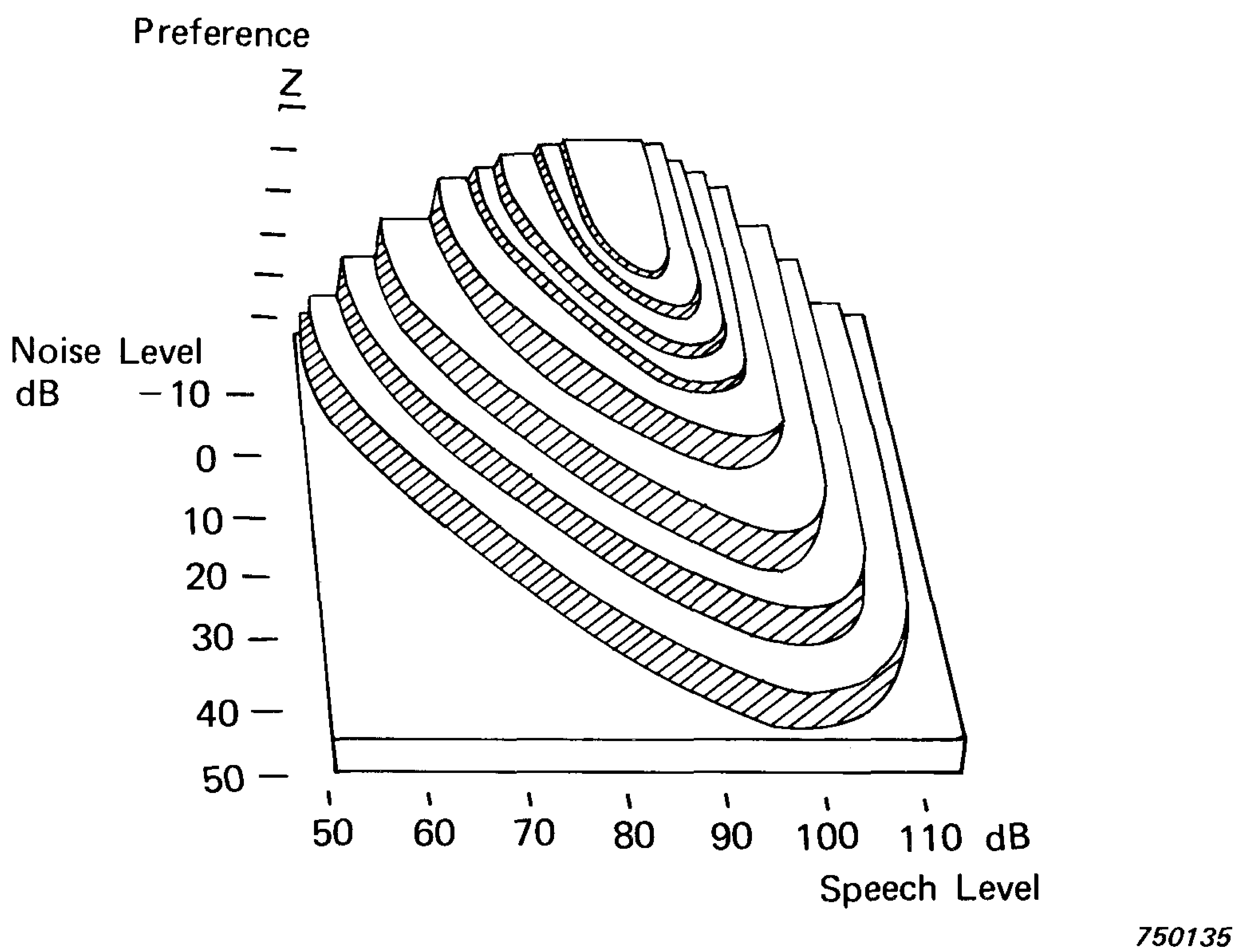


Fig. 3. Three-dimensional representation of "isopreferene curves" for a 3 kHz loudpass transmission system. The contours connect combinations of speech level and noise level that are equally preferred in a listening test (having subjectively equal transmission quality)

speech loudness is only one piece in a complex of problems. Therefore, the benefit of reference equivalent measurements should not be overestimated. The most obvious application is for network planning, but even there it has lost something of its original importance.

Recent research (4) has indicated that the perception of speech transmission quality is two-dimensional and may be divided into two independent factors: **intelligibility** and **naturalness**. The first factor represents the semantic content of the speech signal, and the second factor represents information about the speaker (his identity, state of mind, etc.). Naturalness seems to be more dependent on the presence of the fundamental and first formant in the reproduced speech spectrum, and intelligibility more on the higher formants.

Earlier telephone systems often had a large amplitude/frequency distortion, which together with the rather limited bandwidth and high line attenuation resulted in poor intelligibility. For modern telephone systems the syllable articulation is 98-99%, which means that the sentence intelligibility is for all practical purposes 100%. Hence it is seldom possible to differentiate between systems having somewhat different bandwidth and frequency response by means of articulation tests alone. Instead there is much more to be gained by considering the naturalness factor, especially when further improvements on high quality systems are aimed at.

A brief presentation of the factor-analytical model may underline this statement. The model assumes that every judgement is composed of a number of psychological factors which to a varying extent contribute to the total score, depending on how much importance is attached to them by the judge. For the judge i and system j the test score will be

$$Z_{ij} = a_i X_j + b_i Y_j + \dots$$

X, y, ... are the factors (which may be more than two). X_j, Y_j are the scores that system j gets on these perceptual factors, and the coefficients a, b are the so-called factor loadings representing the relative weight attached to the factors by the judge.

It is seen that according to the model the factor loading is invariant over systems and the factor scores invariant over judges. This relation makes it possible to separate between the subjective and objective parts of a measuring result.

Although the factors X, Y represent sensational dimensions and consequently the factor scores X_j, Y_j can only be assessed by "subjective" methods, the factor scores represent properties of a system that are objective in the sense that they are independent of the observers making the judgements. The subjective effect lies in the factor loadings, which depend not only on the judge but also on the matter transmitted by the systems; that is, in telephony, they depend not only on the listener but also on the talker.

Both factor scores and factor loadings can be extracted from opinion tests by correlation calculations. The next step is to search for relations between physically measurable characteristics and the psychological factors. By that the validation of new "objective" measuring methods would be much facilitated.

Conclusions

Evidently, subjective methods are necessary within certain parts of measuring technique in telephony. The methods involve many risks of errors but these can be reduced by the use of acknowledged experimental procedures and by a careful design.

On the other hand, the trend in the development of measuring technique is to replace subjective judgements by electrical and acoustical measurements. This will be successful at least for routine measurements. The condition is that the relation between data obtained by different methods is fully known, so that the interpretation of test results is univocal.

References

1. L. L. BERANEK: The design of speech communication systems Proc. IRE Vol. 35, 1947.

2. D. L. RICHARDS and R. B. ARCHBOLD: A development of the Collard principle of articulation calculation. Proc. IEE Vol. 103B, 1956.

3. N. R. FRENCH and J. C. STEINBERG: Factors governing the intelligibility of speech sound. J.A.S.A. Vol. 19, No.1, 1947.

4. M. HECKER and N. GUTTMAN: Survey of methods for measuring speech quality. J. Aud. Eng. Soc. Vol. 15, No.4, 1967.

Proposals for the measurement of Loudness Ratings of Operators' Headsets

by

R. B. Archbold

Introduction

One field of telephonometric measurements that appears to have received relatively little attention is that related to Operators' Telephone Circuits (OTCs). For example, no CCITT Recommendation exists on the measurement of Reference Equivalents of these circuits or for the objective determination of sensitivities. On the other hand Recommendations relating to Reference Equivalents of Local Telephone Circuits (LTCs) have existed for many years and considerable progress has been made of recent years leading to proposals for a method of measuring Loudness Ratings as a preferred alternative to Reference Equivalents (Question 15/SG XII, Vol. 5 Green Book pp 216 — 259). With LTCs, proposals are also included for measurement of objective sensitivities appropriate to calculation of Loudness Ratings and other measures of performance (e. g. Speech Voltage, Conversational Opinion Scores).

In view of this the author of this paper, as rapporteur for CCITT Question 3/XII, recently presented first tentative suggestions to CCITT Study Group XII for corresponding Loudness Ratings and objective measurements of CTCs. These will shortly appear as a formal CCITT Contribution but meanwhile opportunity is taken here to extract the main points of interest and problem areas of the measurements.

Characteristics of Operators' Headsets which give rise to measurement problems

Modern headsets take a variety of forms which can generally be considered as falling into two broad classes, i. e. those with an external earphone frequently associated with a horn loaded transmitter, boom microphone or voice tube and those of lightweight form, often with insert ear-

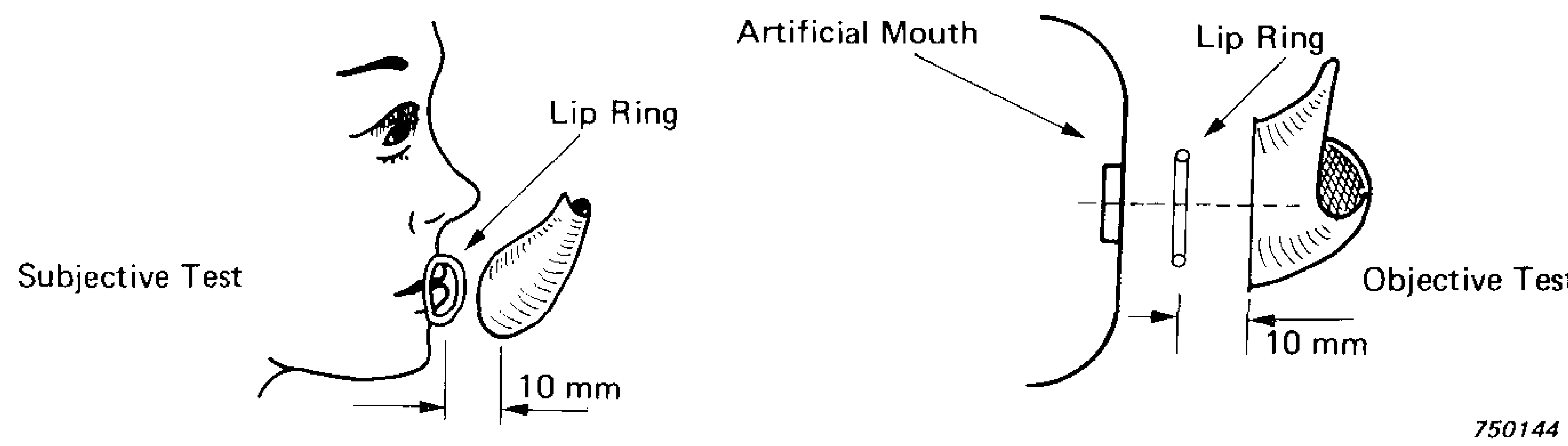


Fig. 1. Microphone Testing Position used formerly as a simply defined position for inter-laboratory tests
Not suitable for some modern Operators' Headsets

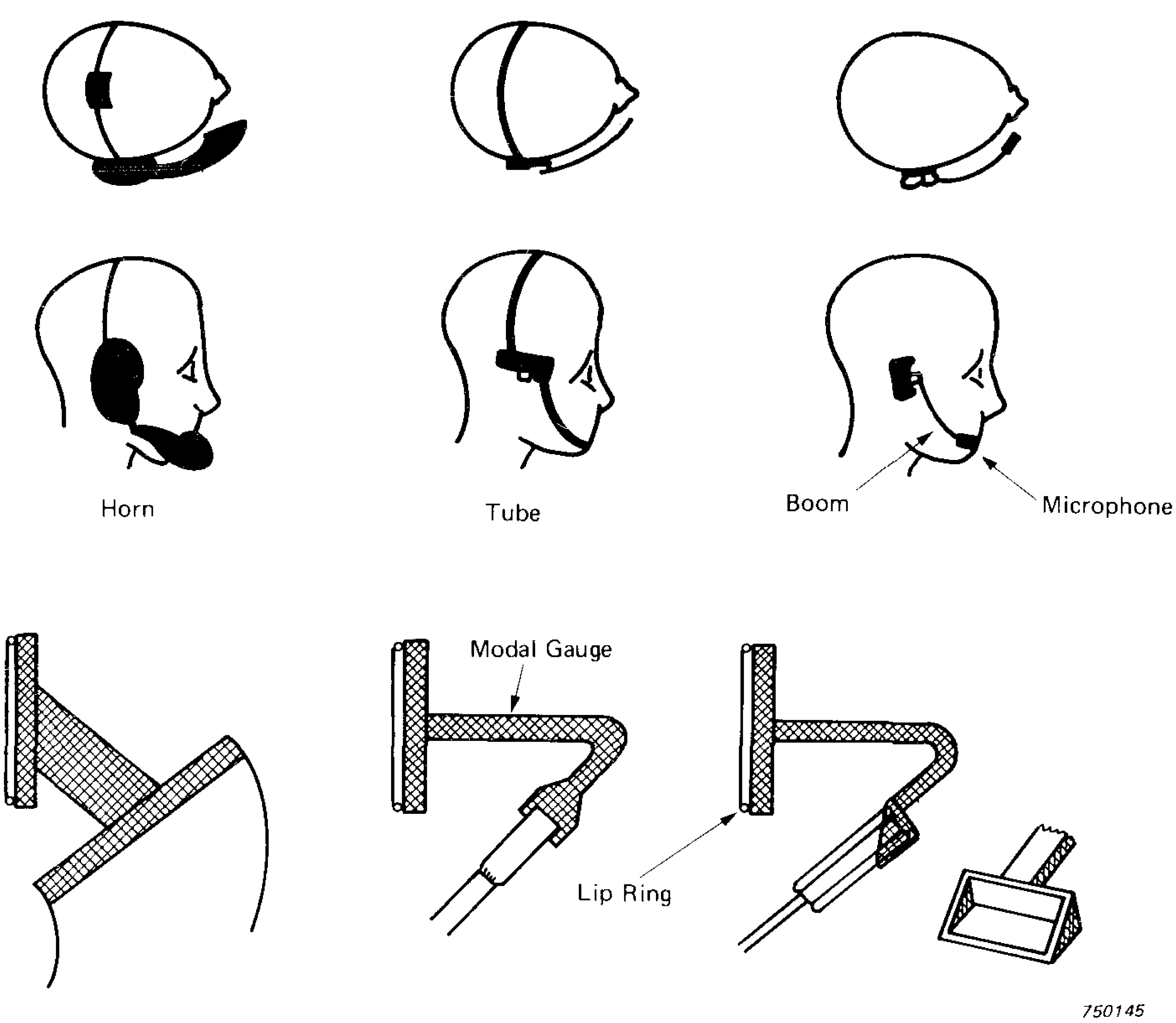


Fig. 2. Pictorial Representations of possible forms of Modal Positioning Gauges
(NB. Although the lip ring and microphone/horn faces are shown in the same plane normal to the paper they will in practice be angled)

phones and voice tubes (Fig. 2 illustrates some of these). With the latter type appreciable variation arises from the form of coupling to the ear, sometimes a true insert is used which enters the ear canal, in other cases a slightly bulbous tip is designed to rest on the entrance to the ear canal or is only partly inserted in the canal. Whilst the former 'insert' type frequently, but not necessarily, ensures a good seal with the ear the latter in many cases does not.

One further restriction is that the microphones of some modern sets react to direct breath pressures by giving rise to an effect known as 'blasting', this does not permit a simply defined talking position such as that shown in Fig.1 to be used even for control tests between laboratories.

Problems of measurement
Some of these problems arise directly from the above characteristics and others from the strong need to align the testing methods for OTCs as far as possible with those for LTCs in order to simplify calculations of transmission performance.

Subjective Loudness Rating
For LTCs loudness balances are made of circuit combinations of the LTC under test and a stable Intermediate Reference System (IRS) employing handset transducers against the CCITT NOSFER circuit (Green Book Vol. 5 pp 236— 238) as illustrated in Fig.3. For Operators' circuits the LTCs can be conveniently replaced by OTCs using specified lines and feed circuits. Problems arise from

a) Difficulty in defining methods of wearing and positioning headsets. With OTCs at the 'send end' the position of the sound 'pick-up' point or area may need to be located to a set of rules provided by the manufacturer or telephone administration. For practical purposes this can be achieved by means of a modal gauge as in Fig.2 using the lip ring of the Fundamental Reference System as datum.

   It may be that ultimately a simply defined compromise standardized position can be adopted for location purposes at least as a control position for inter-laboratory tests. Such a choice would automatically exclude a simple position as in Fig.1 for the reasons given in Section 2 of 'blasting' effects.

b) The need for more complex form of loudness balance when insert earphones are involved, because the simple practice used for LTCs of holding two receivers in one hand for alternate listening is no longer possible.

A contra-lateral method is proposed which requires balances to be made with the insert receiver on one ear against the NOSFER receiver on the other and the process repeated with the two receivers interchanged. Fig.3 gives the balance relationships used to correspond to those for LTCs. The method assumes reasonably normal hearing in both ears on the part of the speech test crew members.

Objective Sensitivity measurements appropriate to calculation of Loudness Ratings
Problems in this area again arise in part from the need to seek compatibility with the corresponding LTC measurements but more particularly from the known, and in some cases probable, inadequacies of measuring equipment, e. g. artificial mouths and ears. More specifically some of the problem areas are

a) With modern headsets the sound pick-up region of the microphone is often closer to the face than with a handset microphone and not directly in front of the lips. Indeed the situation may be that existing determinations of sound pressure distributions around the real mouth do not yield sufficient information in the area of interest. Similarly the distributions around existing artificial mouths may need to be studied.

   For the calculation of loudness ratings it appears that the Mouth Reference Point (MRP) used for LTCs to which the sound pressures are referred can be conveniently retained on the axis of the artificial mouth 25 mm in front of the Equivalent Lip Position.

b) For 'receive' measurements studies are still required on the adequacy of existing forms of artificial ears/couplers to simulate the real ear impedance as seen by various forms of 'insert' receivers. (As an interim measure tests will be made with the 200 coupler to IEC Recommendation, Publication 126).

   In parallel with the above, methods of coupling the various forms of so-called 'insert' type receivers to the artificial ear may need to be studied.

   Jointly the two facets may give rise to considerations of compromise solutions on the form of artificial ear/coupler to use or even to simple means of adaptation to take care of the variation in earpiece types.

c) Calculation methods would be greatly simplified if a fixed reference point for measurement of sound pressures in real ears could be used or sound pressures readily related to it. For LTCs this point is the Ear Reference Point (ERP) which is defined as the centre of the plane of a circular telephone earcap when it is placed comfortably against the ear.

With insert receivers the situation is not so straightforward. If say a fixed reference point were chosen, for example 2 mm in front of the ear-drum, an almost impossible situation arises on the practicality of measuring pressures at this point.

The author proposes an alternative which appears to be worthy of study, namely that of using the same ERP as that used for LTC as a Hypothetical Ear Reference Point in the following manner.

In addition to the conventional loudness balance the same contra-lateral transfer technique should be used for determining loudness balances for narrow bands of noise between the NOSFER and 'Insert' receivers. In this way the equivalent sound pressure at the ERP of the NOSFER receiver can be found as a function of frequency. This in turn can be used in the calculation process.

If the method proves to be tractable it is a useful artifice in that

a) It avoids the difficulties of the real ear pressure measurements.

b) It includes in the determination any leak effect which may be present in the insert receiver and which, as previously stated, may be highly variable.

c) If the bands of noise are suitably selected to match those used for the calculation technique it can yield sound pressure/frequency characteristics directly applicable to the calculation process.

Objective determinations of Loudness Ratings
It will be seen from the foregoing that there will be a number of problems to solve before it can be said with certainty that a type of Objective Reference Equivalent Measurement (OREM) can be readily defined. If however studies of the real mouth/artificial mouth compatibility show that the former can be readily simulated, and similarly a satisfactory artificial ear/coupler arrangement can be proved, then objective instrumentation should be a reasonable possibility.

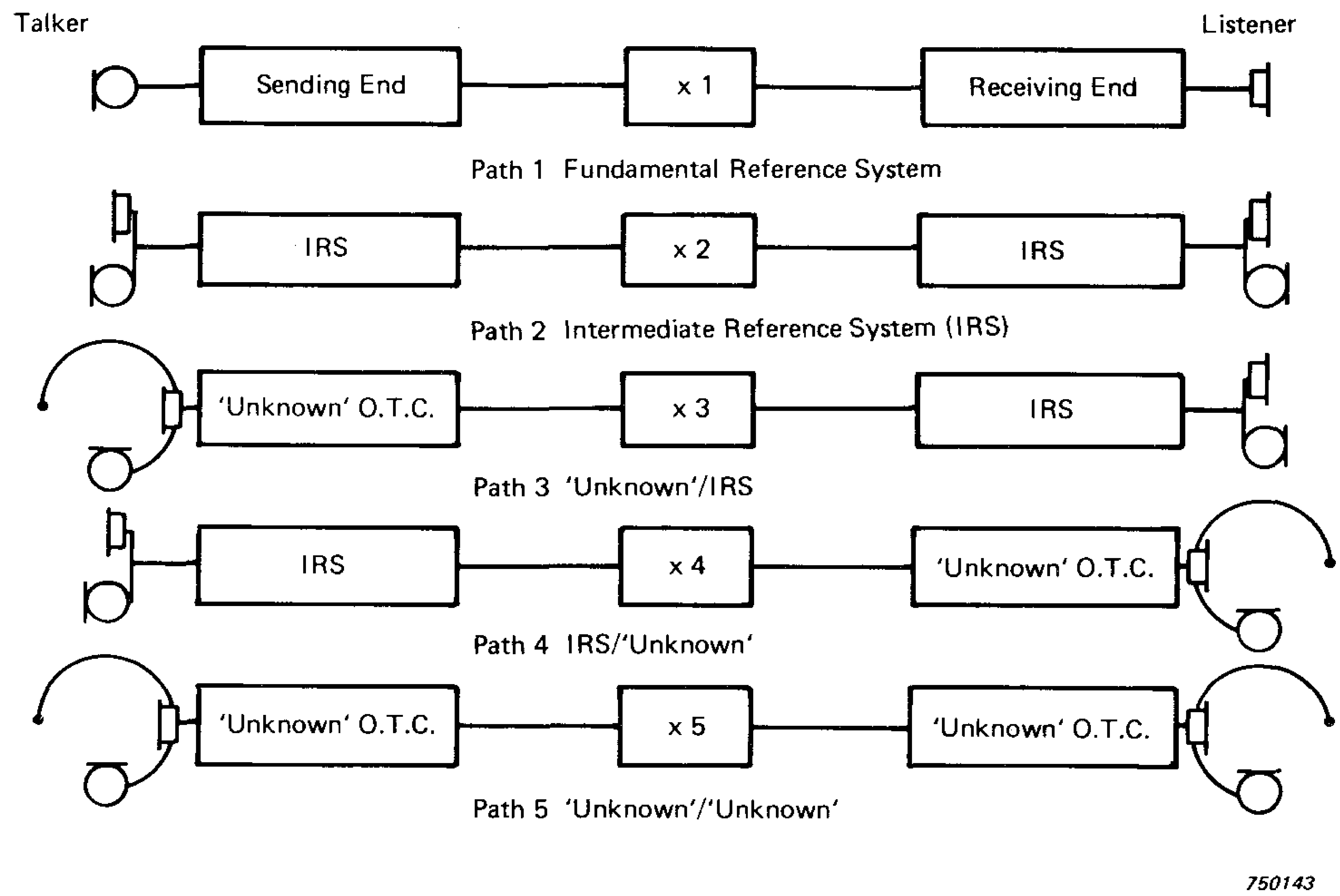


Fig. 3. Measurement of Loudness Ratings of Operators' Circuits (Based on Fig. 2.1, p. 238 CCITT Green Book, Vol. 5, for LTCs)

Notes:

1. *x_2 — x_5 adjusted for loudness balance against Path 1*
2. *The OTCs of Paths 3-5 replace the LTCs*
3. *Tests may show that Path 5 need not be included*
4. *For Operators' Circuits with insert receivers two balances will be required for Paths 4 and 5 against Path 1*

i. e. x_4' with Operators' Headset on Left Ear (Path 4) and Fundamental Ref. Receiver on Right Ear (Path 1) x_4'' with receivers interchanged

Then the Loudness Ratings relative to the Fundamental Reference System will be identical to those of CCITT Green Bk. Vol. 5 p 236 for LTCs by takin

$$x_4 = 1/2\,(x_4' + x_4'')$$

and $$x_5 = 1/2\,(x_5' + x_5'')$$

Subjective ratings of the Unknown OTC relative to the IRS are

Send Rating $= 1/2\,[x_2 + x_4 - (x_3 + x_5)]$

Receive Rating $= 1/2\,[x_2 + x_4 - (x_3 + x_5)]$

Interaction $= 1/2\,[x_2 + x_4 - (x_3 + x_5)]$

Comparison of Results obtained by Subjective Measuring Methods

by

Ib Gilberg

Introduction

Reference equivalent values for a transmission system or for the sending or receiving part of a transmission system, are normally given relative to the master reference system NOSFER. Until 1962 the master reference system was the SFERT, and in Denmark reference equivalent values have been determined since 1951 by the aid of a working standard system the so called SETED equipment. The SETED has been calibrated as well against the SFERT as against the NOSFER, and therefore it is possible to compare the ratings of the three systems.

It is my impression that some laboratories use slightly different measuring procedures and conventions, and what I want is to point out some of the sources which may cause discrepancies when results from different laboratories are compared.

SETED Filter

The SETED speech-path (figure 1 and 2) contains a 300— 3400 Hz filter having an in-band attenuation of 6 dB. In the CCITT papers the filter

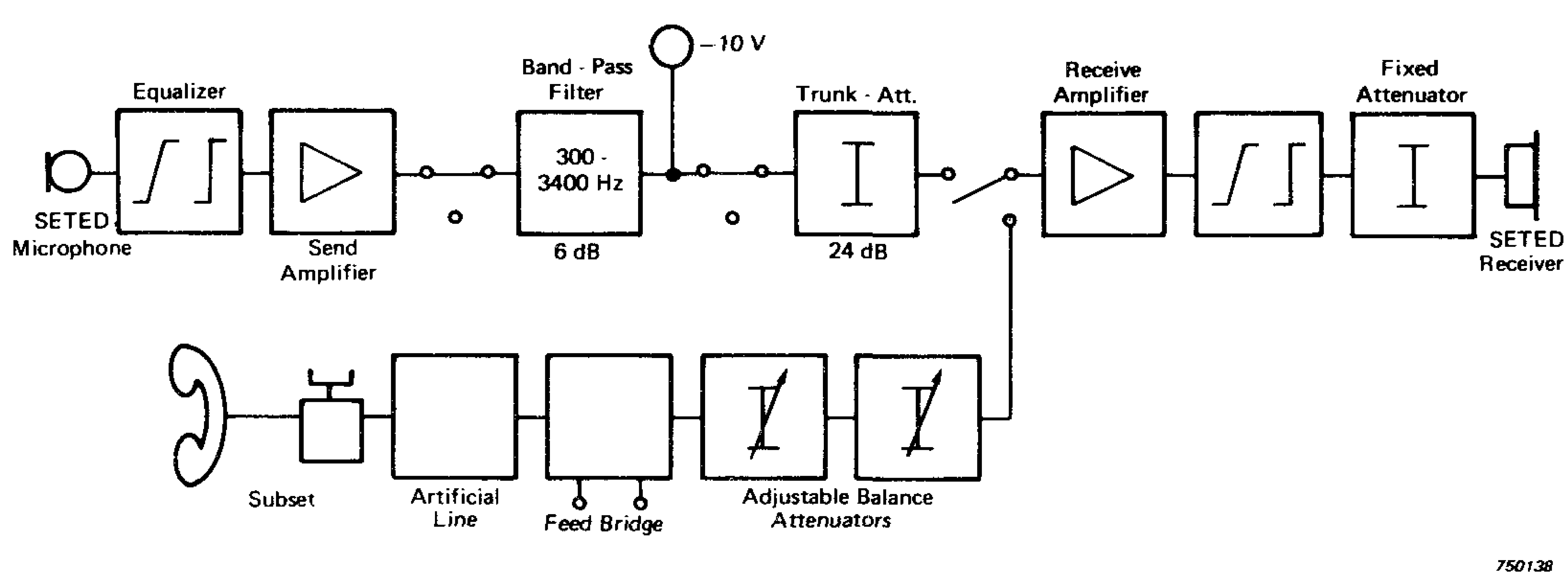


Fig. 1. SETED Sending

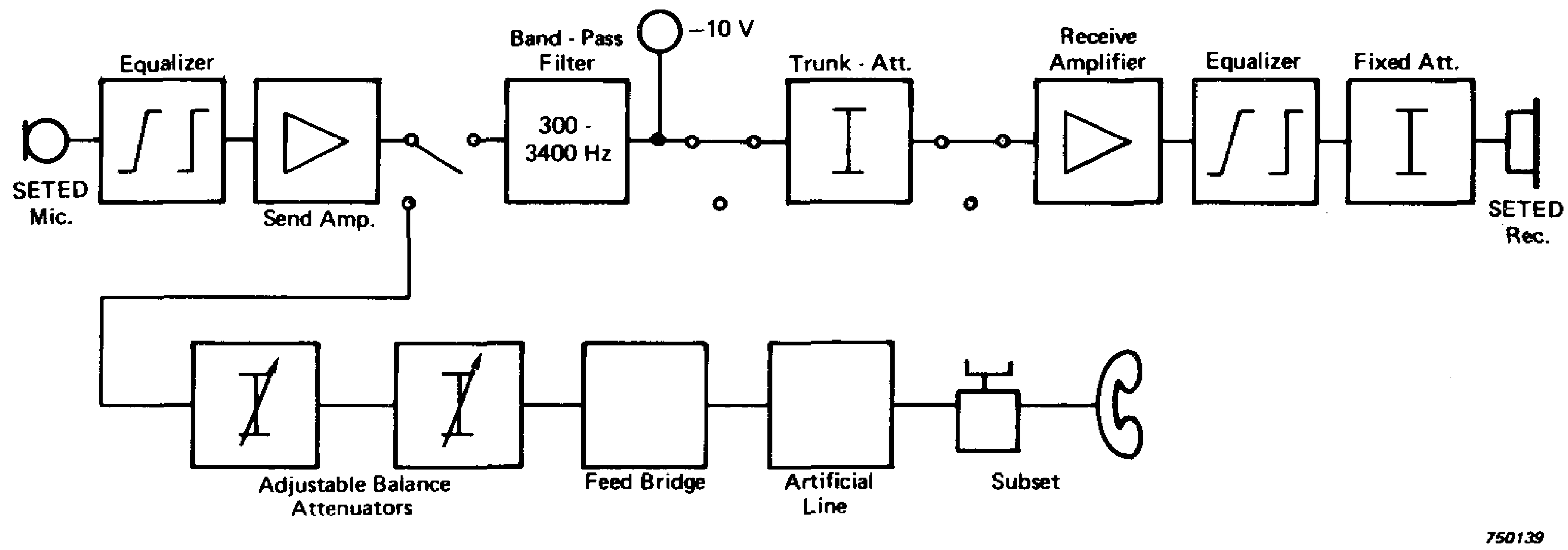


Fig. 2. SETED Receiving

attenuation until 1972 is considered as a part of the trunk attenuation, but in the Green Book, Vol. V pg. 66 it is stated, that the filter is considered as a part of the SETED circuit, and therefore when comparing SETED ratings it must be specified whether the filter is considered as a part of the SETED circuit or as a part of the trunk.

Bandwidth
During SETED measurements in our laboratory the telephone set under test normally is band-limited to 300 — 3400 Hz. The reason for doing this is that then only that part of the frequency spectrum, which can be transmitted in a carrier frequency system is subject to assesment. For telephone sets with linear microphones the frequency band limitation may cause differences in ratings of 2 — 3 dB, and again for comparisons the band limitation must be specified.

Conversion SETED — SFERT and SETED — NOSFER
To make SETED ratings to reference equivalent values, the SETED ratings have to be corrected with corrections obtained by calibration of the SETED working standard system against the SFERT or the NOSFER master reference system.

When the NOSFER replaced the SFERT in 1962 it was intended that the two systems should give equal ratings. The SETED has been calibrated as well against the SFERT as against the NOSFER, and the calibration shows that the ratings obtained by the two master reference systems are **not** identical. The values given in the White Book, Vol.V, pg. 26 are the following:

SETED sending
1953 3,3 dB louder than SFERT
1966 0,1 dB louder than NOSFER

SETED receiving
1953 1,1 dB louder than SFERT
1966 2,8 dB louder than NOSFER

The above mentioned values are with the 6 dB SETED filter loss considered as part of the trunk attenuation. In 1973 the SETED was recalibrated against the NOSFER, and in the Green Book, Vol.V, pg. 66 the following values are found:

SETED sending
1973 7,8 dB quieter than NOSFER

SETED receiving
1973 4,5 dB quieter than NOSFER

These values are with the filter in-band loss considered as part of the SETED speech-path, so to compare them with the 1966-values a 6 dB correction must be used:

SETED sending
1966 0,1 dB louder than NOSFER
1973 1,8 dB quieter than NOSFER

SETED receiving
1966 2,8 dB louder than NOSFER
1973 1,5 dB louder than NOSFER

Again for comparison of results between laboratories it must be specified which corrections have been used.

A special problem is the difference between the SFERT corrections and the NOSFER corrections especially for sending. The difference between the 1973 NOSFER value and the SFERT value is about 5 dB.

The Danish network-planning is based on SFERT reference equivalents and the nominal send- and receive reference equivalent values for the Danish telephone sets really are SFERT values even if they are quoted as NOSFER values.

The Danish telephone system actually works quite well, and with the network-planning already established it is impossible for exonomic reasons to change reference equivalent values from SFERT values to NOSFER values. Therefore we close our eyes and take the official standpoint that SFERT **is** equal to NOSFER.

NOSFER measurements P.72 and "R 25"

NOSFER measurements normally are carried out in accordance with the CCITT recommendation P.72. Two test procedures are described known as the "Two operator, hidden loss method" (figure 3 and 4) and the "Three operator, without hidden loss method". For both methods the listening level at the receiver depends of the sensitivity of the circuit under test and for the first method of the setting of the "hidden loss" attenuator.

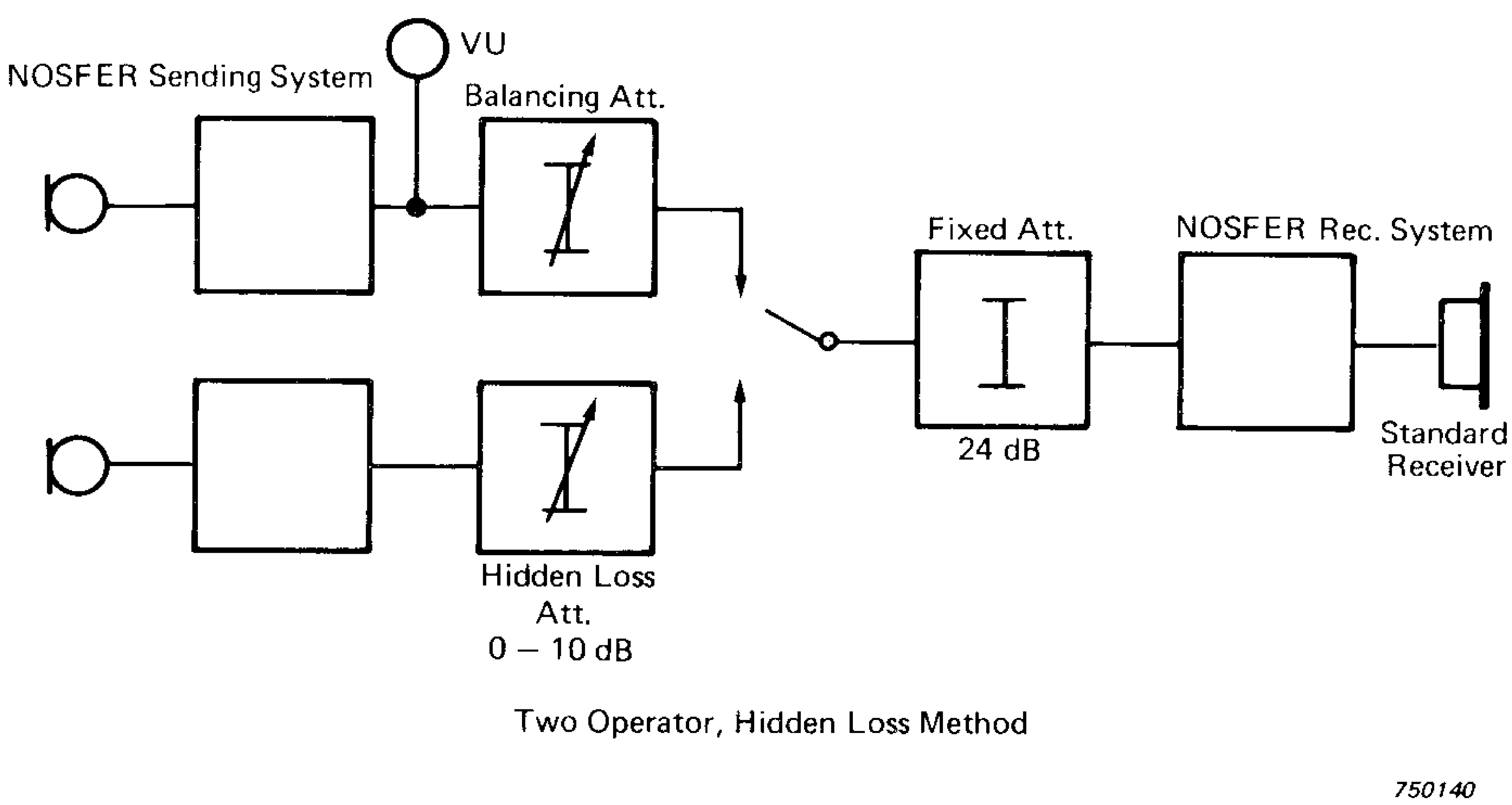


Fig. 3. NOSFER Sending

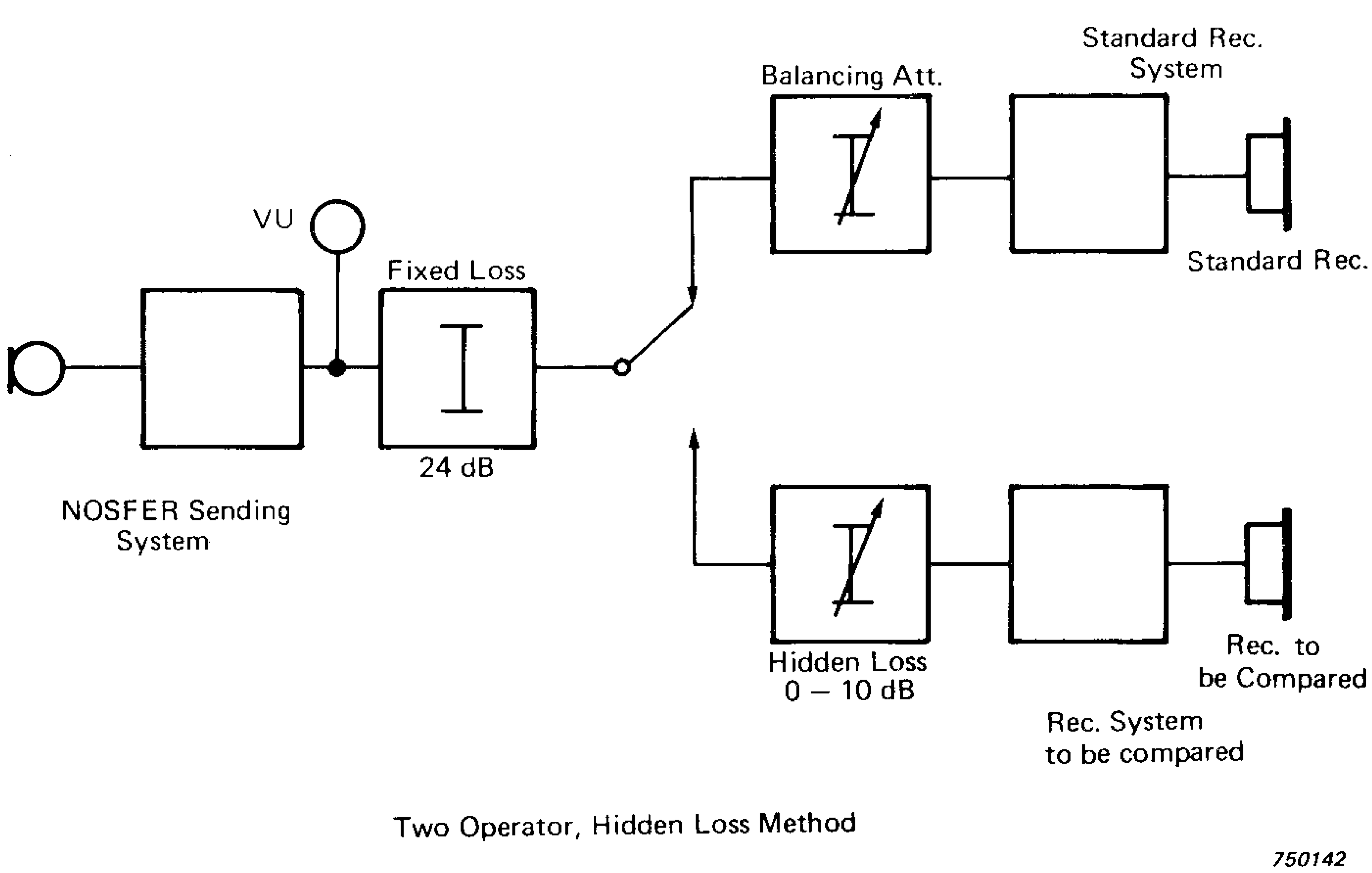


Fig. 4. NOSFER Receiving

In some recent CCITT papers (COM XII-No.3, 1973-1976, point I.2.3) a NOSFER measuring method referred to as "R 25" is described. The "R 25" method uses a constant listening level corresponding to a fixed NOSFER trunk attenuation of 25 dB.

A number of stable subsets have been measured by both the P.72 method and the "R 25" method (COM XII-No.36, 1973-1976) and Bell Northern Research has analysed the results and found a very interesting connexion between the sending reference equivalent values obtained by the two methods (figure 5).

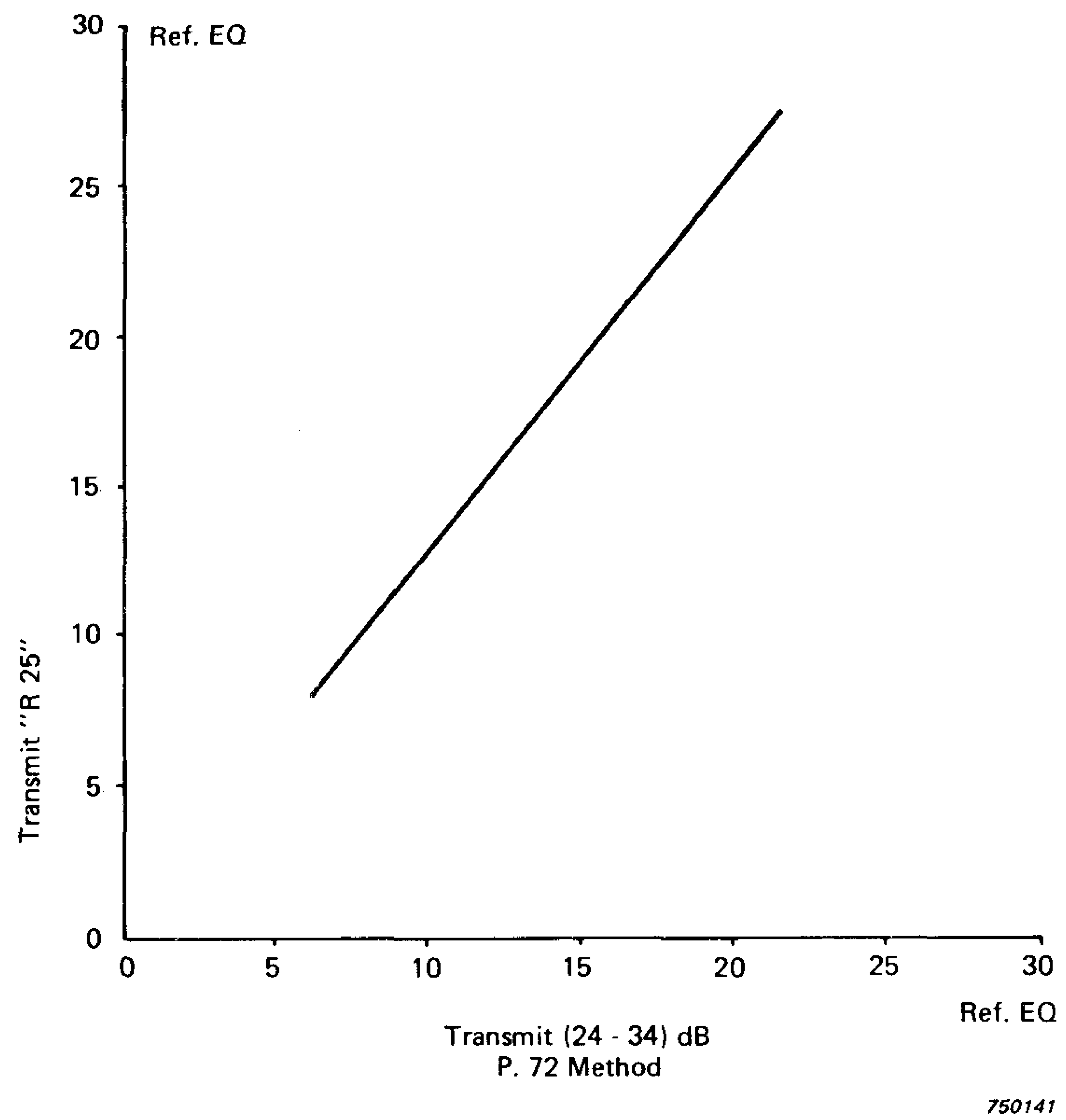


Fig. 5.

The connexion is linear but the slope of the line is not 1,0 but 1,27 indicating that differences between results obtained by the two methods are level dependent i. e. depending on the reference equivalent values obtained.

As the SETED measuring method and the "R 25" method both use constant listening level this seems to indicate that the SETED — NOSFER/P.72 corrections for sending also may be dependent on the reference equivalent values measured.

Conclusion

There are other factors acting upon the results of a subjective measurement. The speech material i. e. the words spoken during the test, the composition of the test crew, male — female, and other factors may influence and cause small differences between the results.

Therefore if we want to compare results obtained by subjective measurements in different laboratories it is very important that the measuring equipment and the measuring method to a very high degree is described.

Repeatabilities in electro-acoustic measurements on telephone capsules

by

R. E. Walford

To a large extent the days of several-dB tolerances in acoustic measurements have disappeared and most engineers now expect their objective work to yield results which are repeatable to within a few tenths of a dB — typically to within half a dB or so.

At Brüel & Kjær we are often asked to comment on telephone measurements in which one-tenth dB is said to be critical. Most of us here today are familiar with the practice of averaging test results from several measurements so as to show dB down to the hundredth place.

We know that a hundredth of a dB is a ridiculous quantity to argue about, but the practical question facing many of us is, just how close can electro-acoustic measurements get to some target value and how closely can this value be measured repetitively?

In an attempt to throw light on this, a tour was made recently of a mixed collection of telephone-manufacturing factories all operating the 3352 or 3353 Electro Acoustic Telephone Transmission Measuring System in their in quality control departments. Some stable telephone capsules were carried to all locations and were measured there by a standard method. The results, shown in Table 2, display an astonishingly good repeatability. However, to understand the implications of this the test conditions must be explained.

All the factories chosen are in Germany. This is the country having the largest number of 3352/3353 systems operated under the same test standard. The factories are widely dispersed around the country, and the tests were conducted over a period of four weeks, including a one/

week gap midway. The systems were mostly of the 3352 type, and their age varied from just over 4 years (from the very first production batch of the 3352/3353) to just under four months. Most of the systems were installed in factory quality control departments and most were in constant use during the working day with a few being used also on a second 8-hour shift. All were operated by semi-skilled staff with a qualified supervising engineer available for special work. Two systems were operated by skilled laboratory staff.

The test began with a check of the system. Reference voltages, and oscillator frequency calibration (at 50 Hz and 1000 Hz) were checked, and the routine calibration described in the 3352 Manual was done. In one case the 285 mV reference voltage was found to be wrong; it was corrected before the measurements were made. In another case the 1022 Oscillator was delivering the wrong sweep frequencies (approximately 40 Hz shift) and this too was corrected before measurements were made.

An important point was that the same 4230 Calibrator was used in all locations; it was used first to calibrate the systems and then the local Calibrator was compared with the single ("travelling") Calibrator to see what difference there was. Calibrator results are shown in Table 1, and it is obvious that differences between final test results would be greater if only local Calibrators were used. The histogram in Fig.1 displays the Calibrator values and it shows that Calibrator deviation could be a major part of differences between locations if only local Calibrators are used.

The actual test consisted of putting three moving-coil receivers Type 902 dyn III (from the firm of Fernsig, Essen) into the approved test jig and performing the standard OREM B test. These three capsules are exceptionally stable as shown by other tests performed earlier.

The tests on microphones were made by putting two Siemens transmitters Type TS 71 into the special test jig WA 0040 used in Germany.

It was known that local air pressure could have an effect on test figures so at each location this local pressure was measured. (The 3352/3353 systems are not affected by local air pressure differences but the test capsules do change slightly). In addition it is known that the Siemens piezo capsule has a slight dependence on room temperature so this too was checked. It was found to be so similar in each factory that it has been left out of the accompanying tables.

Three tables show the basic results.

First, (Table 1, Fig. 1) the behaviour of the locally-used 4230 Calibrators is shown. Each of these Calibrators was checked against one carried from location to location by the measuring engineer. It is not claimed that this travelling calibrator is any better than those used locally: it merely forms a convenient reference. The histogram shows the distribution of output pressure from the 24 calibrators measured. One is obviously outside the specified limits. This particular calibrator, however,

Location	B & K 4230 = +0,6 dB on 4904 scale		
	Local 4230: 4904 scale reading.	Sound pressure re B & K = 94,0 dB SL.	Sound pressure re FTZ 316066
1 449262	+0,1	94,5	94,3
2 260922	+1,0	93,6	93,4
2 316073	+0,5	94,1	93,9
2 371999	+0,3	94,3	94,1
3 as above			
4	+0,5	94,1	93,9
5	+0,5	94,1	93,9
6 298423	+0,7	93,9	93,7
7 as 6	+0,7	93,9	
8 396719	+0,4	94,2	94,0
9 459356	+0,21	94,4	94,2
10 356495	+0,59	94,0	93,8
11 372310	+0,6	94,0	93,8
11 332584	+0,6	94,0	93,8
12 372500	+0,6	94,0	93,8
13 same	+0,6	94,0	
14 372516	+0,5	94,1	93,9
15 356313	+0,5	94,1	93,9
16 316063	+0,6	94,0	93,8
17 316242	+0,6	94,0	93,8
18 same	+0,6	94,0	
19 385342	+0,6	94,0	93,8
20 316068	+0,3	94,3	94,1
23 372311	+0,1	94,5	94,3
24 419320	+0,2	94,4	94,2
Additional			
14 396739	+0,5	94,1	
15 316247	+0,5	94,1	
Repeat on 30/9			
1 449262	+0,1	94,5	
1 316066	+0,4	94,2	

750038

Table 1.

was not in use by its owner who possessed a second unit. All others are within + or —0,3 dB of the mean value with 21 lying within + or —0,25 dB of the mean value. A 0,25 dB shift becomes a serious matter when the final test results are examined so in all these factories it had already been agreed that each Calibrator should be assigned a correction factor measured against one unit held by a central agency. In practice the Calibrators are stable (if treated carefully) and so the apparent 0,5 dB spread is considerably reduced. For the tests performed here, of course, only one Calibrator — the travelling Calibrator — was used to check the test systems so that there is no correction required for Calibrator deviation.

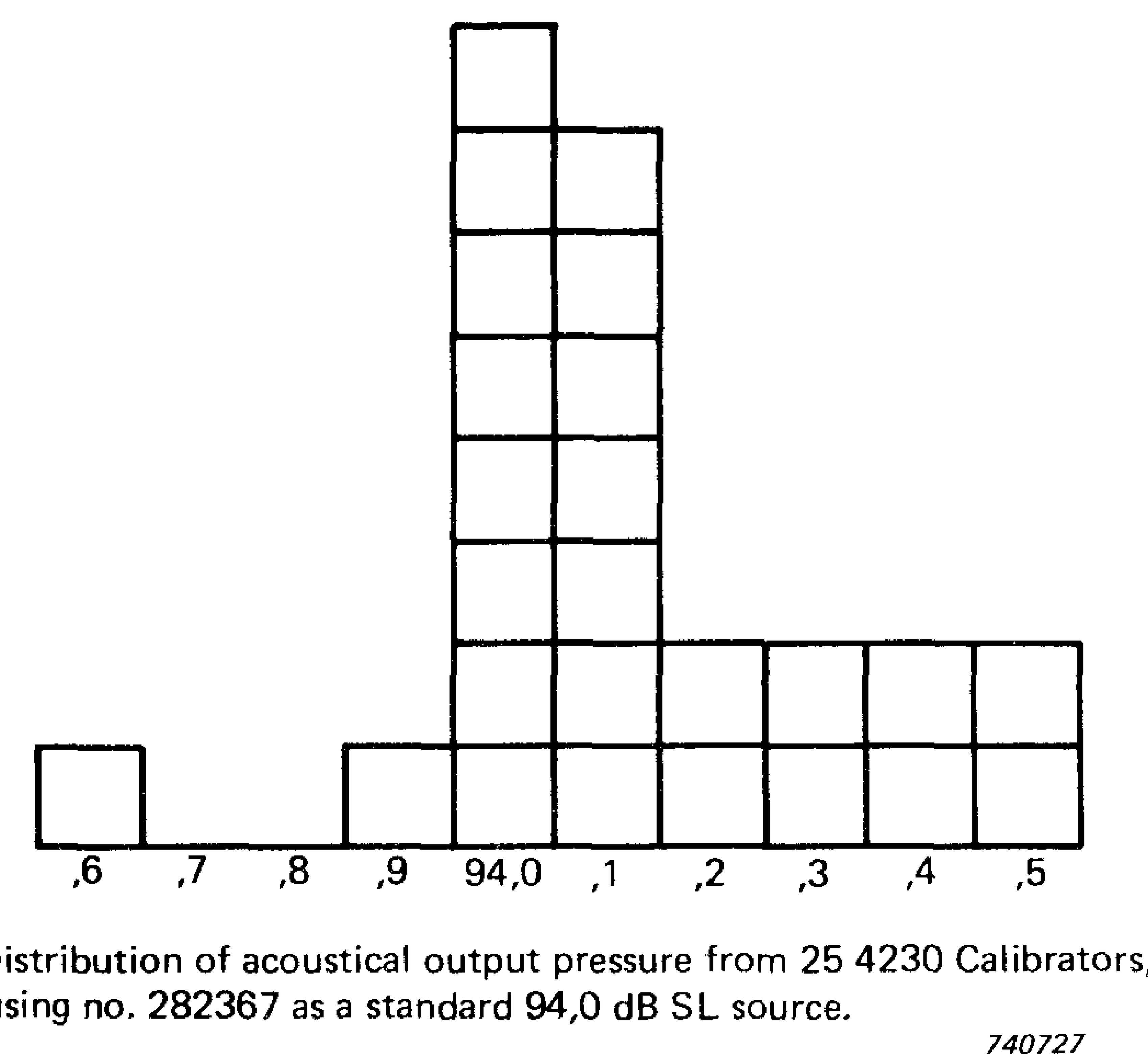


Fig. 1. Distribution of acoustical output pressure from 25 4230 Calibrators, using no. 282367 as a standard 94,0 dB SL source

The second group of results (Table 2, Fig.2) shows the behaviour of the three receiver capsules. Each was measured twice, first at normal working levels (Transmitting Attenuator set to —10 dB) and again at a level 10 dB higher than normal. The accompanying histogram shows results. One is not really justified in calculating standard deviations etc. because a study of the table shows that differences in readings are systematic — i. e. the systems occupy almost the same rank order in all tests. For example, systems 2, 3 and 19 are always high, lying 2 or 3 tenths of a dB above the accepted standard system (number 1), whereas systems 17, 21 and 23 are always low, lying below standard again

	Location number,	Air pressure mm Hg	Receiving OREM B. Transmitting atten = –10 dB Receiving atten = 0 dB Capsule number 1	2	3	Uncorrected meter readings. Mean dB	FTZ mean minus Local mean dB	Transmitting atten = 0 dB Receiving atten = –10 dB 1	2	3	Mean dB	FTZ Mean minus Local mean dB
	1 FTZ	750	3,8	3,8	3,4	3,67	–	3,9	3,9	3,6	3,80	–
	2	723	4,1	4,1	3,8	4,00	–0,33	4,2	4,3	3,9	4,13	–0,33
	3	723	4,0	4,0	3,8	3,93	–0,26	4,2	4,3	4,0	4,17	–0,37
	4	723	3,7	3,7	3,5	3,63	+ 0,04	3,8	3,9	3,6	3,77	+ 0,03
	5	723	3,8	3,9	3,6	3,77	–0,1	4,0	4,1	3,7	3,93	–0,13
	6	720	3,8	3,9	3,5	3,73	–0,06	4,0	4,1	3,7	3,93	–0,13
	7	734	Old system Type 3350,			See bottom of page.						
	8	734	3,8	3,9	3,5	3,73	–0,06	3,9	4,0	3,6	3,83	–0,03
	9	734	3,8	3,9	3,4	3,70	–0,03	4,0	4,1	3,5	3,87	–0,07
	10	724	3,6	3,7	3,4	3,57	+ 0,1	3,8	3,9	3,6	3,77	+ 0,03
	11	761	3,9	3,9	3,5	3,77	–0,1	4,0	4,0	3,6	3,87	–0,07
	12	761	3,9	3,9	3,6	3,8	–0,13	4,1	4,0	3,7	3,93	–0,13
	13	761	3,8	3,8	3,6	3,73	–0,06	3,9	3,9	3,7	3,83	–0,03
	14	760	3,7	3,6	3,3	3,53	+ 0,14	3,7	3,7	3,4	3,60	+ 0,20
	15	760	3,7	3,6	3,2	3,50	+ 0,17	3,8	3,8	3,4	3,67	+ 0,13
	16	760	4,0	3,9	3,6	3,83	–0,16	4,1	4,1	3,7	3,87	–0,07
	17	755	3,6	3,5	3,1	3,40	+ 0,27	3,7	3,6	3,3	3,53	+ 0,27
			No Fe Ap 61 available.									
	19	736	4,1	4,0	3,7	3,93	–0,26	4,0	3,9	3,6	3,83	–0,03
	20		4,0	3,9	3,7			4,2	4,1	3,8		
	21		3,4	3,4	3,0	3,26	+ 0,4	3,5	3,5	3,1	3,36	+ 0,44
	22	760	3,8	3,8	3,4	3,66	0	3,9	3,9	3,5	3,76	+ 0,04
	23	755	3,5	3,5	3,2	3,40	+ 0,27	3,6	3,6	3,3	3,50	+ 0,30
	24	755	3,9	3,8	3,5	3,73	–0,06	4,0	4,0	3,7	3,90	–0,1
30/9	1 FTZ repeat	745	3,8	3,7	3,4	3,63	+ 0,04	4,0	3,9	3,5	3,80	0
	7 Old 3350/ 4901	734	4,25	4,1	3,7			4,35	4,2	3,9		

750039

Table 2.

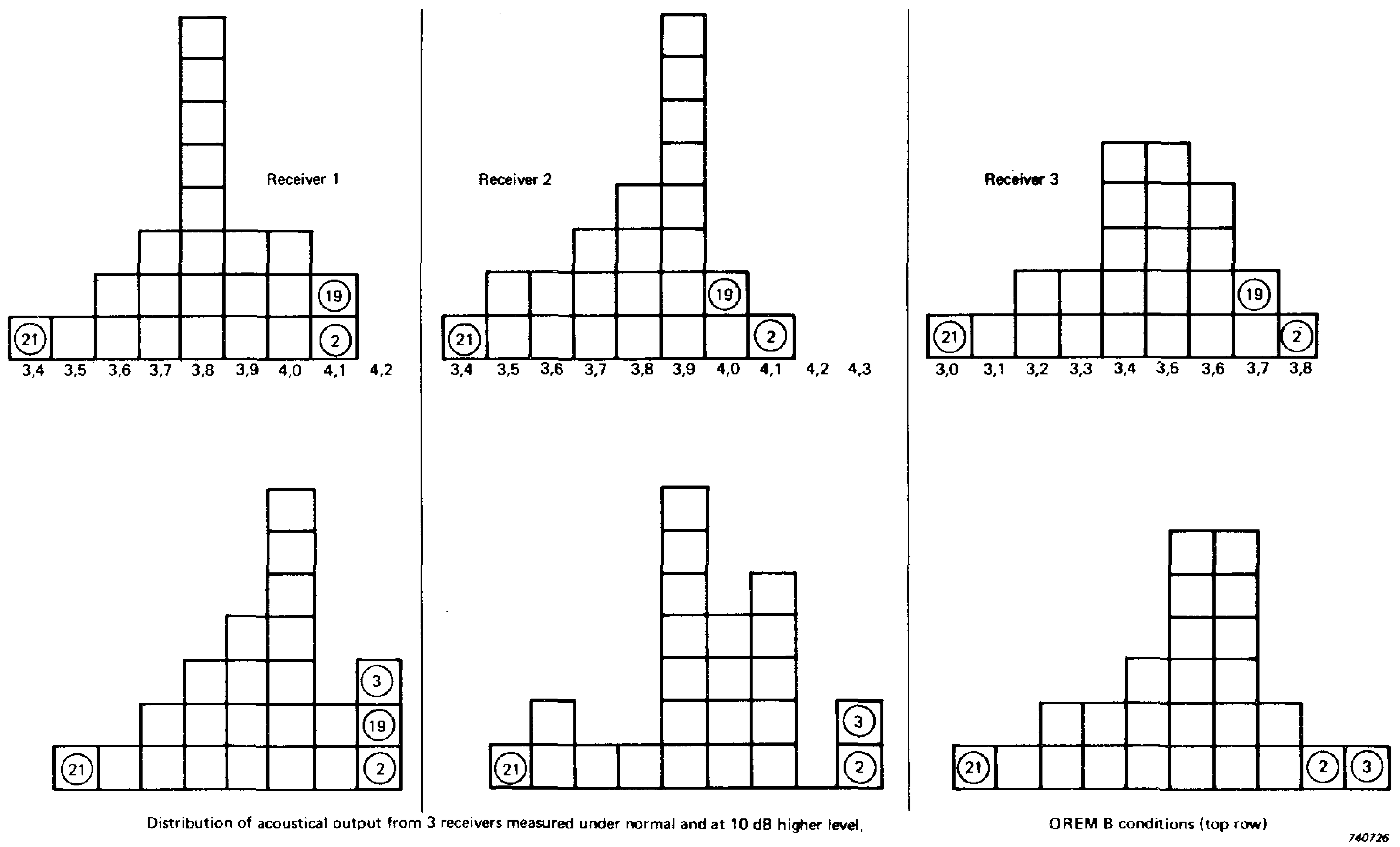


Fig. 2. Distribution of acoustical output from 3 receivers measured under normal OREM B conditions (top row) and at 10 dB higher input (bottom row)

by 2 to 4 tenths of a dB. It should be explained that the standard system is one held by the PTT which is maintained carefully and against which all others are checked; the arithmetical mean of all results is not taken as a standard, although in fact the central system gives results which lie exactly on this mean — a rather pleasant state of affairs. The fact that the rest of the systems follow a repeatable rank order suggests that if necessary each could be assigned a very small correction factor of the order of not more than three-tenths of a dB which would bring all systems to within a tenth of a dB of the central system. Even without this correction factor it can be seen that the total spread, measured for the worst case of single capsule readings, is + or —1/3 dB.

For transmitters (Table 3) the case is complicated by the necessity of applying a correction for local air pressure. The particular value of correction is not at the moment too clear, but it is around 0,2 dB for the worst case — that is, those locations with air pressures of around the 725 mm Hg range, when compared with the standard system working at 750 mm Hz. It is emphasised that this correction applies only to the capsules and not to the B & K systems which retain their calibration at all reasonable air pressures. Using an approximate correction factor

Transmitting OREM B. dB'			
Location number	Air pressure mm Hg	Capsule number. 1	2
1 (FTZ)	750	–2,2	–2,0
2	723	–1,8	–1,7
3	723	–2,0	–2,0
4	723	–2,2	–2,1
5	723	–2,1	–2,0
9	734	–1,5	–1,7
10	724	–1,9	–1,8
11	761	–2,1	–2,0
12	761	–1,7	–1,7
16	760	–1,6	–1,5
17	755	–	–
18	755	–	–
19	736	–2,3	–2,2
20			
23	755	–1,7	–1,8
24	755	–1,9	–2,0
30/9 1 (FTZ) (Repeat)		–2,2	–2,1

750037

Table 3.

then, the system spread is around + or —1/3 dB, but with the same standard system (no. 1) lying appreciably above the mean of all systems. Again, the systems can be ranked, but the rank order is not the same as in receiving tests. This suggests that it is not system calibration which is responsible for systematic deviations from the norm, and in fact, tests now in progress are beginning to focus attention on the exact details of the construction of the measuring jigs as factors in causing these small differences between systems.

Several people, seeing results like these, have asked what is the point of such precision? Under the best listening conditions people have difficulty in detecting a one-dB change in level and most telephone subscribers wouldn't notice a three-dB change, so why chase tenths of a dB?

The answer lies more in the field of economics than engineering. Telephone subset manufacture is extremely competitive. Costs must be cut if any profit is to be made. Amongst other things it means that as few sets as possible must be allowed to fail quality control checks since such failures represent wasted manufacturing time. A quality control

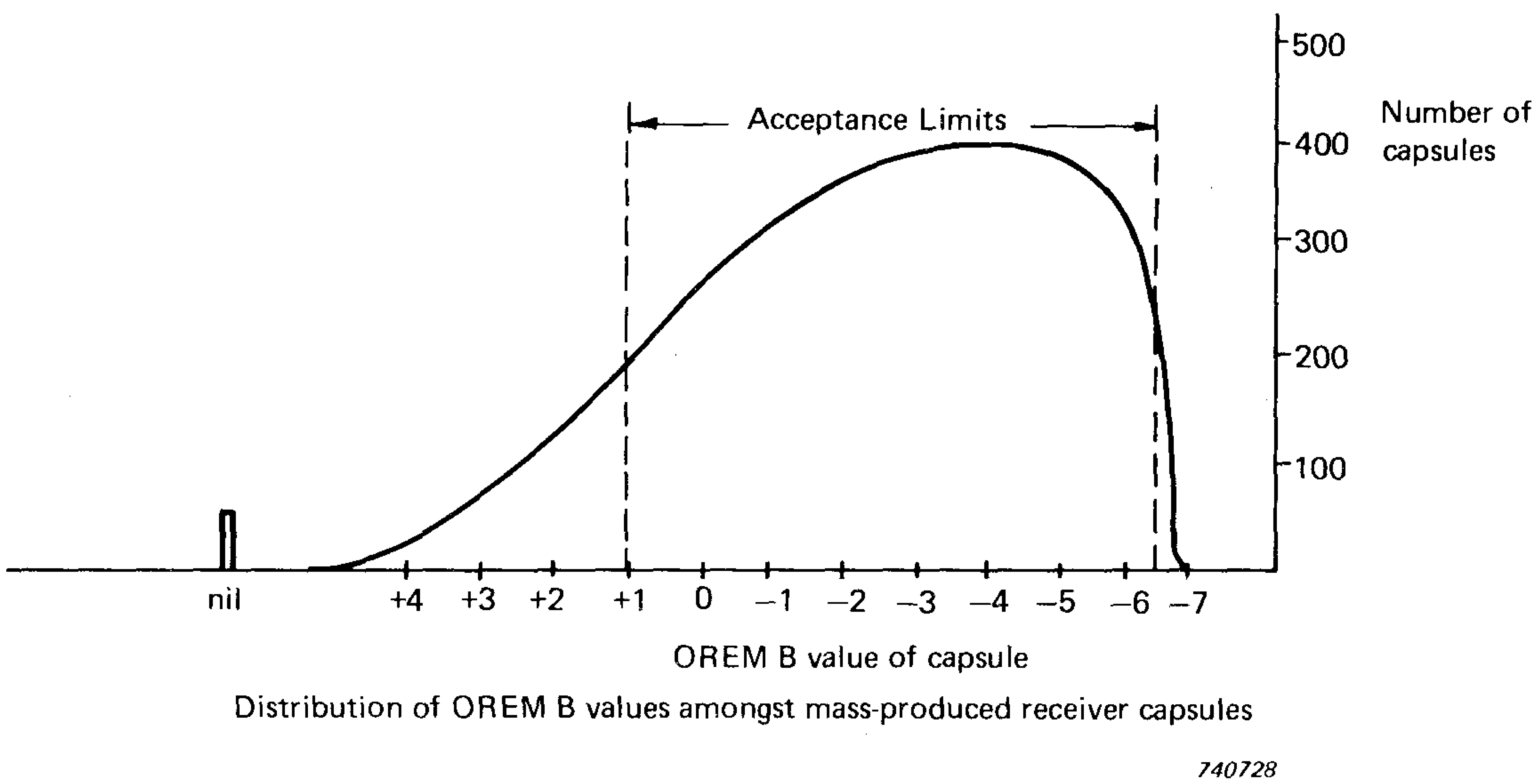


Fig.3. Distribution of OREM B values amongst mass-produced receiver capsules
Figures shown are imaginary and are not related to the actual figures used by any factory or PTT

check used throughout the world is that involving the measurement of OREM values, and it is interesting to look at these OREM values in more detail.

In a factory mass-producing telephones the OREM values of the sets form an approximately Gaussian distribution with disturbances caused by such things as sets which have no output and which appear as a sharp hump at the low end of the distribution curve. Figure 3 shows an imaginary distribution of telephone sets manufactured plotted against their OREM values. The buying agency will typically mark off upper and lower limits on this curve. Sets having OREM values outside these limits will be rejected and will represent a loss of money to the manufacturer. If the manufacturer guards against this by concentrating his sets in the middle of the curve the narrower manufacturing tolerances become increasingly expensive and defeat his attempts at cost-cutting. In general, the wider the acceptance "window" the easier it is for the manufacturer to reduce costs. Ideally he would arrange his sets as in Fig.4 to have a distribution which is exactly rectangular — that is he would spread the sets over the widest possible range but with none falling outside the limits. In practice a good factory may achieve a near-rectangular OREM/output curve whereas a poor factory would produce a curve having a long flat hump. But if the good factory gets the rectangu-

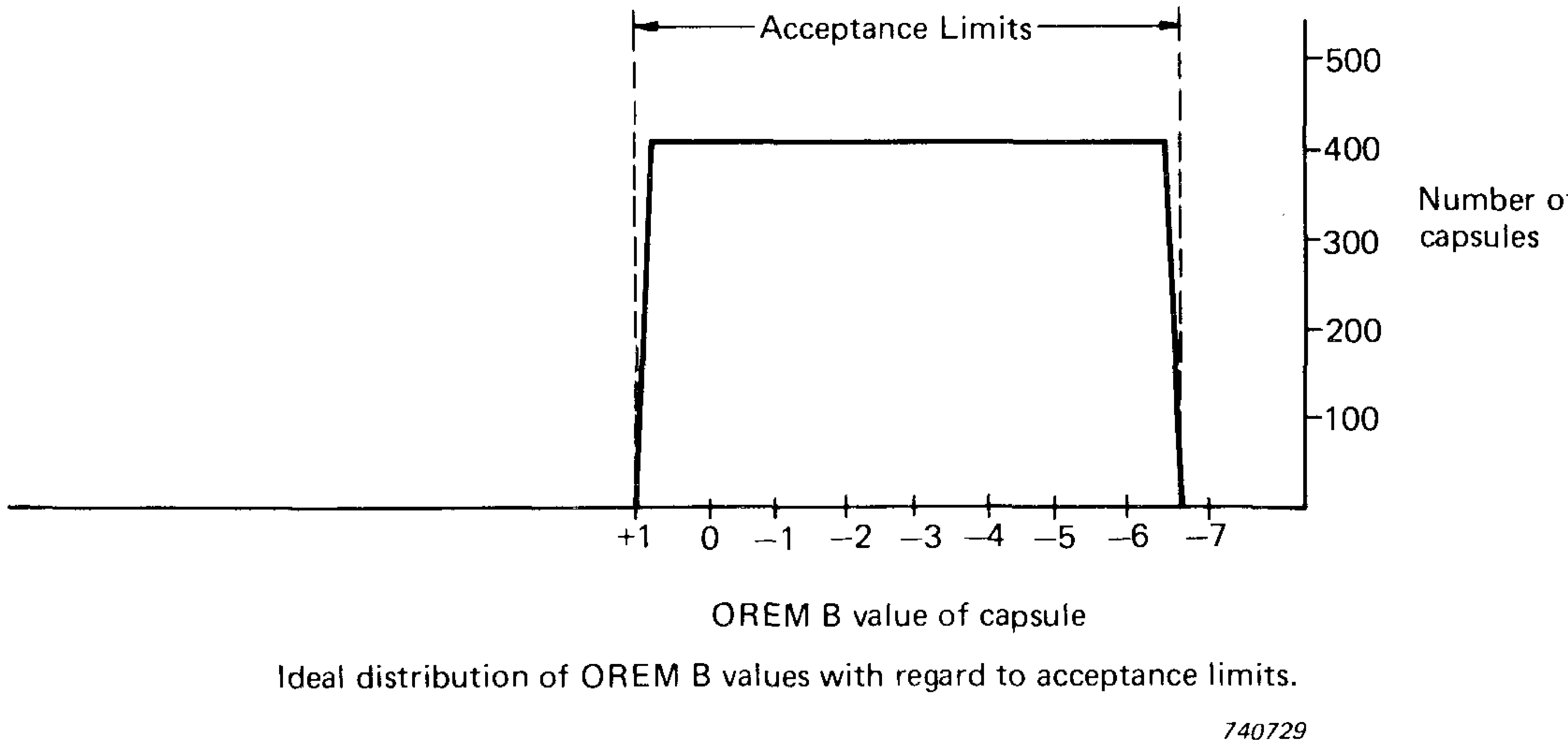


Fig. 4. Ideal distribution of OREM B values with regard to acceptance limits

lar distribution what happens if the test equipment suddenly shifts its zero? The manufacturer loses many sets — in fact the better the factory the more sets it loses if the test equipment changes its reference. (In contrast the weak manufacturer loses just a tiny fraction more sets than usual — so few more that it is unlikely that it would be noticed). Hence the better the manufacturer the more care he must take to have absolutely stable test equipment. Moreover, if the buying agency has a test set with a different reference from that of the manufacturer more sets will be lost. The steeper the skirts of the distribution curve the more is the number of sets lost by shifts in the test equipment reference marks in either the factory test section or in that of the buying agency. When such shifts occur (as indeed they do) there usually follow fierce and expensive arguments between factory and buyer. If these are situated in the same district then it may be possible to settle the argument by many tests on the two measuring systems. But if they are far distant, as in different countries, then common tests become difficult or impossible and the arguments continue until the manufacturer reluctantly reduces the width of the "window" to encompass shifts in test equipment reference marks, so increasing his production costs.

This means that any company working in the field of factory test equipment must aim for the highest standards of stability and repeatability in its test equipment. Referring to well-known classifications of tests, it can be said that factory tests are aiming not at validity but at repeatability. Valid tests are made on prototypes by laboratories, repeatable tests are made on the million production copies by factories. The more repeatable the factory equipment is the more it will bring cost benefits to the advanced manufacturer. Some of these benefits are obvious, such as those resulting from smoother control of production quality. Others are hidden, such as those resulting from having fewer arguments with the buying agency.

That is why we feel that we are justified in aiming at such high standards of repeatability, and that is why we welcome the co-operation of manufacturing groups in helping us to see just how good or bad this repeatability is in practice.

Our grateful thanks are due to the firms listed below (in alphabetical order which is **not** related to the location number in the test tables). In every place the engineers not only tolerated disruption to their important test schedules but actively helped with the work and showed great intererest in the whole exercise. It was a pleasure to work in such atmosphere of keen interest.

Firms taking part were, in alphabetical order (**not** in order of numbering on the list) as follows:

FTZ, Darmstadt.
Fernsig, Essen.
Hagenuk, Kiel.
Friedrich Merk GmbH, München.
Telefonfabrik Reiner, München.
SEL, Straubing.
SEL, Stuttgart.
Siemens AG, Bocholt.
Siemens AG, München.
Telefonbau & Normalzeit, Frankfurt.

The tour was arranged by Reinhard Kühl KG, Quickborn.

Stable Subset Measurements with the 73 D

by

K. Damsgaard

During the development of a new telephone subset it is normally necessary to perform a subjective test (using a working standard such as SETED) to assess the NOSFER Reference Equivalents.

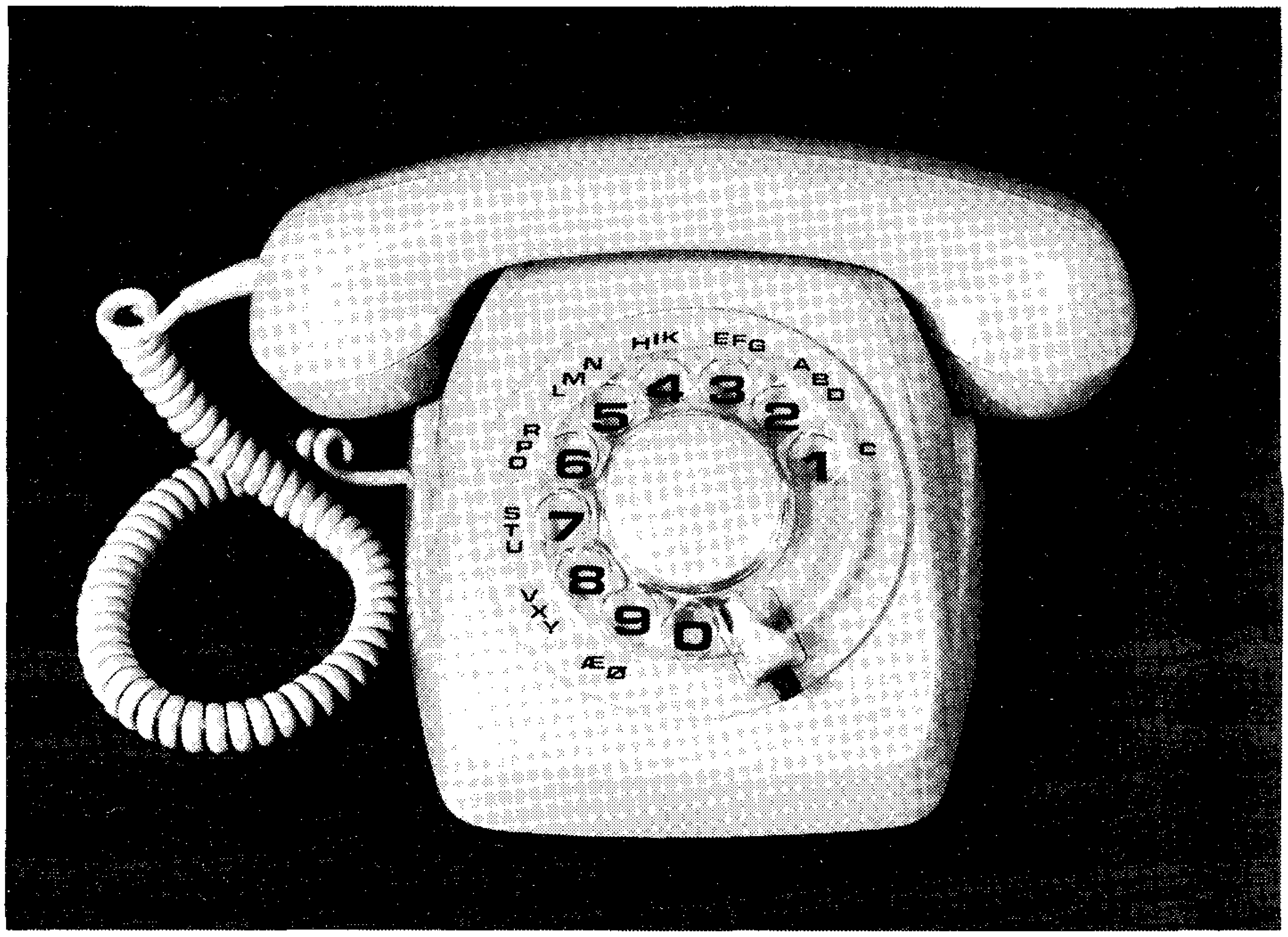
Laboratory work is done with an objective measuring system and so is the process control in the manufacturing phase. There are several reasons for this, one of them being the closer tolerances of objective measurements. An objective system in common use throughout the world is the B & K Telephone Transmission Measuring System Type 3350 and 3352.

Once the correlation between OREM and NOSFER for a particular subset is established future measurements normally will be OREM. In order to take full advantage of the closer tolerances it is important to check the calibration between cooperating measuring systems. Between manufacturer and administrations it is of vital interest to know the correlation between various measuring systems, especially where the measurements concern an advanced subset with close tolerances itself.

A very helpful instrument in this situation is a stable subset. Periodical cross-checks between one or two stable subsets and the OREM system are a good supplement to the OREM systems recommended calibration procedure. Also stable sets can detect differences between different measuring groups and even between operators in a single group.

When constructing a stable subset it should be remembered that ease of use and travel are important features. So it could be a light-weight subset in a normal subset case, which in most cases would be appreciated by the user. It must not be fitted with a carbon microphone and special attention must be paid to the handset which should be mechanically stable and of a shape giving an uncritical position in the measuring systems test head.

The specifications for the 73 D subset enable the set to be made in a stable subset version. 73 D is an easy to handle subset with a handset like the German "Assistent" (see Fig. 1). The subset has been developed as a constant output subset so that its reference equivalents are independent of the line current. This is not quite true for sending, but is close enough for the influence of feeding tolerances on the equivalents to be eliminated and for the influence of contact resistance in the hook switch to be minimized.



750132

Fig. 1. 73 D Stable Subset

The handset is equipped with two identical dynamic transducers. The transducers and their treatment are very important for stable set. The type used is selected to avoid a frequency response with peaks which could affect the long time stability. After the magnetizing process the transducers are demagnetized to a predetermined level in order to stabilize the magnetic system. Finally they are submitted to a number of temperature cycles.

After this treatment the 73 D subset is ready for use as a Test Telephone. The initial KIRK measurements on a B & K 3352 system are supplied with the subset. If the initial measurements are to be valid over a long period rough handling of the subset should be avoided, especially extreme temperature variations, vibrations and mechanical shocks.

The relation between OREMA and NOSFER is of minor interest for a Test Telephone. It is worth noting, however, that if the relation is established through subjective measurements against a SETED system where two or more test teams are involved, there will normally be a tolerance of ± 3 dB or more at the 95% confidence level.

Using the stable set it is very important that the position of the handset in the Test Head is correct. For this purpose the 73 D Stable Set is supplied with a matching disc which locates the 73 D earcap on the B & K 3352 Test Head. The measurements are made without the Voice lipring. While the distance between the Voice and the microphone is rather uncritical, attention should be paid to the handset axis, which must be parallel to the Test Head Platform so as to place the microphone correctly in front of the Voice.

During a day the room temperature in a laboratory will often change 5 C° or even more. This change may influence the ORE measurements by several tenths of a dB. So, very precise measurements should be performed only in a narrow temperature interval as stated in the test report. Initial measurements on 73 D are made at 22°C.

The same attention should be paid to the atmospheric pressure if measurements from geographically different places are compared.

In the manufacturing process in the Kirk factories in Horsens two B & K Telephone Measuring Systems Type 3352 have been in service for some time. The system is slightly modified for automatic testing (see Fig.2). There has been added an AKU-unit including limit detectors in order to perform a go — no go test.

These Measuring Systems have been checked regularly against two stable subsets.

After a burn-in period of some months and initial calibration of the system the readings from periodical checks on one of the systems are shown on Fig.3.

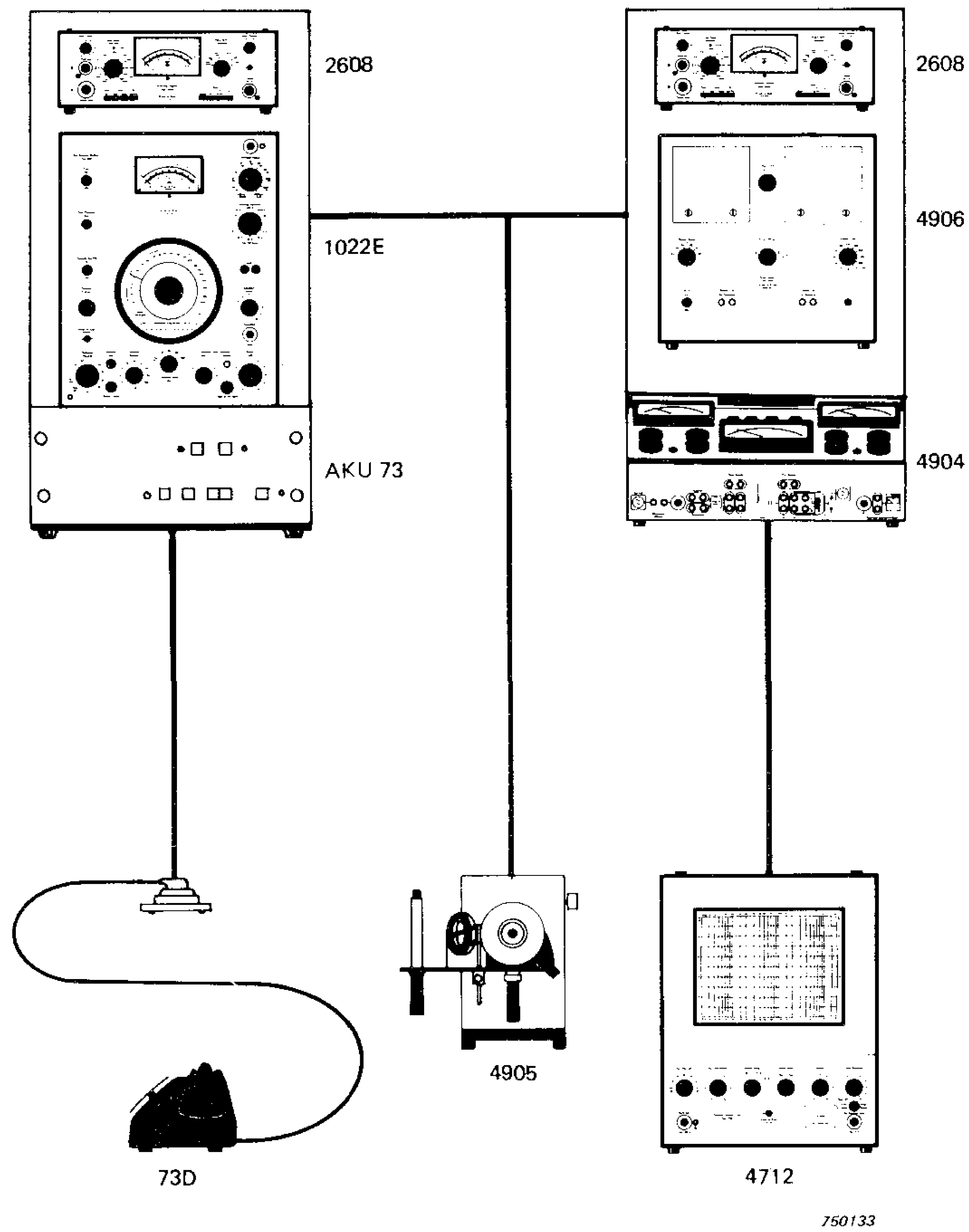


Fig. 2. The 3352 system modified for automatic testing

In order to increase reading accuracy the ORE-meter on the 4904 was modified to have a mirror-scale. It was decided not to recalibrate a system unless a change between the measuring system and the stable sets exceeded 2/10 of a dB.

Except for a slight change on SRE during the first two months the curves demonstrate an excellent stability of the system, and no recalibration has been necessary during nearly one year.

In order to check variations on external measurements a test telephone was measured at two locations in Denmark, two in England, and one in Northern Ireland. All the companies had B & K telephone measuring equipment, and the resulting OREMA values were:

RRE: —2,6 dB —3,0 dB —3,0 dB —2,2 dB —2,7 dB
SRE: + 3,4 dB + 3,9 dB + 3,4 dB + 3,8 dB + 3,6 dB

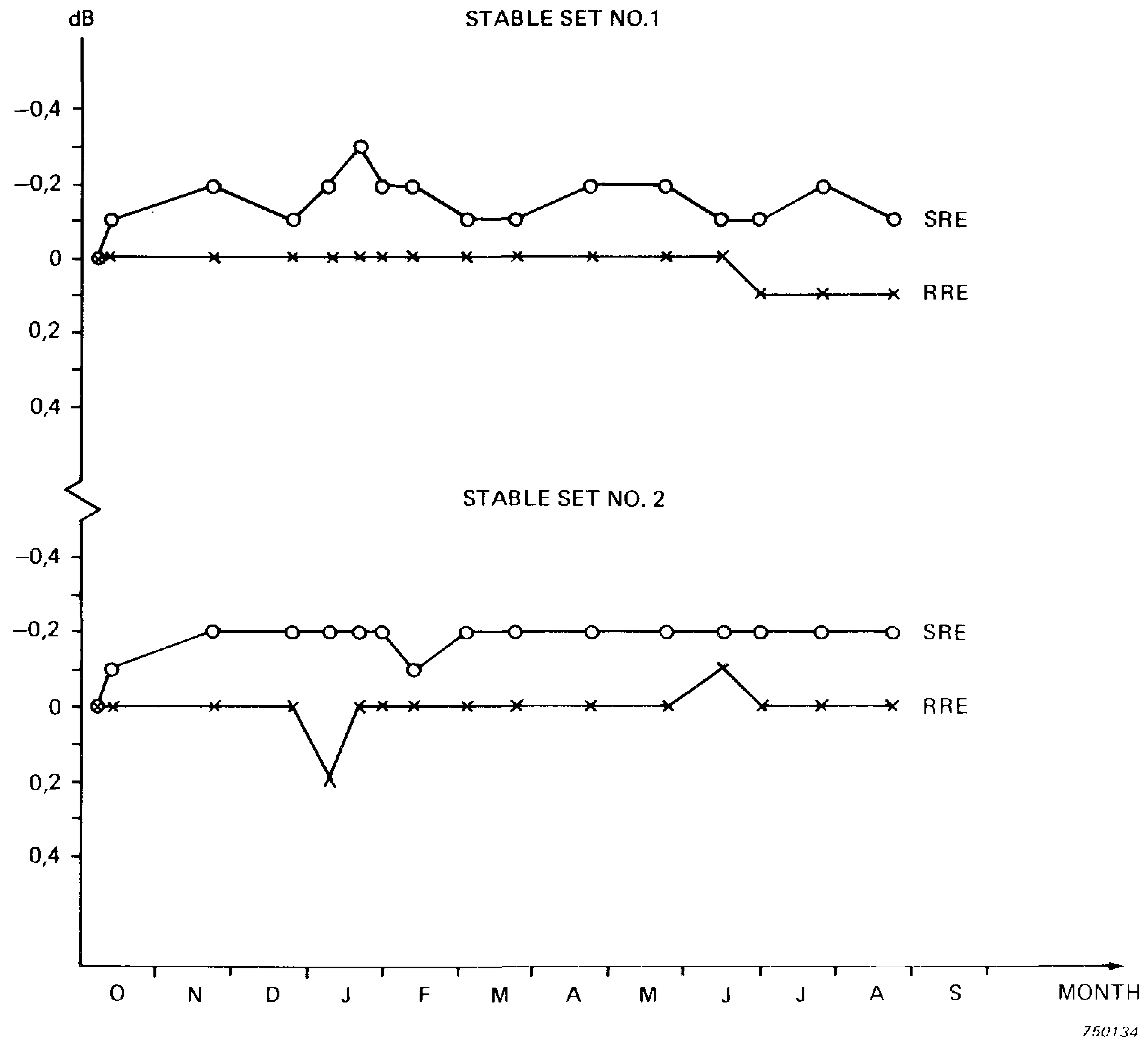


Fig. 3. Check of a 3352 system against two stable subsets

Our conclusion is that with skilled operators, in-house measurements can be repeated with an accuracy of a few tenths of a dB. Between companies, however, errors of 0,5 dB or more are very difficult to avoid.

Vibration Testing of Telephone Equipment

by

R. Fluhr

The only shock and vibration test currently used on telephones consists of dropping a telephone on a concrete floor to see what happens. Seen from an environmental testing standpoint, this is a perfectly valid test.

The whole point of vibration and shock testing, or any other form of environmental testing for that matter, is to try to duplicate as well as possible in the laboratory those conditions to which the test object will be exposed during its lifetime. Often the conditions are exaggerated such that the object is subjected to the "worst case" to see if it can survive. Typically, the worst case for a telephone is when it gets knocked off a table, therefore the aforementioned test is in the best tradition of environmental testing! Still, there's more to vibration testing than just breaking specimens.

One of the drawbacks of testing phones in this part of the world is that they're too good. You can subject a Scandinavian telephone to a vibration test that would ruin any ordinary electronic instrument and the telephone goes right on functioning with hardly an alteration in its characteristics. In one country in South America, the telephone company has many problems with faults such as badly soldered connections that require field servicing. A vibration test at the factory would show up a fault like this. We have also had reports of relays which set up so much vibration that they have to be continuously readjusted. But these problems are unknown in Scandinavia.

More pertinent to the Scandinavian situation is the use of vibration testing as a design tool. Aside from the accelerated life cycle type of test such as endurance and fatigue testing, there is another side to vibration testing and that is investigative testing. This is where one uses vi-

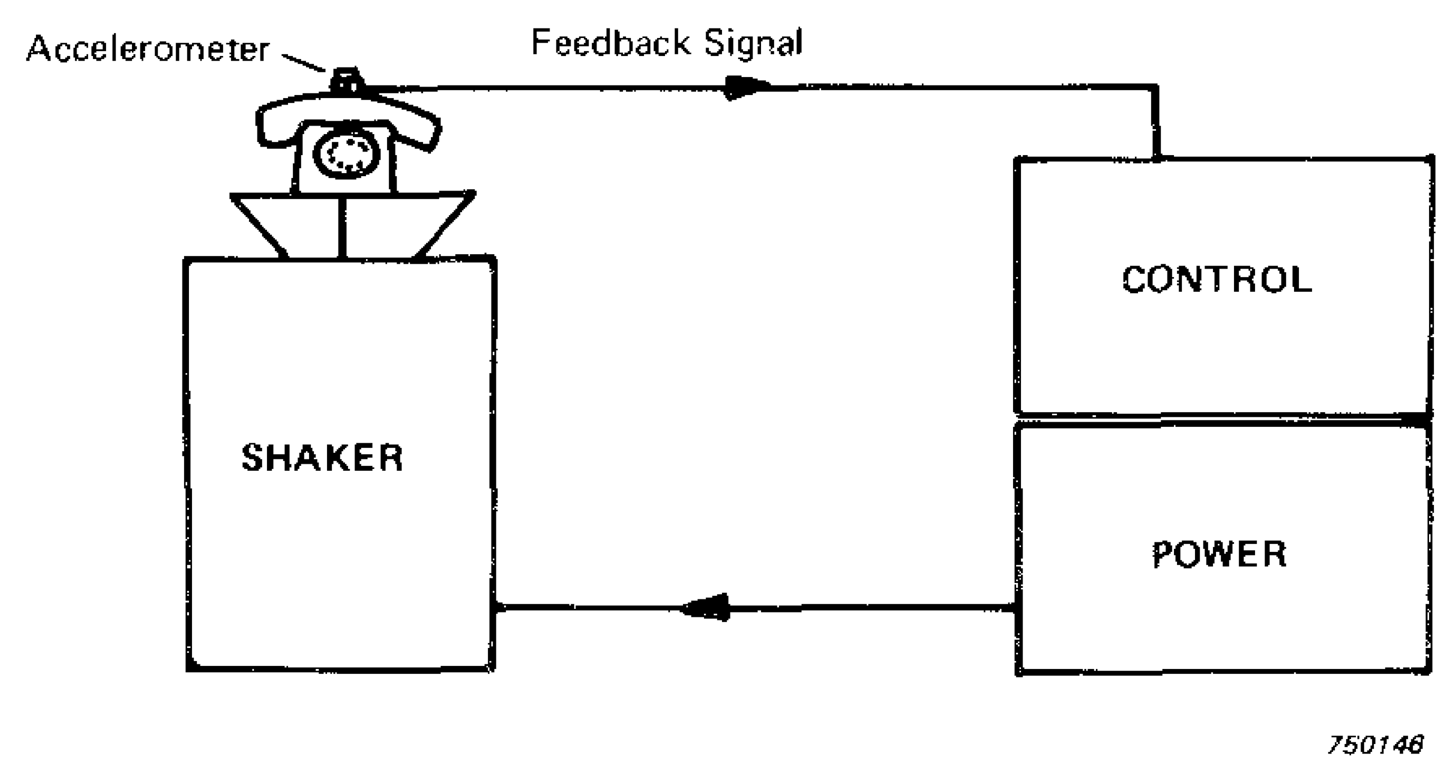


Fig. 1. Basic vibration test setup

bration to explore the mechanical properties of the test object. It can be done in three ways:

1) Resonance test with a stroboscope
2) Vibration signature using two accelerometers
3) Impedance measurements

Fig. 1 shows a basic vibration test setup. It is similar to a telephone test system setup in that a feedback signal from a transducer is used in a compressor circuit to control the level of the excitation. The difference here is that an accelerometer is used instead of a microphone.

In the first investigative method mentioned above, the control is set for some reasonable vibration level and then the frequency is swept manually. A signal, taken from either the feedback or the generator is used to trigger the stroboscope. The vibrating object appears to be standing still when the stroboscope is synchronized with the motion. If the strobe is then adjusted slightly off the frequency of vibration, the object seems to be moving in slow motion. In this way, both the frequency and type of motion can be observed. This type of investigation is called a resonance search because that is the major bit of information that can be obtained by it. When a part or component reaches resonance,

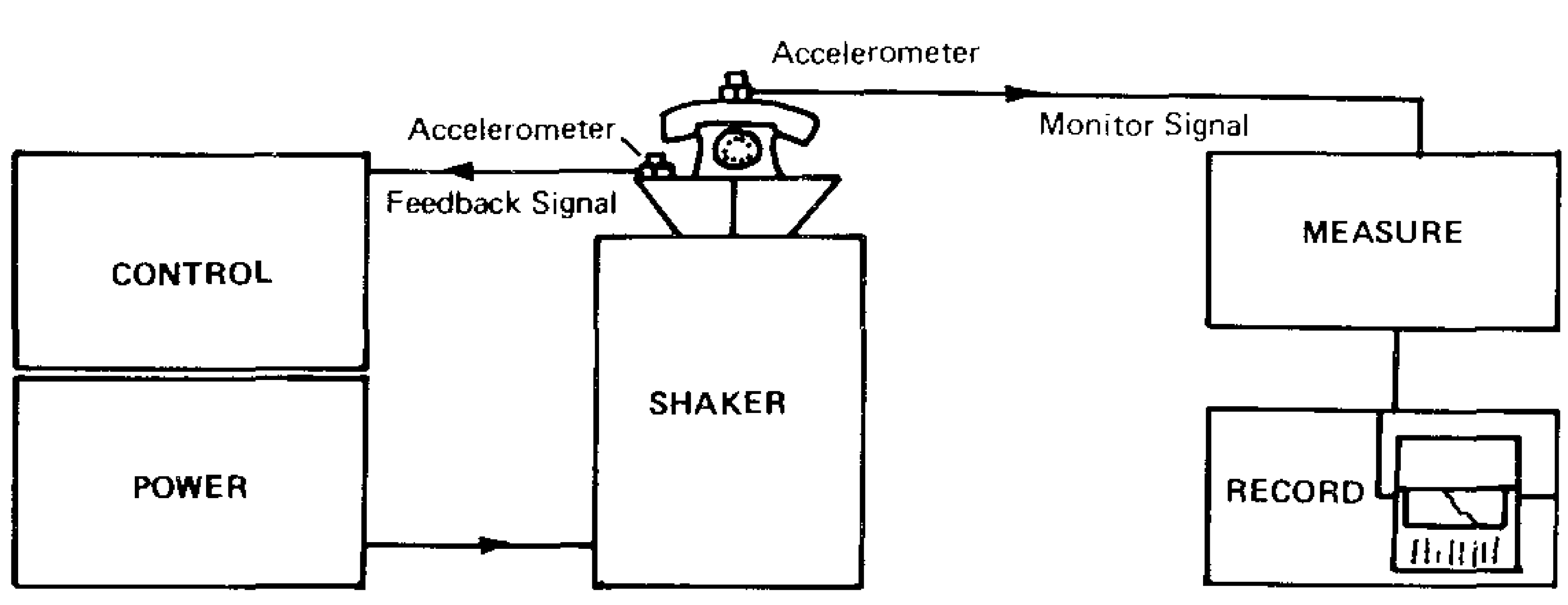


Fig. 2. Vibration test setup using two accelerometers

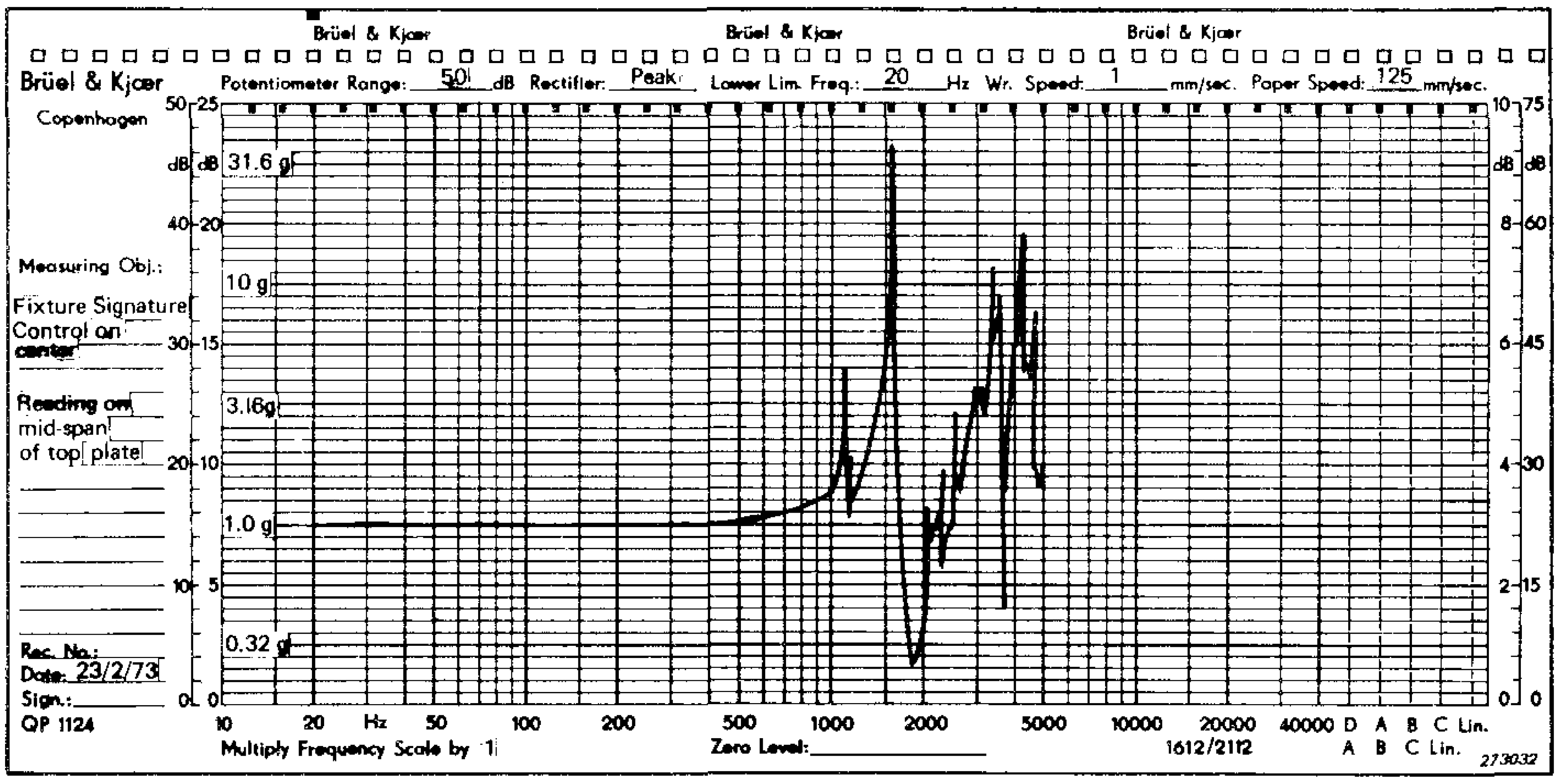


Fig.3. Vibration signature of a test fixture

its motion is greatly exaggerated and under the strobe light it appears to be moving while everything else is standing still.

A more sophisticated method is shown in Fig.2. Here the vibration level of the shaker is controlled with one accelerometer while another is used to monitor the vibration at some point of interest. This monitor signal is then plotted out on a recorder. The curve of relative amplitude vs. frequency thus obtained is known as the vibration signature of the object. It is, in fact, a plot of transmissibility, or output vs. input. The frequency at which the transmissibility peaks is the resonant frequency. Fig.3 is the signature of part of a vibration fixture. Note that the curve is completely flat (transmissibility = 1) until it nears the resonance, at which point it climbs quickly. Note also that before the main peak at 1600 Hz there is a smaller one at 1100 Hz. This was caused by the transmission of the resonance of another part through the system.

Impedance Measurements	
Apparent Mass	F/a
Mechanical Impedance	F/v
Dynamic Stiffness	F/d
Inertance	a/F
Mobility	v/F
Compliance	d/F

750012

Table 1.

To avoid this interference from other resonances, a third method may be used, which tends to isolate the particular part under study. This is the mechanical impedance method.

The term mechanical impedance covers a number of similar measurement concepts which all involve holding one frequency-dependent parameter constant while measuring the other. The six types of impedance measurements are shown in Table 1.

The one used throughout these tests was the inertance, a/F. The set-up for this is shown in Fig.4. Here the force is held constant while the resulting acceleration is measured. The acceleration reaches a peak at the resonance.

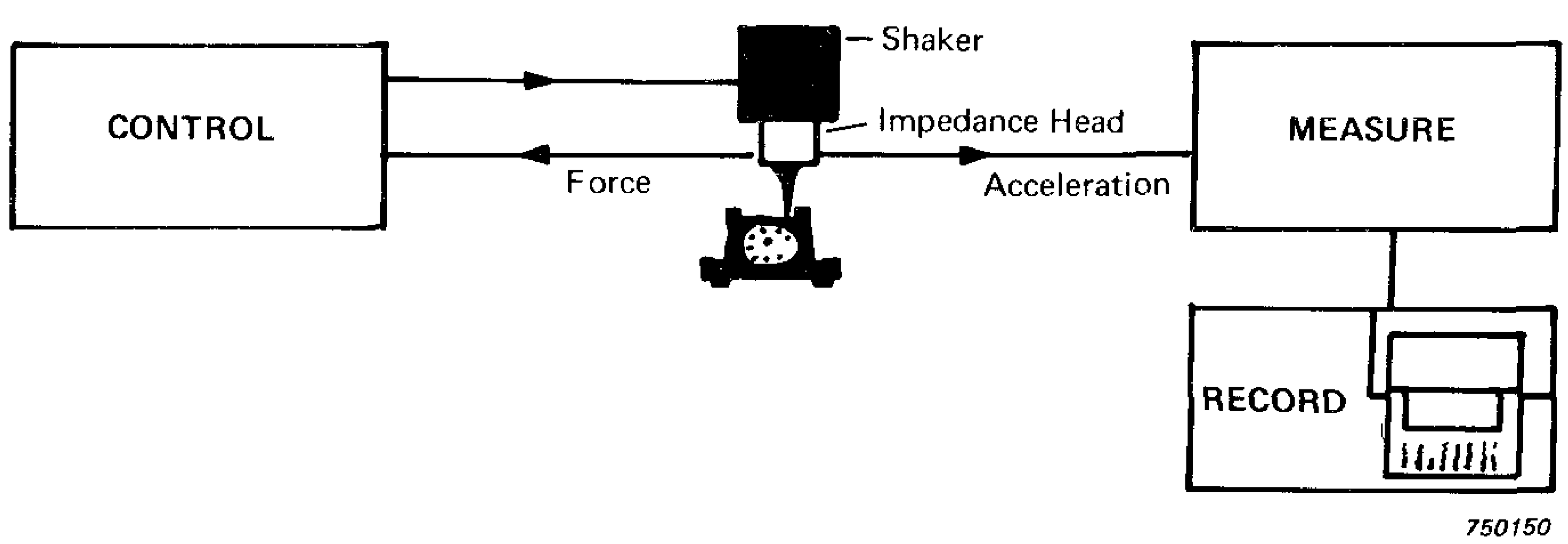


Fig.4. Basic mechanical impedance setup

Testing Conditions and Results

There were four telephones used in the test program, numbered 1, 2, 5 and 7. Numbers 2 and 7 were from one manufacturer and 1 and 5 were from another. 2 and 7 had dynamic microphones, 5 had a carbon granule type and number 1 had a piezoelectric.

Investigation

Number 7 was examined with a stroboscope while being vibrated on a shaker. The casing and handset were removed and the base fastened to the shaker. The frequency was adjusted manually and motion of the different components studied with the stroboscopic light.

Number 7 was also examined with a mini-shaker and probe. Table 2 shows some of the results obtained with the two methods.

Figs.6 to 9 show the frequency response curves obtained with the probe. Since the probe itself has a resonance at 2 kHz, the influence of

Telephone Number 7			
Component	Frequency obtained with Stroboscope		Frequency obtained with probe
Dial	183, 331		310
Spring under cradle	213	1g	
Contacts under cradle	143 - 157		
Bell rings	97		
T 1 rocks	149, 539		
C 8	177, 320		
Bells and magnet rock on housing	214 - 229	3g	
Magnet rocks	377		400
C 9	480		
C 2	480		
C 20	480		
C 1	363		340
C 4	546?	5g	720

750014

Table 2.

this must be subtracted from the curves. Fig.5 shows the frequency response of the probe. The solid lines in Figs.6 to 9 show the overall curves and the dotted lines show the curves with the effect of the probe removed.

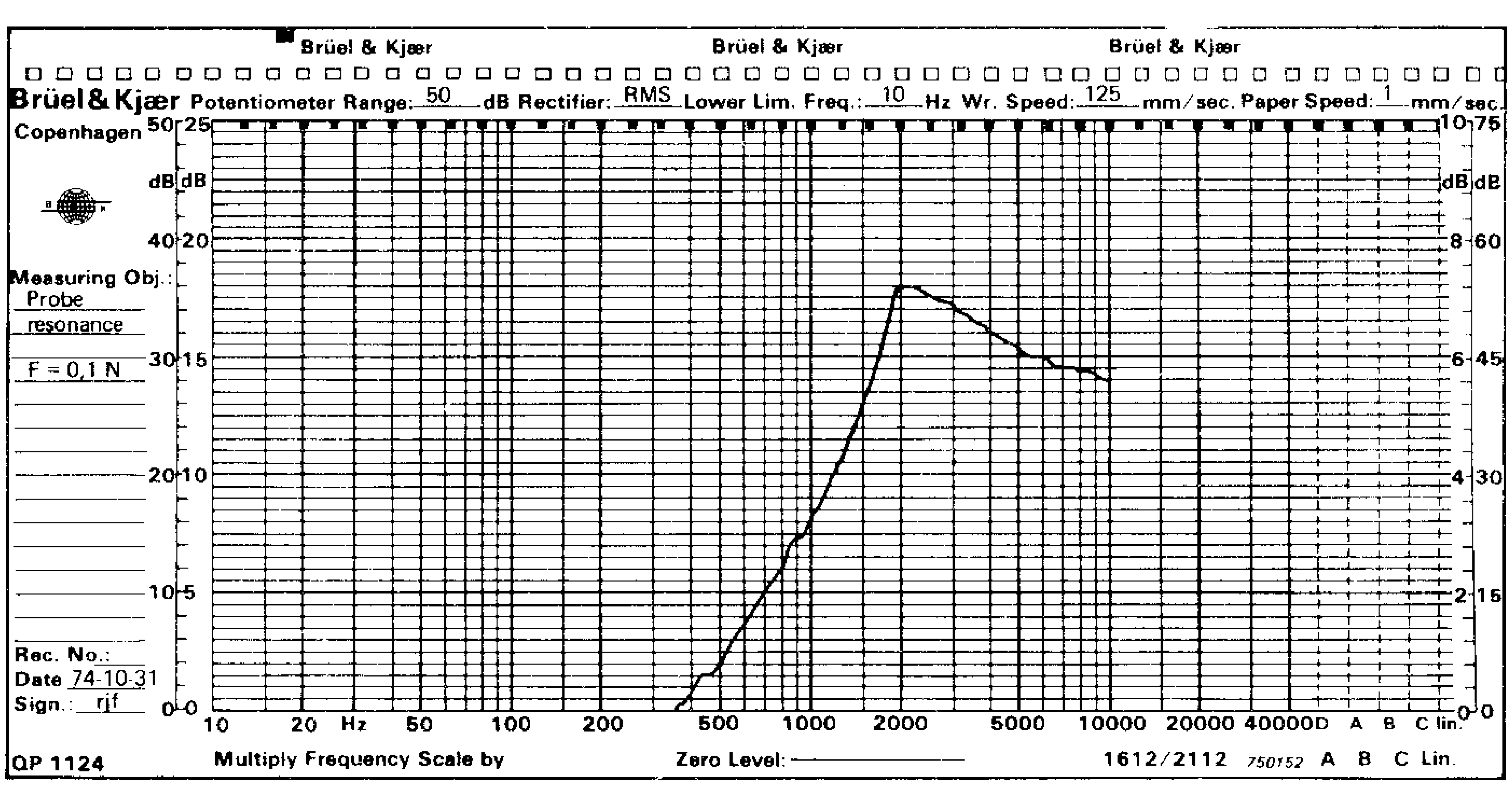


Fig.5. Frequency response of the probe

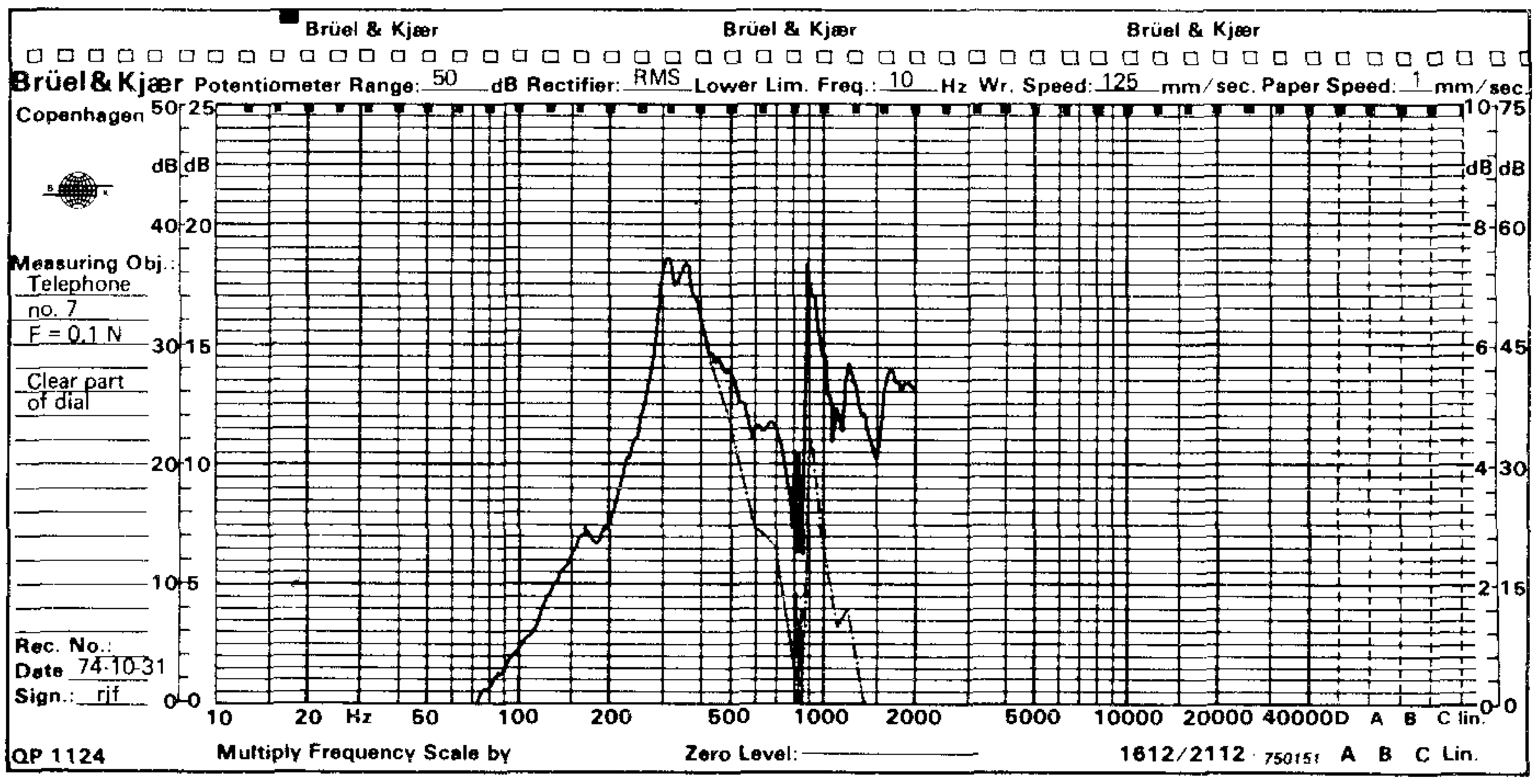


Fig. 6. Inertance plot: dial

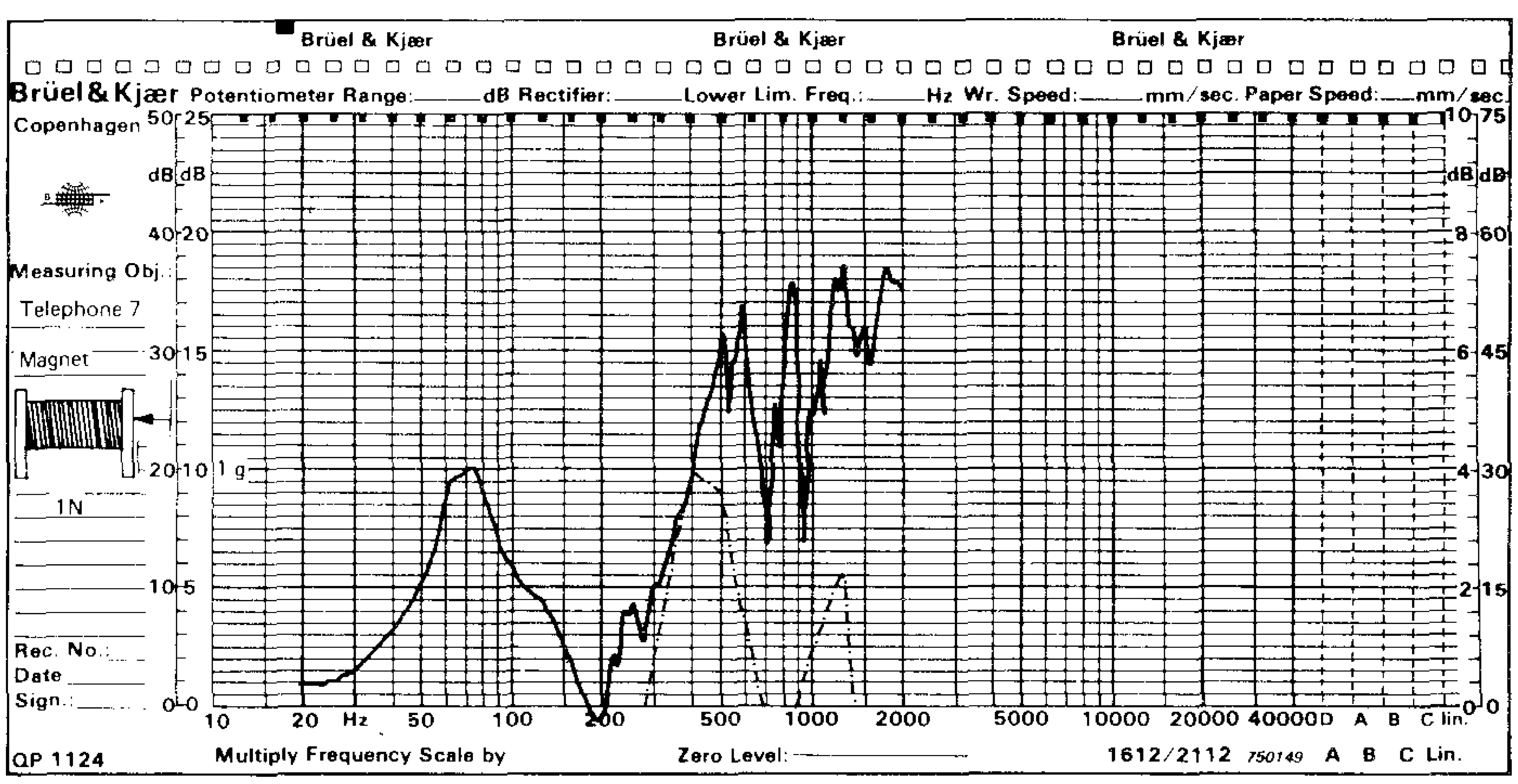


Fig. 7. Inertance plot: magnet

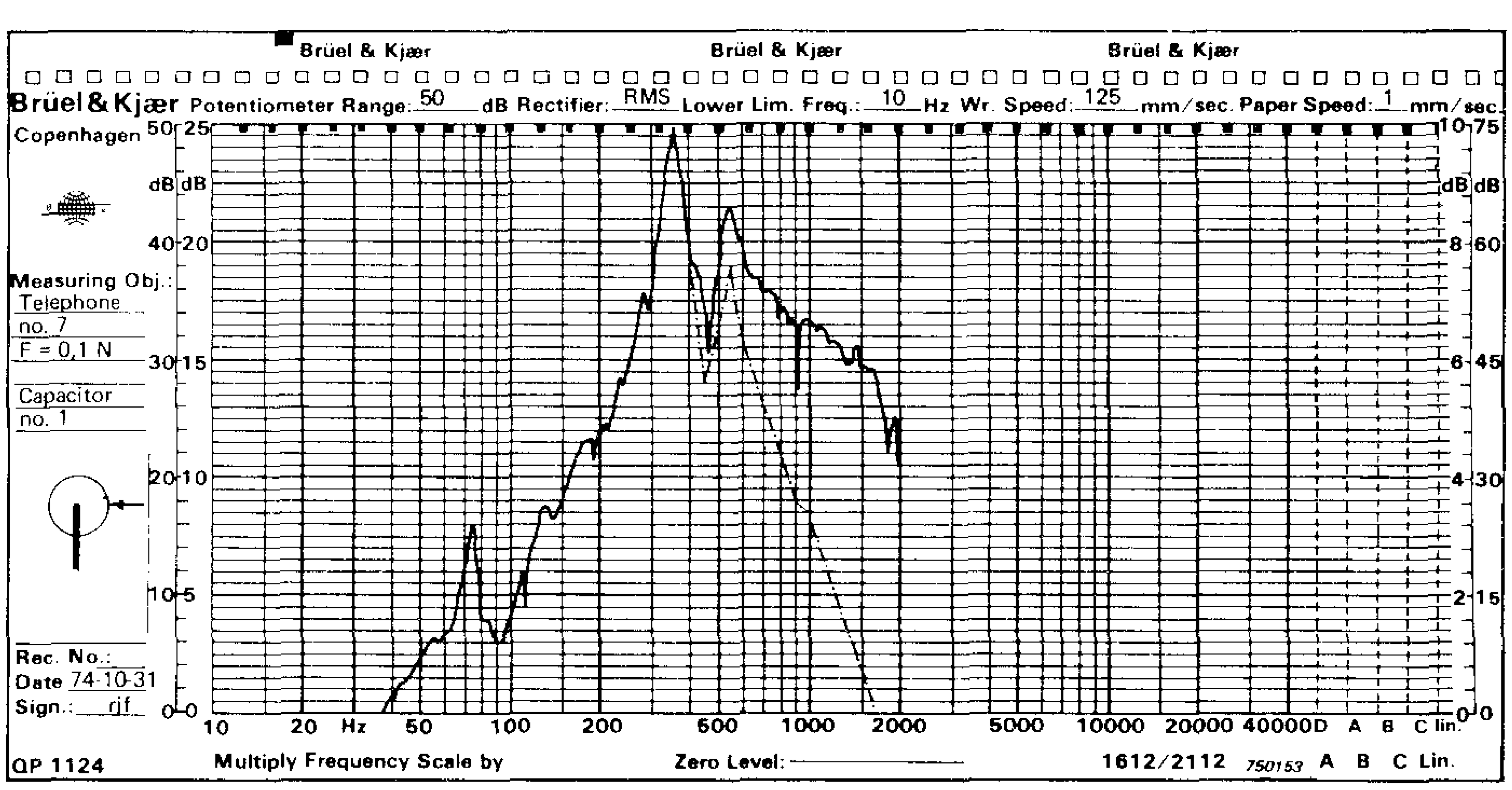


Fig. 8. Inertance plot: Capacitor no. 1

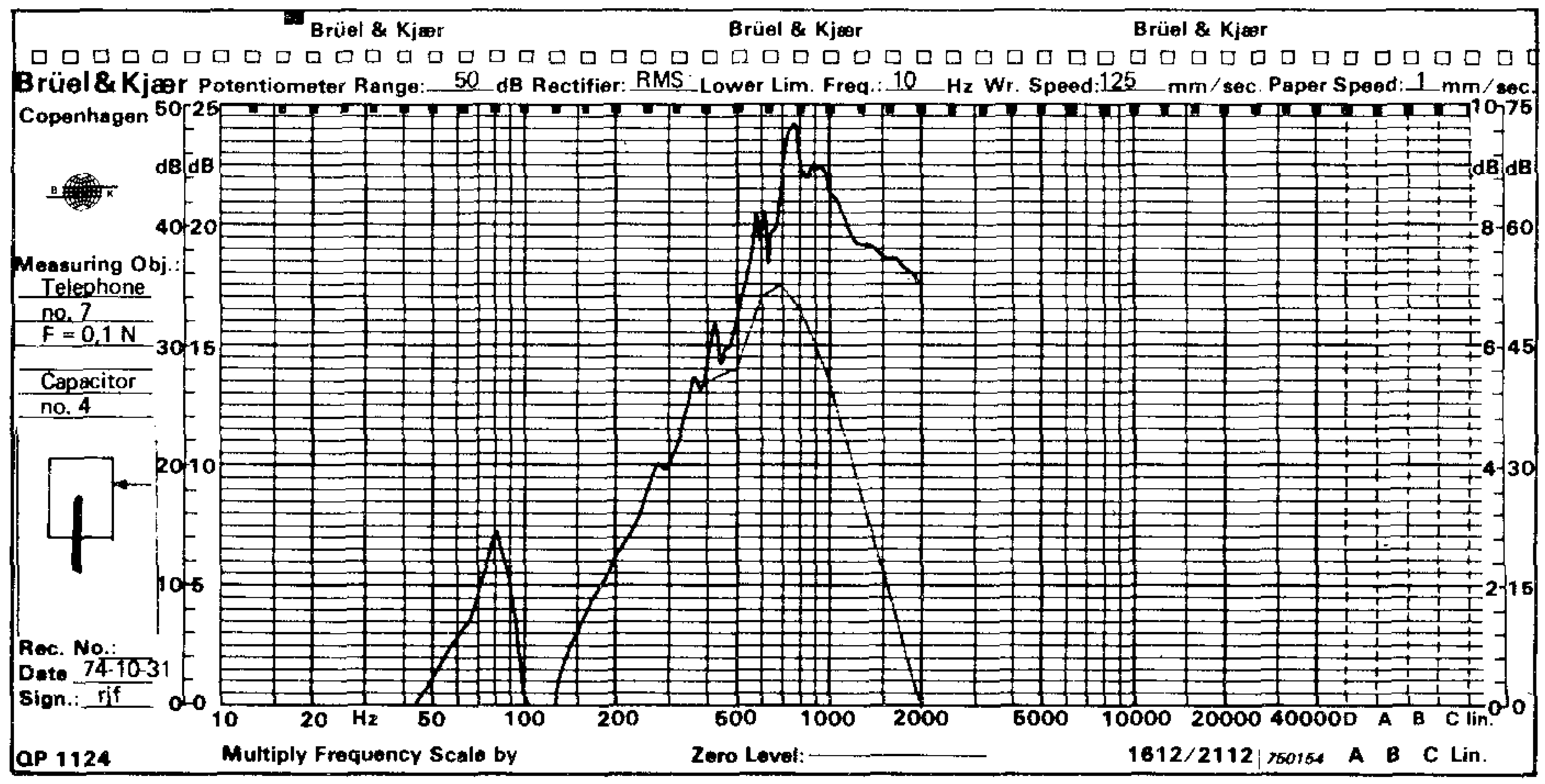


Fig.9. Inertance plot: Capacitor no. 4

Telephone number 1 was examined by attaching an accelerometer weighing 2 grams to various parts of the phone with beeswax. Results are shown in Table 3.

Telephone Number 1	
Component	**Frequency**
Clear part of dial	88
Entire dial unit rocks about supports	161
Clapper, up and down	85 (Strobe)
Clapper, side to side	93 (Strobe)
Transformer, rocking	126, 218
High tone bell	252, 642
Low tone bell	247, 509, 613
Electromagnet, bell	458

750011

Table 3.

Endurance testing

Telephones 1 and 7 were each in turn placed on a shaker and swept from 10 to 2000 Hz at 5 g for 6 hours. It should be pointed out that electronic instruments at the B & K factory are normally tested for two hours at two g, so this represents a considerable escalation on normal

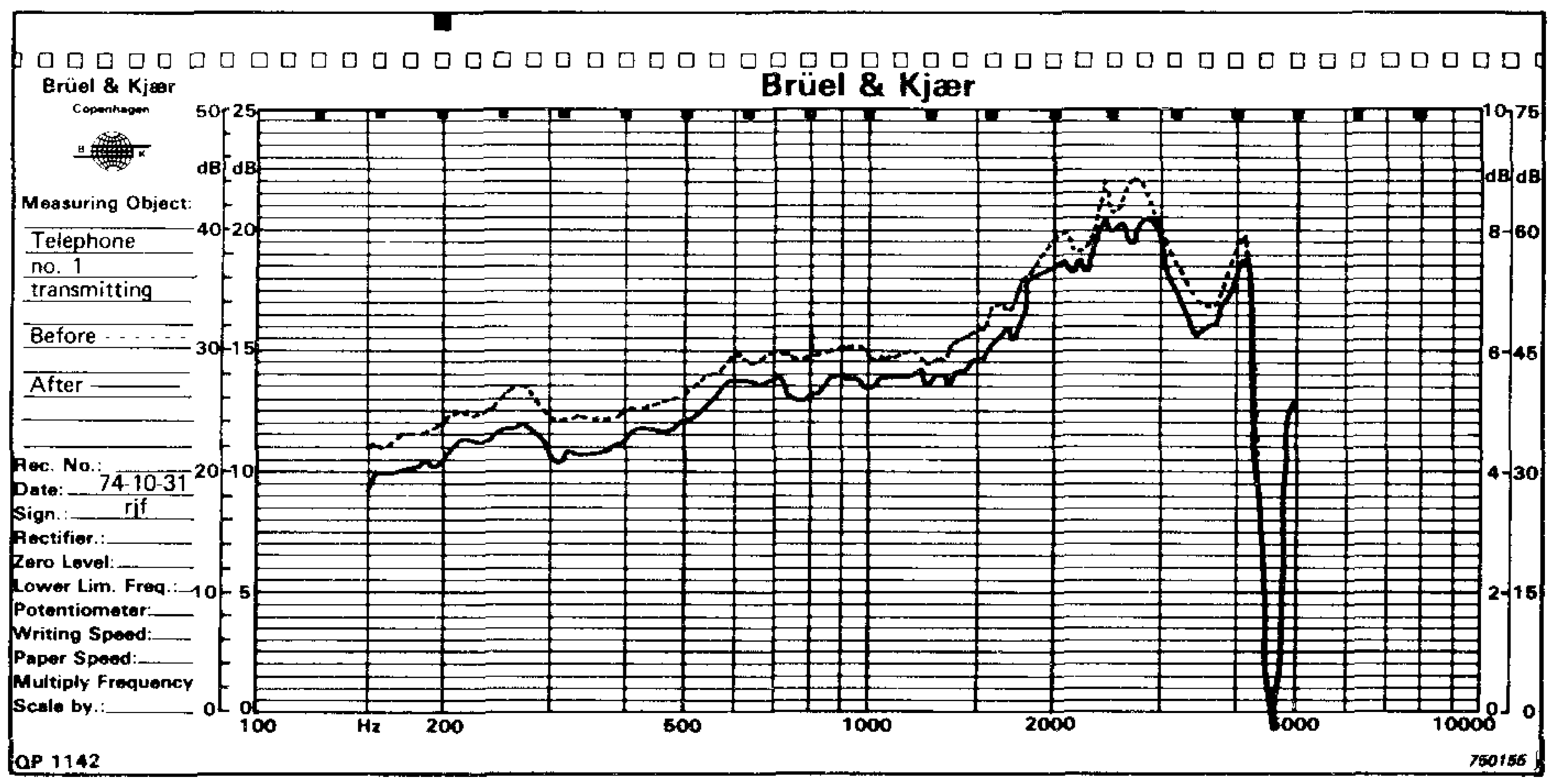


Fig. 10. Telephone no. 1 before and after testing

testing levels. In spite of this, both telephones stood up to the test. They were both still functioning afterwards and, as the curves in Figs. 10 and 11 show, their characteristics had not changed that much. In fact, the shapes of the before and after curves are almost identical in both cases. They have only shifted, no. 1 downwards and no. 7 upwards. This would tend to indicate that the microphones themselves have not altered, only the amplification system.

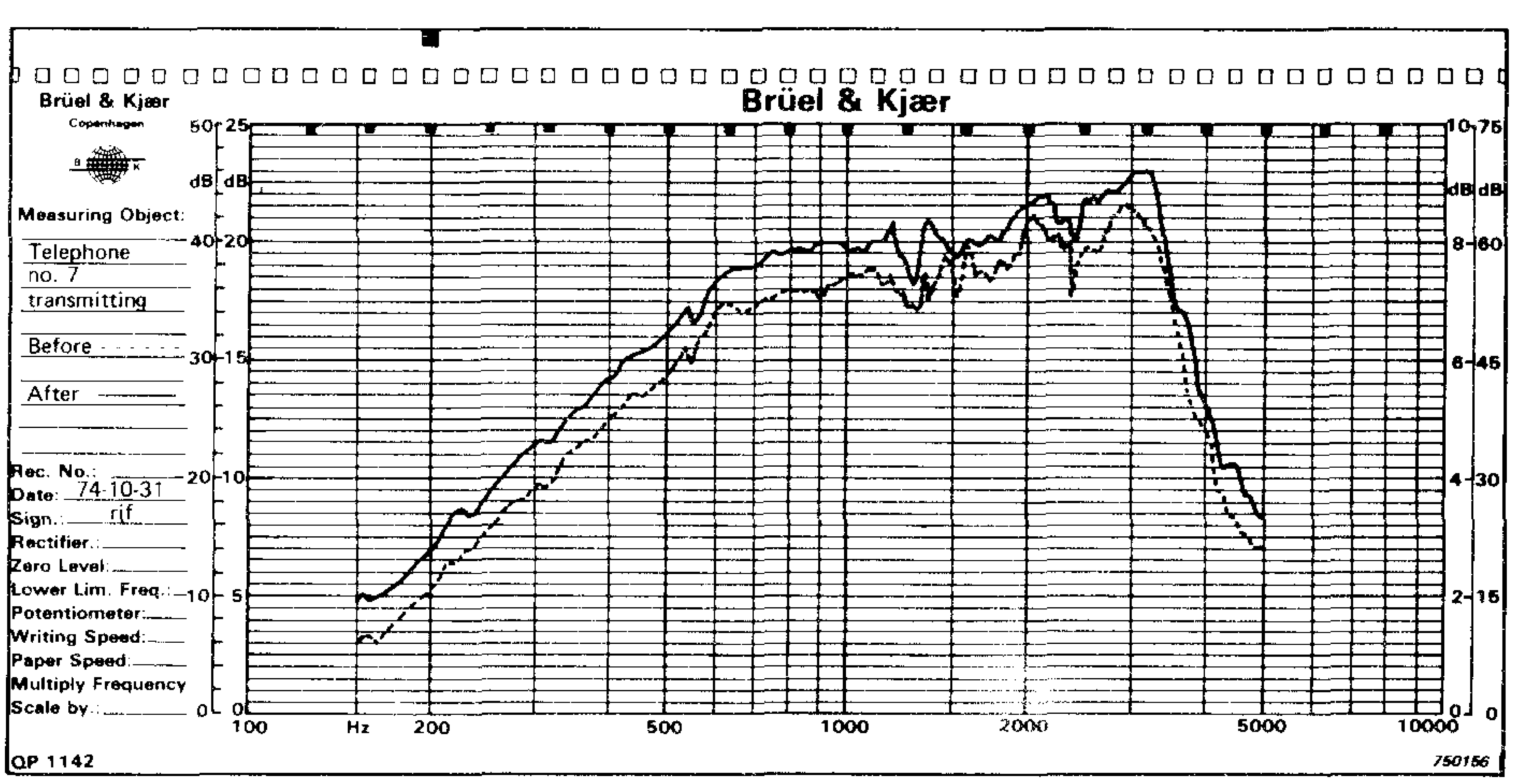


Fig. 11. Telephone no. 7 before and after testing

Note also, that in this test and in the random test, only the transmission curves were altered, not the receiving curves.

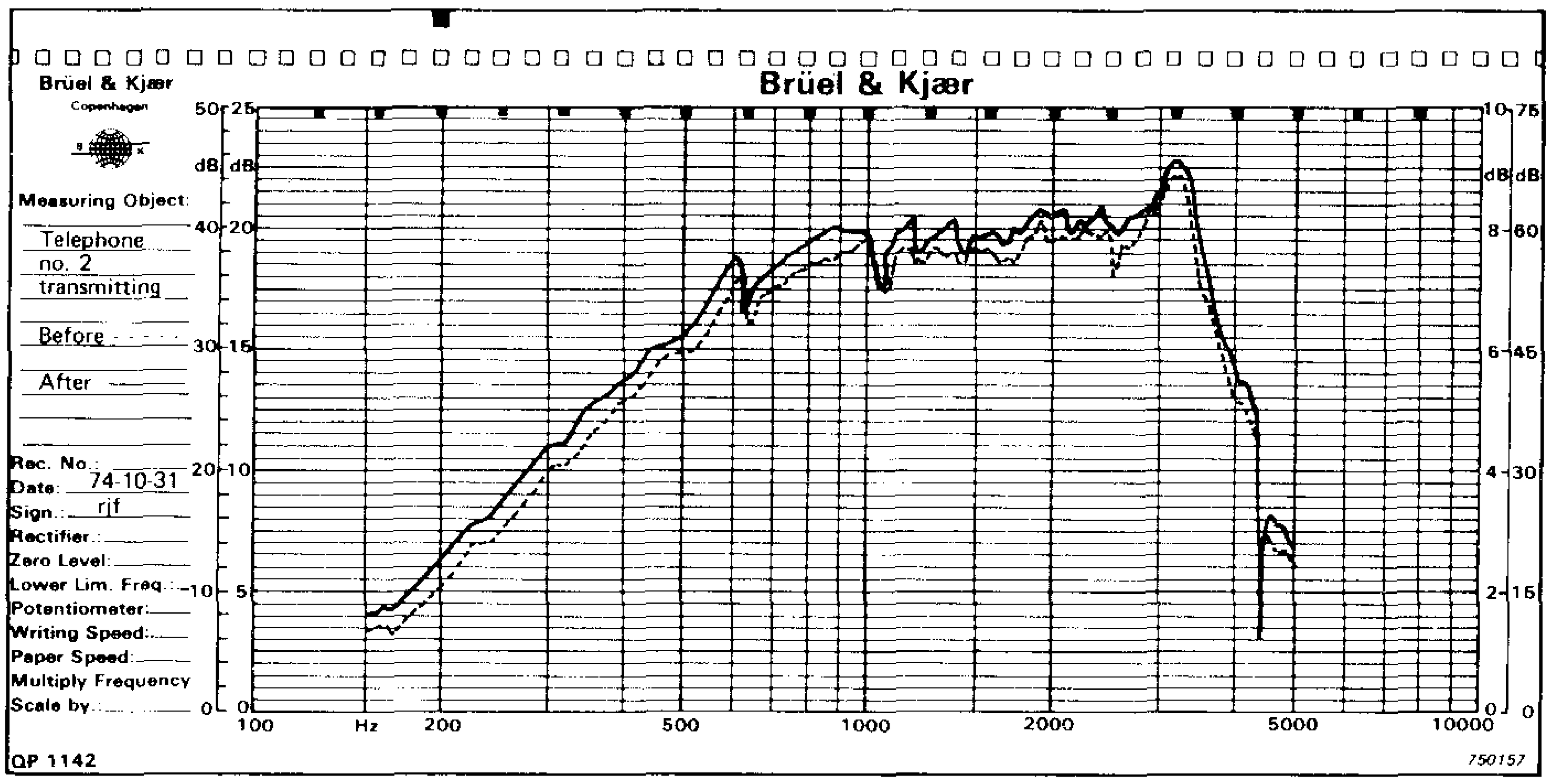


Fig. 12. Telephone no. 2 before and after testing

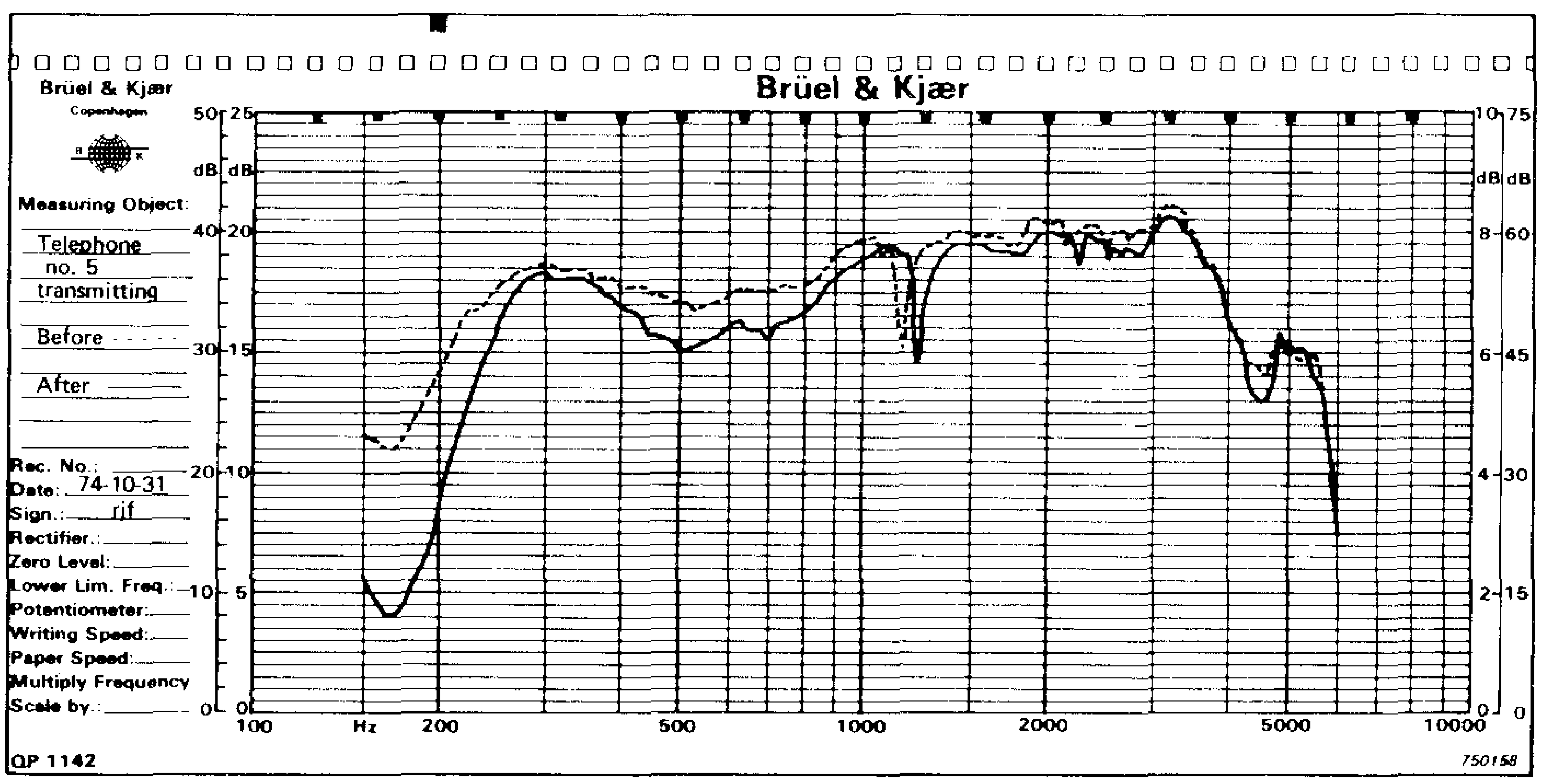


Fig. 13. Telephone no. 5 before and after testing

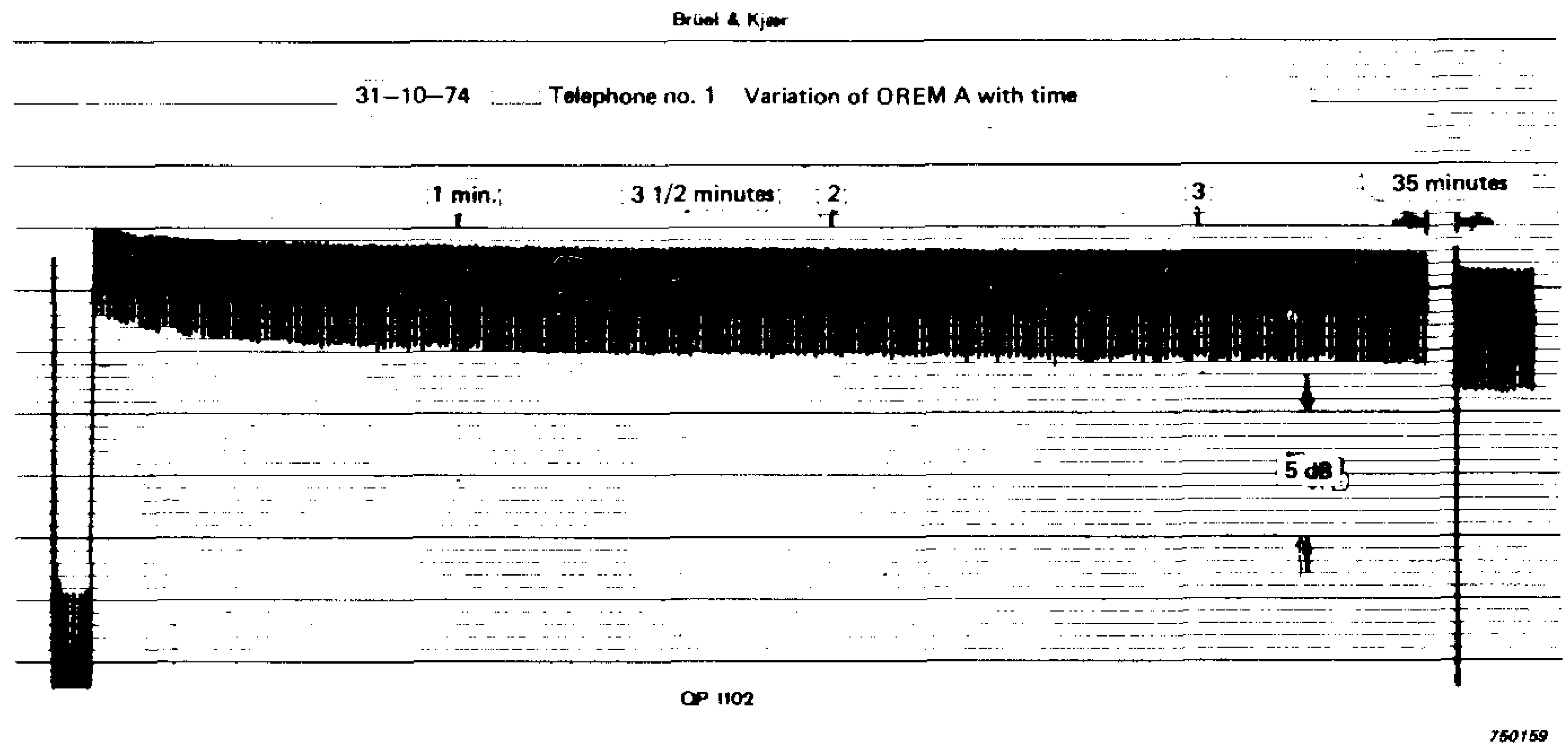


Fig. 14. Instability in telephone no. 1 as a result of vibration testing

Telephones 2 and 5 were each shaken for two hours with a 5 g RMS wide-band random (20 Hz to 2000 Hz) excitation. Figs. 12 and 13 show the results. The transmission ability of no. 2 has gone up slightly while that of no. 5 has changed radically in the lower frequencies.

There was one singular thing about the effect of vibration on no. 1, mentioned earlier. The transmission curve had not altered but the level not only went down, but became unstable. Fig. 14 shows the variation of OREM A with time over a period of 35 min. There is a drop of almost 2 dB in this time. This happens every time this phone is tested.

PREVIOUSLY ISSUED NUMBERS OF BRÜEL & KJÆR TECHNICAL REVIEW

(Continued from cover page 2)

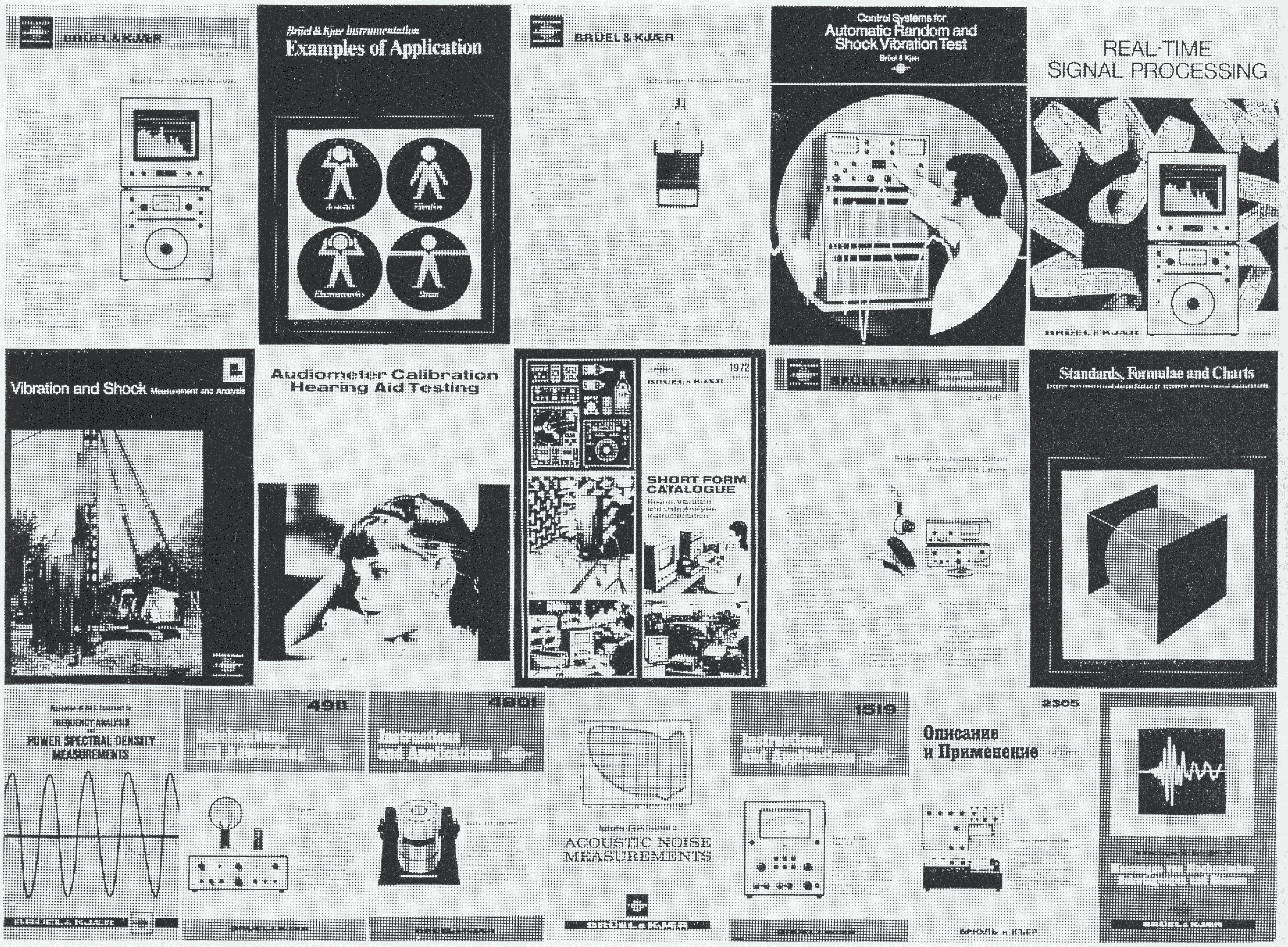
SPECIAL TECHNICAL LITERATURE

As shown on the back cover page Brüel & Kjær publish a variety of technical literature which can be obtained free of charge.
The following literature is presently available:

- Mechanical Vibration and Shock Measurements (English, German, Russian)
- Acoustic Noise Measurements (English, Russian), 2. editon
- Architectural Acoustics (English)
- Power Spectral Density Measurements and Frequency Analysis (English)
- Standards, formulae and charts (English)
- Catalogs (several languages)
- Product Data Sheets (English, German, French, Russian)

Furthermore, back copies of the Technical Review can be supplied as shown in the list above. Older issues may be obtained provided they are still in stock.

PRINTED IN DENMARK K LARSEN & SØN A/S DK-2860 SØBORG



Brüel & Kjær

DK-2850 NÆRUM, DENMARK · TEL.: (02) 80 05 00 · CABLE: BRUKJA, COPENHAGEN · TELEX: 37316 BRUKA DK