the fourth
international
congress on
acoustics

Congress Report II

copenhagen
august
21-28 1962

Published by
the Organization Committee of
the Fourth International Congress on Acoustics,
Copenhagen, 1962.
L. C. 58-30838

Index

Preface 5

Avant-Propos 7

Vorwort 9

Committees 11

BERANEK, L. LEO
Rating of Acoustical Quality of Concert Halls and Opera Houses 15

BRANDT, OVE
Sound Insulation Requirements between Dwellings 31

CHERRY, COLIN
On Listening with Both Ears 55

DAVID, EDWARD E. Jr.
Computer-Catalyzed Speech Research. 69

DOGNON, A.
Importance de diverses conditions expérimentales dans l'étude des effets ultrasoniques en milieu liquide 89

FURRER, WILLI
International Standardization Work in the Field of Acoustics 99

GRANDJEAN, E.
Biological Effects of Noise 109

JACOBSEN, E. H.
Microwave Ultrasonics and Studies of the Solid State 115

MEYER, ERWIN
Neuere Analogien zwischen akustischen und elektromagnetischen Schwingungen und Wellenfeldern 139

ROBINSON, D. W.
Recent advances in the Subjective Measurement of Noise 157

ROSENBERG, L. D.
Einige physikalische Erscheinungen, die in hochintensiven Ultraschallfeldern entstehen 179

SCHOUTEN, J. F.
On the Perception of Sound and Speech 195

WIENER, FRANCIS M.
Rocket noise of large space vehicles 209

YOUNG, ROBERT W.
A Decade of Musical Acoustics 231

Preface

The present volume represents the second part of the report on the Fourth International Congress on Acoustics, which was held in Copenhagen, Denmark, from August 21st to August 28th. It contains the invited papers read at the Congress. The first part of the report comprises the contributed papers distributed at the opening of the Congress. The Congress was the fourth of the international congresses on acoustics held under the auspices of the International Commission on Acoustics. This Commission was appointed by the International Union of Pure and Applied Physics and is thus connected with The United Nations Educational Scientific and Cultural Organization (UNESCO). Financial aid from this organization and the foundations and firms mentioned on page 12 enabled the Acoustical Society of Scandinavia to act as host of the Congress. The practical arrangement was handled by the Acoustical Society of Denmark, the board of which acted as organization committee, assisted by the advisory committee.

The Congress had 864 delegates from 33 countries. The technical activity consisted of 14 invited papers and 316 contributed papers, divided into 10 sections.

The succesful development of the Congress depended on a frictionless cooperation between many groups and individuals. It is a great pleasure for me to thank all, who in some way or other contributed to this, to our conditions, demanding arrangement.

Special thanks are expressed to Dr. G. v. Békésy, the Nobel Prize Winner, who by his paper lent additional lustre to the opening session of the Congress, and to the invited lecturers, whose manuscripts form the basis of this book, and who submitted with great understanding to the restrictions necessitated by technical and editorial conditions.

Copenhagen, February 1963

F. Ingerslev
President of the Congress.

Avant-Propos

Ce volume-ci constitue la deuxième partie du rapport du IVe Congrès International de l'Acoustique, tenu à Copenhague, Danemark, du 21 au 28 août. Il comprend les conférences générales tenues lors du congrès. La première partie du rapport contenant les conférences particulières fut publiée lors de l'ouverture du congrès.

Ce congrès était le quatrième dans la série de congrès internationaux de l'acoustique organisés sous les auspices de l'International Commission on Acoustics. L'International Commission on Acoustics est établie par L'Union Internationale des Sciences Physiques Pures et Appliquées et est ainsi reliée à The United Educational Scientific and Cultural Organization (UNESCO).

L'aide financière de cette organisation et des fonds et firmes mentionnés p. 12 a permis à la Société Nordique de l'Acoustique d'agir en qualité d'hôte du congrès. L'arrangement pratique fut l'oeuvre de la Société Danoise de l'Acoustique, dont le conseil forma le comité organisateur aidé par le comité consultatif.

Le Congrès compta 864 participants venus de 33 pays. L'activité professionelle se répartit en 14 conférences générales et 316 conférences particulières, divisées en 10 sections.

L'heureuse réalisation du congrès dépendait de la collaboration parfaite de nombreux groupes et de particuliers. A cet égard j'ai la joie de remercier tous ceux qui, d'une façon ou d'une autre, ont pris part ou contribué à cette organisation importante, eu égard à nos conditions.

Je remercie tout spécialement le Dr G. von Békésy, lauréat du prix nobel, dont la conférence donna de l'éclat à la réunion inaugurale du congrès, ainsi que les conférenciers invités dont les manuscrits forment le fonds de ce livre. Les conférenciers ont été très compréhensifs en acceptant les restrictions nécessitées pour des raisons techniques et rédactionnelles.

Copenhague, février 1963

F. Ingerslev
Président du Congrès.

Vorwort

Der vorliegende Band umfasst den zweiten Teil des Berichtes von Dem IV Internationalen Kongress für Akustik, der am 21.–28. August in Kopenhagen, Dänemark, abgehalten wurde. Er enthält die auf dem Kongress gehaltenen Plenarvorträge. Der erste Teil des Berichtes umfasst die Kurzvorträge und wurde bei der Eröffnung des Kongresses verteilt.

Der Kongress war der vierte in der Reihe internationaler akustischer Kongresse, die unter den Auspizien der International Commission on Acoustics abgehalten worden sind. Diese Kommission ist von The International Union of Pure and Applied Physics ernannt worden und steht hierdurch in Verbindung mit The United Nations Educational Scientific and Cultural Organization (UNESCO).

Geldunterstützung seitens dieser Organisation und seitens der auf Seite 12 genannten Fonds und Firmen hat es der Nordischen akustischen Gesellschaft ermöglicht, als Gastgeber für den Kongress aufzutreten. Die praktischen Veranstaltungen wurden von der Dänischen akustischen Gesellschaft ausgeführt, deren Vorstand als Organisationskomitee wirkte, von dem beratenden Komitee unterstützt.

Der Kongress hatte 864 Teilnehmer aus 33 Ländern. Die fachmässige Aktivität bestand aus 14 Plenarvorträgen und 316 Kurzvorträgen, in 10 Sektionen geteilt.

Die erfolgreiche Durchführung des Kongresses war von einer reibungslosen Zusammenarbeit vieler Gruppen und Einzelpersonen bedingt. Es ist mir eine Freude hierdurch allen danken zu können, die in irgend einer Weise zu der nach unseren Verhältnissen anspruchsvollen Veranstaltung beigetragen haben.

Einen besonderen Dank richte ich an den Nobelpreisträger, Dr. G. v. Békésy, dessen Vortrag der Eröffnungssitzung des Kongresses einen besonderen Glanz verlieh, und an die eingeladenen Redner, deren Manuskripte die Basis dieses Buches sind, und die mit grossem Verständnis die Beschränkungen befolgt haben, die aus technischen und redaktionellen Gründen erforderlich waren.

Kopenhagen, Februar 1963

F. Ingerslev
Präsident des Kongresses.

Patrons

Viggo Kampman	Prime Minister
Carl P. Jensen	Minister of Housing
K. Helveg Petersen	Minister of Education
Dr. Niels Bohr	Professor, Chairman of the Danish Committee of The International Union of Pure and Applied Physics
Dr. Esther Ammundsen	Chief Medical Officer
Dr. Carl Iversen	Professor, Rector of the University of Copenhagen
Urban Hansen	Lord Mayor of Copenhagen
A. Skalts	Head of Department, Ministry of Housing
Miss A. S. Vøhtz	Head of Department, Ministry of Education
Dr. Eggert Knuth-Winterfeldt	Professor, Rector of the Royal Technical University of Denmark
Dr. J. L. Mansa	Professor, President of the Academy of the Technical Sciences
Dr. A. N. Neergaard	Director, Chairman of the Danish Council for Scientific and Industrial Research

The International Commission on Acoustics

Dr. W. Furrer, Professor	Switzerland, President
Dr. F. Ingerslev, Professor	Denmark, Secretary
Dr. N. N. Andreyev, Professor	U.S.S.R.
Dr. L. Cremer, Professor	Germany
Dr. F. V. Hunt, Professor	U.S.A.
Mr. D. W. Robinson	Great Britain
Dr. J. Slavik, Professor	Czechoslovakia
Corresponding member:	
Dr. N. Mokhtar, Professor	The United Arab Republic

Advisory Committee

Finland: J. J. Jauhiainen, Professor
Dr. A. Sovijärvi, Professor
P. Arni, of the Finnish Broadcasting Corporation

Norway: Dr. J. Holtsmark, Professor
Reno Berg, Lecturer

Sweden: O. Brandt, M. Sc.
Dr. G. Fant, Ass. Professor
Dr. Erik Wedenberg, Lecturer
Erik Mattson, Director

Denmark: H. Højgaard Jensen, Professor
Dr. R. E. H. Rasmussen, Professor
Dr. Fritz Ingerslev, Professor

Acknowledgements

The Fourth International Congress on Acoustics has received financial aid from the following funds:
The Danish Government Fund for Scientific and Industrial Research
The Thomas B. Thrige Foundation
The Otto Mønsted Foundation
and from the following private firms:
A.E.G., Dansk Elektricitets A/S
American Danish Oticon A/S
Bang og Olufsen A/S
Brüel og Kjær A/S
Danavox A/S
A/S Dansk Glasuldfabrik
Dæmpa A/S
L. M. Ericsson
Evers og Co. A/S
General Radio Comp.
A/S Hellesens
IKAS A/S
Otto S. Jensen
Philips A/S
A/S Rockwool
Dr. A. Stankiewicz, G.m.b.H.
Tøpholm og Westermann A/S

The Board of the Acoustical Society of Scandinavia

Denmark: Dr. Fritz Ingerslev, Professor, Chairman
Finland: P. Arni, of the Finnish Broadcasting Corporation
Norway: Reno Berg, Lecturer
Sweden: Dr. Gunnar Fant, Ass. Professor

The Board of the Acoustical Society of Denmark

Dr. Fritz Ingerslev, Professor, Chairman
Dr. P. V. Brüel, M. Sc.
Dr. H. W. Ewertsen, Chief Physician
Dr. V. L. Jordan, M. Sc.
Dr. H. K. Kristensen, Professor
Dr. A. Kjerbye Nielsen, M. Sc.
This Committee is also acting as the Organisation Committee of the Congress.

Ladies' Committee

Mrs. D. Brüel
Mrs. B. Ewertsen
Mrs. E. Hauschildt
Mrs. E. Ingerslev
Mrs. E. Jordan
Mrs. R. Kristensen
Mrs. L. Kjerbye Nielsen
Mrs. L. Boel Pedersen
Mrs. B. Juhl Pedersen
Mrs. I. Petersen
Mrs. H. Rasmussen
Mrs. I. M. Ring
Mrs. E. Schlegel

Officers

Secretary-General: Peter Plum, M. Sc.
Secretary to the Exhibition Committee: O. Ring, M. Sc.
In charge of the lecture arrangements: Dr. A. Kjerbye Nielsen. M. Sc.
In charge of the technical excursions: Knud Rasmussen, M. Sc.
In charge of meeting rooms: S. Boel Pedersen, M. Sc.
Secretary to the President: O. Juhl Pedersen, M. Sc.

Rating of Acoustical Quality of Concert Halls and Opera Houses[1]

LEO L. BERANEK

Bolt Beranek and Newman Inc., 50, Moulton Street, Cambridge, Massachusetts.

I. Introduction

A detailed study of the acoustics of concert halls and opera houses was launched in 1955 for the purpose of discovering why most halls built in the twentieth century have failed to satisfy musicians, music critics, and audiences. This investigation has ranged over twenty nations and from it detailed architectural and acoustical data have been obtained on 54 halls for music. Twenty three internationally known conductors and musicians were interviewed in depth—some interviews taking up to three hours. [2] Twenty one music critics of the United States, Canada, Great Britain, and Germany were interviewed in the same manner. [3] A spread of ages and backgrounds were represented.

The acoustical measurements that were performed and technical details assembled include (a) reverberation time as a function of frequency with the audience and orchestra present, (b) reverberation time as a function of frequency with the hall empty, (c) pulse measurements (in some halls only) designed to show the decay of sound in detail at several positions

1. The material in this paper is abstracted from a book by the author, *Music, Acoustics, and Architecture*, John Wiley and Sons, Inc., New York and London (1962).

2. The artists include: Sir John Barbirolli, Sir Adrian Boult, Leonard Bernstein, Eleazar de Carvalho, Alexander Gibson, Tauno Hannikainen, Irwin Hoffman, Herbert von Karajan, Erich Leinsdorf, Igor Markevich*, Dimitri Mitropoulos, Pierre Monteux, Charles Munch, Eugene Ormandy, Fritz Reiner, Sir Malcom Sargent, Hermann Scherchen, Stanislaw Skrowaczewski*, Izler Solomon*, William Steinberg*, Isaac Stern, Leopold Stokowski, and Bruno Walter. (*Interviewed by Russell Johnson.)

3. The music critics and their newspapers are: Thomas Archer*, Montréal *Gazette;* Roger Dettmer, Chicago *American;* Cyrus Durgin, Boston *Globe;* Alfred Frankenstein, San Francisco *Chronicle;* Albert Goldberg, Los Angeles *Times;* Jay Harrison, *Music Magazine;* Peter Heyworth, London, *The Observer;* Frank Howes, London, *The Times;* Paul Hume, Washington *Post;* Irving Kolodin, New York *Saturday Review;* Paul Henry Lang, *New York Herald Tribune;* Eric McLean*, *Montreal Star;* Robert C. Marsh*, Chicago-*Sun Times;* Colin Mason, Manchester, *The Guardian;* Antonio Mingotti, Munich, *Abendzeitung;* Harold Rogers, *The Christian Science Monitor;* Max de Schauensee, *Philadelphia Bulletin;* Harold Schonberg, *New York Times;* Desmond Shawe-Taylor, London, *The Sunday Times;* Howard Taubman, *New York Times;* Jules Wolffers, *Boston Herald.*

in the hall, (d) accurate architectural drawings, (e) dimensional measurements of stage, pit, stage enclosure, and the hall itself, generally made from existing architectural plans, with portions checked to insure accuracy, (f) the number of seats on each floor level, number of standees and number in boxes, and (g) details on seats, carpets, wall and seating surfaces, stage and hall floors, draperies, curtains, proscenium opening, etc.

Much information was collected during personal visits as the architectural drawings seldom show up-to-date detail. Also, it was found that many halls have been modified: the seats changed, balconies rebuilt, stairways relocated, stage extended or shortened. As a result, great care was used to check all factors rather than assuming blindly that the hall was in conformance with existing drawings. All technical material was sent to appropriate people for checking for accuracy.

II. Results of the Interviews and Listening Judgments

In order for acoustical science to provide directions to architecture in the interests of music, a scale is needed against which the quality of a hall can be gauged, and the importance of a particular acoustic attribute weighed. The evaluations of 53 of the halls were limited to two types of performances, namely, symphony orchestra concerts, or non-Wagnerian opera, or both. With the help of the interviews with musicians and critics and my own notes made while listening in almost all the halls, I have been able to assign them into five categories of quality: A+, excellent; A, very good to excellent; B+, good to very good; B, fair to good; and C+, fair. The 54th hall, Bayreuth Festspielhaus, is especially suited to Wagnerian opera and must be rated on a different basis. This study has purposely eliminated poor halls as their gross defects are easily identified.

It is probable that any particular hall could be ranked one category higher or one category lower than its present assignment. An error of three categories is almost impossible, and an error of two categories is highly unlikely.

But, the initial assignment of the halls into categories was not depended on alone. After the compilation was completed, it was sent with a request for criticism to most of the men interviewed. A small number of shifts to an adjacent category resulted from their comments, but the majority wrote as did Howard Taubman of the *New York Times*, "On the whole I agree with you about the halls that I know anything about."

A. *Definitions of the Categories*

Category A+: The halls in this category were rated "excellent" for symphony orchestra concerts or non-Wagnerian opera by nearly every musician or critic interviewed. The halls received very few adverse comments. Six concert halls and seven opera houses fell into this category.

Category A: The halls in this group were generally rated "very good" to "excellent." They are all deficient to some extent in one or two of the important acoustical attributes. Each hall is highly treasured by the community it serves and is spoken of favorably by the majority of those musicians who perform there. There are nineteen concert halls and eight opera houses in this category.

Category B+: The halls in this category were rated "good" to "very good." They received more favorable than adverse criticism. These halls are not mediocre; any negative criticism received came in response to a request for an assessment of both the good and bad qualities. Some of the halls are better suited to the music of the Baroque period that to that of either the Classical or Romantic period. Some are very large. All but three were built for all types of use—music, speech, dance, and so forth—calling for compromises that generally degrade the acoustics of the hall for concerts or opera. There are fourteen concert halls and four opera houses in this category.

Category B: The halls in this group were rated "fair" to "good." They received more adverse than favorable criticism in the interviews, but they definitely are not poor or unsatisfactory halls. Of the seven concert halls in this category all but two were built to serve many different purposes.

Category C+: There is only one hall in this category. The chief factor that adversely affects its acoustics are its large lateral dimensions and its huge cubic volume. It is quite satisfactory for organ and choral music of the cathedral type and for ceremonial orchestral and band music.

B. *Cubic Volume and Age*

During the interviews, musicians often commented that: (1) small halls generally sound better than large ones; (2) halls built to serve many purposes are inferior to halls built especially for concert or opera; and, (3) old halls sound better than new ones. Investigation reveals that the median cubic volume of the ten best-liked concert halls is 620,000 cu ft (17,600 cu m) compared to 770,000 cu ft (21,800 cu m) for the 10 least-liked concert halls. This difference of 25% in cubic volume is hardly significant by itself.

In opera houses, size is very important. An artist can sing more easily in a small house than in a large one. Singers who sound one way at the Metropolitan Opera House in New York, sound different in smaller European theaters. In small houses, singers are less inclined to force their voices and so are more relaxed. Obviously, for equal reverberation times, a voice is louder in a small house than a large one.

In regard to age, neither acoustical data nor reviews by music critics indicate that any hall has changed acoustically with time unless architectural changes were made. The eight least-liked 20th Century halls of this study differ from the seven best-liked 19th Century halls in the following ways: The median floor area occupied by seats and orchestra is 65 per cent greater in the newer halls than in the older halls. The median ceiling height is 8 ft lower. The median mid-frequency reverberation time is 0.2 sec lower. The interiors in the new halls, in general, contain thin wood (two contained large areas of sound-absorbing tiles). Most damaging of all to the 20th Century halls, the median width is 83 per cent greater, yielding a greater difference between the times of arrival at listeners' ears of the direct sound and the first reflections. Architectural differences, not age, are clearly responsible for the acoustical differences.

C. *Importance of Reverberation Time*

Let us see if there is a high correlation between the subjective rank orderings of concert halls and their mid-frequency reverberation times (See Table 1). In Category A+ there are six

halls with reverberation times between 1.7 and 2.05 seconds—a median of 1.9 seconds. In the three categories of A, B+, and B the reverberation times range between 1.0 and 2.0 seconds, with median values of 1.5 and 1.6 seconds.

Table I

Correlation Between Mid-Frequency Reverberation Times and Rank-Order Categories for 47 Fully-occupied Concert Halls

Rank order category	Mid-Frequency reverberation times—sec.	Median sec.
A+	1.7, 1.8, 1.8, 2.0, 2.05, 2.05	1.9
A	1.2, 1.3, 1.4, 1.4, 1.5, 1.5, 1.5, 1.6, 1.6, 1.6, 1.6, 1.7, 1.7, 1.7, 1.7, 1.7, 1.8, 1.9, 2.0	1.6
B+	1.0, 1.3, 1.4, 1.4, 1.5, 1.5, 1.5, 1.5, 1.5, 1.6, 1.7, 1.7, 1.7, 1.9	1.5
B	1.3, 1.4, 1.5, 1.6, 1.6, 1.6, 1.8	1.6
C+	2.45	2.45

Four tentative conclusions can be drawn from Table 1.

1. In order for a concert hall to fall in the A+ category, its reverberation time must be at least 1.7 sec.
2. A reverberation time as long as 1.7 sec does not ensure that a concert hall will fall in category A+ and, therefore, reverberation time alone does not distinguish an excellent hall from an inferior hall.
3. Let us assume that the only reason the seven halls in Category A whose reverberation times are shorter than 1.6 sec do not fall in A+ is their short reverberation time since, if they were deficient in additional ways, they would fall in a still lower category. Then it follows that the difference between the median of the A+ group (1.9 sec) and the median of the seven halls with short reverberation time (1.4 sec) is one rating category.
4. Thirteen halls in Categories A, B+, and B have reverberation times between 1.7 to 2.0 seconds. Perhaps deviation in only one acoustical attribute is primarily responsible for their falling into a lower category below A+.

III. The Positive Acoustical Factors Affecting Musical Quality

Let us examine Item 4 above to see if there is one acoustical attribute that, when the other attributes are alike, is responsible for the observed differences in the rank orderings.

A. *Initial-Time-Delay Gap*

The acoustical data for the 20 concert halls of Table 1 with reverberation times at least as great as 1.7 seconds reveal that all but three halls have about the same ratios of low frequency and high frequency reverberation times to those at mid-frequencies; there is no observable tonal distortion; there is almost no noise; balance and blend are good; sound diffusion is adequate; and those listeners in the middle of the main floor are about 60 ft from the concert master. The only clearly observable variable among the remaining 17 concert halls is the time difference between the arrival (at the listeners' ears) of the direct sound and the first reflection. Let us call this difference the initial-time-delay gap (See Figs 1 and 2).

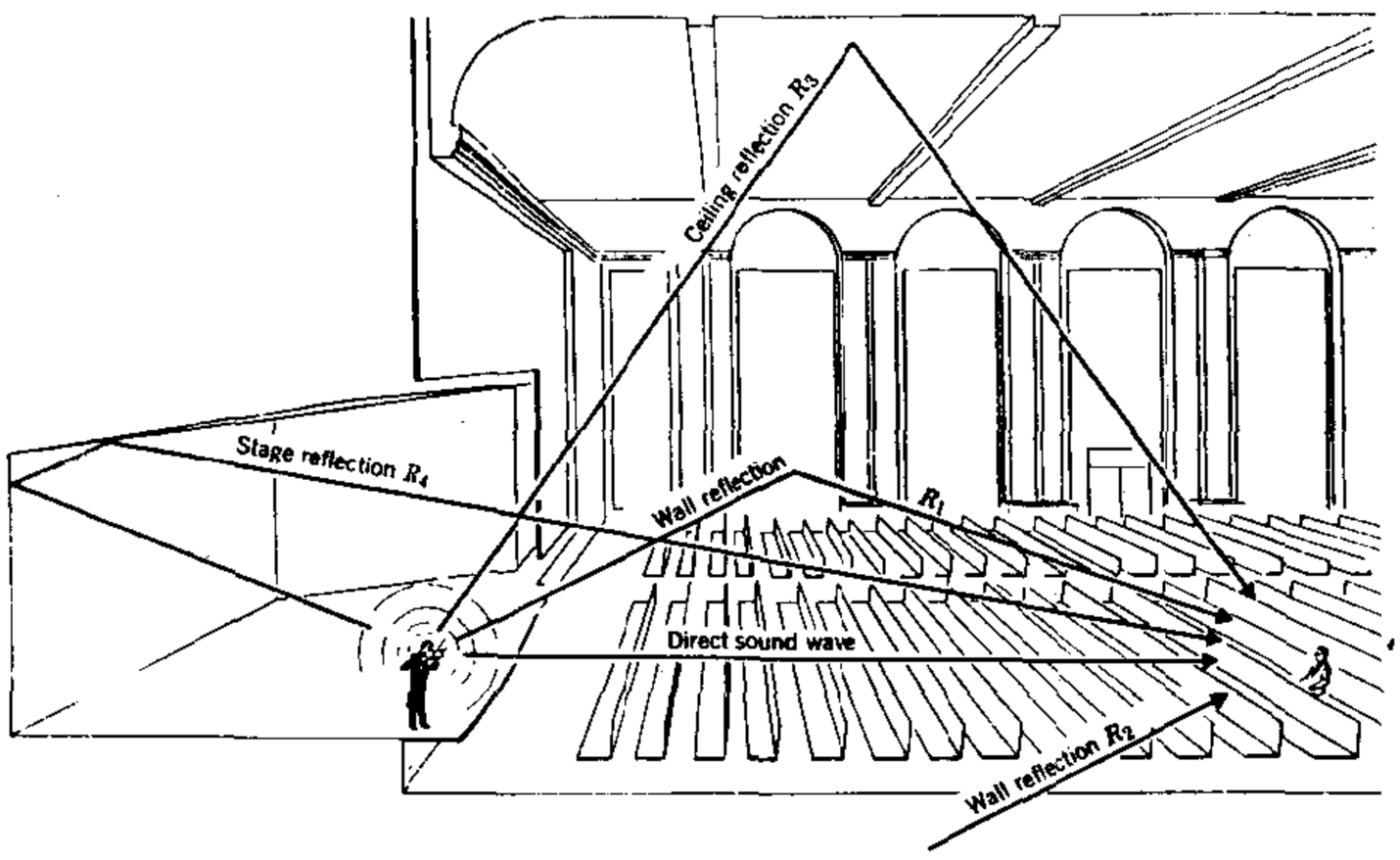


Fig. 1. Sketch showing direct and four reflected sound waves in a concert hall.

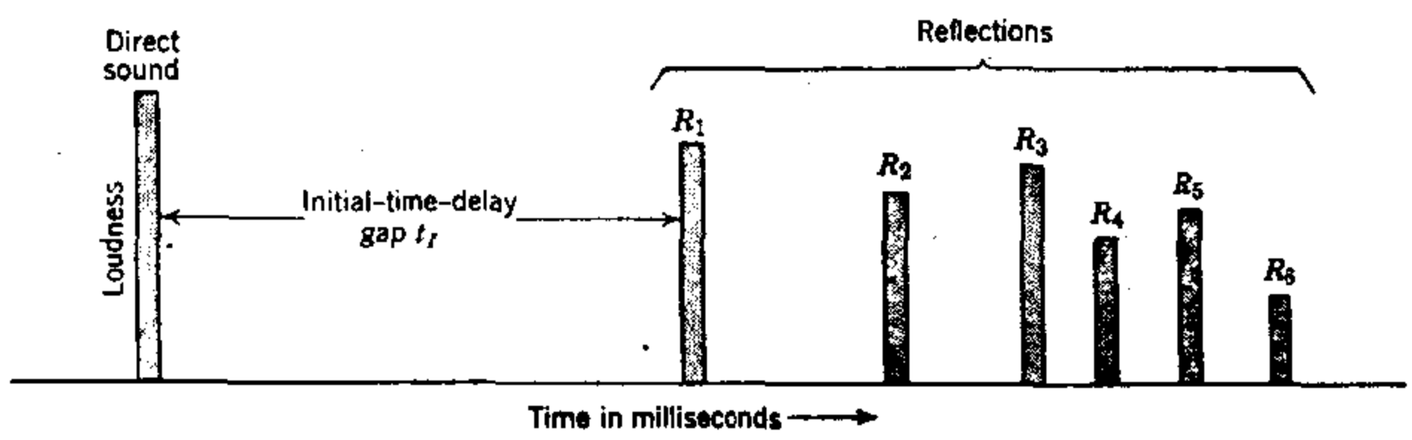


Fig. 2. Time diagram showing that at a listener's ears, the sound that travels directly from the performer arrives first, and after a gap, reflections from the walls, ceiling, stage enclosure and hanging panels arrive in rapid succession. The height of a bar is related to the intensity of the sound. The initial-time-delay gap t_1 is indicated.

Concert Halls: The initial-time-delay gap at the ear of a listener on the main floor of a concert hall is easily measured from architectural drawings. The initial-time-delay gaps for the 47 concert halls of this study are plotted within the blocks with solid boundaries on Fig. 3. The 17 halls that differ primarily in initial-time-delay lie within the blocks with solid boundaries. One other hall with slightly low bass reverberation also falls in this group. Of the 18 halls, the six in Category A+ have initial-time-delay gaps, measured at the center of the main floor, of between 8 and 21 milliseconds; the 8 in A, between 22 and 34 milliseconds; the 2 in B+, between 35 and 46 milliseconds; the one in B, between 47 and 58 milliseconds; and the one in C+ between 59 and 71 milliseconds. Each solid block is about 12 milliseconds

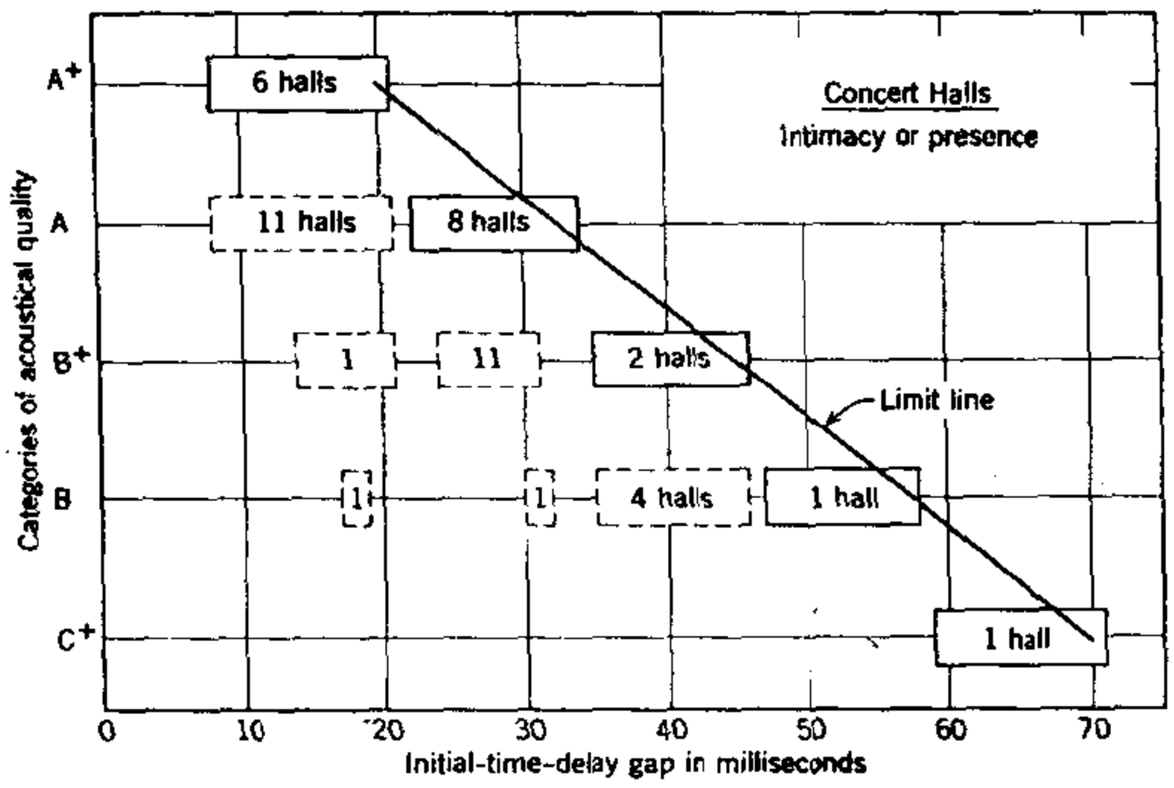


Fig. 3. Chart of rank-order categories for 54 halls as a function of initial-time-delay gap. The heavy diagonal line indicates the highest category that a hall with a given initial-time-delay gap can achieve.

long. The remaining 29 concert halls are plotted in blocks with dashed boundaries. These 29 halls fall one or more categories below what would be expected if the initial time delay gap alone could determine the excellance of a concert hall. Apparently, these 29 halls have deficiencies in some or all of other acoustical attributes.

Of the 26 Concert Halls of Fig. 3 that fall one category below the solid blocks, 9 have short reverberation times at all frequencies, 7 are slightly deficient in a number of acoustical attributes, 5 have somewhat short reverberation times at mid and high frequencies and are somewhat deficient in the ratio of bass reverberation to mid-frequency reverberation, 4 have distortion, combined with fairly short reverberation times at mid and high frequencies; and one is seriously deficient in the ratio of bass reverberation to mid frequency reverberation.

Of the 2 Concert Halls of Fig. 3 that are two categories below the solid blocks, both have short reverberation times at all frequencies, both have low ratios of the reverberation times at low frequencies to those at mid frequencies, and both are wanting in balance and blend.

The 1 Concert Hall of Fig. 3 that falls three categories below a solid block has a short reverberation time, a deficiency in bass reverberation, poor balance and blend, some tonal distortion, and some echo.

From this recital of deficiencies one can make certain deductions about concert halls:

1. A relatively short mid-frequency reverberation time, (the average of R.T.'s at 500 and 1000 cps), of about 1.4 sec with full occupancy, causes a drop of about one rating category. This was also suggested by tentative conclusion 3, relating to Table 1.
2. A serious deficiency in ratio of bass to mid-frequencies reverberation produces a drop of one category.
3. Poor balance and blend account for a drop of about one half a category.
4. Distortion to the extent found in these halls causes a drop of about one-half a category.

Obviously, a short initial-time-delay gap is necessary to assure excellance in a concert hall, but it alone is not sufficient to guarantee it.

Opera Houses: A similar graph for opera houses [4] has been drawn. It is found that about 20 per cent longer initial-time-delay gaps than those shown on Fig. 3 appear to be permissible in each rating category.

In Fig. 3, a heavy, diagonal line intersects the right-hand end of each of the five solid blocks. This line delimits the highest category into which a concert hall can fall for any initial-time-delay gap. It appears to have basic psychoacoustic significance.

Haas [5], and Muncey, Nickson and Dubout [6] have investigated the addition of a delayed reflection to original music. The Australian experiments were performed by playing music through loudspeakers to a group of twenty listeners and adding to it the same music delayed in time by up to 1000 milliseconds. The subjects, laboratory personnel, were asked to decide whether they were disturbed by the added reflection. The results showed they were most disturbed by reflections added to fast string music, the kind of music that forms the foundation of present-day symphonic concerts. When the intensity of reflections was not more than 5 decibels below the intensity of the direct sound, fewer than 20 per cent of the subjects were disturbed by delays of 15 to 30 milliseconds. Nearly all the listeners were disturbed by reflections that were delayed by 100 milliseconds.

If we assume that those 10 to 20 per cent of the group who found delays of 15 to 30 milliseconds disturbing were critical listeners, then it should be reasonable to conclude that an initial-time-delay gap in excess of about 20 milliseconds detracts from musical quality. The result of the tests imply also that, the greater the initial-time-delay gap the less pleasant the music becomes. It seems safe to conclude that acoustical intimacy cannot survive an initial-time-delay gap of 70 milliseconds, particularly if the first reflection has an intensity within 10 decibels of the direct sound. Hence, the findings of the Australian group and the results shown in Fig. 3 are in general agreement.

According to Fig. 3, since an initial-time-delay gap of 70 milliseconds or more corresponds to a ratio of C+ or lower, it contributes no points to the rating of a concert hall. How many points should be assigned to a change of one category? It appears from the interviews and the author's judgment that those halls in Category B are subjectively about 70 per cent as good as those in A+. Hence, it seems that the right number is about 10 points per category on a 100 point scale. (See Fig. 4) Hence, if an initial-time-delay gap of 70 milliseconds is zero, then one of 20 milliseconds or less, being four categories higher, must equal 40 points. Rating scales so derived, for the attribute of initial-time-delay gap for concert halls and opera houses, are shown in Fig. 5.

B. *Reverberation Times at Middle and High Frequencies*

The interviews clearly show that for today's typical orchestral repertoires, musicians and most music critics consider that reverberation times at mid-frequencies of under 1.6 seconds are too short. The optimum reverberation times, for full occupancy, at mid-frequencies

4. Twelve of the halls used for deriving rating scales for concerts are also used for operas. The musicians' and conductors' ratings of a hall are generally different for opera than concerts and this fact was taken into account in the ratings.

5. H. Haas, *Acoustica*, *1*, pp. 49–58 (1951)
6. R. W. Muncey, A. F. B. Nickson, and P. Dubout, *Acoustica 3* pp. 168–173 (1953)

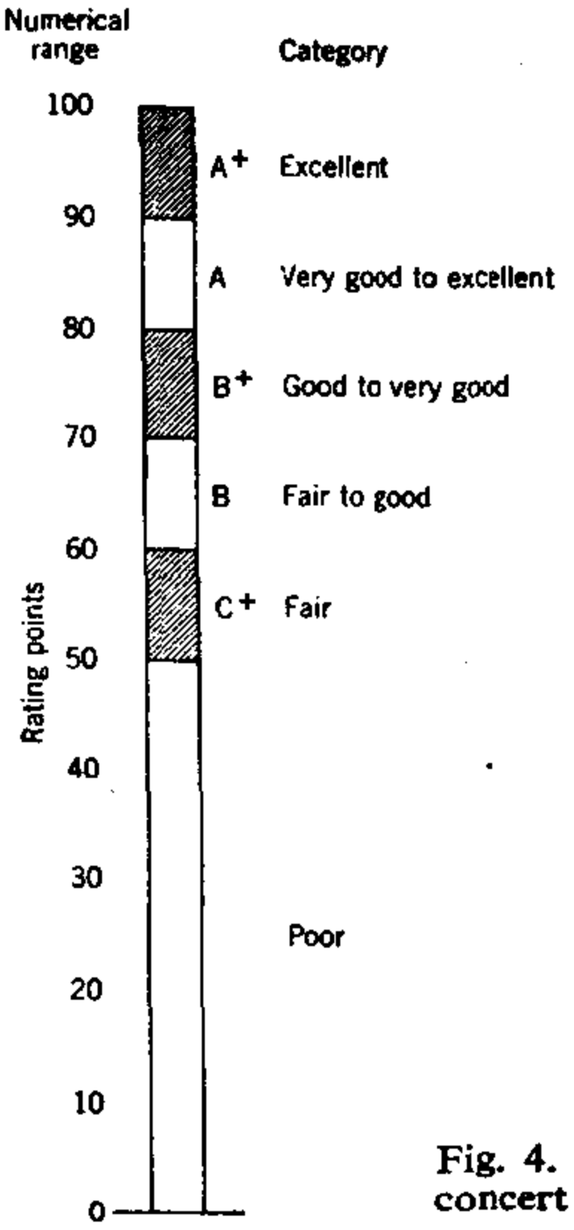


Fig. 4. Range of numerical values arbitrarily assigned to the five categories into which the concert halls of this study are divided, plus the category "poor".

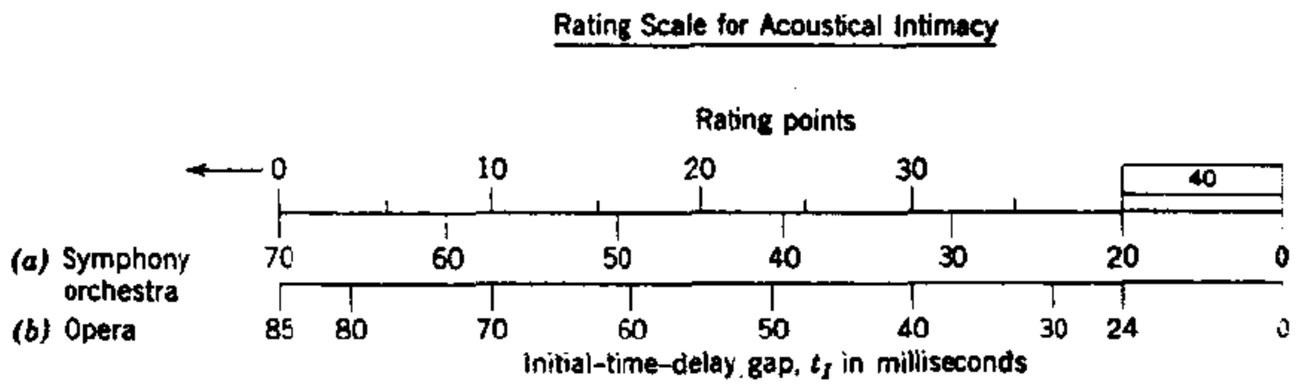


Fig. 5. Rating scale for the acoustical attribute of initial-time-delay gap: a) concert halls. b) opera houses.

(average of R.T.'s at 500 and 1000 cps), are found to be 2.1 to 2.3 sec for music of the Romantic period; 1.6 to 1.8 for music of the Classical period; and 1.4 and 1.6 for music of the Baroque period. The compromise optimum reverberation time for today's symphonic repertoire is found to be 1.8 to 2.0 sec. These times seem independent of cubic volume, for the range of volumes studied, (200,000 to 1,500,000 cu ft.).

Measurements in the 54 halls showed that, if there were no significant amounts of draperies or other sound-absorbing materials on the ceiling or walls, the reverberation times at 2000 cps was about 0.9 that at mid-frequencies; that at 4000 cps was about 0.8 and that at 6000 cps was about 0.7 [7]. Reverberation times apply to full occupancy.

We learned from the previous section that in a concert hall when the reverberation time is low (about 1.4 sec), compared to its optimum of about 1.9 sec, the rating decreases by about one rating category, equivalent to 10 rating points. Also, it is well accepted that the reverberation in a concert hall adds very little to symphonic music if it is shorter than about 1.1 sec. One judges that if a change in the region from 1.9 sec down to 1.4 sec is 10 points, then an additional drop of 0.3 sec to 1.1 sec amounts to about five points. The overall range, therefore, of the scale for mid-frequency reverberation times is 15 points.

From the evidence at hand, it seems that an opera house for Italian music should have about the same reverberation time as a concert hall for Baroque music and one for Wagnerian music should have about the same R.T. as a concert hall for Classical music.

Rating scales for mid-frequency reverberation time for concert halls and opera houses are given in Fig. 6.

7. Leo L. Beranek, *J. Acoust. Soc. Am.*, *32*, pp. 661–670 (1960)

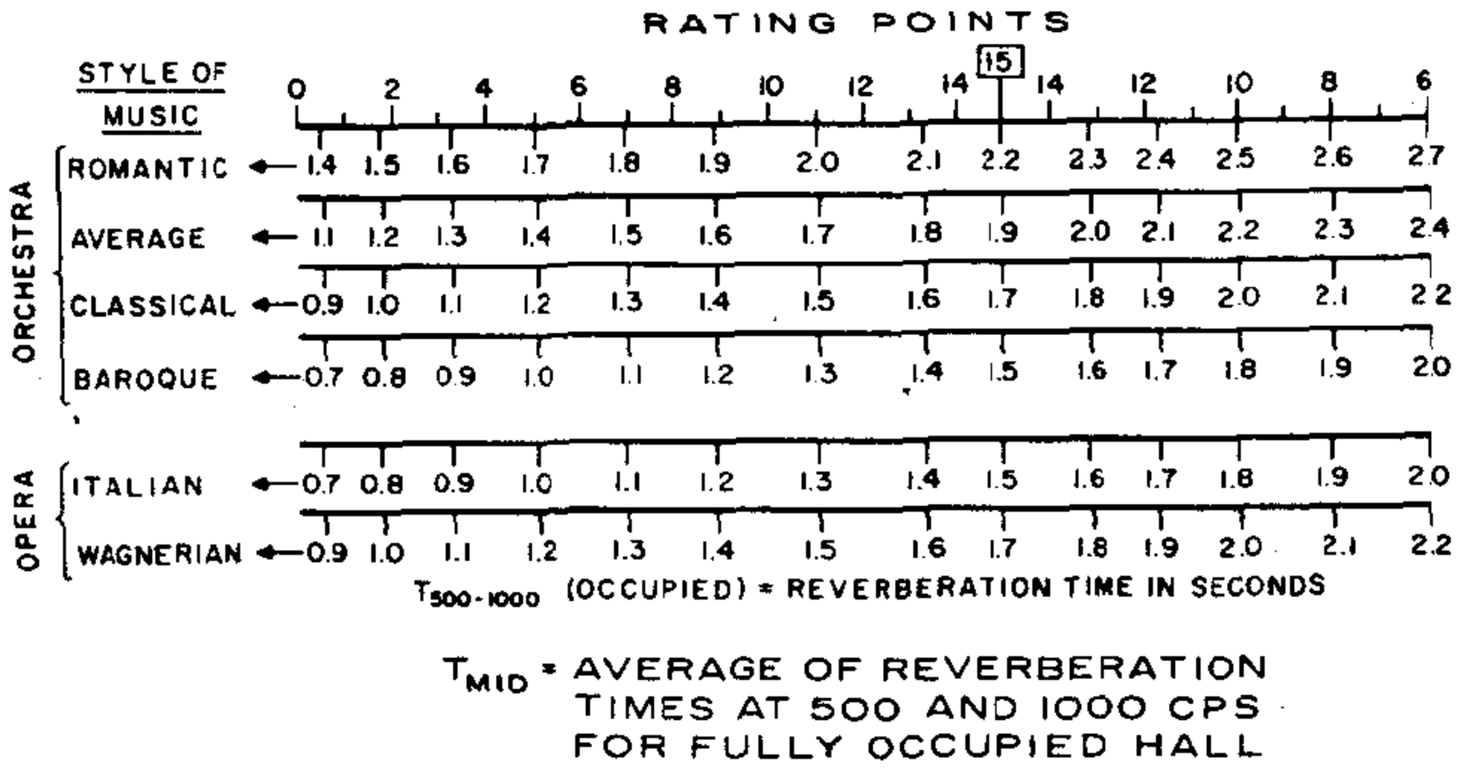


Fig. 6. Rating scale for the acoustical attribute of mid-frequency reverberation time for fully occupied halls.

C. *Bass to Mid-Frequency Reverberation Time Ratio:*

Many concert halls were criticised by musicians and critics as being deficient in bass. If one divides the concert halls into four groups: (1) excellent bass, (2) good bass, (3) fair bass, and (4) poor bass, as judged by those interviewed, he finds that the *bass ratio*, the average of the reverberation times at 125 and 250 cps divided by the average of the R.T.'s at 500 and 1000 cps, is for (1) between 1.2 and 1.25, for (2) about 1.05; and (3) about 0.96, and for (4) about 0.88. In no hall was there a specific complaint about too much bass even though the bass ratios are as high as 1.4.

In Item 2 of the section on initial-time-delay gap, we said that a serious deficiency in ratio of bass to mid-frequency reverberation corresponds to a drop of about one category. Also, it seems that the bass ratio is about of the same importance as the mid-frequency reverberation time in setting the musical quality of a hall. Hence, let us assign, for concert halls, 15 points to an average bass ratio of between 1.2 and 1.25. For opera houses, the bass ratio need not be as great as in concert halls, because of the relative unimportance of the bass singing part.

Rating scales for concert halls and opera houses, as derived from this study for the attribute of bass ratio are shown in Fig. 7.

D. *Intensity of the Direct Sound*

There is an optimum listening level for music and this applies separately to the direct sound and the reverberant sound. If the direct sound is too weak, it may get lost in the reverbera-

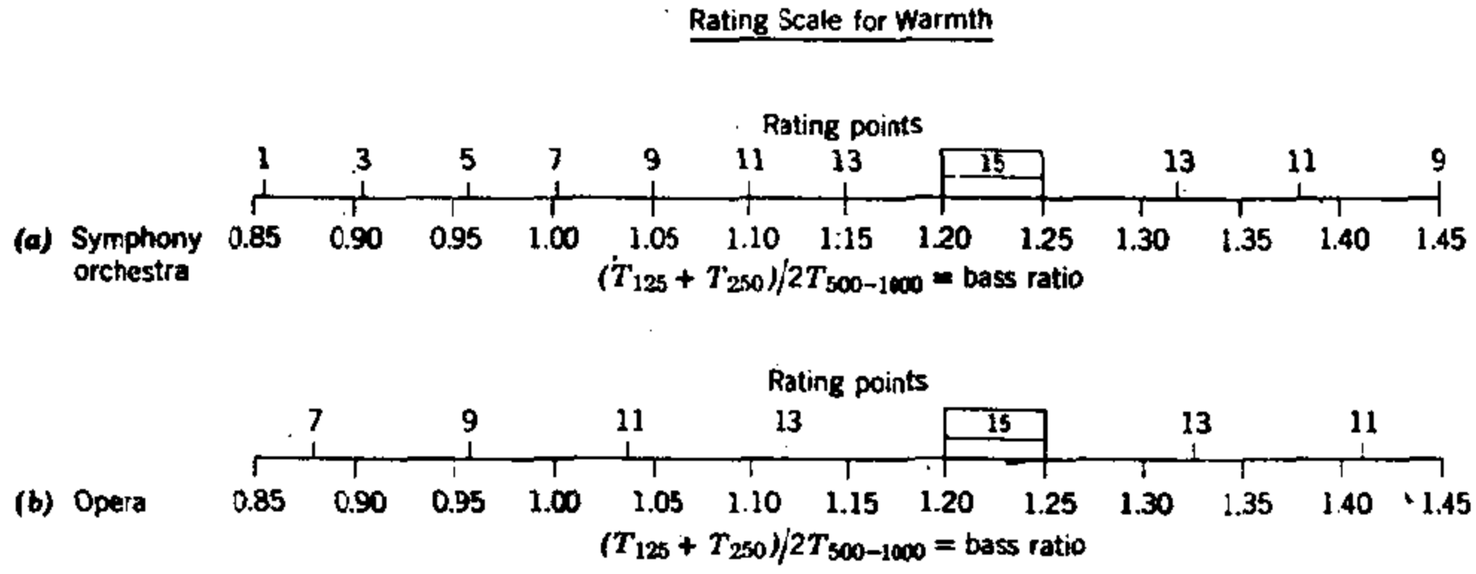


Fig. 7. Rating scale for the acoustical attribute of bass ratio, the average of the reverberation times at 125 and 250 cps to the average of the R.T.'s at 500 and 1000 cps (mid frequencies), for: a) Symphony Orchestra. b) Opera.

tion or in general audience noise. Reflecting surfaces enhance the level of the direct sound by shortening the initial-time-delay gap—but this contribution has already been rated. Furthermore, the design of the stage affects the balance and blend of the music, a contribution that is to be rated later.

Investigation of people's preferences for seats on the main floors of halls for music, indicate that the optimum distance for an auditor from a large symphony orchestra is about 60 feet measured from the concertmaster or, in an opera house, from the average singers' position. In halls with a very large main floor and no balcony overhang (e.g., Tanglewood Music Shed at Lenox, Massachusetts [8]), the direct sound decreases noticeably in intensity at the rear of the hall. It is the author's observation that at about 160 ft on the main floor, the intensity of the direct sound is weak enough to have deteriorated the musical quality by about one category or 10 rating points. In balconies, the situation is different. The sound does not have to travel directly over the heads of as many people to reach an auditor, and, further, additional reflecting surfaces usually come into play. Hence, in balconies, corrections to the main floor ratings are necessary.

The rating scale for intensity of the direct sound on the main floor of a hall for music is shown in Fig. 8. The correction chart for balconies is given in Table 2. It should especially be noted that the rating system cannot be used in any situations where the direct sound does not travel in a customary manner from the stage to the listeners. For example, the method will not rate the quality of the acoustics of seats at the front of a hall and to one side of the proscenium where the listeners are cut off from the direct sound.

Table 2

Corrections to the Rating Scale for Intensity of Direct Sound for Seats in Balconies

Conditions	Incremental Rating Points
Favorable reflections (under 35 msec delay) From a pair of balcony fronts	2
Favorable reflections (under 35 msec delay) From a pair of side walls	4
Highly directive ceiling	4
Maximum possible increment	8

Note: The combined direct sound rating, i.e., the sum of the rating from fig. 8 and the correction above, must not exceed 10.

8. F. R. Johnson, L. L. Beranek, R. B. Newman, R. H. Bolt, and D. L. Klepper, *J. Acoust. Soc. Am., 33* 475–481 (1961)

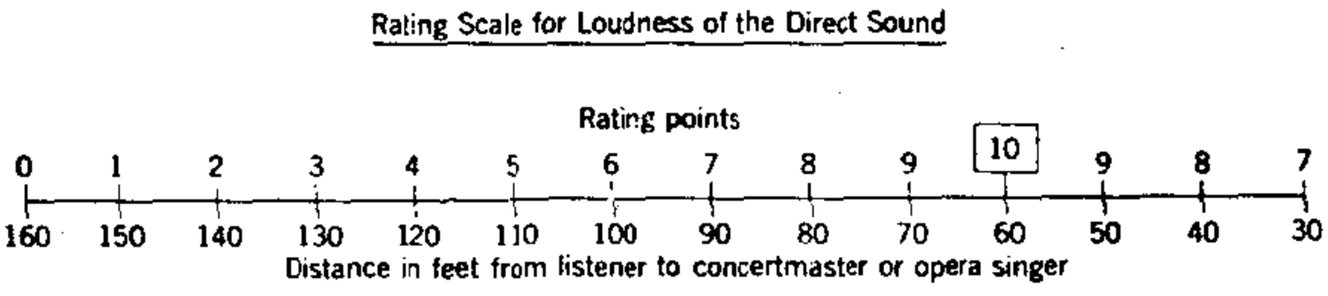


Fig. 8. Rating scale for intensity of the direct sound. This scale is applicable to seats on the main floor only. Balcony seats must be handled by a combination of the scale above and correction numbers from Table 2. The combination, however, must not exceed a total rating of 10 points.

E. *Loudness of the Reverberant Sound*

The reverberant sound is too loud in some halls and too weak in others. Following L. Cremer [9] the loudness is related to,

$$L = (T/V) \times 10^6$$

where T is the mid-frequency reverberation time with full occupancy and V is the volume in cubic feet.

Using 14 halls with acceptable loudness as examples, it is found that optimum loudness, L, should have a value between 2.5 and 3.5. A rating scale, based on the available data, is given in Fig. 9.

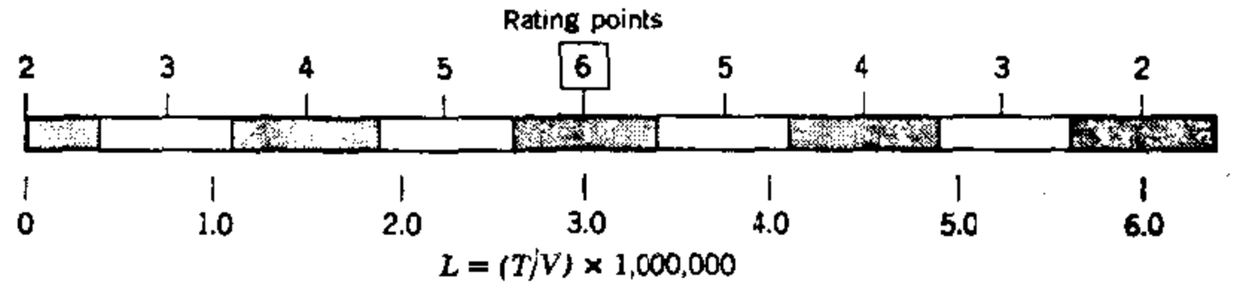


Fig. 9. Rating scale for loudness of the reverberant sound *vs* the quantity $L = (T/V) \times 1{,}000{,}000$, where T is the reverberation time in seconds and V is the volume in cubic feet.

The rating system does not apply to seats that are under balconies or are deep in boxes where the listener is shielded from the reverberant sound in the hall.

F. *Diffusion*

It has been recognized that diffusion is of importance in a concert hall [10]. The listener observes that with adequate diffusion, sound arrives at his ears from all directions. Without diffusion, sound seems to come straight at him from the front of the hall. The present study reveals that except for its strong relation to reverberation time, diffusion *per se* is not a very large factor in the rating of a hall, so that visual observation of the nature of wall and ceiling irregularities is an adequate measure. The rating scale for diffusion in concert halls is given in Fig. 10.

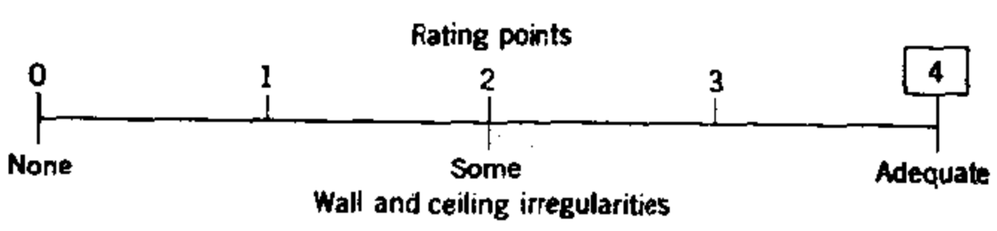


Fig. 10. Rating scale for sound diffusion in concert halls. Since sound diffusion is not important in opera houses, this scale is not used in rating them.

These studies indicate that diffusion is unimportant in opera houses. Hence, it receives no consideration in the rating scheme for them described here.

9. L. Cremer, *Statistische Raumakustik*, S. Hirzel, Stuttgart, p. 231 (1961)
10. E. Meyer, *J. Acoust. Soc. Am. 26*, 630–636, (1954)

G. *Balance and Blend*

By balance in concert halls, we mean the relative loudness of the various sections of the orchestra as heard in the hall. In opera houses, we mean the relative loudness of the singers *vs* the orchestra. By blend we mean the combination of the sound from the sections of the orchestra into a pleasant whole. If the stage or pit for the orchestra is too wide or too deep, time delays destroy the blend for the listeners who sit off the centerline of the hall. No numerical measure of balance and blend has been derived. An experienced observer can judge the quality of the balance and blend from architectural drawings. A musician or other trained listener can hear the quality of the balance and blend immediately.

In concert halls, the attribute of blend and sectional balance is rated a maximum of 6 points. In opera houses, balance and blend between singers and orchestra is so important that this attribute is rated a maximum of 10 points. The rating scales are given in Fig. 11.

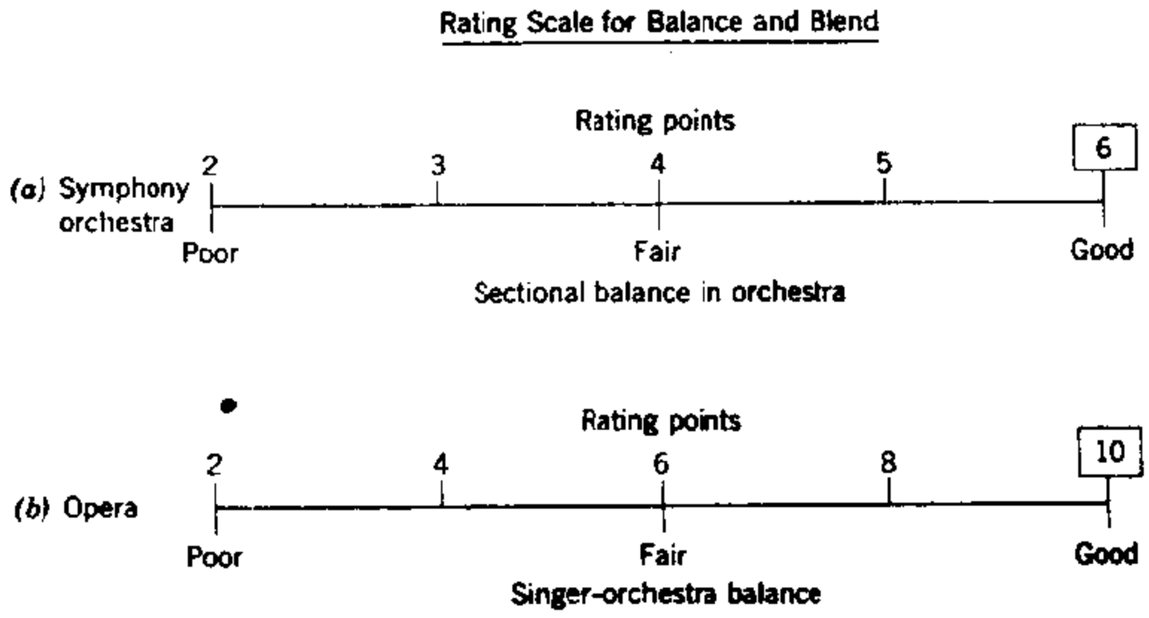


Fig. 11. Rating scales for balance and blend:
a) Between sections of symphony orchestra.
b) Between opera singer and pit orchestra.

H. *Ensemble*

Good ensemble results when the performers play together in perfect unison. Ensemble is partly a matter of the skill of the conductor and the performers and partly a matter of the design of the stage enclosure or the reflecting surfaces to the sides and above the stage. Insofar as the hall, (as distinct from the performers) contributes to ensemble the rating scale is that shown in Fig. 12.

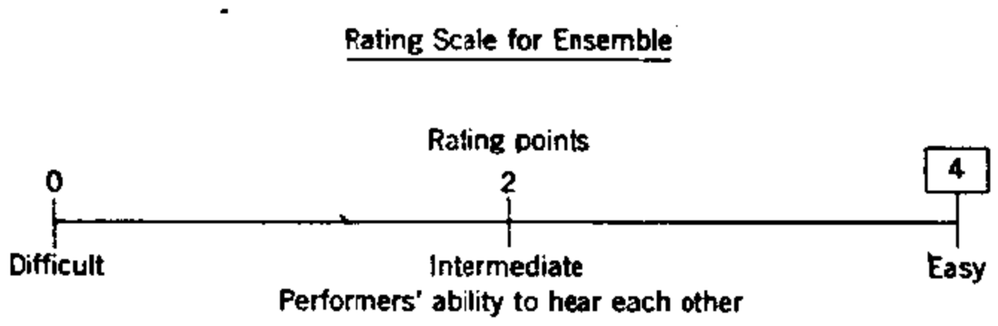


Fig. 12. Rating scale for ensemble, i.e., performers' ability to hear each other.

We still have to take into account acoustical blemishes. These comprise echo, noise, tonal distortion and hall non-uniformity and are handled as deductions. They contribute nothing positive to the quality of the music.

IV. Acoustical Blemishes

A. *Echo*

In this paper, an echo is defined as a reflection delayed by more than 70 milliseconds, sufficiently intense to be annoying. Echo arises from a rear wall whose smooth surface is not

broken up by balconies or diffusion. It is intensified by inadequate ceiling and side wall diffusion and low reverberation time. Without ceiling and side wall irregularities, the first reflection that returns to the front part of the hall from anywhere is from the rear wall and will stand out in the front part of the hall like a black spot on a white surface. The worst potential source of echo is a long curved wall at the rear of a wide fan-shaped hall. Two procedures for control of echos in such halls are published elsewhere. [8, 11]

B. *Noise*

Noise may arise from ventilation openings, machinery, subways, trains, aircraft or traffic. Vibration from machinery of subways may also be troublesome. Criteria for acceptable noise levels in concert halls and opera houses (measured without audience) are shown in Fig 13. Only if economics is a serious factor, should the levels be permitted to approach the upper curve.

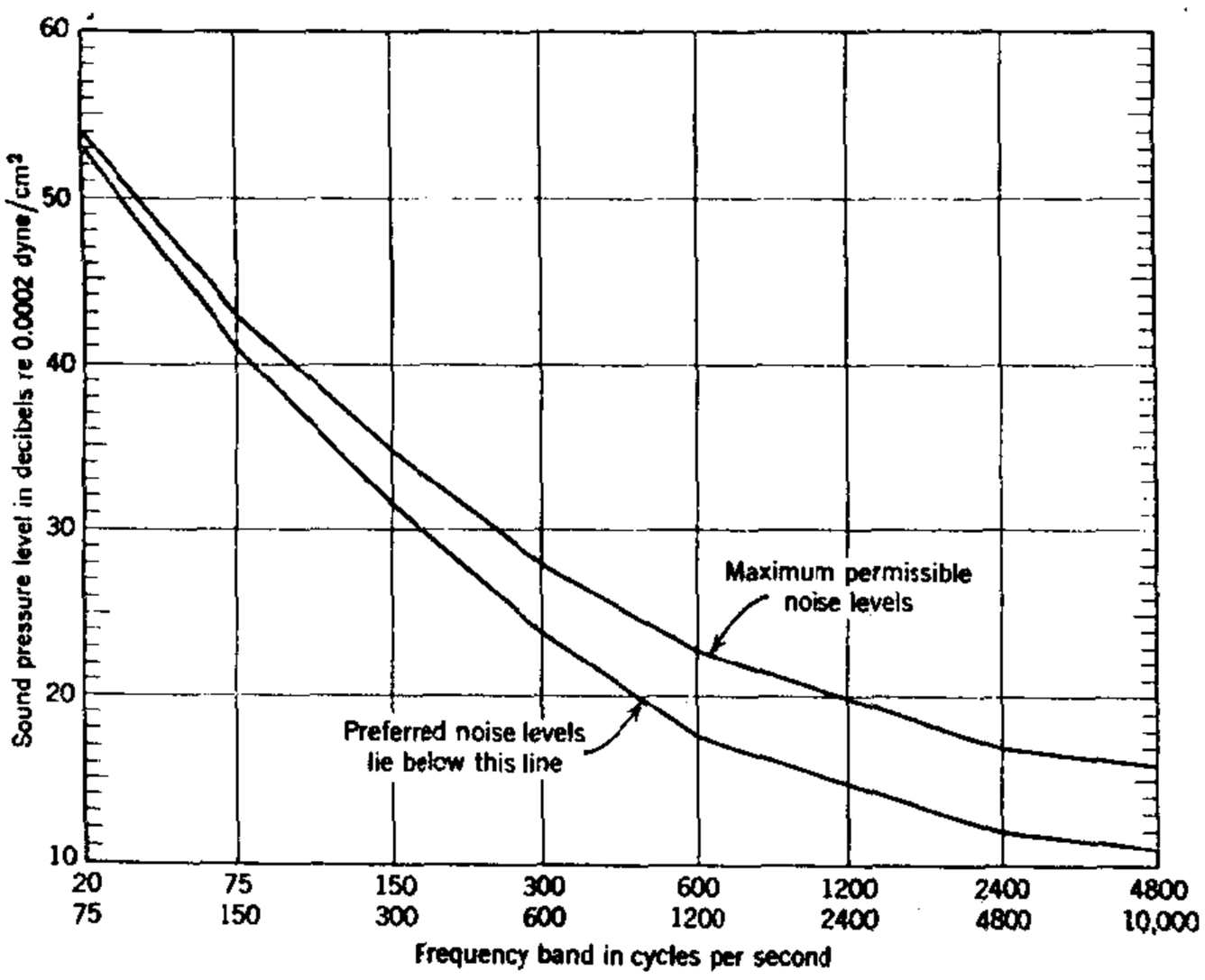


Fig. 13. Criteria for acceptable noise levels in concert halls and opera houses. The lower curve (NC-15) is recommended. Economy considerations should not permit the levels to exceed the NC-20 curve under any circumstances. Measurements are made with an American Standard Sound Level Meter and Octave Band Analyzer. The noise levels are measured without audience.

C. *Tonal Distortion*

Tonal distortion may arise from several sources: (1) selective sound absorption in the hall, e.g., wave coincidence effects in thin, stiff materials, etc.; (2) sympathetic ringing tones, e.g., decorative grilles in front of pipe organs; (3) diffraction grating effects, e.g., from orderly vertical strips on side walls; (4) flutter echo, e.g., as between parallel surfaces; and (5) poor balance among sections of an orchestra due to improper design of stage enclosures.

D. *Hall Nonuniformity*

Nonuniformity of sound in a hall may arise from balcony overhangs that are too deep, boxes with too small openings, better reflections of sound into a balcony than to the main

11. J. B. C. Purcell, NOISE REDUCTION, L. L. Beranek, Editor, McGraw-Hill Book Company, New York, (1960), pp. 408–409

floor, or vice versa, poor sight lines (the interfering heads block off sound), seats near the front of a wide hall and to one side of the stage, and seats too near the stage.

E. *Effect of Blemishes on Rating*

Table 3 gives a suggested means for de-rating a hall if it possesses one or more of the three blemishes shown. Hall nonuniformity is handled by performing the overall calculation for a number of seats and then striking an average. The calculation scheme is not valid for seats under balconies, deep in boxes, or to one side of the proscenium near the front of the hall.

V. Validation of the Rating System

The major part of the job of rating the musical quality of a hall has now been completed. A perfect hall would rate 100 points. A hall that rates under 50 points is unsuited for musical performances.

To validate the rating system, the physical data on each hall were assembled. Interviews had already placed the halls in their respective rating categories, as explained earlier. The calculations of acoustical quality were carried out for two positions in the hall (1) a seat to one side of the center line of the main floor and about half way between the most protruding balcony front (if any) and the stage and (2) a seat in the balcony (if any) and about half way between the front rail and the rear balcony. The ratings for the two seats were averaged to yield the overall rating for the hall.

Table III

Corrections to the Rating Scales for the Negative Attributes, Echo, Noise, and Distortion:

A Separate Correction is Applied for Each Blemish to the Rating.

Amount of Negative Attribute	Correction to Rating Points
None	0
Some	–5
Substantial	–10
Bad	–15 to –50

The best concert hall of the A+ group in this study calculates 96 points on the main floor and 95 points in the balcony. The worst hall in the B group calculates 58 points on the main floor and 71 points in the balcony for an average of 65 points. The one hall in the C+ group is difficulate to rate because it is very non-uniform and very large.

Comparison of the subjective ratings with the calculated ratings showed that 50 of the 54 halls were calculated to fall in the correct one of the five subjective categories, A+, A, B+, B, and C+. The calculations for the four missed their categories by only a small amount. This rating system is probably not the ultimate one although it rank orders existing halls

well. Experience and laboratory experiments will suggest refinements. The author will be satisfied if these studies lead to better concert halls and opera houses immediately and stimulate further research throughout the acoustical world.

VI. Acknowledgment

The author wishes to acknowledge the help of more than two hundred people—scientists in many countries, musicians, orchestra and opera managers, music critics, architects, hall owners, colleagues at Bolt Beranek and Newman Inc. and others who supplied me with data and judgments and criticised my efforts along the way.

Author's Note Added in Galley Proof:

Recent experience shows that the attribute of acoustical *warmth* is effected by the ratio of the bass energy to mid-frequency energy in the sound reflections that arrive within 50 milliseconds of the direct sound, as well as by the reverberant bass ratio defined in III-C. Thus the rating scale of Fig. 7 must be modified in those cases where the early reflections are deficient in bass energy.

Sound Insulation Requirements between Dwellings

by OVE BRANDT

In a number of countries it has, during more than the past two decades, become necessary to introduce acoustic insulation specifications for flatted dwellings. The reasons for this are several. One is that modern flats get poor insulation if such directives are not enforced one way or another. In many countries flats are no longer built the traditional way with thick and heavy floors and walls but instead they are erected by modern prefab methods which usually imply reduced mass and thickness for the sound insulating barriers between the flats. Even then a good insulation may be obtained but only by a very careful planning of the buildings. However, many building designers have little or no acoustic training to solve this problem and it is simply ignored in most cases if no acoustic requirements exist.

It is not necessary to remind the readers that the number and power of acoustic sources in flats have grown tremendously also and thus stress the need for insulation between neighbours.

We do not expect this problem to be taken so seriously in countries where most people live in their own house. But in England where only 5% of dwellings were built as flats between the two great wars, acoustic recommendations were issued during the 1950-ies nevertheless and they seem to be developing into strict requirements in Scotland where a tradition for living in flats exists. Such is also the case in the colder climates of Scandinavia—it is not at all surprising that Sweden where 73% of the dwellings produced are flats (1961) was among the first countries to introduce insulation requirements.

If we do not want our cities to grow enormously we simply have to build flats in place of houses. But people will not want to remain in their flats if we do not solve the sound insulation problem.

For such reasons and others acoustic specifications have now been introduced in at least 13 countries. I shall try to review the international situation within this field.

Do the insulation requirements give us enough protection?

When the first proposals for acoustic requirements were made in Germany in 1938 [1] little was known as to how much insulation is required between two flats. Our theoretical and experimental knowledge was to a great extent limited to laboratory conditions for partitions and floors. It became necessary to estimate what was required.

As to *airborne sound* the choice fell on the insulation equivalent to that provided by a 25 cm plastered brickwall. Thus, the first requirements were expressed as minimum average figures

principally based on laboratory measurements on this brickwall. The frequency range chosen was nearly the same as we have today: 100–3000 Hz. In Scandinavia the same estimation was also made and the same expressions used when requirements were introduced here shortly after the war.

However, the brickwall was often replaced by other types of partitions, very often lightweight double walls in lighter prefabricated buildings. It was then easy to get a very high average figure, especially if it was measured in a laboratory with good craftmanship and no flanking transmission. But, the result in the field as experienced by the tenant was not judged to be equally good. It was thought necessary to express the required insulation not as an average figure for the whole frequency range but as a curve, based on octave or $^1/_3$-octave intervals, a *grading curve*. Thus constructions with a high average insulation based on the insulation curve of the double wall as in fig. 1 would not be permitted. Also the realities of field conditions were taken care of in introducing requirements based on field results and intended for field control.

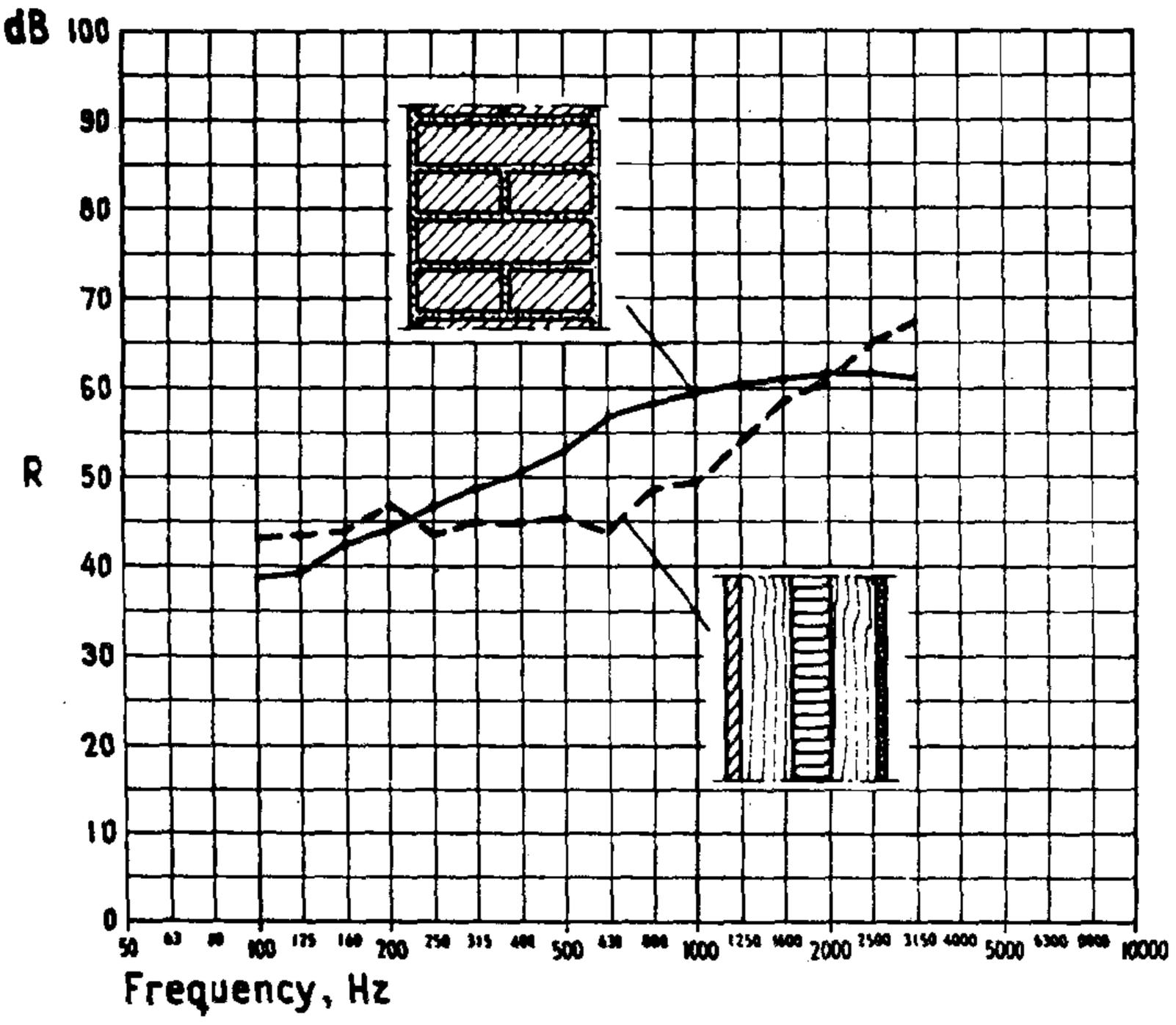


Fig. 1. Airborne sound insulation of a 25 cm plastered brickwall ($\overline{R}$ = 52 dB) and a double wall consisting of two leaves of 8 cm plastered stone and Rockwool ($\overline{R}$ = 50 dB).

In Germany, a new single figure, *the Schallschutzmass*, was proposed to replace the average arithmetical figure [2]. For airborne sound, the figure Luftschallschutzmass (LSM) was based on the proposed grading curves: it is the number of dB's that a measured curve has to be lifted or lowered in order to satisfy the required grading curve. LSM becomes 0 if the requirement is exactly satisfied, has positive and rising figures for accepted insulation curves but negative for insulation below the grading curve. Similar figures were proposed for the impact sound insulation, Trittschallschutzmass (TSM).

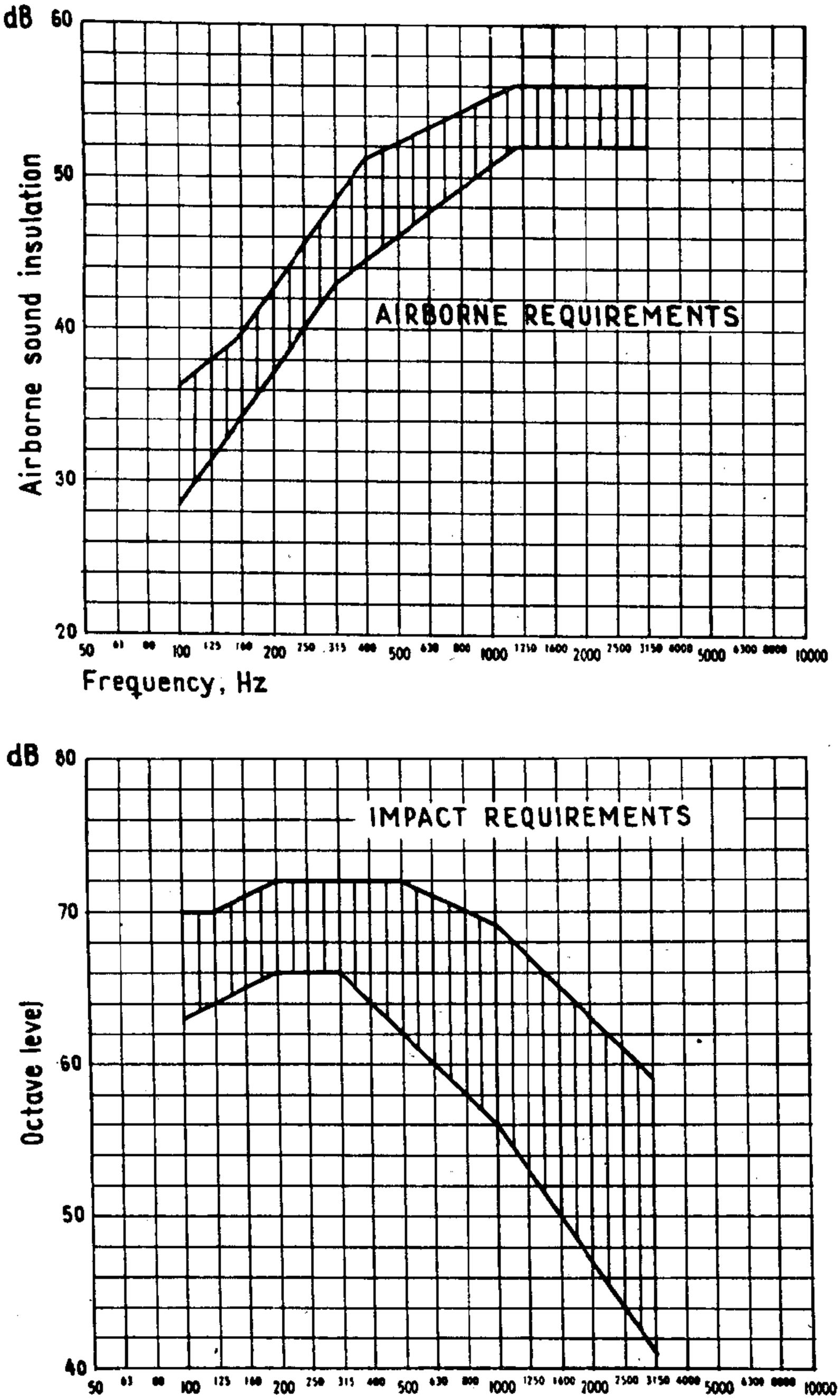


Fig. 2. Present grading curves for airborne sound (A) differ less than the curves for impact sound (B).

Even with these refinements, the background was still the same assumption that the 25 cm brickwall had sufficient insulation. The grading curve, first introduced in Germany after the war, was based on a number of laboratory and field measurements on this type of wall. However, with changing building technique towards prefabs in some countries one might ask why the insulation provided by this brickwall should be a divine answer to the need for acoustic protection as interpreted in the laboratory as well as in the actual buildings in the form as average figure and as minimum curve with the correct value at all frequency bands. We have had a similar development for the requirements on impact sound insulation. However, in this case different countries have apparently not had a common construction to suppose was adequate as with the brickwall for airborne insulation. It seems that in each country a choice has been made between current floor constructions and the better of them have become the standard and this has lead to a much greater spread in requirements for impact insulation compared with airborne insulation, fig. 2. So even more for impact insulation the question may be raised: "Which is the "right" answer for adequate protection against impact noises?"

The direct method to find out an answer to these questions is simply to ask people living in flats what they think about the acoustic insulation against the noise in the other flats and then make an objective measurement of the insulation in order to find out what the answer means in dB-requirements. It sounds very easy, but in fact it is not the easiest way to do it. The need for acoustic insulation may vary much from family to family. Some families produce a lot af sound with radios, TV:s, children and many more sources and do not care much about the noise they may hear from the neighbours in pauses between their own noises, and they may be honestly surprised if they get noise complaints from their neighbours. Some families may be at the other extreme: producing very little sound themselves and thus creating no masking to the neighbours' noises which may upset them very much and perhaps disturb rest and sleep thereby leading to strong complaints about the insulation. Of great importance is also the outside background noise level, with traffic as the main source: a high level leads to masking of the interior noises and thus an impression that the sound insulation is good.

For these and many more reasons it is of no use to make such a survey on a little scale if anything useful shall be concluded. The survey must comprise several hundreds of flats, carefully selected to give a typical picture of the numerous variations in the human reaction and activity and the objective sound insulation. In practice it is not really possible to get enough material to answer all the questions one might like to have answered.

Such social surveys have been carried out in England, Holland, Norway and Sweden [3, 4, 5, 6, 7]. The English surveys shall be briefly reviewed. In a flat survey the material was divided in 3 groups of flats with a difference in floor insulation of roughly 5 dB between each group, but having the same insulation in the horizontal direction. In a similar survey for row houses the material was divided in 2 groups, one having an average airborne insulation between neighbouring houses of 50 dB, the other with an insulation of 55 dB. These dwellings were all chosen amongst local authority houses or flats which, as I understand, means that they are built in an economic way in order that people with a low income can afford to live there. The results are therefore, as pointed out by the investigators, not necessarily valid for other sorts of dwellings with higher rent and standard.

In the *row houses* only the airborne sound insulation in the horizontal direction was measured. The two groups, comprising 250 pairs of houses, each had, as mentioned, an average insulation of 50 and 55 dB, for a single, plastered 25 cm brickwall and for a double wall of two leaves of 11 cm brick and an airspace of 5 cm, respectively. The insulation curves reported from field measurements on these two walls are given on fig. 3. It was found that there was no distinguishable difference in the disturbance in the two groups of houses. As the difference in insulation is found primarily at high frequencies it was concluded that

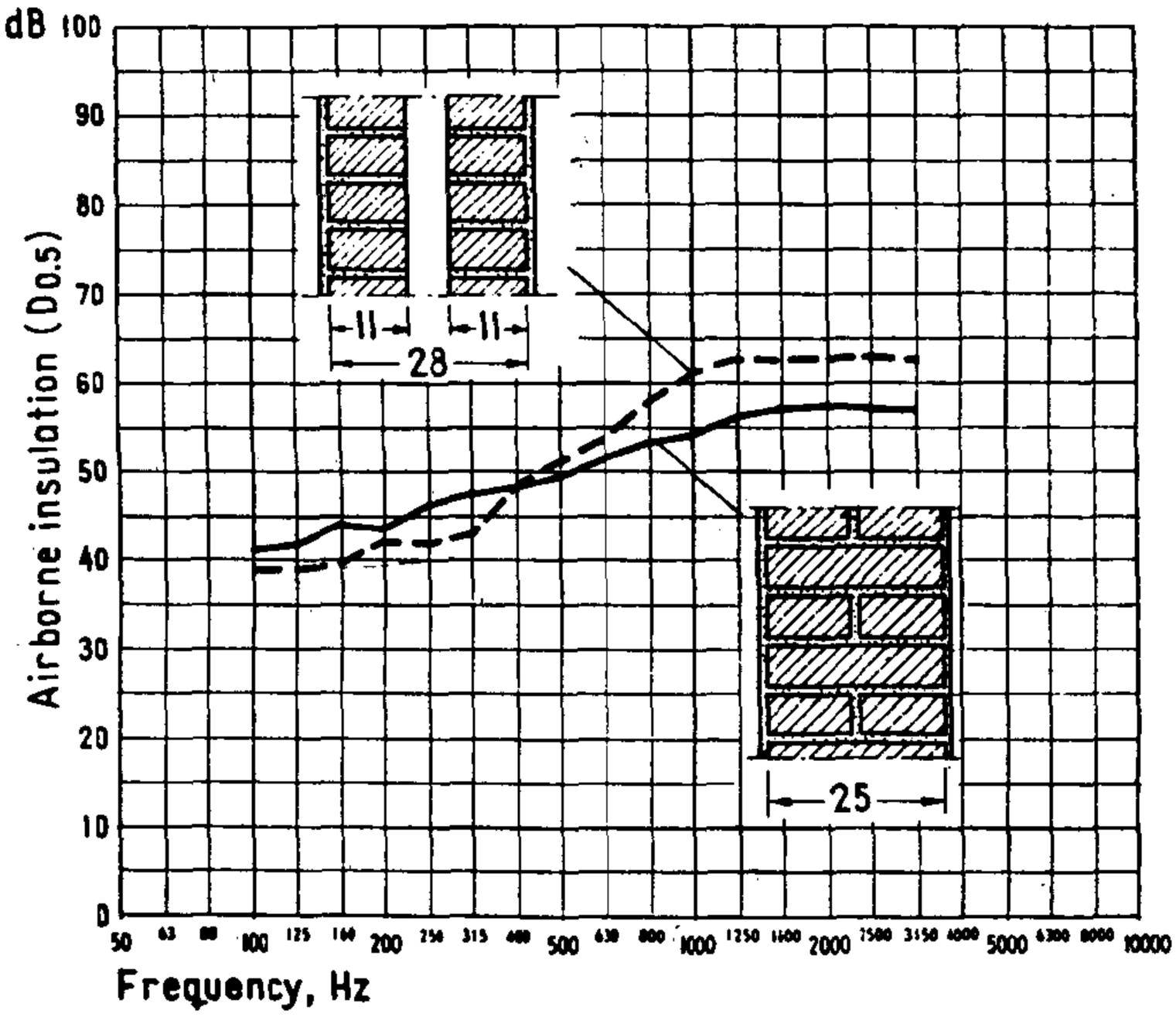


Fig. 3. Airborne insulation ($D_{0.5}$) for the English party walls in the social survey. Average of twenty-one 23 cm solid brick walls and five 27 cm cavity brick walls in houses.

better high-frequency insulation, obtained with a double wall, gives no appreciable advantage for the tenants. This is explained by the fact that it is the low and medium frequencies that are heard through walls as such frequency components dominate in the source which is verified by other investigations.

These results were ready at about the same time as the first grading curves, still based on the insulation of the 25 cm brickwall, were proposed in Germany. As the same type of wall was concluded to be sufficient for row houses in England, even here a grading curve was used based also on the brickwall. The two grading curves do not agree very well as seen from fig. 4.

The English social surveys in *flats* comprised 3 groups of about 1500 flats arranged according to different floor insulations for both airborne and impact sound. As mentioned before the average floor insulation differed 5 dB between each of the 3 groups while the horizontal airborne insulation was equivalent to a 25 cm plastered brickwall, i.e. roughly 50 dB in average. Group I had an average airborne insulation of 49 dB, Group II 44 and Group III 39 dB. Insulation curves for the Group I-III floors are given in fig. 5. The difference be-

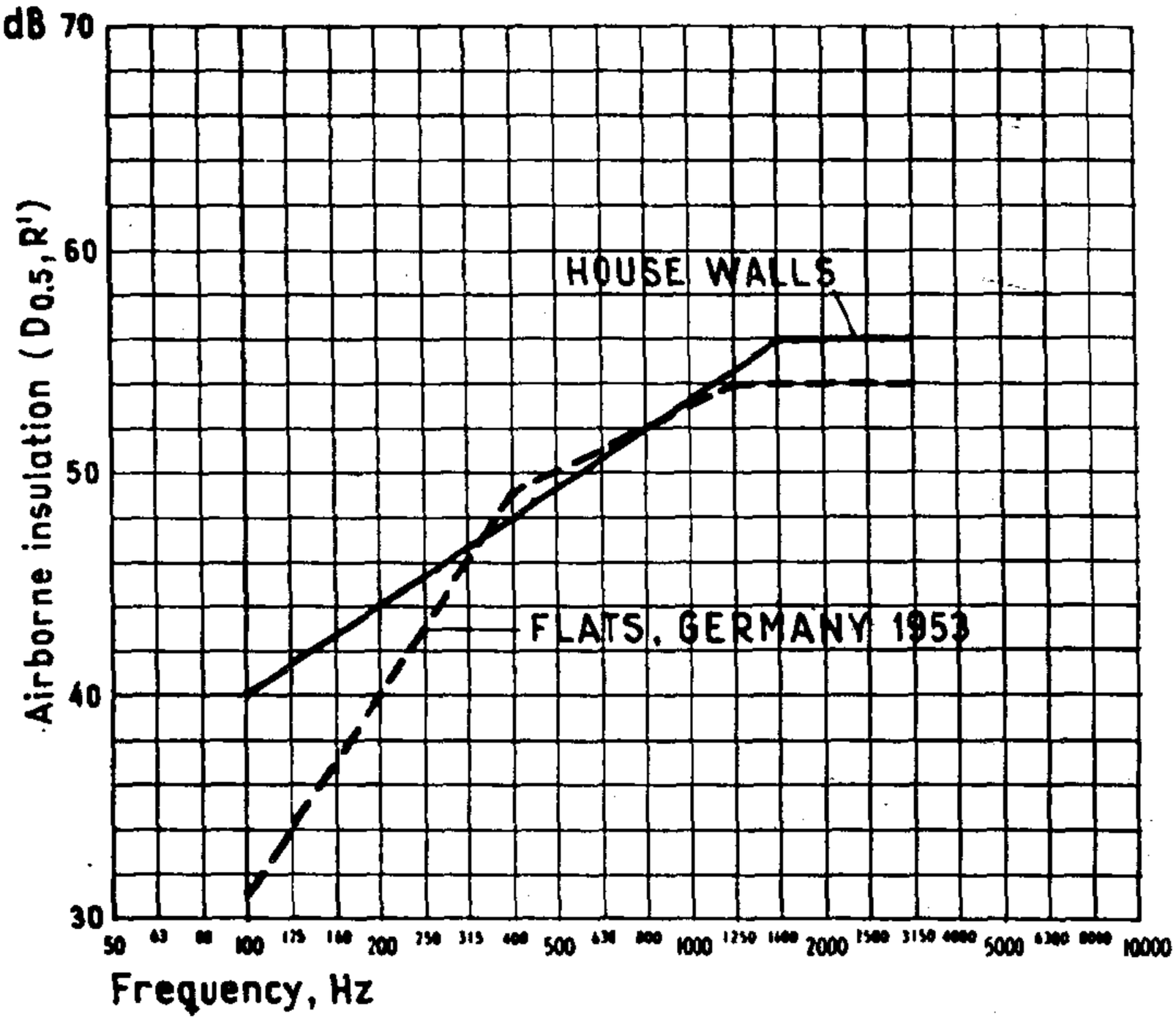


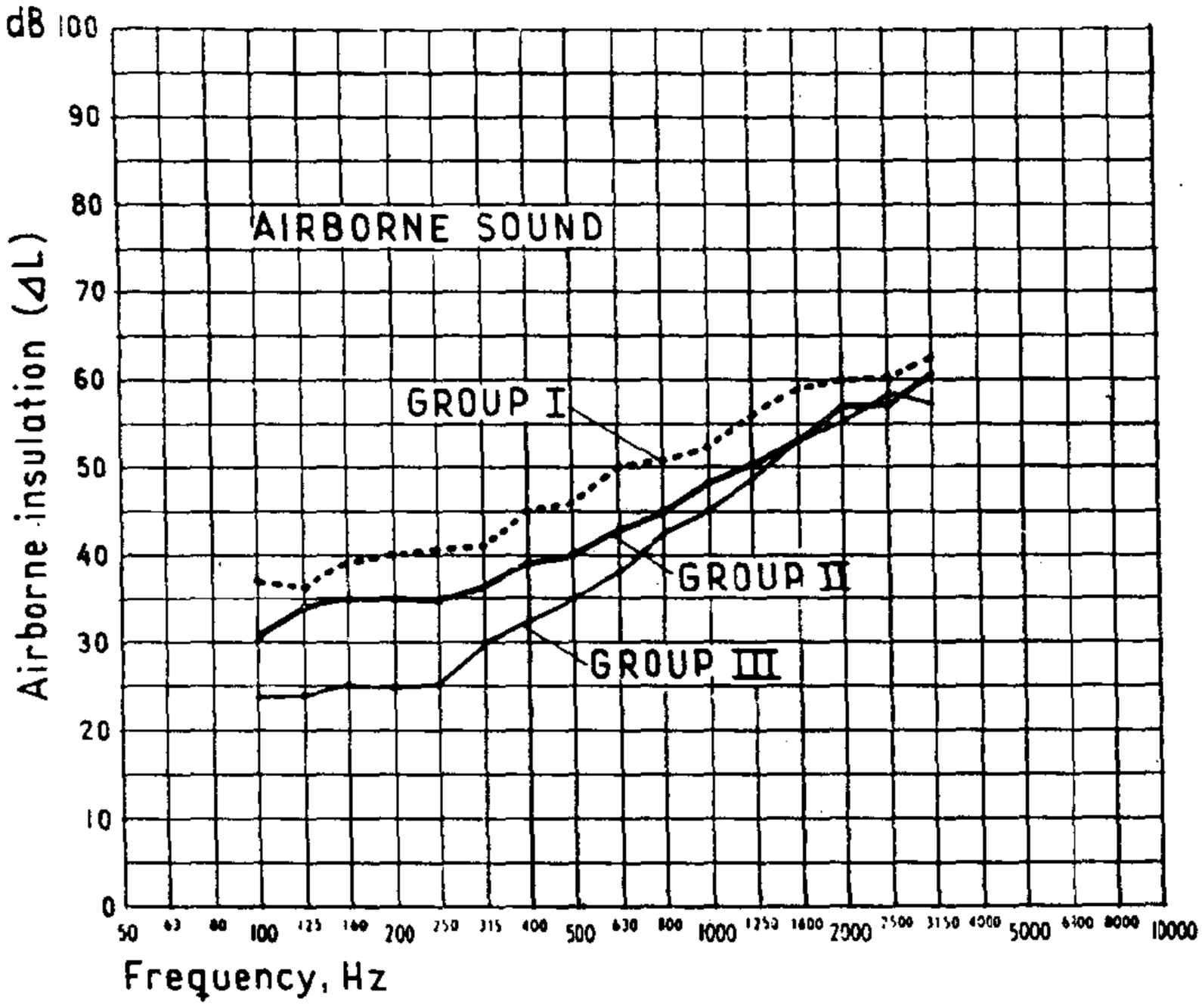
Fig. 4. The English row house recommendation and the original German requirement (1953), both based on the brickwall insulation do not coincide.

tween these insulation Groups is so big that one expects a clear indication of annoyance, least in Group I. The results of the survey did also verify this expectation for the Groups I and II: In the first Group 22% said they were disturbed by the noise, in the second Group the number of disturbed increased to 36%. In Group III this number surprisingly decreased to 21%. This unexpected relative satisfaction with acoustic insulation was explained by the fact that the tenants in Group III previously had had very bad dwellings and still seemed to compare the present improved conditions with their preceding living conditions.

In Group I noise from the neighbouring flats was no more annoying than so much else attached to living in a flat—as mentioned before England is not a country where it is considered a natural thing to live in a flat in place of a traditional house. In Group II flats noise was one of the biggest disturbances. Another measure for these Groups is that in Group I only 7% did not complain of anything, while this figure in Group II increased to 14%, and in the immune Group III these uncomplaining people were no less than 42%. This last Group was not used as a basis for recommendation as its tenants were uncritical in general.

It was concluded from this survey that the insulation obtained with the floors in Group I flats should be used as a minimum recommendation for flatted dwellings, as these tenants apparently equally complained about noise as so much else in the flats. The average insulation curve was somewhat simplified, fig. 6, and was called Grade I.

A grade II was defined as a 6 dB lower curve at all frequencies. It was stated when employing this Grade that the tenants must be expected to find their neighbours noise the worst thing to endure in the flats.



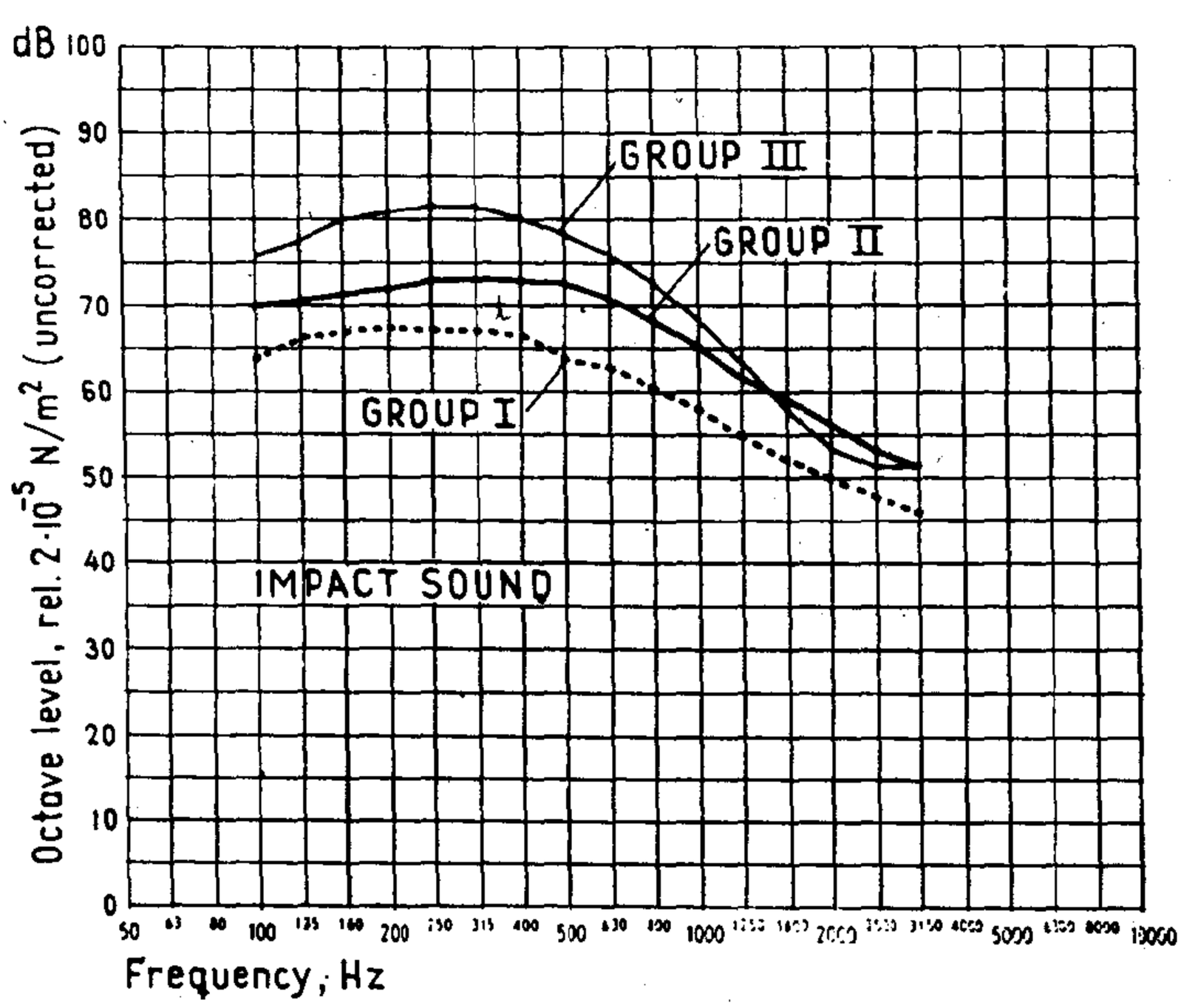
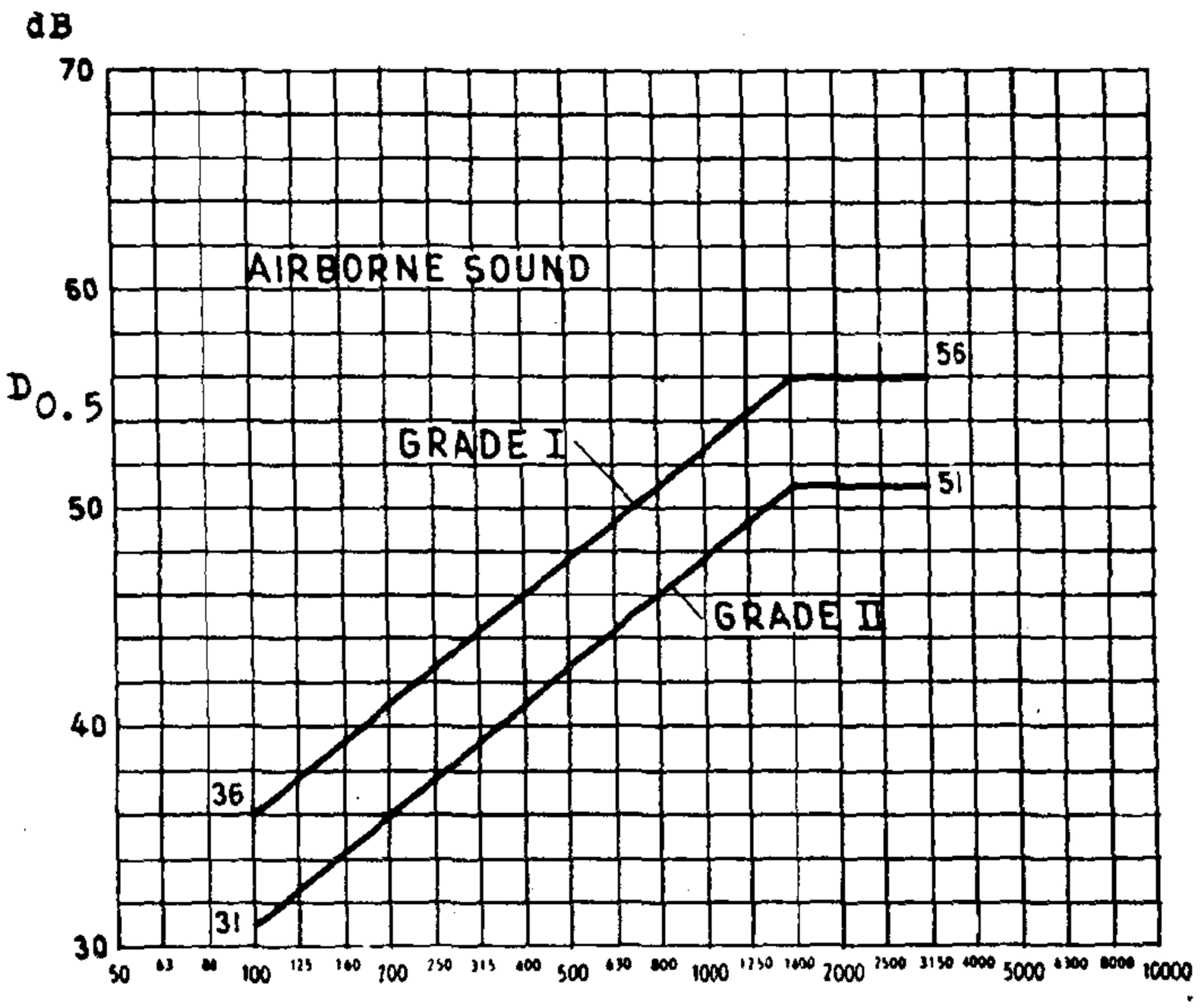


Fig. 15. Measured values for airborne and impact sound in the English flat survey. The values are average figures and uncorrected.



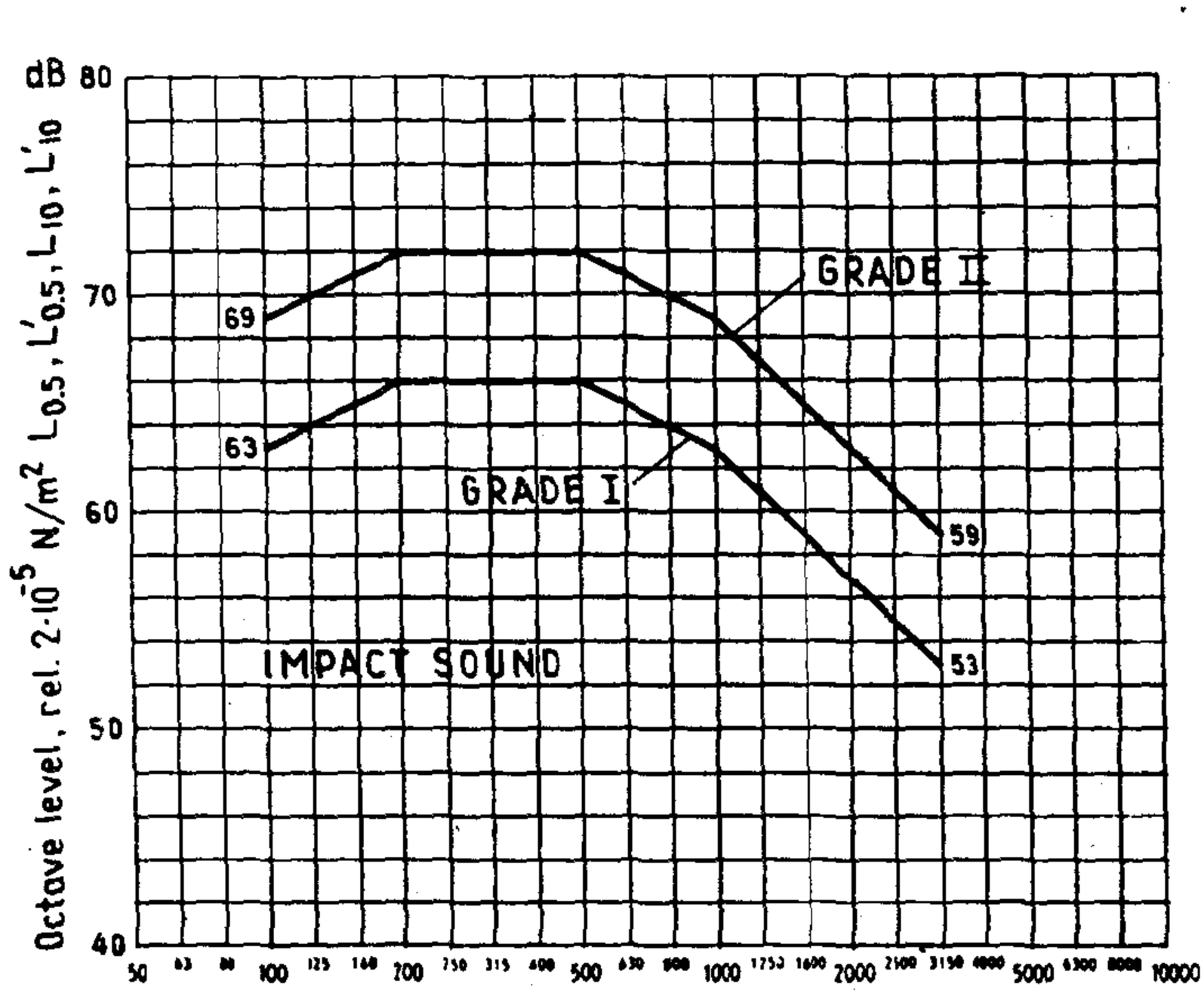


Fig. 6. From the English flat survey the conclusion was to use two grading curves for airborne and impact insulation.

It must be recalled when using Grade I for planning a block of flats that noise then is considered equally bad as draught, dampness, faults in the heating system etc. If we get rid of such shortcomings—which must be quite easy in a modern flat—one must expect that the complaints against the sound insulation increase. Also it should be remembered that this Group of flats was taken amongst local authority flats with, perhaps, relatively uncritical tenants. It must finally be remembered that flats are not the traditional type of dwellings for an Englishman and he may not complain so much because he considers his flat as only a provisional state before finding his definite dwelling in a house. Apparently, the Grade I recommendation cannot be expected to give a very good acoustic protection for the tenants. A few results from the Swedish survey complete this picture. It was carried out in about 500 flats at about the same time independantly of the British surveys. As a criterion for the airborne insulation the average figure in the range 100–3200 Hz was used, which is possible because very few of the walls or floors showed anomalies in the insulation curves as they were heavy, single leaf constructions. It was found that amongst people in flats with an average airborne insulation of about 45 dB 21% were disturbed by the neighbours airborne sounds. For flats with an insulation of 48–50 dB—roughly equivalent to the 25 cm brickwall—16% expressed dissatisfaction with the airborne insulation. At the highest insulation, 50–55 dB, only 7% were disturbed by these sources.

From these surveys we see that a decent protection is gained against airborne noise with the traditional brickwall, but we can hardly expect that this standard of protection is to be considered sufficient when the general standard of flats is raised. This is especially the case in countries where the flatted dwellings tend to dominate and people do not consider a flat as a provisional place to live. Also the noise sources seem to increase in number and power and this increases the need for airborne insulation.

Most specifications for noise protection are now expressed as a grading curve. As stated before a grading curve based on the measured insulation for a 25 cm plastered brickwall is not necessarily the correct answer at all frequencies, even if such an assumption may serve us well for a provisional standard. To find out what is the correct curve is, however, not easy. It can hardly be done with the same sort of social surveys as the ones mentioned, be-

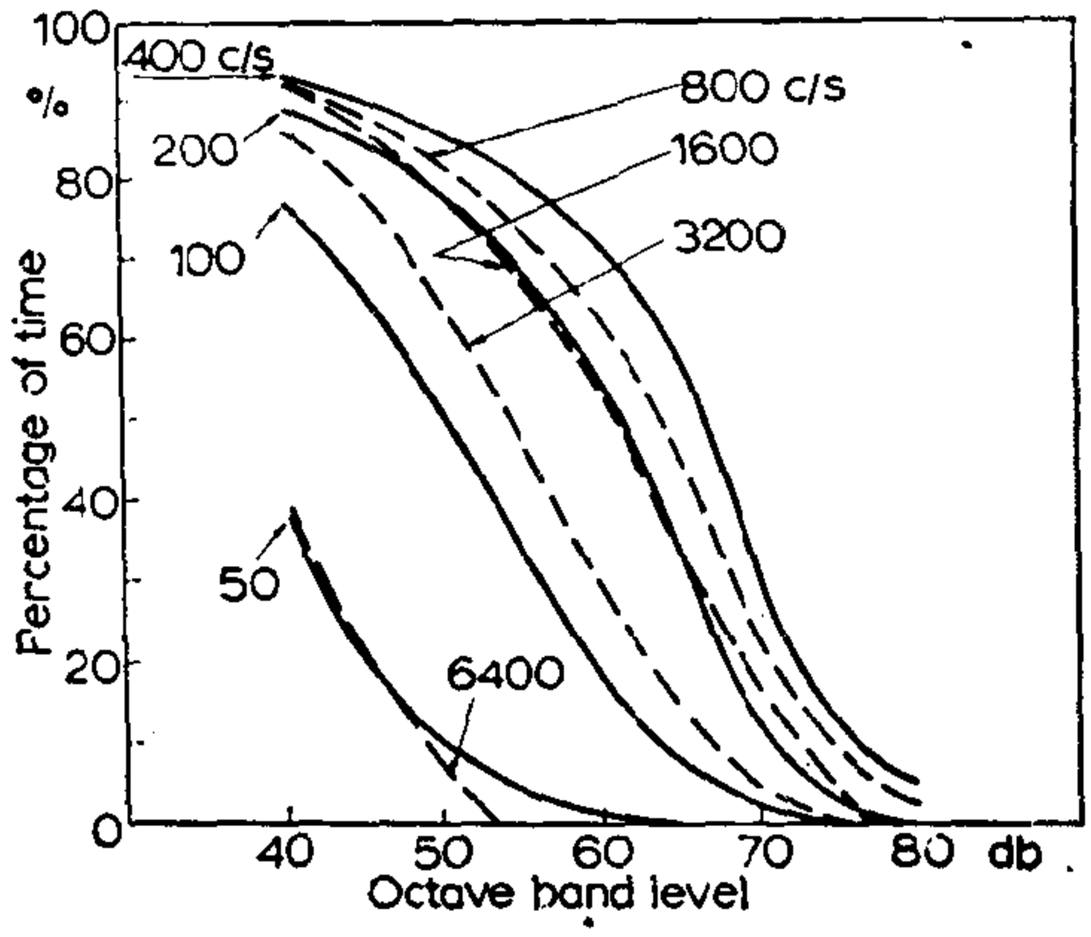


Fig. 7. Statistics of peak levels in radio-programmes. For each of eight octave bands the percentage of the time that certain octave band levels are surpassed is indicated. Based on about 55 hours of mixed radio-programmes (v. den Eijk).

cause we then need a very big material and we should have to ask people about frequency distribution etc. in terms that they are not familiar with. Other methods must be found. One method has been used by v. den Eijk in Holland. [8] He uses the fact that radio and TV-sets are the most annoying noise sources in flats and in order to find out how much insulation is needed he makes field studies on the time and frequency distribution of radio sounds in the source room in dwellings. He presents the results of such studies of 17 mornings and afternoons in fig. 7. Then he requires the level in the receiving room to be 0 phon

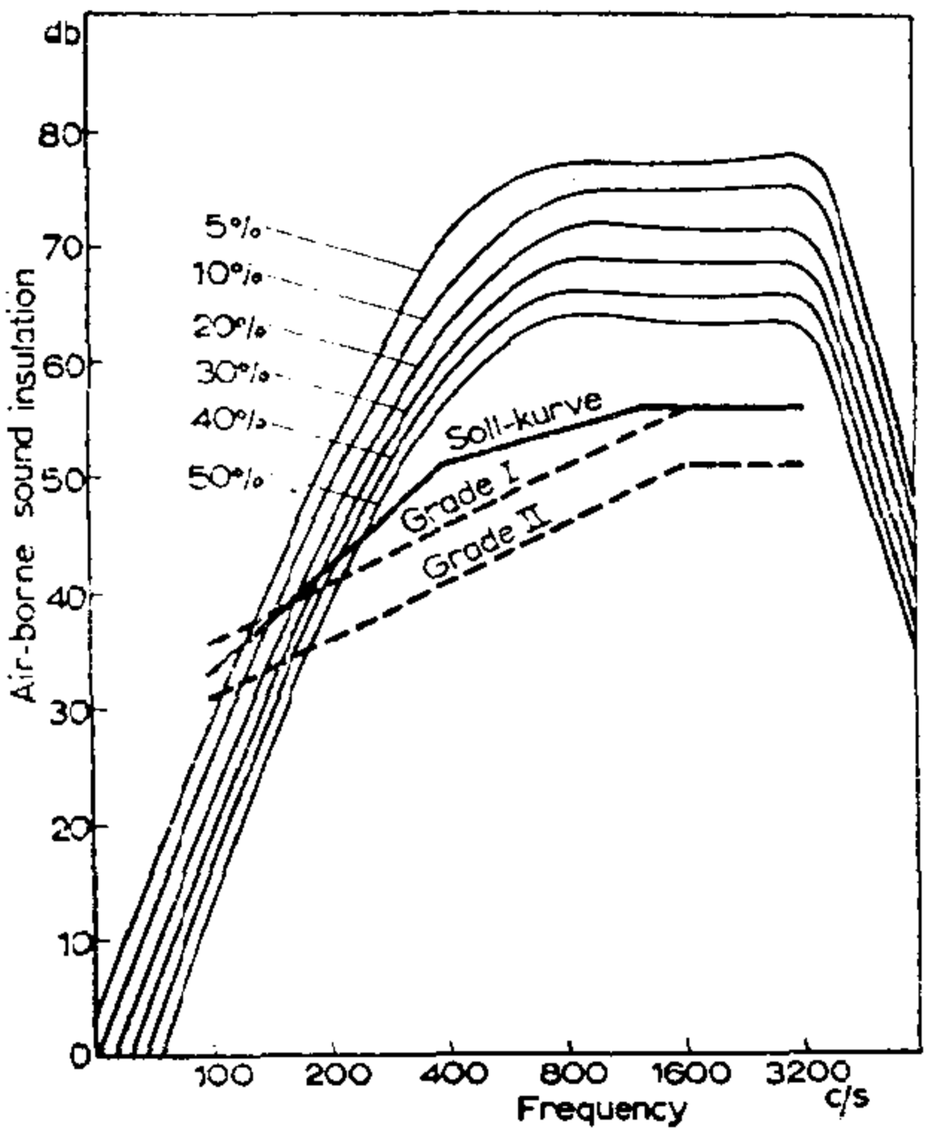


Fig. 8. Required airborne sound insulation based on a disturbing neighbour's radio level surpassing 0 phone during, in the mean, 5, 10, 20, 30, 40 or 50 per cent of the time. For comparison the German (Soll-Kurve) and the Brittish (Grade I and II) requirements for dwellings are added (v. den Eijk).

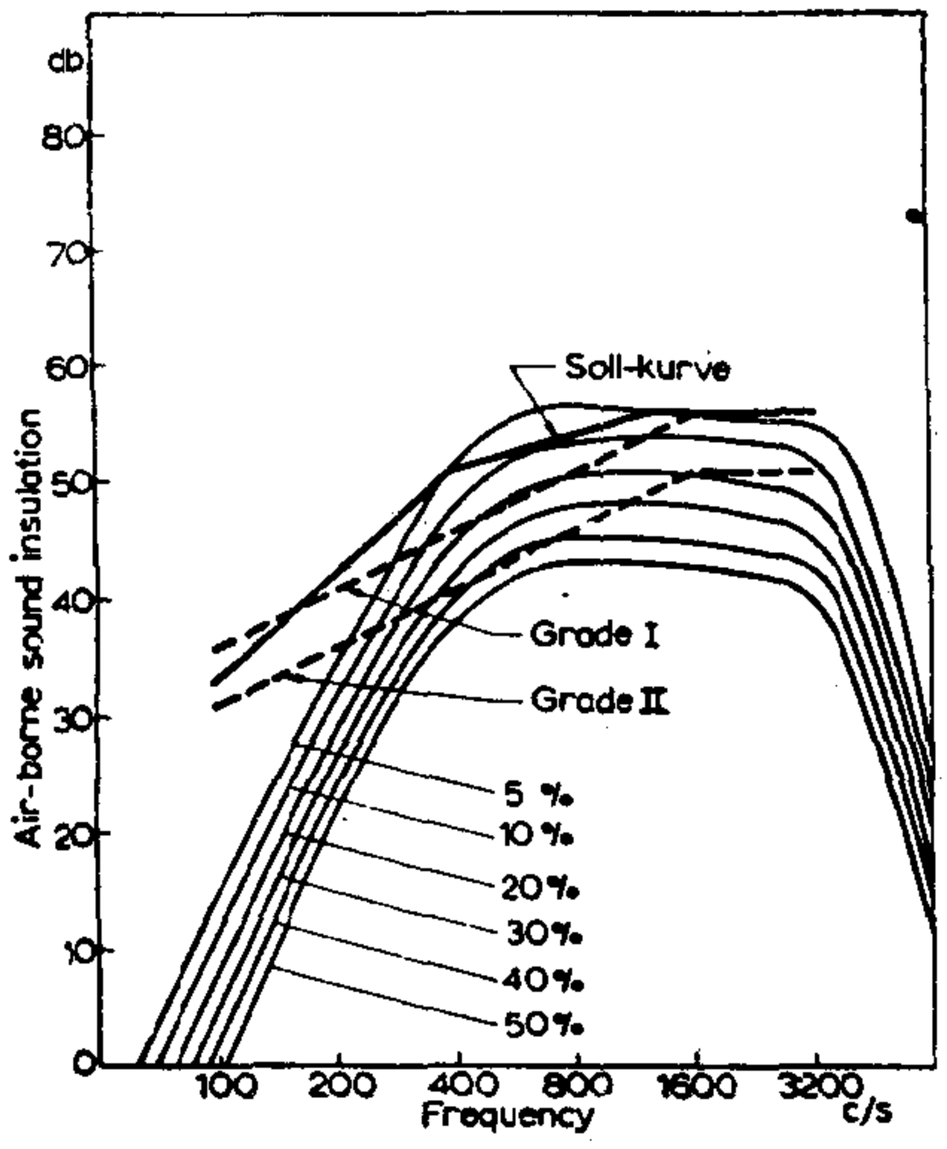


Fig. 9. Required airborne sound insulation based on a disturbing neighbour's radio level surpassing 20 phones during, in the mean, 5, 10, 20, 30, 40 or 50 per cent of the time. For comparison the German (Soll-Kurve) and the Brittish (Grade I and II) requirements for dwellings are added (v. den Eijk).

using the Fletcher-Munson 0-phon contures for pure tones. In this way he can get the shape of required level difference. As this requirement is very high he gets curves that lie very much higher than the present grading curves in Germany and Great Britain, fig. 8. He finds it more realistic to ask for a reduction to the 20 phon-contours. This leads to required level differences which by comparison with the German grading curve can be reached with the traditional brickwall, fig. 9. As normal insulation curves are less steep below 400 Hz and usually increase above this frequency he raises the question if there is any need to have requirements outside the important frequency range 400–800 Hz. Fasold, Germany, gets similar results. [9]

The correct shape of the grading curves have also been studied by Rademacher and Venzke, Germany. [10] They simulate the insulation curves of the walls with electric filters and arrange a receiving room similar to a normal dwelling room in volume and acoustics. The observers enter this room one by one and listen to different complex sounds from loudspeakers, filtered through the "wall" filters, and compare the loudness with a third-octave band of random noise centered around 1000 Hz. The selected source sounds are male and female speech, music and random noise of different band widths—all with little dynamics to make it easier for the observers to compare with the 1000 Hz random noise.
With this technique they demonstrate how different insulation curves influence the loudness of typical sounds in a receiving room. For each type of sound they ask the observers to compare the loudness of the sound filtered through different wall filters. The results of these subjective judgements are then compared with different objective figures such as the average arithmetical insulation and different German Luftschallschutzmass based on a number of grading curves, including the one in use and others proposed in Germany. They find that quite different grading curves can be used as a basis for the Schutzmass without appreciably changing this objective measure compared with the subjective one based on loudness. Even the average figure seems to follow the subjective measure surprisingly well, fig. 10. This fact is further studied and seems to be explained by the phenomena that two frequency ranges with good and bad insulation can compensate each other. This is further studied with the wallfilters as examplified in fig. 11. The higher insulation of K at medium

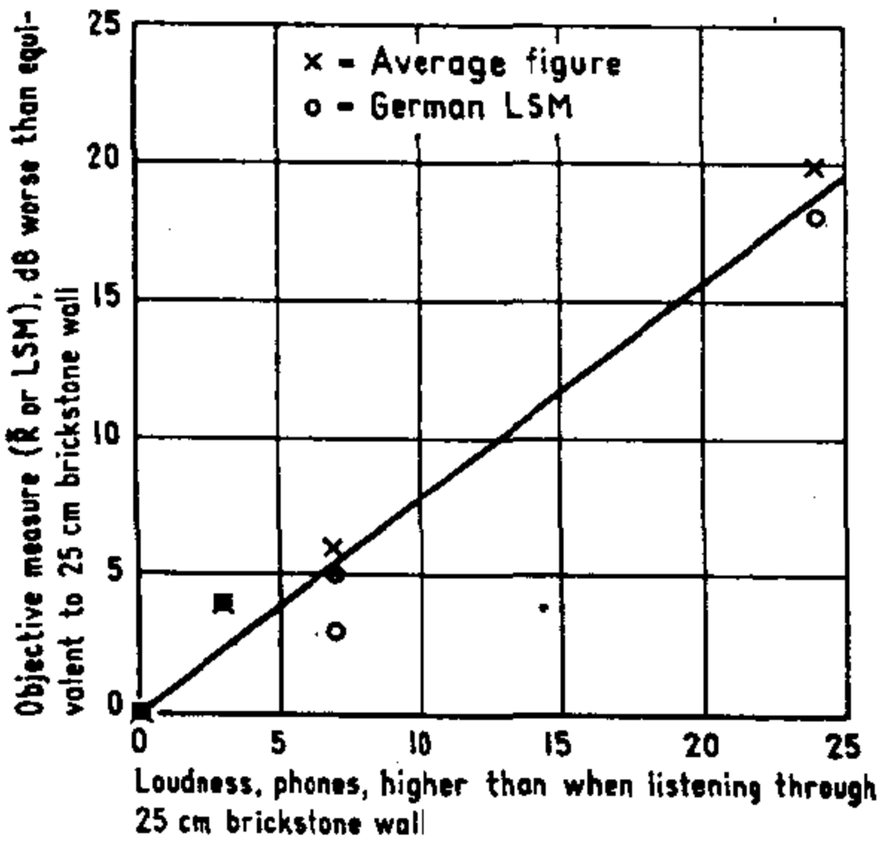
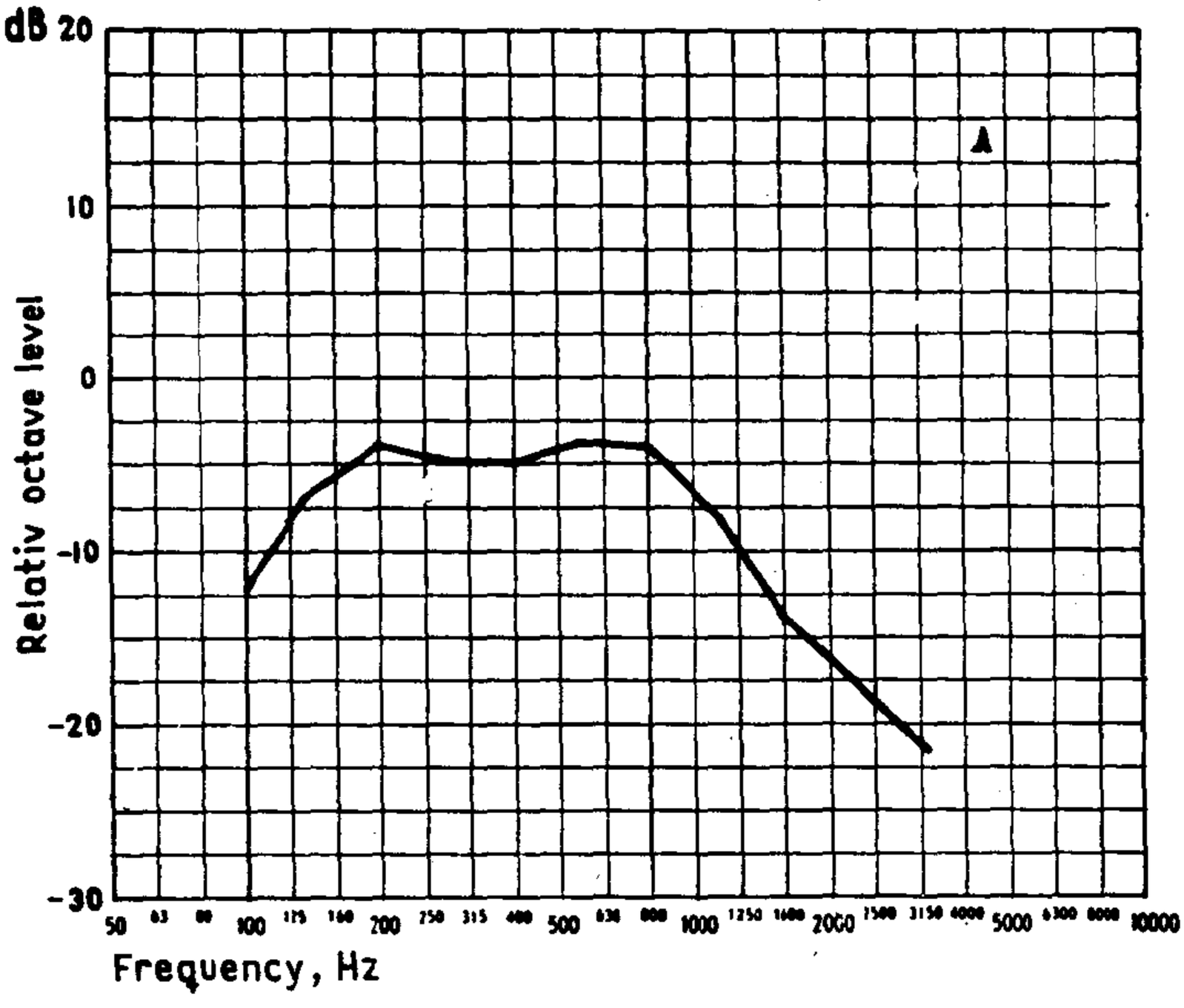


Fig. 10. Example from Rademacher and Ventze's work ([10]). For taped music, listeners have judged the equal loudness (phones) of this sound which they listened to "behind" different walls, evaluated by the average figure $\bar{R}$ or the Luftschallschutzmass LSM. All results are reduced to the case of 25 cm brickwall (0 dB and 0 phone).



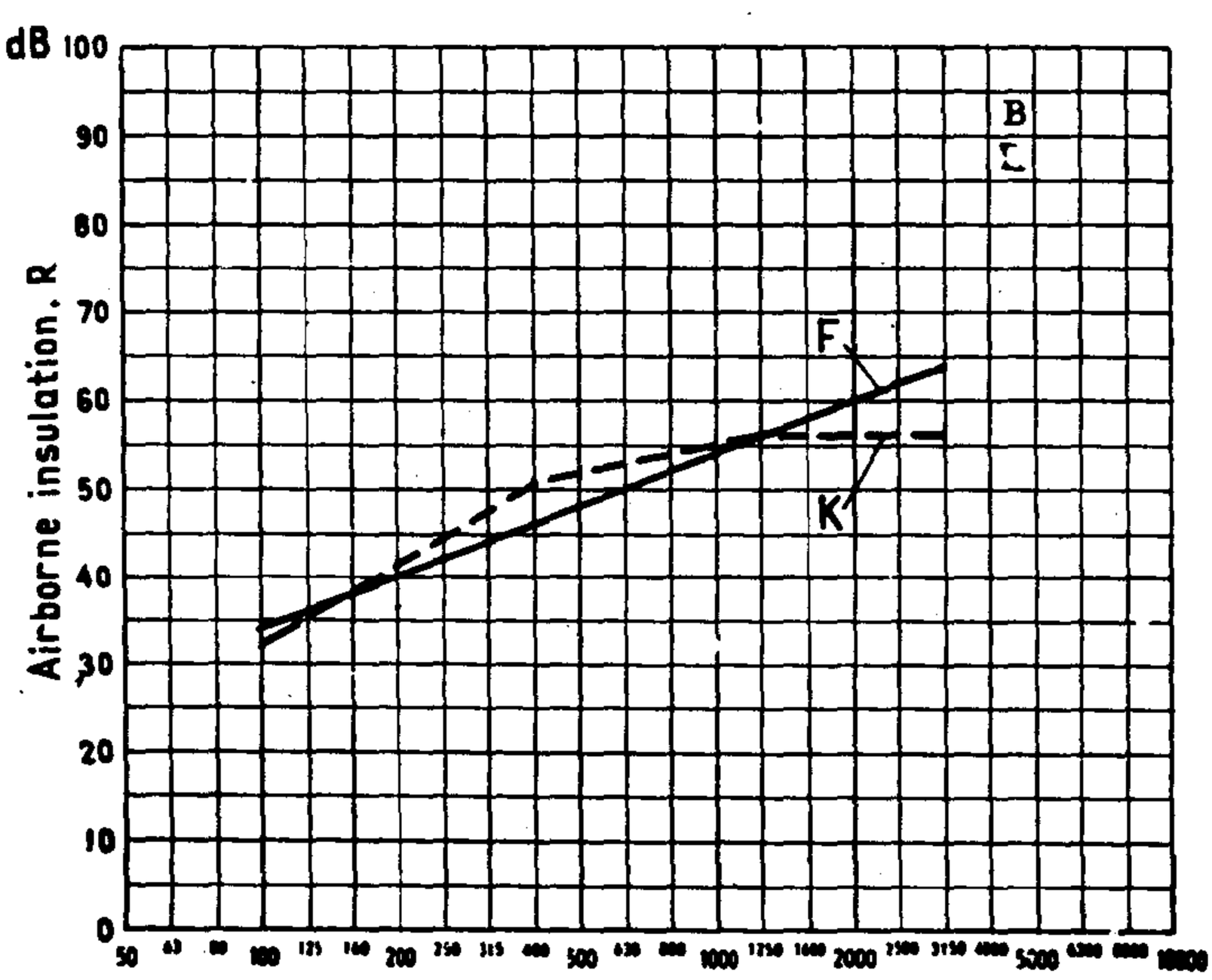


Fig. 11. When the observers listened to "coloured" noise, see A, through the wall filters with the responses F (6 dB/octave) and K, see B, it was judged to be equally loud. The average figure $\bar{R}$ and Luftschallschutzmass LSM are for F: $\bar{R} = 49$ dB, LMS = 0 dB, for K: $\bar{R} = 50$ dB and LMS = 2 dB.

frequencies seems to be compensated by the better insulation of F at frequencies above 1600 Hz so that the two loudnesses are alike. This result is most interesting as the main objections against the classical average figure have been its unrealisticly high values for steep insulation curves. It must be remembered that these results have been obtained according to loudness levels judged at 20–30 dB higher levels in the receiving room compared with what is usually experienced in a dwelling room. When one is exposed to the sound in a building some of the frequency range of the neighbours sound may be masked by the background noise and we do not know the distribution in time and frequency of this masking noise.

That the background noise must be very important for the judgement of the interior insulation is demonstrated for instance in the Swedish social survey mentioned above. Here the flats were put into 3 groups according to the exposure to outdoor noise—the noise was characterized as 1) high level, 2) normal town level and 3) low or very low noise level. The tenants who said they were annoyed by the outdoor noise were 19, 13 and 6% for the Groups 1)–3) respectively. When they were asked about the annoyance caused by noise from other flats the percentage disturbed were 26, 42 and 50% for the same 3 groups 1)–3), a very clear indication of the influence of the outdoor noise on the subjective experience of indoor insulation.

As to *impact sound insulation* our knowledge is so far quite limited. From the English surveys in flats we could draw some conclusions which lead to Grade I and II with similar remarks as for airborne sound insulation. It was also concluded that the light wooden floors had not sufficient impact insulation, even if Group III was not aware of insulations defects. As a matter of fact, in England it was recommended to use floating, concrete floors in order to satisfy Grade I, even if usually a floating floor well done should give more insulation than the required curve. From the Dutch survey we can conclude that the light floors and especially the light wooden floors are not usually sufficient for impact insulation. Finally the Swedish survey indicates that impact sounds do not seem to be a big problem if we use solid concrete floors. For tenants with floors without a separate screeding course only 7% were disturbed by impact sounds. This percentage fell to 2% for floors with a floating course on a mineral wool mat. Remembering that in the same survey the percentage of people who were annoyed by airborne sounds was 16—when airborne insulation requirements were just satisfied—one must conclude that impact sound insulation is not a big problem if the floors are not expecially light as e.g. wooden floors. This is perhaps also the explanation why grading curves for impact insulation in different countries vary so much. It thus seems that the present requirements give us a moderate protection against the neighbours' noise, at least for airborne noise: probably some more insulation is required, especially at the low and medium frequencies, but investigations made on the frequency response have used loudness and not annoyance as a subjective criterion for sound insulation. Further masking has not been considered. We have little evidence about how closely the present grading curves must be followed before the tenants are aware of such a change. Grading curves can hardly be taken as more than a rough indication as to what sort of insulation curves we want. It is probably too early to establish new single figures based on such grading curves as they may have to be changed as new research results appear.

How is „sound insulation" defined?

As mentioned before the first insulation specifications grew out of studies in traditional transmission laboratories where only the direct sound reduction factor for a test panel is measured. For such tests we have a very reliable measuring method which we have agreed upon in the International Standardization Organization. [15] We determine the airborne sound reduction factor R in measuring the level difference ΔL between two neighbouring rooms divided by the test panel of area S the absorption A in the receiving room and thus get R from the formula:

$$R = \Delta L - 10 \log A/S \text{ dB}$$

This formula is valid if all sound in the receiving room is transmitted through the test panel. Also, diffuse fields are required in the rooms. Such conditions are not difficult to satisfy in a stationary laboratory. But we want to make the specifications in building codes valid also for the field. If we could only test or check in the laboratories rules would be of little value and certainly not gain much respect in practice.

But can we expect to have enough diffuse sound fields in normal dwelling rooms, furnished or unfurnished to make sensitive measurements? Can we use the same relationship between level difference and the reduction factor as is used in the laboratory according to the formula above? Or do we have more practical relationships to base our requirements on?

As a matter of fact, it is easier to make reliable measurements in dwelling rooms than one might expect. Of course we do have some troubles at very low frequencies when the room dimensions are of about the same size as the wavelength. Usually not more in a furnished room than in a smaller laboratory as we get some diffusion from the furniture. At higher frequencies we expect to get difficulties as the porous damping of the higher frequencies tend to make the field look like a free field in place of a diffuse field. Gösele [12] in Germany has, however, shown, that we do measure one or two dB higher levels in the pressure field in the receiving room, but if we correct to a constant absorption we get too low values for the absorption determined from the Sabine formula and from short reverberation times, which compensates for the error in the level measurements. He showed that by changing the reverberation time in the receiving room from less than 0.5 seconds to more than 3 seconds the corrected impact sound level changed less than 2 dB at the individual frequencies for the same floor.

In one sense there is a great difference between the laboratory and field conditions: we cannot guarantee that the sound in the receiving room has arrived only through the partition or the floor in the building. Rather it is so, that a good deal is transmitted through flanking elements, *flanking transmission*. Of course, we can still use the same formula above, but then we must include the flanking transmitted sound in the reduction factor (which is then nominated R') if we still take S as the area of the common surface for source and receiving room. This method is used with success in e.g. the German requirements and its advantage lies in its simplicity for the building designers as we shall see.

In some other building codes the level difference is used as a measure for sound insulation in a dwelling, but this magnitude must be normalized in one way or another. If we only used the level difference in a requirement, sound insulation would depend on the acoustics of the receiving room. If we increase the amount of absorption we get an apparent increase

of the sound insulation observed in the diffuse field of the receiving room. We then have the possibility to correct to a certain time of reverberation or to an absorption of the dwelling room. What is to be preferred?

In order to answer this question some reverberation measurements have been made in e.g. England and Denmark. [13] It might be expected that the reverberation time increases with the room volume as we know is the case for classical concert rooms. This is also the case for unfurnished rooms and for rooms with little furniture, but not for furnished rooms. For living-rooms Larris found that the reverberation time varied only between 0.35 and 0.7 seconds with an average value around 0.5 seconds when the room volume varied from 19 to 118 m^3, fig. 12. For the same furnished rooms the absorption computed from Sabine's formula varied from 6.5 to 38 m^2. This is explained by the fact that the principal absorption

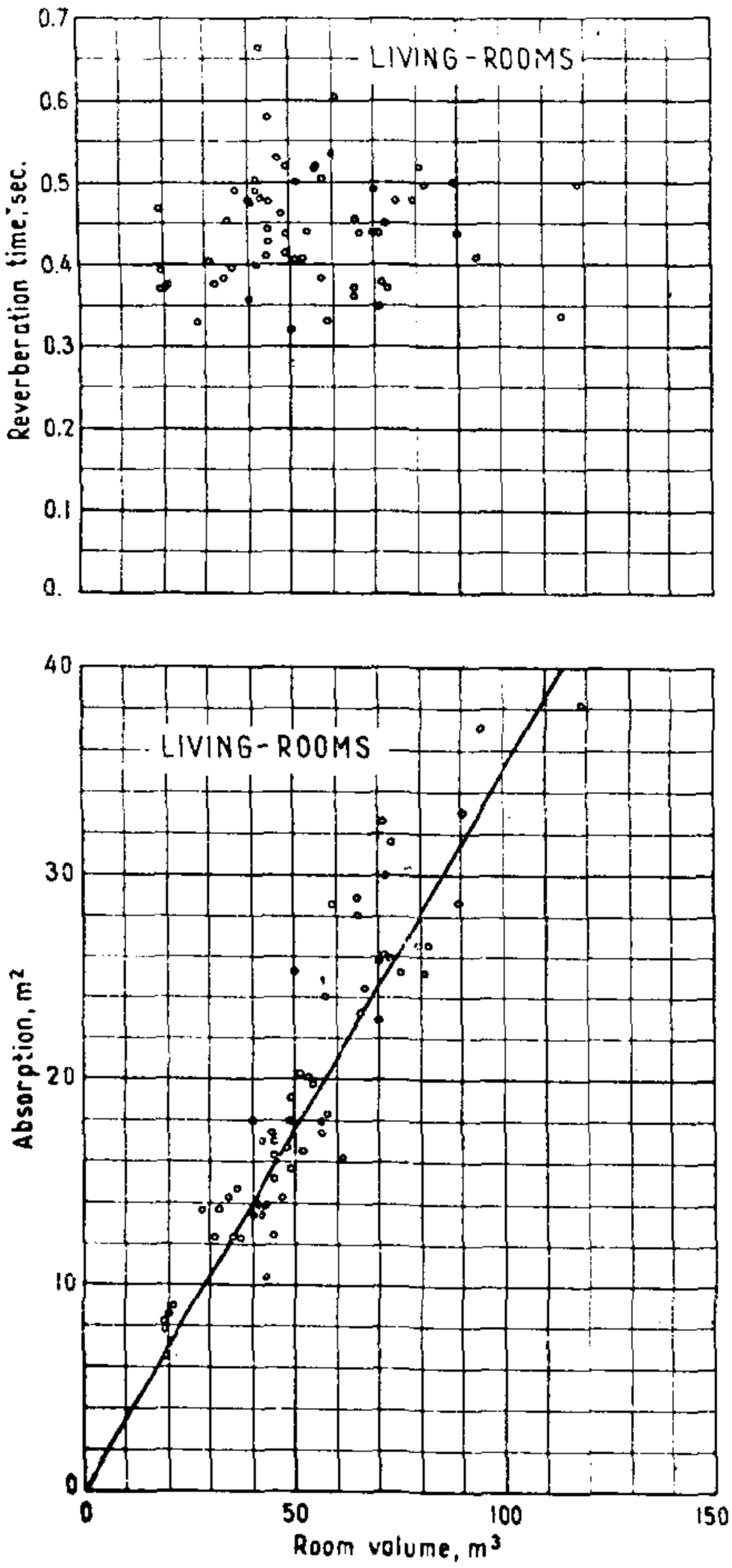


Fig. 12. Reverberation time and absorption in furnished living-rooms. Average values for 125—4000 Hz (Larris).

in living-rooms such as stuffed furniture and mats is connected with the floor. When the floor area increases with the volume the absorption must also increase and thus it is easy to show for a rather constant density of furniture the reverberation time must be quite constant. This is less true in bed-rooms where the total furniture is more constant, fig. 13. The frequency dependance has a peak in the mean frequency range, as the low frequency absorption is procured by panel absorbents and the high frequency absorption by porous absorbents.

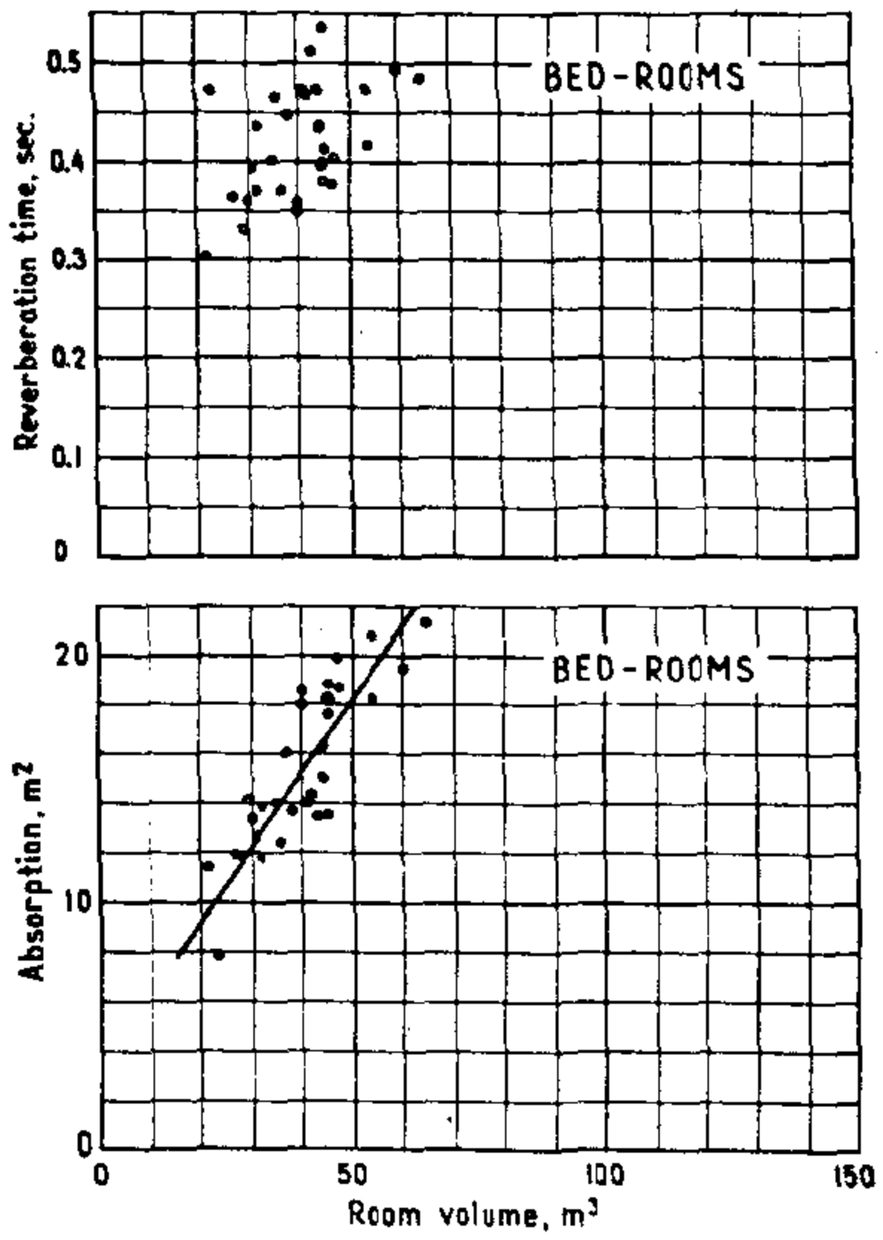


Fig. 13. Reverberation time and absorption in furnished bed-rooms. Average values for 125—4000 Hz (Larris).

If we state that a dwelling room has a reverberation time of 0.5 seconds we must have in mind that this is primarily so for living-rooms, less for bed-rooms—which in some countries tend to disappear in smaller flats—and it is not the case for rooms like kitchens, bath-rooms, halls and similar rooms with little or no furniture where we expect the reverberation time to increase with the volume.

The fact that the reverberation time in a furnished living-room is nearly constant independant of volume has lead some countries to use the level difference normalized to the reverberation time of 0.5 seconds as a basis for insulation specifications. Thus this required level difference $D_{0.5}$ is defined as:

$$D_{0.5} = \Delta L + 10 \log T/0.5 \text{ dB}$$

In this way the required level difference and also the measured one is a true picture of the observed level difference when having a living-room as a receiving room, a very important practical case in flats.

This normalized level difference is then a result of the reduction factor R' of the common

surface S between two neighbouring rooms and the flanking transmitted sound from other surfaces. This is quite easy to understand for building planners without acoustic training, but it is in practice not always so easy to evaluate, not even when flanking transmission can be neglected. The fact is that $D_{0.5}$ also depends on the volume V of the receiving room which we see in expressing $D_{0.5}$ as a function of R':

$$D_{0.5} = R' + 10 \log \left(\frac{0.32 \cdot V}{S} \right) \text{ dB}$$

It will be noticed that this measure is not reciprocal if used between two rooms with different volumes; the building designer may suspect that sound insulation of a structure is not reciprocal. So the direction of the measured level difference must be stated in the reports. If we choose to normalize to a constant absorption we do not get this drawback. This measure D_{10} which has been standardized by ISO for field measurements is then defined as:

$$D_{10} = \Delta L - 10 \log A/10 \text{ dB}$$

thus normalizing the level difference to a reference absorption in the receiving room of 10 m^2. If we express this measure by R' and the common surface of two neighbouring rooms we get:

$$D_{10} = R' + 10 \log 10/S \text{ dB}$$

leaving to the building planner to make his calculations based upon the insulation R' with or without flanking transmission and the size S of the transmitting element.

Of course also this definition has its drawbacks. For instance, for big rooms separated by big surface areas this correction gives a false picture of the real insulation when the rooms are normally furnished. We correct then to a much smaller absorption and neglect that the real absorption is bigger. When we use the same value of D_{10} for all room sizes in dwellings —which we must for the sake of simplicity—the requirements tend to be too rigourous for big rooms and perhaps too mild for small rooms. The trend should of course be in the opposite direction.

Both $D_{0.5}$ and D_{10} are of course a little difficult to handle for the architect or builder with little acoustic training. To simplify this planning it may be better to specify permitted partitions, floors etc. in the building codes, completely omitting acoustic specifications. The drawback of this method is that it may put a brake on the development of building constructions and in many cases it is difficult to give information of all the permitted combinations of e.g. partitions and joining elements. What is usually preferred is *both* an acoustic requirement somehow in dB *and* a number of examples demonstrating how to satisfy the requirements.

Some countries have like Germany preferred to simplify the specifications and also the planning by using the same reduction factor as in the laboratory, here nominated by R'. The planner then need pay no attention to variations in wall surface or room volume, but can simply refer to measuring reports from the identical constructions, even combined with the right joining constructions. Then the requirements must be adjusted to cover even big surfaces. One of the only drawbacks of this principle is that it cannot be used for cases where a common surface S between two rooms are not defined, e.g. between a living-room

and a staircase. It may also be a bit disturbing to attribute all defects of for instance a bad outer wall to the common surface S. In Germany laboratories have been built to measure R' for rooms with flanking walls but still referring to a constant area of the partition, here much better insulated than the flanking construction.
The three existing definitions on airborne sound insulation, R', D_{10} and $D_{0\cdot5}$, may lead to quite different results when the same figures are required as some examples will show.
If we require D_{10} and R' to be equally big for the same case the wall surface must be 10 m^2.
If we even demand D_{10} to equalize $D_{0.5}$ for horizontal insulation the volume of the receiving room must be 31.3 m^3. For a room height of 2.5 m we thus get a standard receiving room with the dimensions $4 \times 3.1 \times 2.5$ m^3, which is quite a normal room in a modern flat. But quite big deviations from these dimensions may occur.
If we look at quite a big room with the floor size of 8.0×3.1 m^2 and standard height of 2.5 m, we get the following differences (vertical insulation):

$$D_{0.5} - R' = 6 \text{ dB}$$
$$D_{10} - R' = 3 \text{ dB}$$
$$D_{10} - D_{0.5} = 9 \text{ dB}$$

If we turn to small rooms the differences are usually not quite as big. A minimum standard floor for a Scandinavian bed-room is about 2.1×3.3 m^2. With the room height of 2.5 m and vertical transmission we get:

$$D_{0.5} - R' = 0.25 \text{ dB}$$
$$D_{10} - R' = 2.8 \text{ dB}$$
$$D_{0.5} - D_{10} = 2.5 \text{ dB}$$

For *impact sound transmission* we have luckily only two alternatives for definitions. One of these is to refer the measured level in the receiving room to 0.5 seconds for the same reason as for airborne insulation. This leads to the definition:

$$L_{0.5} = L + 10 \log 0.5/T \text{ dB}$$

Unlike $D_{0.5}$ we have no such drawback as lack of reciprocity because the direction of transmission is given.
The other alternative which is recommended by ISO for field and laboratory measurements is to correct to 10 m^2 of absorption:

$$L_{10} = L + 10 \log A/10 \text{ dB}$$

Both these alternatives have the drawback that we get a higher figure for decreasing insulation, but this disadvantage does not seem to bother building planners so much as they apparently quickly get used to it.
Obviously, we get cases when these two definitions give different figures, even if the difference is not so big as for the measures for airborne sound. Still we get the same figure if the room volume is 31.3 m^3 which means a floor size of 12.5 m^2 for the room height of 2.5 m. A great majority of modern flats have floor sizes of this order. If the floor increases to 25 m^2 the difference is 3 dB. A small room has the size of about 6 m^2 which still gives us a difference of about 3 dB.

It is easy to show that a correction to a constant absorption is the same thing as to assume a constant power from the ceiling independently of its size. Thus we should get the same results for the same floor construction even if we measure on different floor areas. This seems to be the case for floor sizes in the range from 6–25 m^2 according to German [14] (Gösele) and Swedish measurements. Thus L_{10} would seem to be a good physical magnitude, but with corrections not fitted for the normal acoustics in living-rooms as for D_{10}. We can also show that the correction to a constant time of reverberation as for $D_{0.5}$ is the same as to assume a radiated power from the ceiling growing with its surface. This measure then has the advantage to follow the variation in room volume as is done in furnished rooms but it has as mentioned its physical disadvantages.
Obviously, the three definitions for airborne sound insulations and the two for impact sound have its advantages and disadvantages and it is a matter of taste which is to be preferred. However, it should be a step forward if we could agree internationally on this subject in order to reduce confusion.

Insulation requirements or recommendations in different countries

In the preceding sections we have looked a little at the present background and terminology for insulation requirement. Let us now look at some of the principles used in different countries for acoustic specifications. A detailed report is being prepared by ISO.
In some countries such *specifications* are presented as *requirements*, in others as *recommendations*. There may be little difference in practice. The recommendations *may* have much stronger power than strict requirements which may be only writting table products completely ignored by building designers. The advantage with recommendations is that the real acoustic claim may be expressed without too much compromize with other factors from the very start. An example of this is the British Grade I recommendation for impact noise which is based on floating floors. In Austria a 5 dB higher Luftschallschutzmass (based on the German Sollkurve) is recommended. Germany gives us a good example with requirements which work well and many stationary and mobile labs are available to control the results in practice. In such a case the specifications must be somewhat milder and roughly be intended to cut off the extremely bad cases. The danger in this system is that the standards must be compromised and consequently are only partly sufficient in the majority of cases. Building planners may easily get the impression that all is well if they build just to satisfy the requirements. In fact, it might be better to have a minimum requirement combined with an uncompromised recommendation but this leads to complicated specifications without the simplicity which must characterize rules for building planners with little acoustic training.
Today at least 13 countries have insulation specifications for dwellings. In the great majority grading curves are used to express the minimum values. For airborne sound 10 countries use one of the curves presented in fig. 14 and 15.
To evaluate a measured curve in relation to a grading curve many countries follow the German system of computing the average negative deviations in the whole frequency range and setting positive deviations equal to 0. In Germany this average deviation must not exceed 2.0 dB, based on third-octave frequencies, while e.g. USSR, Bulgaria and

Czeckoslovakia base this average deviation on octave frequencies and add the rule that no single negative deviation may exceed 8 dB. In Great Britain and Scandinavia this procedure is somewhat simplified as only the sum of negative deviations is computed and not permitted to exceed 16 dB.

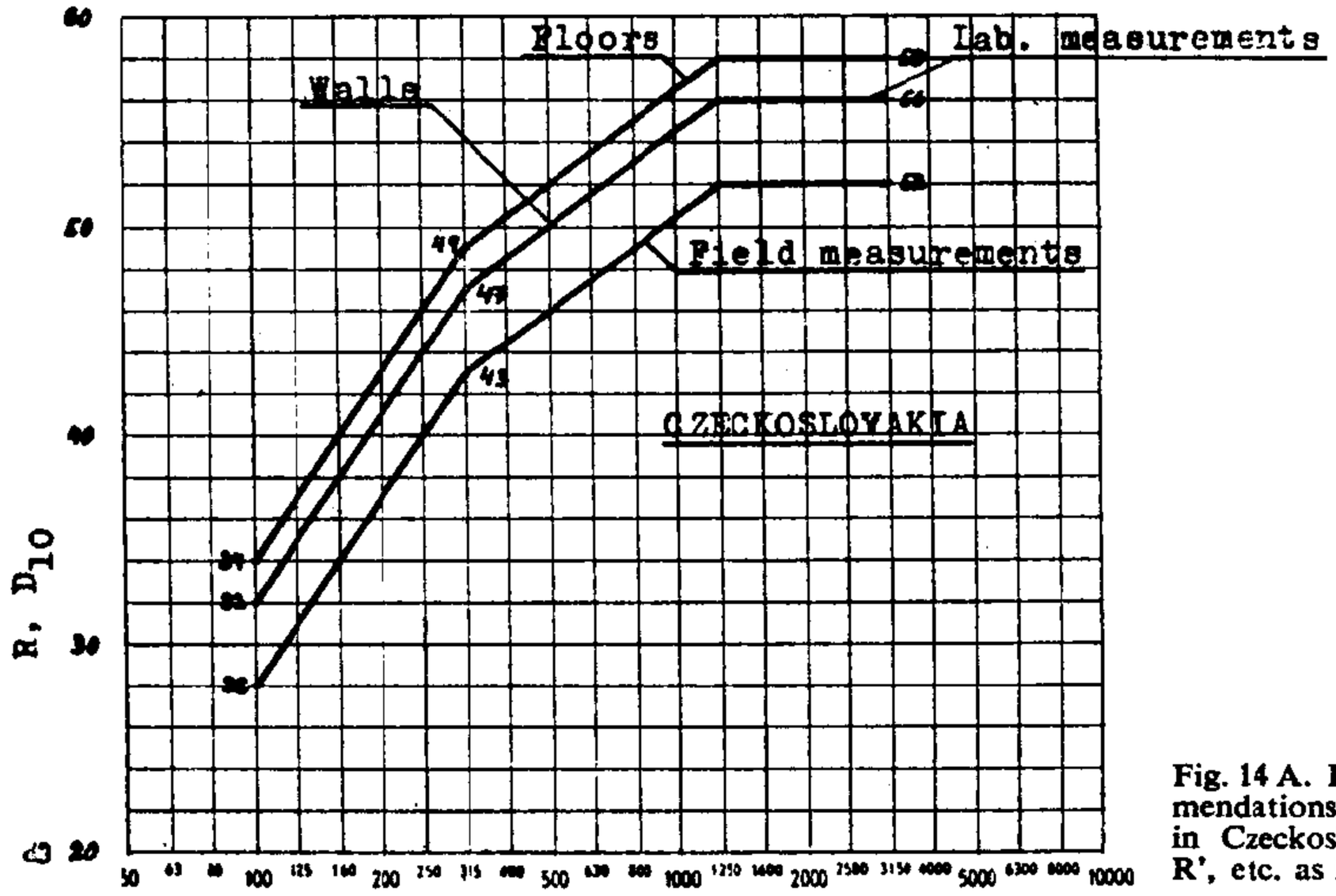


Fig. 14 A. Requirements and recommendations for airborne insulation in Czeckoslovakia. Symbols for R, R', etc. as in the preceeding section.

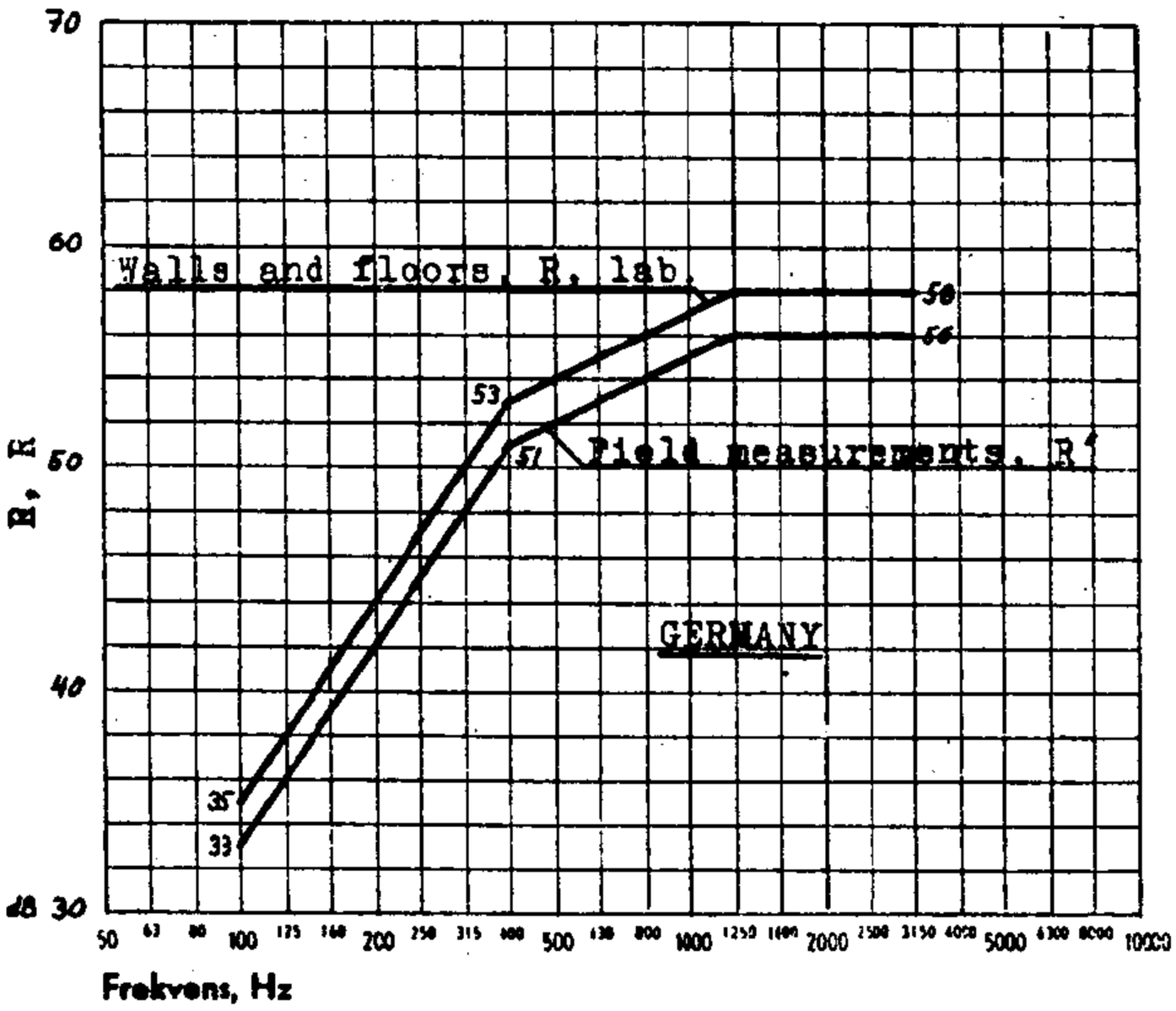


Fig. 14 B. Requirements and recommendations for airborne insulation in Germany. Symbols for R, R', etc. as in the preceeding section.

The present grading curves for impact insulation are presented in fig. 16 and the measured impact insulation should result in a curve *below* the grading curve. We have similar rules as for airborne insulation to decide on cases where part of the measured curve lies above the grading curve. The same 10 countries that have grading curves for airborne sound have such curves for impact sound transmission.

In Canada which was one of the first countries to introduce insulation specifications for airborne sound the average minimum figure of 45 dB has been recommended for the

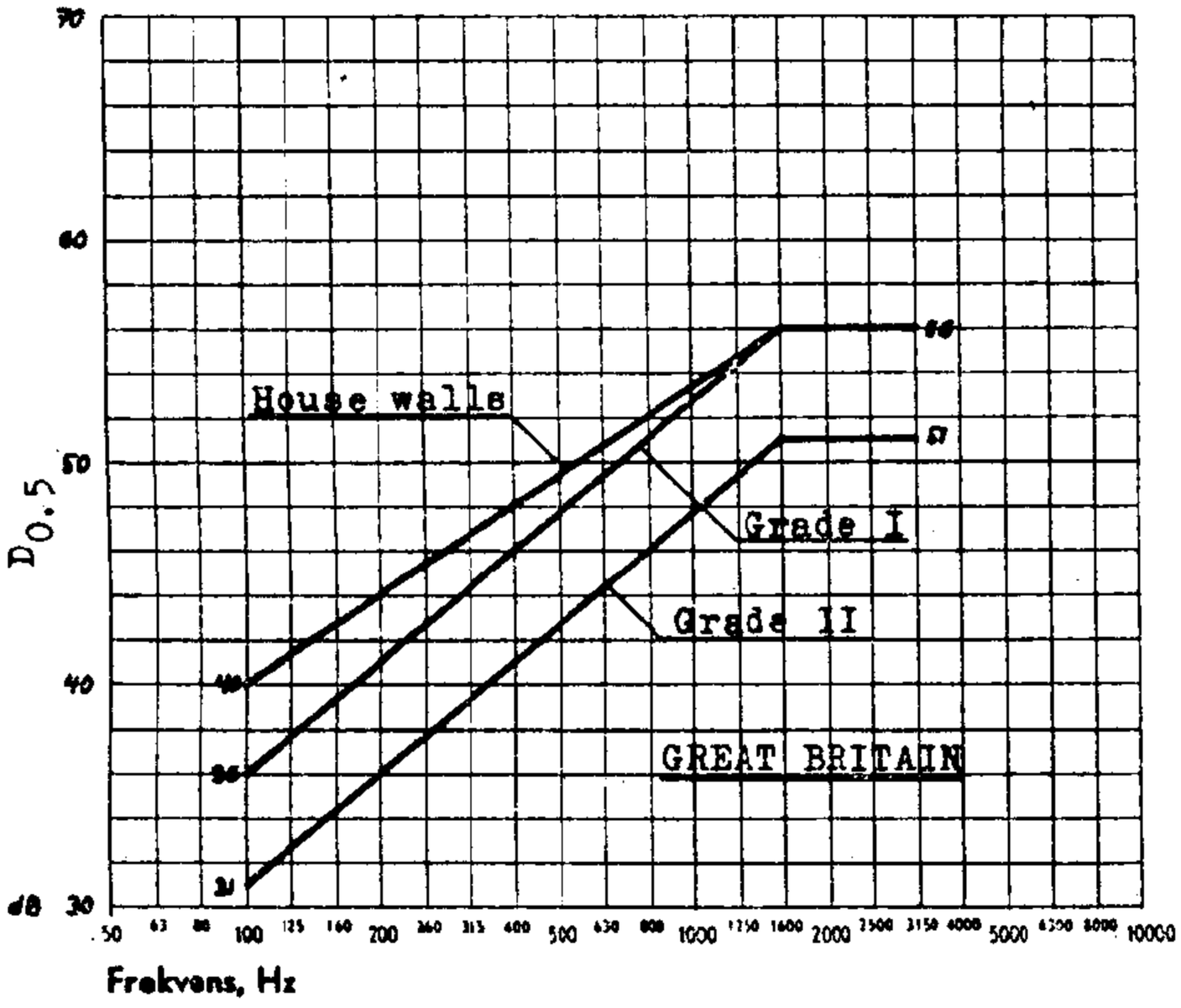


Fig. 14 C. Requirements and recommendations for airborne insulation in Great Britain. Symbols for R, R', etc. as in the preceeding section.

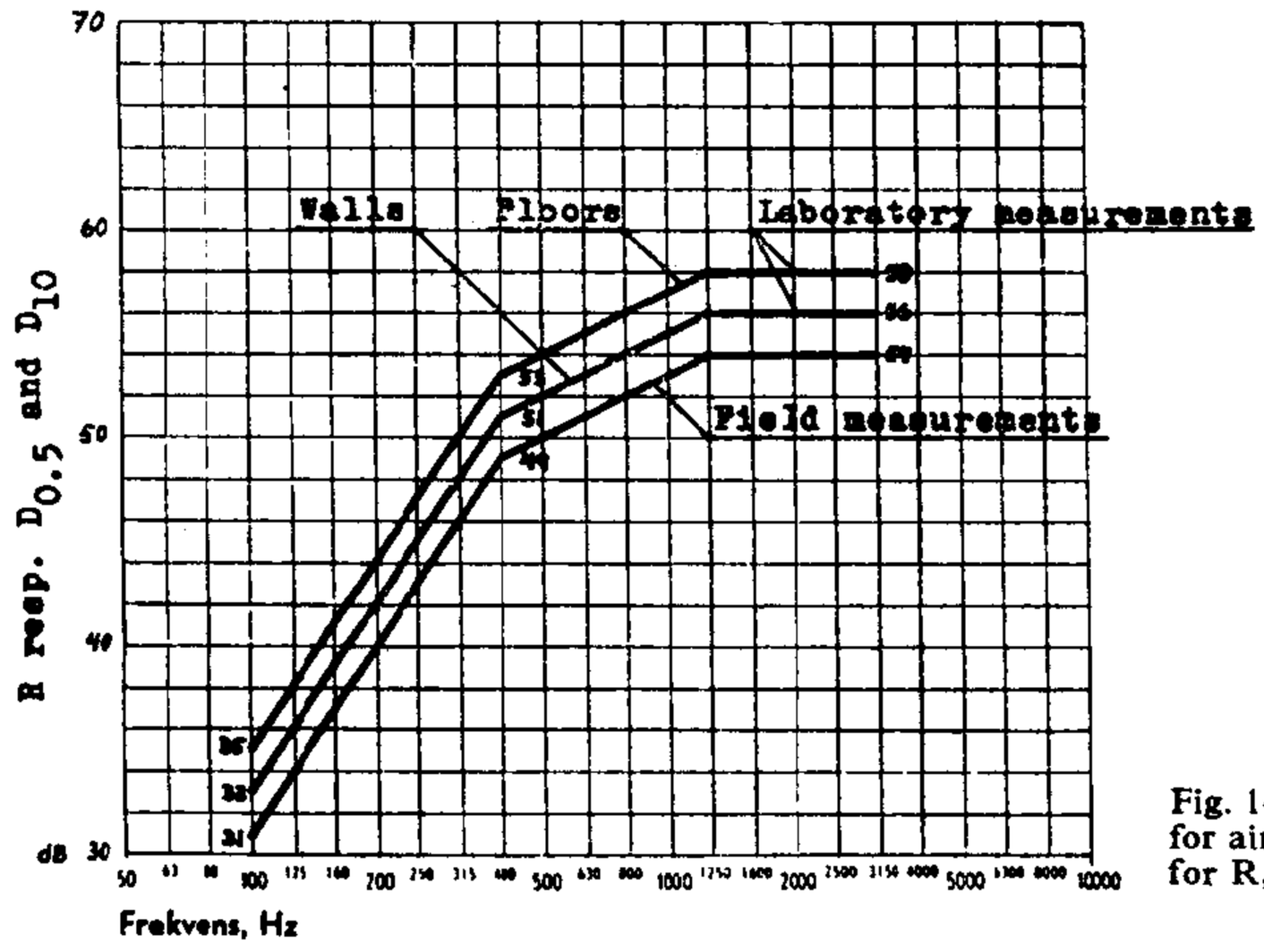


Fig. 14 D. Requirements and recommendations for airborne insulation in Scandinavia. Symbols for R, R', etc. as in the preceeding section.

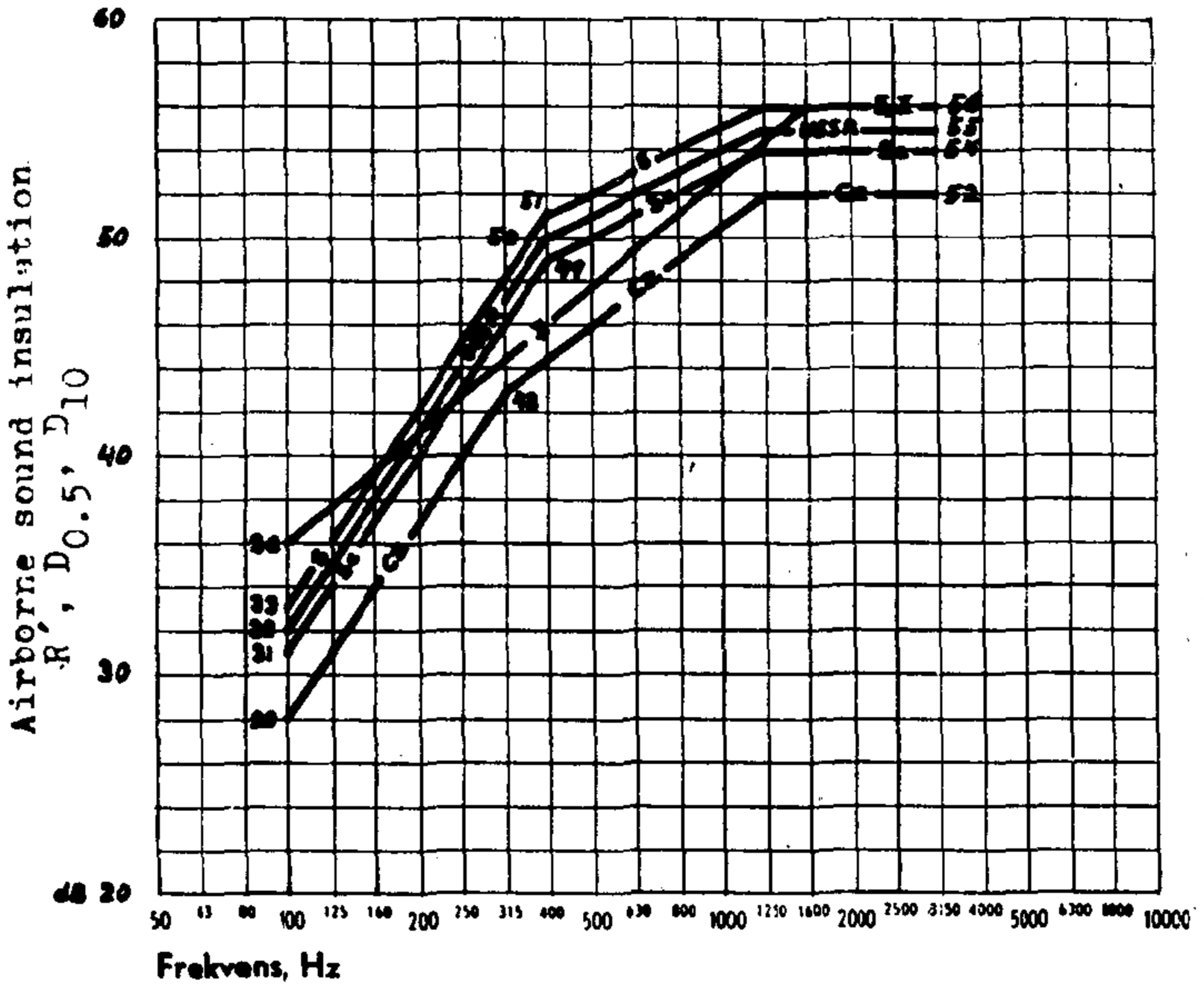


Fig. 15. Airborne sound insulation. Comparison between field requirements for flats.
1 = United Kingdom, grade I ($D_{0.5}$)
S = Germany, Sollkurve (R')
USSR = USSR, curve I (R', D_{10} (?))
Sc = Scandinavian field requirement(D_{10} in Sweden, $D_{0.5}$ in the rest of Scandinavia)
Cz = Czeckoslovakian field requirement (D_{10}).

frequency range of 125–4000 Hz; now a grading curve is being prepared. In France, average figures for airborne and impact sounds are given for the frequency ranges 100–320, 400–1250 and 1600–3200 Hz. This is a very simple principle without troublesome evaluations. Nor does it pretend to more knowledge than we possess.

In some countries, e.g. Scandinavia and France, the specifications comprise both the reduction factor for the bounderies between flats as measured in a traditional laboratory and the normalized level difference in the completed building. This complication is made because one could reach a sufficiently high level difference even if a very small element in the partition has a very low reduction factor. However, if for instance a bed is placed close to such an element very low insulation is experienced.

Some effort has so far been made to get a quality figure for insulation to replace the traditional average insulation. A few countries have followed the German example to introduce a Schallschutzmass for airborne (LSM) and impact (TSM) sound. As mentioned before it is based on the grading curves and is defined as the number of dB's that a measured curve has to be lifted or lowered in order that it can be considered acceptable (average deviation 2,0 dB). The drawback for such a single figure is primarily that it is tied to a certain curve. If this is changed we get new quality figures which must be very confusing for building designers. This is the primary reason why some countries like Scandinavia and England have hesitated to introduce another single figure for sound insulation before we have got an international agreement on such requirements. In the meantime only the sum of deviations from the grading curve is used as a provisional quality figure but with the drawback that it gives the figure 0 for all cases that we get an insulation higher than required.

A grading curve may be difficult to change when finally it has become well established in

a national building code. In place one can use two (like e.g. Great Britain) or more grading curves and require an appropriate curve to be satisfied in the specific case. But it is also possible to have only one curve (like e.g. Germany, for walls) and then require different Schallschutzmasses for different situations which is the same thing as choosing between a great number of parallel grading curves.

In view of these facts one might raise the question whether it is possible to establish some sort of international standardization on sound insulation requirements, a great advantage in the growing international exchange of knowledge and products. One might well be a little pessimistic as to the success of such a work considering the different grading curves already established. Further, we can hardly as acousticians expect to change building traditions in some countries which happen to accept for instance floors with low insulation and have no apparent tenants' reaction. Obviously, other countries with building technique which happens to favour sound insulation—or have strong public opinion on this subject—would not be ready to accept an international standard so compromised. Nevertheless I have some hope for such an attempt at international cooperation on this problem.

This feeling of optimism is supported by the success of a Scandinavian collaboration on this subject. We met five years ago to agree just on the measuring methods, but found it possible also to agree on requirements. These were then shaped as the grading curves shown in fig. 14 and 16. As to airborne sound our first proposal was a grading curve a little different from the British Grade I and the German Sollkurve. However, we found it wrong to introduce another curve and thus increase the international confusion. In place we accepted the German Sollkurve for airborne sound.

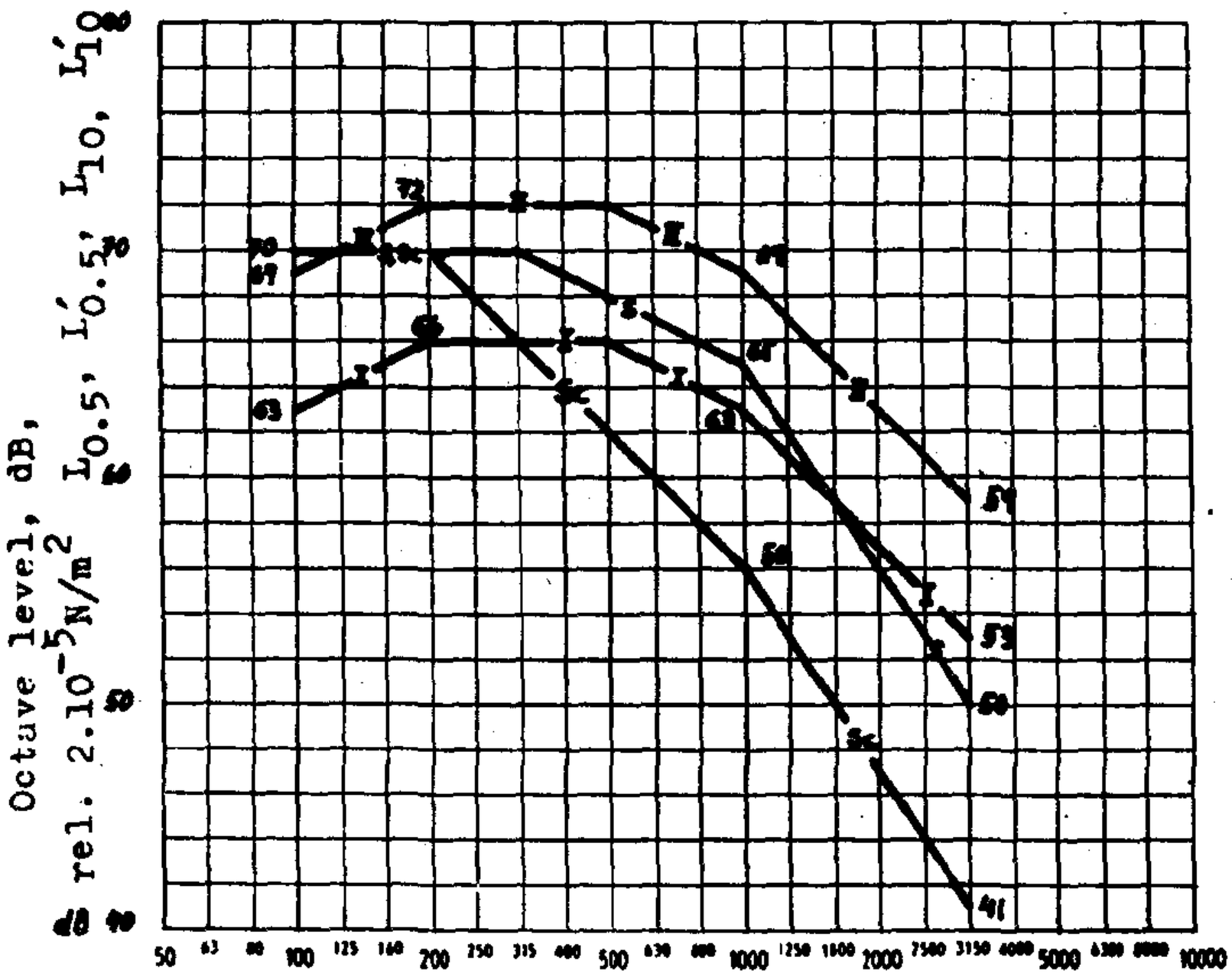


Fig. 16. Impact sound transmission. Comparison between requirements for flats.
I = United Kingdom, grade I ($L'_{0.5}$)
II = United Kingdom, grade II ($L'_{0.5}$)
S = Germany (Sollkurve) and USSR (curve IV) (L_{10}, L'_{10})
Sc = Scandinavia ($L_{0.5}$ and $L'_{0.5}$ except for Sweden where L_{10} and L'_{10} is used).

As we know the existing grading curves lead to very little change in tenants' reaction it should be possible to agree on an international grading curve, at least as a first step for airborne insulation. Also the present French method of having a number of average figures for part bands should be discussed because of its simplicity and leading to no new single figures. Also the appropriate definitions for sound insulation should be discussed and decided on.

While we discuss and perhaps accept such an international provisional recommendation we should organize more research on this subject to see how well the different systems function and also if it is possible to simplify—for instance in limiting the frequency range as suggested by v. den Eijk and others. Such an international discussion which already has started within ISO may also be a great help in countries where such specifications are not yet considered but probably needed.

REFERENCES

1. DIN 4110. Technische Bestimmungen für die Zulassung neuer Bauweisen. Abschnitt D 11, (Schallschütz) 1938.
2. DIN 52211 (Vornorm). Bauakustische Prüfungen. Schalldämmzahl und Norm-Trittschallpegel. Einheitliche Mitteilung und Bewertung von Messergebnissen (September 1953).
3. P. G. Gray, A. Cartwright and P. H. Parkin, Noise in three groups of flats with different floor insulations. National Building Studies. Research Paper no 27. HMSO, London, 1958.
4. P. H. Parkin, H. J. Purkis and W. E. Scholes, Field Measurements of Sound Insulation between Dwellings. National Building Studies. Research Paper no 33. HMSO, London, 1960.
5. C. Bitter and P. Van Weeren, Geluidhinder en Geluidisolatie in de Woningbouw I. TNO, Rapport no 24. September 1955.
6. Gunnar Ö. Jörgen, Stöy i Boliger (Noise in dwellings), Norges Byggforskningsinstitutt, Rapport nr. 16, Oslo 1955.
7. O. Brandt and I. Dalen, Är ljudisoleringen i våra bostadshus tillfredsställande? (Is the sound insulation in our dwellings sufficient?). Byggmästaren B6, 1952.
8. J. van den Eijk, Mijn Buurmans Radio (My neighbours' radio), TNO, De Ingenieur, no. 39, 1960, Gezondheidstechniek 5. see also: J. van den Eijk, My Neighbours' Radio, Proceedings of the 3rd International Congress on Acoustics, p. 1041, Elsevier Publishing Company, 1961.
9. W. Fasold, Untersuchungen zum zweckmässigsten Sollkurvenverlauf den Schallschutz im Wohnungsbau. Proceedings of the 3rd International Congress on Acoustics, p. 1038, Elsevier Publishing Company, 1961.
10. H. J. Rademacher and G. Venzke, Die subjektive und objektive Bewertung des Schallschutzes von Trennwänden und -Decken, Acustica, vol. 9 (1959), p. 409.
11. L. Cremer, Der Sinn der Sollkurven, Schallschutz von Bauteilen. Verlag von Wilhelm Ernst & Sohn, Berlin, 1960.
12. K. Gösele, Zur Durchführung von Trittschallmessungen in bewohnten Bauten, Gesundheits-Ingenieur, 80 (1959), Heft 2.
13. F. Larris, Undersøgelse over Efterklangstiden i Beboelsesrum (Reverberation time in dwellings), Teknologisk Instituts Forsøgsberetning 1946–48, Meddelse Nr. 5, Copenhagen 1948.
14. K. Gösele, Trittschall- Entstehung und Dämmung, VDI-Berichte Bd. 8 (1956), "Schall und Schwingungen in Festkörpern", p. 23.
15. ISO Recommendation R 140. Field and Laboratory Measurements of airborne and impact Sound Transmission. 1960 (E)

"On Listening with Both Ears"

Professor COLIN CHERRY

Imperial College, London.

I want in this lecture to give my views upon the nature of human perception, more especially aural perception, and I want to talk about that most important of all sounds, human speech.

When we study any aspect of a human being we must remember that we are studying an intelligent, articulate, thinking creature, and it is necessary to have some philosophy to guide the design of our experiments and I should like to start by giving an outline of my own; then to follow with a description of a number of experiments which have been carried out by my students and myself during the past ten years, which have arisen from this particular philosophy.

First then, the question of behaviourism. Rather than use the term "behaviourism" (for I believe this to be ambiguous) I prefer to distinguish between two languages, or rather meta-languages. There is first the language of the articulate subject himself and then there is the language of the scientist-observer who describes the experiments. The really important thing is to distinguish between these two. It must of course be remembered that the scientist-observer is himself as biassed and fallible as the subject, being one of God's creatures. But so is any physicist looking at an instrument; like him we must design our experiments so as to be repeatable, describable and reliable as possible–which is what I would mean by the word "objective". My scientist-observer's meta-language might include such words as *motivation*, *purpose*, *past experience* and other factors which are not directly observed but are inferences based upon sufficient observations.

The second belief which I should like to stress is one widely shared; it is in the overwhelming importance of speech as compared to other sounds in the world around us.

We may seem to be sight-dominated so far as our safety and security go, with our eyes wide open; but our thoughts would seem to be far more speech-dominated. Our private musings are wordy, every moment of the day. Our silent soliloquy may be rambling, incoherent and vague, but inasmuch as it uses any rules of syntax, it has form and structure. The scraps and bits of phrases make up a greater part of our formulated thoughts and as this ceaseless stream flows on our decisions become clear and our actions determined; not by any logical reasoning process, for the greater part of our everyday experience, but by our verbal habits, the scarcely conscious use of our engrained rules of language. That is to say we have such thoughts as our particular linguistic heritage permits us. Whatever may

be our views upon free-will, our thoughts must be closely channeled by the speech habits of our culture.

Surely, it is this faculty of speech, with its power of setting our concepts into relationship, and formulating our thoughts, that raises us above the animals and entitles us to the name *homo sapiens?* Descartes might well have said, not "Je pense, donc je suis" but "Je parle, donc je suis".

The greater part of classical acoustics has aimed to abstract out of human experience by deliberately using non-speech-like stimuli, such as pure tones or clicks and so to become "objective" by avoiding the effects of a persons linguistic habits. Perhaps this is true of many researches reported upon at this Congress. So-called physical acoustics is explainable in terms of properties of the stimulus (for example, spectral properties, or intensities) and perhaps of the environment. Many psycho-physical experiments are, too. But we number ourselves among those who aim to be equally objective by going to the extreme other end of the scale, deliberately involving a listeners' past experience, such as his speech habits, and aiming towards (though not always achieving) as natural, real-life, conditions as possible. When we speak, we normally speak *to* someone. Yet nearly all studies of speech relate to one person's voice. It is the *conversation*, surely, which we should seek to understand, as the basic unit. Very few studies of conversation seem to have been made, in the scientific and objective sense. We have done a little, but the British Post Office has done far more, [(1)] distinguishing particularly three different conditions under which the perceptions are working. 1) one person only speaking, one listening; 2) both persons speaking at once; 3) neither person speaking, but both waiting.

Again, conversation is rarely uninterrupted. Usually others are going on around us. How does the brain perform this remarkable feat of separating the speech of one person from the buzz going on around? Some years ago, I referred to this question as "the cocktail party problem", a name which seems to have stuck! We have been experimenting upon this problem for several years, and I shall say a little about it later.

I believe the answer to the cocktail party problem rests upon two facts. The first is our possession of two ears; the second is our basic faculty of *mimicry*, to which I shall return in a moment.

Why do we have two ears? Of course it gives us the facility of directional hearing. But this is not the primary faculty that two ears gives us, which is rather the greatly enhanced ability to separate two or more sources of sound into distinct images, or *gestalten*. A person deaf in one ear complains, not that he doesn't know where you are standing but that he can't listen to what you are saying when someone else is interrupting. Image separation must precede directional hearing; thus two sources of sound cannot be heard to lie in different directions, one *here* and the other *over there*, unless they first are separated and given identity. Theories of stereophonic sound systems have been confused by missing this point, since most have considered one point source of sound only; yet an orchestra has many instruments.

Let me return to my point concerning mimicry. A child learns the shape of common objects by handling them; it learns to speak, not only by listening to its mother's voice, but by imitating the movements of her mouth and making the same sounds. In each case, overt bodily action is called for. Why should this imitative action ever desert us? I believe it does

not, but that, in adult life, we are *prepared* to do the same; we may inhibit actual overt action, but we are prepared. I see a tea-cup on the table and I run my finger round the rim; or I may merely be prepared to; that is, I have perceived it as circular. All perception may not be of this nature; that of colour, for example, is not. If I see the tea-cup as yellow it is a judgement forced upon me, but no imitative action is involved.

I believe this mimicry faculty to be basic to the perception of speech, especially. For speech forms a very special class of sound, since we can both hear it and make it; in this sense speech is distinct from all other sounds of this world—footsteps, motor horns, things clattering. I include singing and all vocal sounds. If you sing a note I can sing it too. But, further, before making any sound I can set my vocal organs into nearly the right configuration; that is, I am *prepared* to imitate it.

So too with speech. I believe that our deeply engrained speech habits are called upon every time we listen to anyone speaking; that we are, subthreshold, imitating what they say. Our overt mimicry actions of childhood are normally suppressed, but we are readily stimulated into mimicry by the sound of another person's voice, as I will demonstrate. (2)

Demonstrating of "shadowing" at this point.

We have called this experimental technique "shadowing". Notice that my mimicry was not perfect, since my own speech habits do not conform exactly to those of the text read by the speaker. For example, if an American shadows an Englishman he uses many American words like *railroad, gotten, airplane.*

During or after shadowing in this way, a listener has almost no idea of what he has said; he cannot even give a summary of the text or outline the arguments. That is to say, the action is limited entirely to the motor level, and the overlaid semantic and intellectual processes which normally accompany speech perception are shut largely off.

This view of speech perception as a suppressed motor activity is tantamount to regarding it not as a recognition of the sounds of speech but as a recognition of those movements of the vocal organs necessary to produce these sounds. This is close to regarding speech perception as a Pavlovian conditioned reflex, of which the phenomenon of shadowing, which I have just shown, is a clear demonstration.

Let me get back to binaural hearing. We can listen with one ear, if we have the misfortune to be deaf in the other, or we can listen with both. As I have stressed already, two ears are better than one for clear formation of sound images, as when picking out one voice in a crowd. When listening with both ears we hear only one world, because a process of fusion takes place which is not under our voluntary control. But the sounds reaching each ear are never exactly alike. The brain carries out a continual analysis to determine the degree of "correlation" between the two; if closer than some limit the fusion operation occurs, but if not then the sensations at each ear are preserved as two quite separate gestalten.

Let me illustrate this. If you wear a pair of headphones, and drive the two earpieces from two tape-recorders, upon which are recorded two quite different readings, you can listen to one speaker at your left ear, or to the other at your right; no fusion occurs. Further than this, you can shadow either of them, in the way I demonstrated, without making errors, but

if you do this you will have no idea whatever concerning the message at your other ear; perception of it is almost entirely inhibited.

I say "almost". You know whether or not it is a human being speaking, but you will usually not know the language. (2) Again, there is evidence that certain words of high significance, like your own name, will "pop out" at you. But perception at one ear almost totally inhibits perception at the other, when the sounds are quite independent.

This question of "dependence" is not entirely a statistical one, measurable by long-term correlation functions, but is a matter of moment by moment dependence, assessable by very short-term correlations.

A simple experiment illustrates this. Suppose you wear headphones and listen to continuous speech. You can shadow it, as I demonstrated. But suppose you use a twin-track recorder and, every now and then, split one word into two synonyms, one to the left ear and one to the right. For example, you might hear in your left ear: "He walked along with a smile..." and in the right: "He walked about with a smile...". When shadowing you will speak only one of these synonyms, but it might be from either the right or left ear. There seems to be no evidence here of a dominant ear.

Thus the brain has detected the lack of relationship between the left and right ear stimuli in a time shorter than it takes to say a single word. In fact it is a very short time analysis; our own experiments suggest the time is only about 6 or 7 milliseconds. (3)

This time of observation, by the brain, of the signals at each ear is very important. It ensures, for example, that binaural fusion takes place when there is an inter-aural time difference, as when we listen to a source of sound not directly ahead of us. The maximum delay that can ever occur between the ear, in real life, is about half a millisecond, which happens when the source is directly to one side of the head. But, using a tape recorder with two reproducing heads, and by wearing earphones, we can artificially introduce any time-interval we wish between the ears. Suppose we play one speech into the left ear and a delayed version of the same speech into the right ear; as the delay time is increased from zero the fused sound image moves across the head. But the time delay can be increased beyond half a millisecond and we find nevertheless that the fused image remains. In fact it remains fused into a coherent image until the interaural delay time is as much as 6 or 7 milliseconds, the correlation time I mentioned, after which it starts to disintegrate and become two "mirror" images, one lying to the left hand side of the head and the other to the right. (3)

It must be appreciated that, when wearing headphones, we hear the sound image *inside* our heads and can only say whether it lies to the right or left of our nose. Because head-turning is of no help, the sounds at the ears are unchanged and do not provide the brain with any evidence of angular direction. Such simple right/left judgements provide an excellent objective basis of measuring a number of properties of the subjective image. For example, a reversing switch under the control of the experimenter enables him to move the subjective sound image in the listener's mind, to the right or left by varying amounts according to the time-delay setting on the dual tape recording. Both the extent and the sign of the time delay can be set to successive random values, so that the listener hears the image moving right and left across his head in random positions.

Suppose, first, we simply use the same speech recording presented to both ears, but with

this varied time delay. The listener may be asked to guess only whether the fused image appears to lie to his right or his left–not by how much. Then his % correct guesses may be plotted against the time delay, as shown [3] in figure 1.

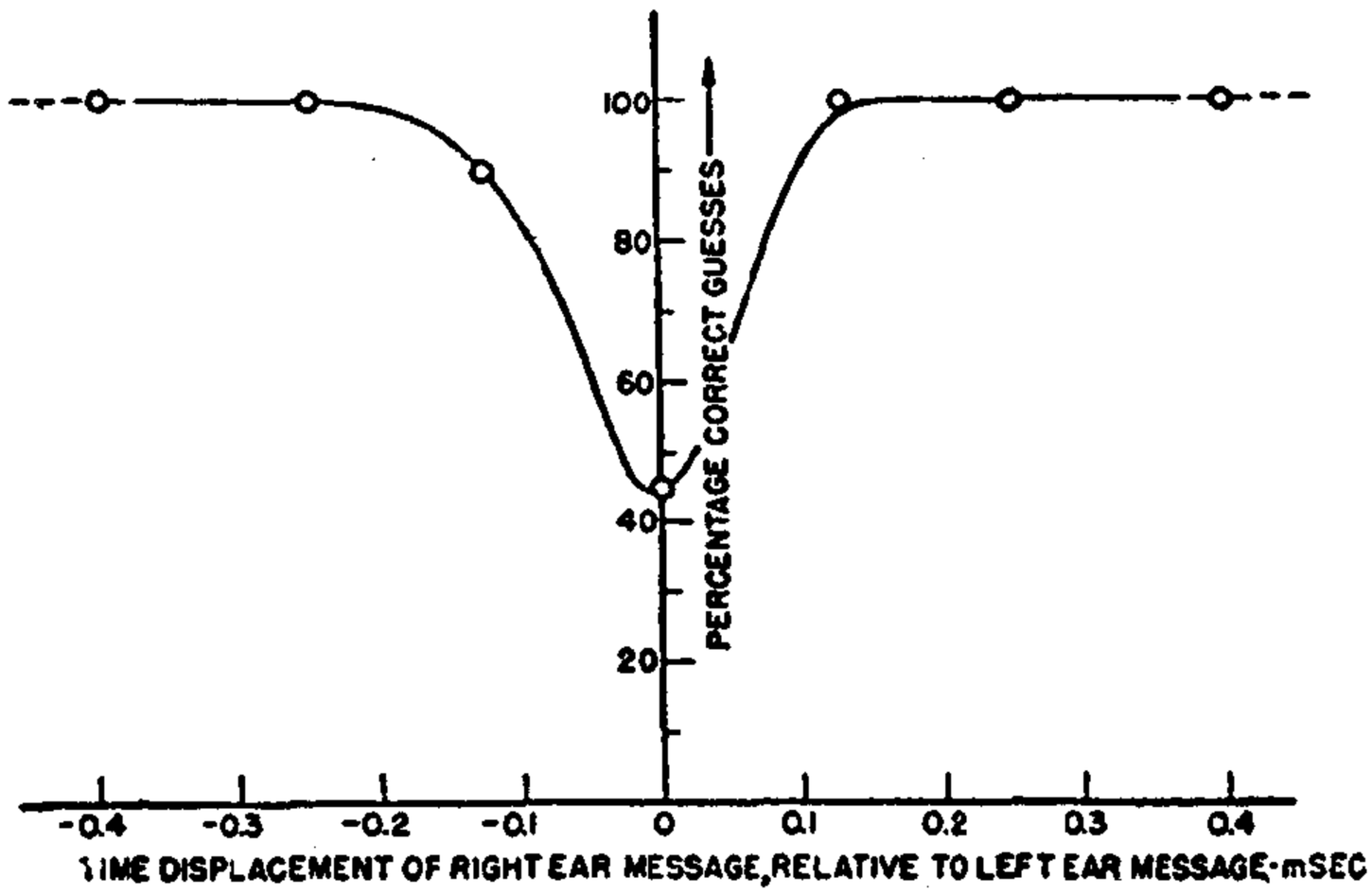


Fig. 1. Binaural perception of speech direction (right or left).

The dip to 50% here, when the image is exactly in the middle of his head, shows that his guesses are then random. He simply "cannot say". The width of the dip here (a few micro-seconds) measures his uncertainty due to 1) the finite subjective size of the image and 2) absence of any indication of where "the centre of his head" is (the intra-cranial plane). The time delays here are no greater than half a millisecond. Suppose we increase this, as I described just now, until the image fusion fails, and disintegrates. The result is that, after

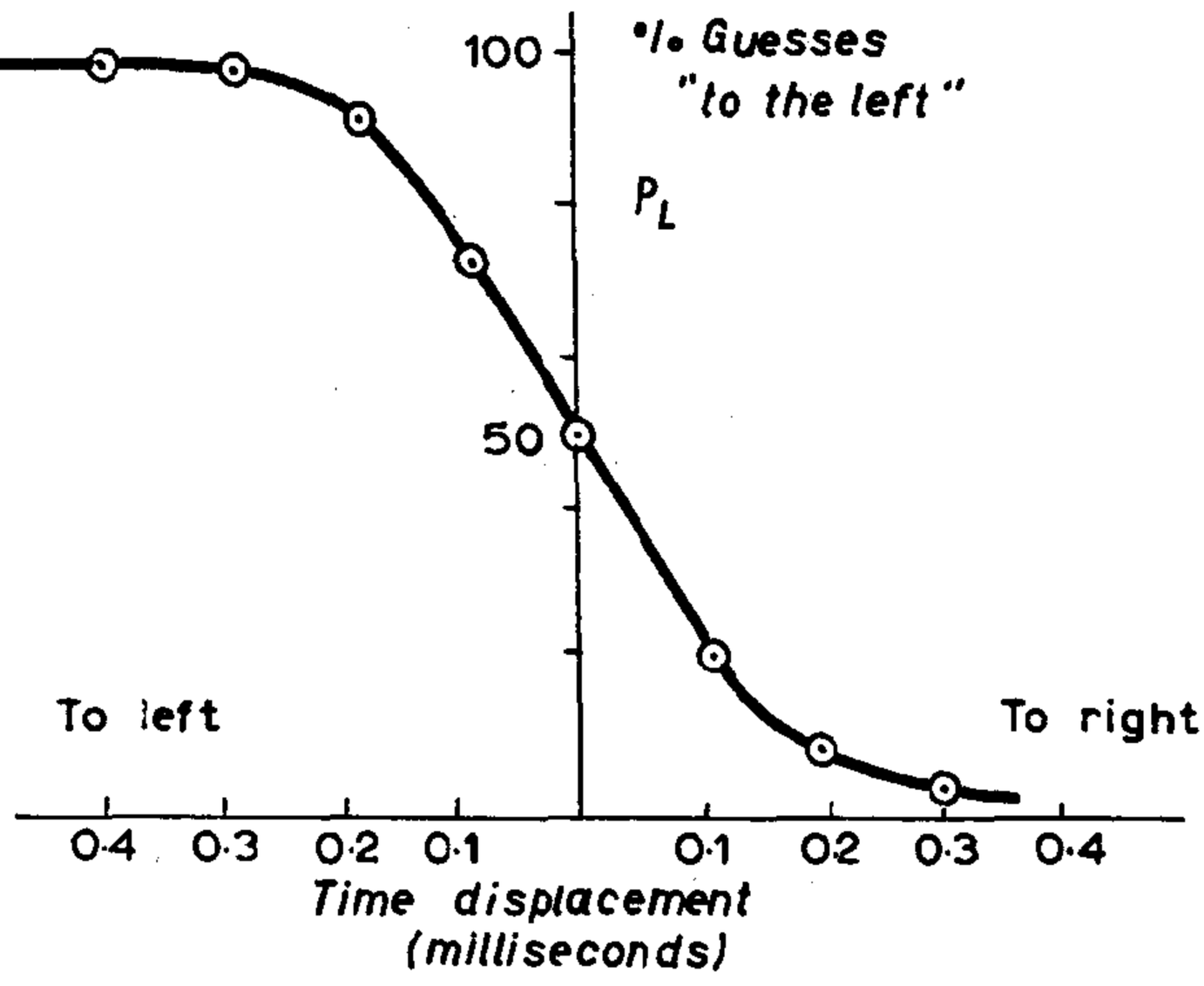


Fig. 2. The Binaural Fusion Curve plotted as % guesses to one side.

6 or 7 milliseconds delay, the subject is unable to guess *right* or *left* correctly because the disintegrating image is ceasing to have exact location. (3)

It will be appreciated that .he subjective location of the centre of our heads is vague. We know our right and left sides, but where is the dividing line? To avoid this difficulty, the same curve that I have just shown can be plotted another way, as in fig. 2:

Here we plot the subjects guesses, "to the left", not his correct guesses right or left, as the sound image is moved from side to side at random. We call this P_L.

This "forced judgement" technique has proved to be a powerful way of examining the mechanism whereby the brain analyses the sound data reaching each ear and decides whether or not to fuze the two fields. As I have already said, if the two sounds are closely similar, fusion occurs; if they are not, we perceive one of two fields, either at the right or the left ear. Then how close is "close"? How does the brain assess the degree of similitary of the sounds reaching each ear?

We have sought a description of the processes involved, in mathematical terms, based upon experimental results, using this forced-judgement technique. Thus we can adjust the sounds reaching the left and right ears, using headphones, to be as similar or dissimilar as we wish, in several different ways. Thus we can add noise to the sounds at one ear, or we can filter their spectra in different ways; or, again we can add non-linear distortion to either side. My colleague, Dr. Sayers, has carried out a very extensive set of measurements and the nature of the analysis performed by the brain has gradually been exposed (4). As you might imagine, it is a process of correlation, but not a simple one; I can best describe its nature by a diagram; fig. 3.

There is indeed a process of cross-correlation carried out centrally; as mentioned earlier its integration time is approximately 6 to 7 milliseconds. Where is this carried out in the nervous system? I should like to know. Fibres from the right and left auditory systems are cross-connected at more than one place, certainly at the olivary nucleus and the inferior colliculus, before the united data is routed to the right and the left auditory cortex.

But this is not all, for a very good reason. Before this central cross-correlation, the sounds at each ear are first auto-correlated, with an integration time of about 1 millisecond. This auto-correlation (called "transformation" in fig. 3) gives, amongst other things, a running average value of the sounds at each ear. Why do we need this? It enables us to assess the direction of very high-pitched sounds, sounds whose wavelength is less than twice the distance between the ears–say, higher than about 1200 cycles a second. With such sounds there would be complete ambiguity of direction if no evidence was supplied to the ears other than the inter-aural phase or time difference. But, with this mechanism for observing a running average, the *envelopes* of the two sounds become available for cross-comparison and an unambiguous fusion results.

A very simple experiment illustrates this. If, when wearing headphones, a pure tone of 2000 cycles a second is played into one of your ears and 2100 into the other, no fusion occurs. The sounds are heard lying to opposite sides of your head. If now, both tones are modulated with a one or two hundred cycle tone then immediately fusion occurs, and the image is heard to lie in the centre of your head. The modulation has provided each tone with a running average value and this is passed to the central cross-correlation process, where the similarity of the envelopes is observed.

Where does this auto-correlation occur, at each ear? There is the possibility that it is carried out neurally, at each cochlear nucleus, but I have also wondered whether the cochlea itself may not possess suitable mechanical properties to execute auto-correlation. Numerous theories of the mechanics of the basilar membrane have been written, each making certain simplifying assumptions; they have been well summarised by Wansdronk [5]. Is it possible that we have two motions, one travelling over the other, being the compression waves in the fluid and the transversal vibrations of the membrane? Then can the stimulus of the hair-cells be some product of the two, which will vary with time and with distance along the basilar membrane?

I am not asserting this here. Rather, I should like to take the opportunity of asking those of you here, more familiar with the anatomy of the cochlea than I, whether this suggestion is nonsense.

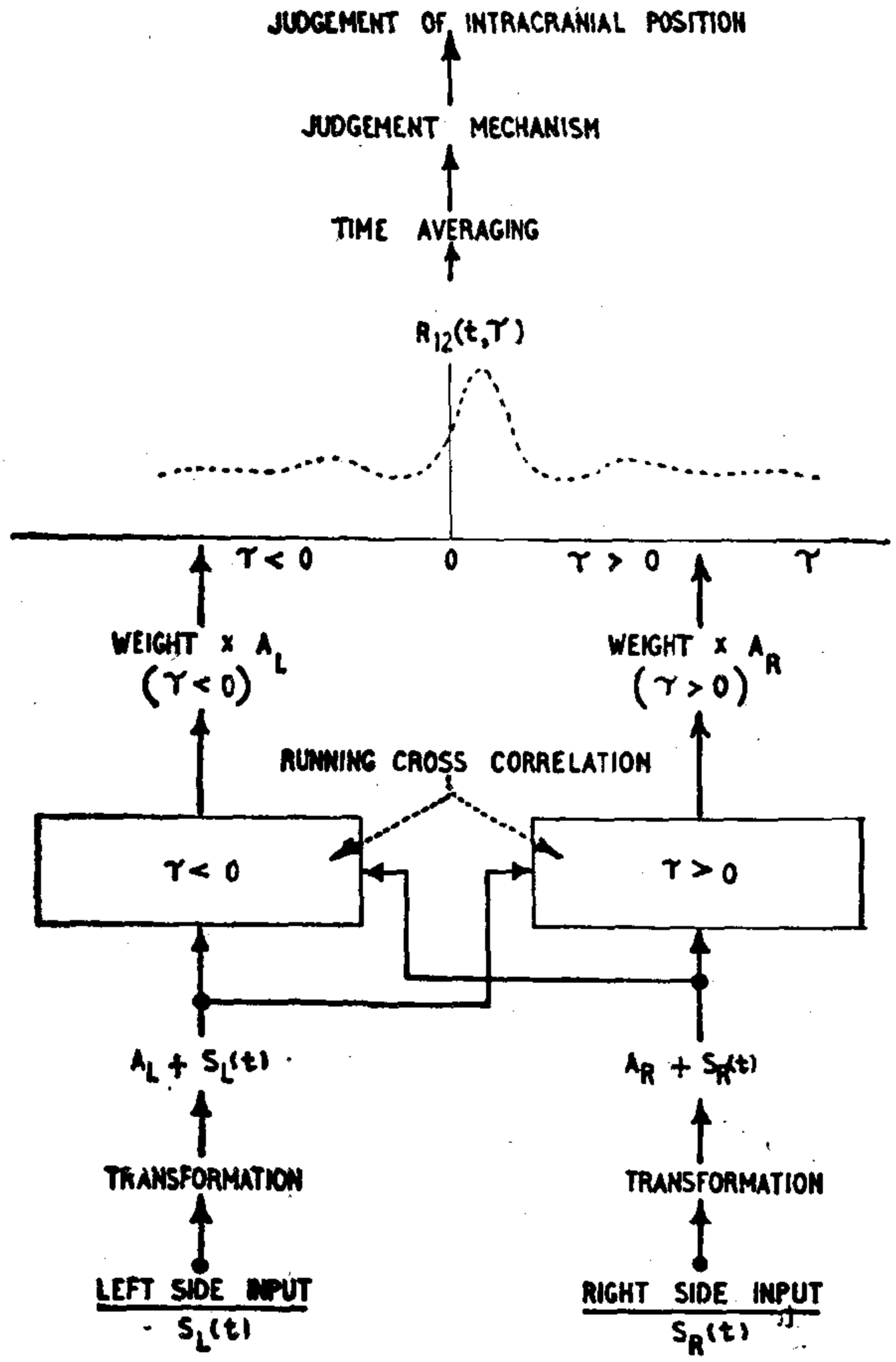


Fig. 3. Model of the simple cross-correlation theory of binaural fusion. Note that the running cross-correlation unit has been divided in two parts only for clarity of representation.

Before leaving this subject of binaural fusion, I should like to show a few results of our forced-judgement experiments, and compare them to calculations based upon our correlation model. I have shown earlier a slide of the subjects' guesses as the fuzed sound image is moved left or right by random amounts, when the signal is speech, or noise, or other statistical source. Suppose instead we use a pure tone, 800 cycles per second; the result is on fig. 4.

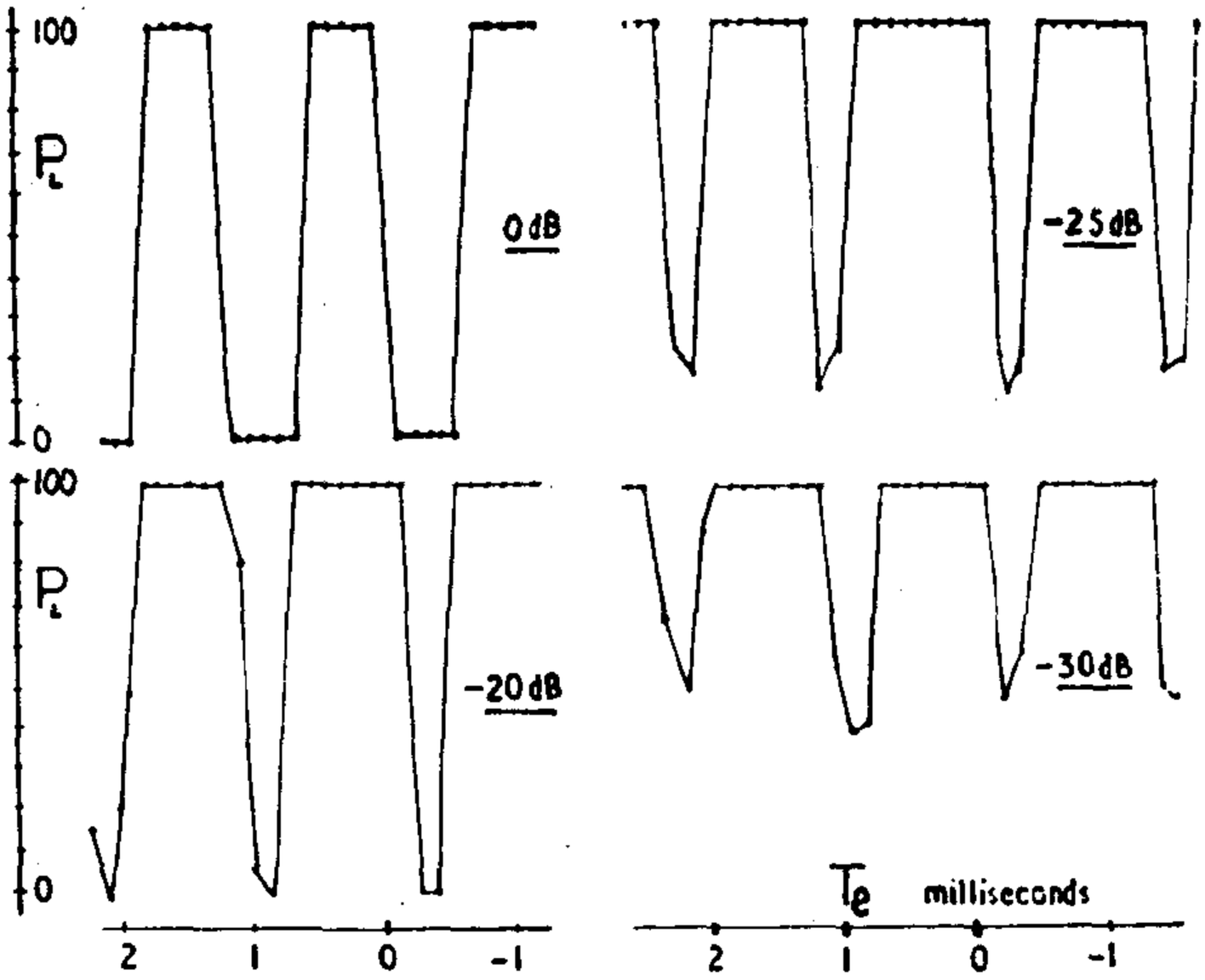


Fig. 4. The effect of interaural amplitude difference on coherence curves for a binaural pure tone (about 800 cps). The parameter in this figure is the level of left-ear signal $S_L(r)$ referred to right ear signal $S_R(r)$.

As the inter-aural time delay is varied the fuzed image appears to move back and forth across the head; a time shift of one period is inconsequential, so that this "fusion curve" is itself periodic.

A most interesting thing occurs if the intensity of the sound at one ear is increased or decreased, which we see in fig. 4. It is as though the fusion curve is bodily displaced, up or down, without changing its shape, but is "sawn-off" at the 100% and 0% limits on the vertical scale P_L. This is always the case, whatever the type of signal. Fig. 5 shows the subjects guesses "to the left" when both ears are stimulated with two sinusoids, of 600 and 800 cycles per second, equal in intensity.

On fig. 5 there is shown dotted the calculated fusion curve, based upon our auto-correlation model and the spectra of the sounds. The agreement is seen to be quite close.

Figure 6 shows the curve when both ears are stimulated by an intoned vowel (the vowel *a* as in *ask*) noise masked at the left ear only. Again, as the relative intensity between the right and left ear stimuli is varied, so the curve is bodily moved up or down between the 100% and 0% limits.

The effect of noise masking is to bring out the sharp dips seen here. These dips correspond to the formants of the vowel *a*; again, calculated results are shown dotted here. Many such fusion curves have been measured, using different vowels, with and without noise masking,

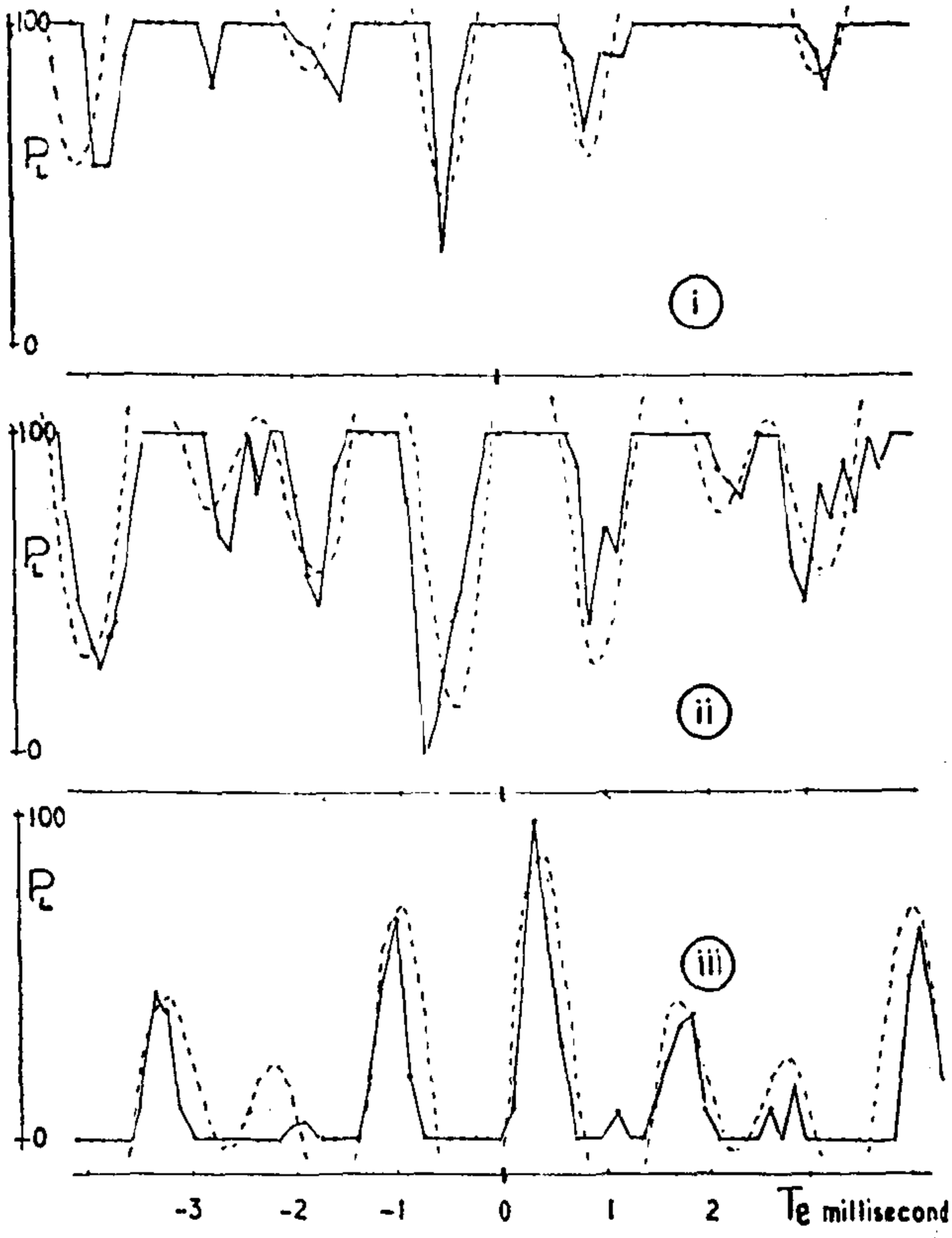


Fig. 5. Coherence curves for a two-sinusoid component binaural signal, showing influence of interaural amplitude difference. Lower frequency components 600 cps –2 db with respect to higher frequency component 800 cps, each channel. Left-ear signal above right-ear signal by 10 db curve (i), 6 db curve (ii) and –6db curve (iii). Theoretical curves dotted.

and even using running speech, together with corresponding calculated curves [4]. But I have no time to show these today.

However, out of this work one further experimental finding resulted which may be of great importance in real life, when we listen with both our ears. It is illustrated by fig. 7; if both ears are supplied with sound from some source (say speech or a sung vowel) and if then the sound at one ear is masked with noise, or is distorted, the fusion mechanism will show a preference; it will fuze the image completely when the inter-aural time delay is such that the unmasked or undistorted sound leads in time, otherwise only partial fusion results [4]. Fig. 7 illustrates this "preference effect"

This complete lack of symmetry arises over a variety of noise-masking and distortion conditions, but not over all. It must have its effect upon our perceptions, though I am not sure exactly what effect. It is unlikely to assist the brain in reflective or reverberant situations, as I had once thought, because these reflected sounds are normally delayed by more than 6–7 milliseconds behind the primary source, so that the fuzion does not occur. However, it may assist the brain to solve the problem I have called the "cocktail party problem", or the

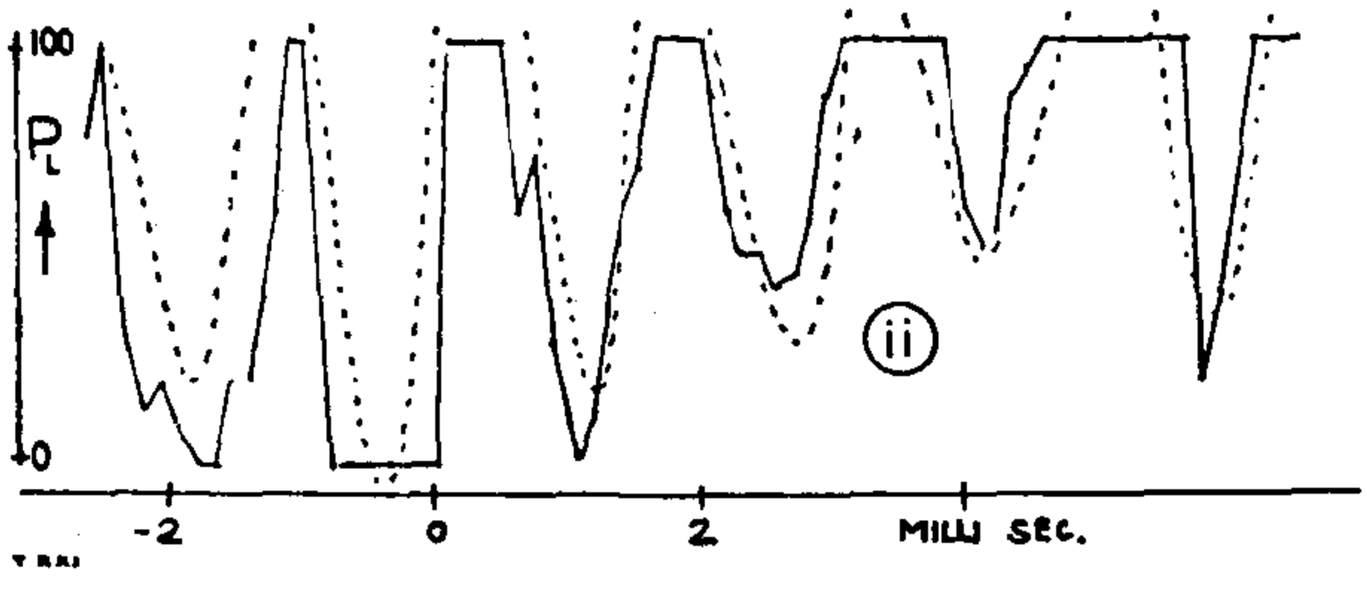


Fig. 6. Coherence curves for an intoned vowel sound "a" noise-masked on one channel only (left); S/N = — 12 db. Theoretical curves (dotted) computed, neglecting all but the two most energetic components. Note that certain samples of intoned vowel sounds with low "voicing" noise, and having closely repetitive patterns, require little or no noise-masking to bring out periodicities in the coherence curve. These may be treated directly as multiple-sinusoid component signals.

separation of one wanted voice from others chattering nearby. But I am not yet sure. At least one would expect such a pronounced effect to have some psychological consequences which distinguish, in some way, a pure source from the same source corrupted by other sources. More research needs to be dohe.

How do we listen to one voice in a noisy crowd? Do we face the speaker or not? We carried out a simple experiment to find out. In this, a listener sits on a swivel chair between two loudspeakers; two different readings are played into these and the listener is required to listen to one of them. To ensure he obeys these instructions, every few seconds a word is

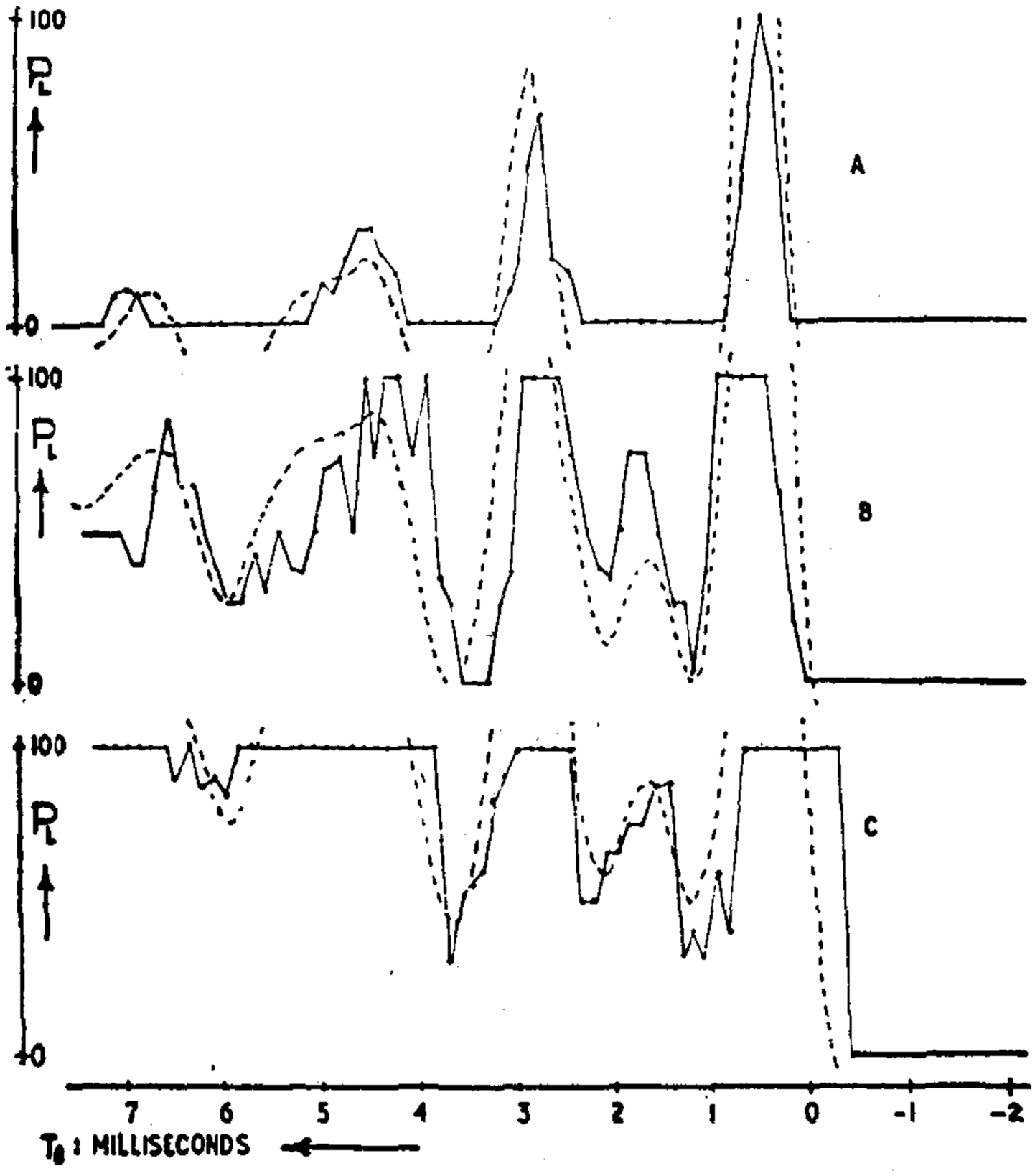


Fig. 7. Coherence curves for speech sample *A*, noise masked on the left-ear channel. The three curves shows the effect of increasing the level of the right-ear channel signal, moving from the lowest curve upwards.

struck off the tape recording and a spoken random number is inserted, which he is asked to write down. In every case the listener turns his head sideways so as to maximize the time difference of the sounds at his two ears. In this way the wanted sound image is subjectively moved as far away as possible from the unwanted; the two internal representations, in the head, are isolated to a certain extent in the sense that they are two overlapping, though not separated, sets of data. This preparation of the acoustic evidence, by turning the head, presumably enables the subsequent act of inductive inference to be more readily done.

This cocktail party problem is one upon which we are still working. It is a phenomenal act of inductive inference–phenomenal in speed and accuracy. To my mind it is facilitated largely by our faculty of subjective mimicry, as I stressed earlier in this lecture and illustrated by my little "shadowing" experiment. That is, we have stored in our memories a vast representation of prior probabilities, which we call our speech habits; such probabilities of course facilitate inductive inference.

However not only are we able to separate a voice from different noises, such as a background of street traffic, which have completely different probabilities, but we can separate one voice from another. Two people, both speaking English say, have very similar probabilities! Nevertheless the separation of the voices is far less ready or accurate.

To illustrate this we can make a recording of two simultaneous speeches, perhaps using the same speaker. Head-turning and directional hearing cannot help now. Nevertheless it is possible to listen to this mixture and to piece together one of the messages—though you will make mistakes.

RECORD DEMONSTRATION

Here your speech habits, your prior probabilities, or the syntax structure, call it as you wish, are helping to an enormous extent. But it is possible to compose messages having very curious syntax structures by using clichés, strung together in random order, perhaps with conjunctions in between. Such messages cannot be separated in this ready manner. For example, this message is composed of English clichés:—

"*It is ridiculous to the extreme* that *in this day and age* we should *decide upon the spur of the moment* and *act upon an impulse. Our political opponents* cannot *count the cost* but *in a rash moment base their arguments upon a fallacy. We are in troubled waters, fishing for time*, and our *standard of living hangs upon a thread...*"

This sounds much like a party political broadcast. It has some semantic content. Each cliché is a highly probable, almost certain, string of words, yet at the juncture of any two clichés the probabilities are uniform. Let us listen to two such simultaneous cliché speeches. You will pass from one to the other [2] without realizing it.

RECORD DEMONSTRATION

Concerning this cocktail party problem our main interest has been to try and isolate the specific acoustic cues which the brain uses. For example, how important are the variations of larynx pitch, or the speech formants, or the turbulent breath sounds? We have again used the forced judgement technique, which I illustrated earlier with several figures; in this

case, we force a listener to say whether a speaker's voice lies to his right or left side, when he wears headphones and a random delay between the ears is introduced but (and this is the difference now) when an interfering voice is present.

Suppose first we fix one voice, subjectively, in the centre of the listener's head (that is, using no inter-aural time delay) and then place a second voice right or left, by various amounts. We find that the listener can guess the position of the wanted voice just as accurately as when the fixed, interfering voice is not present. In other words, separation of the voices is complete.

Throughout most of this lecture I have referred to experiments which involve the listener in wearing headphones. Head turning then gives his brain no evidence of direction, so he does not hear the sound source in any particular direction. Rather he hears it lying inside his head, to the left or right. In real life, when we listen to sounds we hear them in some direction, but such directions can only be the consequence of the brain solving the cocktail party problem and separating the various sources so that one appears to lie over *here* and one over *there*. I would go further and say that our sense of direction is nothing more than the subjective effect of the better image separation we execute by virtue of listening with both ears.

It is interesting to see how the brain could, in theory at least, determine the direction of a *single* sound source entirely from the inter-aural time difference; it does not need to be provided with either the velocity of sound, v nor with the distance between the ears, h.

Let the single sound source be orientated at an angle α to the listener's front and elevated by an angle β. Then it can be shown that the inter-aural time T is given by:

$$T = \frac{h}{v} \operatorname{Sin} \alpha \operatorname{Cos} \beta$$

$$\text{So that } dT/d\alpha = \frac{h}{v} \operatorname{Cos} \alpha \operatorname{Cos} \beta.$$

where h = distance between ears

v = velocity of sound.

$$\text{Hence Tan } \alpha = \frac{T}{dT/d\alpha}$$

which is independent of h, v and elevation β. Hence, in theory, the brain needs only to assess the inter-aural time difference T and also, by slight head rotation $dT/d\alpha$, using the kinesthetic sense of the neck muscles. With two or more sources, the story is totally different–the brain must first solve the cocktail party problem and separate these sources. The brain is extremely sensitive to slight inter-aural time differences–as I have shown earlier, sensitive to only a few microseconds. It is of course also sensitive to amplitude differences at the ear, but I think time difference is dominant.

This faculty most of us possess, of binaural hearing is phenomenal in its precision and in its psychological value. With it we bring in sound to aid us in formation of the numerous images which make up the world around us, and give it movement too, for sources of sound are rarely stationary.

Acknowledgement

Some of the diagrams reproduced here have been abstracted from the Author's paper, together with Dr. B. M. Sayers as co-Author, and due acknowledgement is hereby given to the Acoustical Society of America.

REFERENCES

1. D. L. Richards. General Post Office Research Station, Dollis Hill, London. (Private communication).
2. Colin Cherry. "Some Experiments on the Recognition of Speech, with One and with Two Ears". Jour. Acoust. Soc. Amer. *25* pp. 975–979. Sept. 1953.
3. Colin Cherry. "Some Further Experiments upon the Recognition of Speech, with One and with Two Ears". Jour. Acoust. Soc. Amer. 26. pp. 554–559. July 1954.
4. B. McA. Sayers and Colin Cherry. "Mechanism of Binaural Fusion in the Hearing of Speech". Jour. Acoust. Soc. Amer. *29* pp. 973–987. Sept. 1957.
5. C. Wansdronk. "On the Mechanism of Hearing". Philips Research Laboratories. Eindhoven. Nov. 1961.

Computer-Catalyzed Speech Research

by EDWARD E. DAVID JR.

Bell Telephone Laboratories, Incorporated, Murray Hill, New Jersey

Introduction

Over the past two decades, the sound spectrograph has become the instrumental staple of speech research. [1] There was good reason, for a spectrogram presents the acoustic aspect of speech in a convenient and interpretable form. Here are to be seen all of those elusive features; formants, bursts, friction, etc. which before sound spectroscopy could not readily be examined and measured. In the past few years, however, speech research has been rapidly outgrowing the spectrograph. Not that it will cease to be a useful or even vital instrument, but greater versatility and sophistication are required. Speech researchers are realizing that the physical acoustic theory of speech production and linguistic theory, to name only two influences, are not mere abstract figments, and must underlie successful analysis and synthesis.

No longer are we satisfied to define formants as peaks in the frequency spectrum. Rather, we define them as vocal tract resonances and identify them as those resonance curves which, when added, accurately approximate the spectrum. No longer do we visualize a speech recognizer as operating inflexibly on its acoustic input. Rather, we think of its operation as contingent upon its internal structure—structure reflecting linguistic constraints and reaching at least to the syntactical level. Our speech synthesizers, too, partake of these constraints. Rather than merely copying acoustic patterns, we find ourselves manipulating symbols on the phonetic or phonemic level and worrying about interactions of their acoustic counterparts. This increasing sophistication is mediated in no small way by a new tool for speech research, the digital computer.

Digital computers are, in some ways, similar to the familiar desk-top adding or calculating machines. They deal exclusively with numbers; adding, subtracting, multiplying, and dividing them. Computers are unlike the desk calculator in two ways. First, the detailed sequence of operations to be performed must be specified in advance by a sequence of written instructions. This sequence, called a program, is inserted into the computer memory with the data to be processed so that once a computation is begun, it proceeds without human intervention. Second, computers can compare one number with another, and select the larger, or smaller, one. Subsequent operations can be made contingent upon the outcome of such a test. This ability gives the computer the facility for logic—if A then B, if not

A then C, etc. Herein lies the power of digital computation, for numbers so manipulated can symbolize any real world situation. In our case, the numbers can represent the amplitudes of successive samples of a speech waveform, the smoothed and sampled output from a spectrum analyzer, or phonetic symbols from the IPA alphabet. The computer, then, is a powerful symbol-manipulating device; potentially a speech-processor of great versatility or a speech generator incorporating linguistic constraints. It can handle masses of data which are stored in easily accessible form in its electronic memories. Thus, computers are much more than super desk calculators or slide rules.

However, it is certainly true that computers will not do things that cannot be done in principle by other equipment. They do make it possible to experiment in ways not previously considered feasible. There are many similar examples in history. The automobile is much more than a super horse. Yet given enough time and horses, a person can do just about everything that an automobile can. Still the automobile has changed our way of life profoundly because it brought mobility within easy grasp of the populace. A necessary ingredient was a good road system so that people could master driving with reasonable effort. We know, too, that automobiles can be misused; they can bring ills as well as benefits. Computers, also, must be accessible if they are to be effective, and regrettably they, too, can be used for nonsenscial purposes.

The key to accessibility is not physical proximity, but convenient programming systems so that the user does not have to be an expert programmer. In addition, versatile input and output–display equipment is required. The responsibility for worthwhile computer usage lies, of course, with the individual researcher.

The remainder of this paper discusses briefly both programming methods and input-output facilities tailored for speech research, then turns to a number of research results from computer studies.

Input-Output and Programming for Speech Processing

Speech must be represented in numerical form if it is to be computer-processed. The details of this conversion need not concern us here. Suffice it to say that speech either as a time waveform or a spectral distribution can be represented by a succession of close-spaced samples, the height of each being specified by a number. This succession of numbers can then be recorded on magnetic tape as binary pulses along with the punctuation marks necessary for computer input. This tape can then be played directly into the machine. In some computers the speech samples can be inserted directly into the machine's memory without an intermediate stage of recording. In either case, automatic equipment to accomplish the coding has been available now for several years. [2, 3]

The equipment works in reverse also, turning a succession of numbers into a speech sound wave with equal ease. Thus, if the computer is simulating a speech processor, a vocoder for example, its input and output can be compared directly by listening. Other available output displays include oscillograms, amplitude-frequency spectra, sound spectrograms, and indeed any other which the user chooses to program. These outputs are commonly prepared from a magnetic tape, obtained at the output of the computer, feeding a high-speed typewriter or printer, a pen galvonometer, a microfilm printer, or a cathode-ray tube. The out-

put data rate can be adjusted to match a particular display organ and is independent of the input rate. Thus a low-speed oscillograph can trace out a speech waveform which was originally sampled 10,000 times/second. In any case, computer–driven displays are capable of presenting results in forms well suited for human examination.

Most people outside the computing field don't know how to use and program a computer, and are loathe to learn, since programming can be tedious and exacting. A program must ultimately be a list of numbers, for this is the machine's language. It is not the language of people, and people find it exasperating to write numerical instructions. Also, their error rate tends to be unacceptably high. Such difficulties hardly encourage computer usage.

This obstacle can be overcome by setting up a language which the human programmer finds to easy to use. It commonly contains letters, punctuation, and pseudowords. A fixed correspondence between pseudosentences made up of these, and the numerical instructions required by the computer can be established.

This correspondence itself can be set up as a program, which is called a *compiler*. A computer with compiler will translate a program written in human-oriented language into machine language. The translated program can then be used to run the problem at hand.

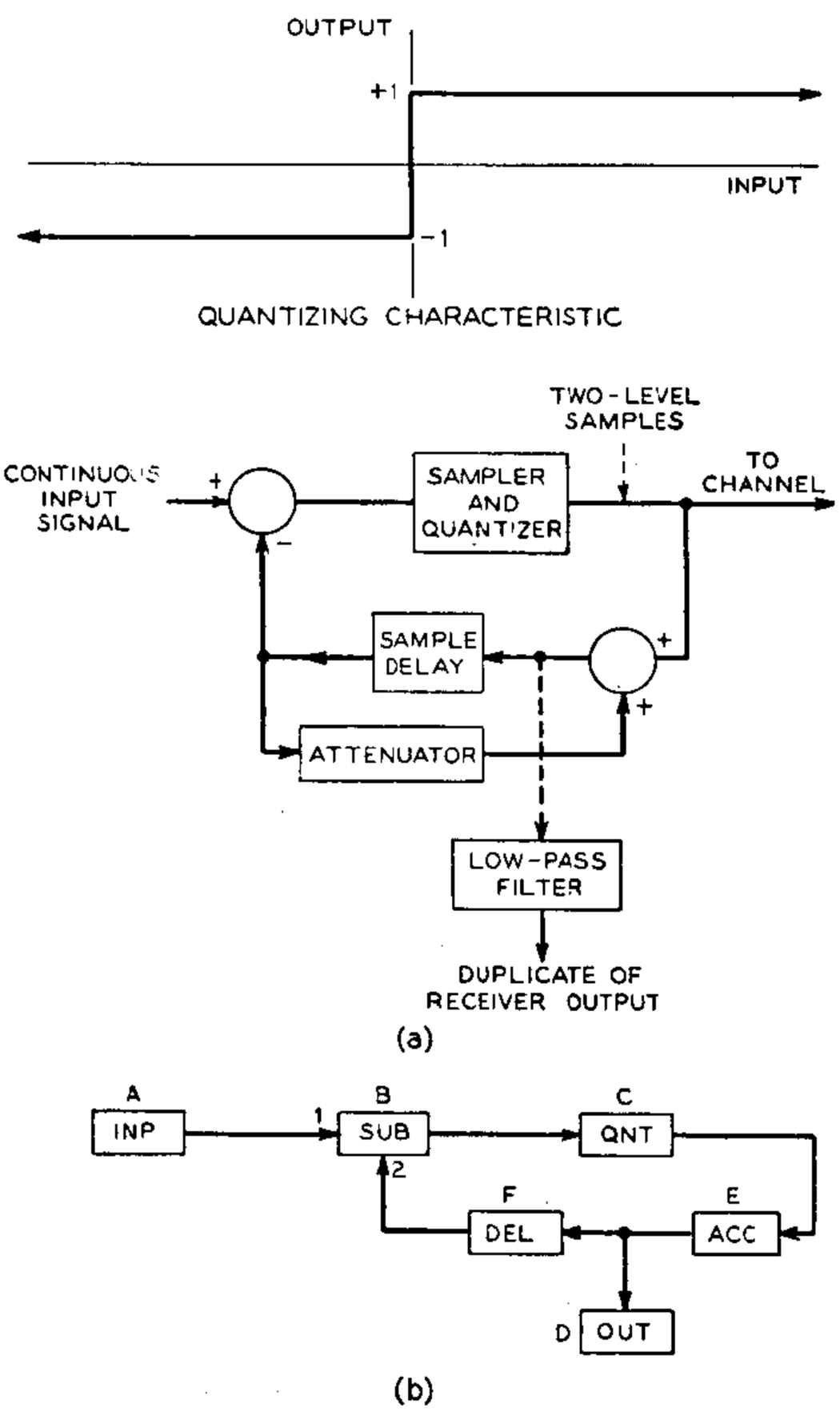


Fig. 1. Block diagram of a delta-modulation coder with an equivalent for simulation programming.

A compiler especially well suited for speech work is the so-called BLØDI program. [4] It assumes that the programmer has a block diagram describing the desired processing. The program he writes is merely a description of that block diagram. Consider a simple example. Suppose we want to investigate delta-modulation speech-coding. [5] A block diagram of a delta-modulation coder is shown in Fig. 1(a). For programming purposes, it can be redrawn as shown in Fig. 1(b). Input and output blocks have been added and each block labelled with an arbitrary letter. The program consists of a three-column listing of the blocks. Column 1 contains the block label, Column 2 is a mnemonic specifying the block function, and Column 3 lists the block parameters and the destinations of the block's output connections. The order of the listing is immaterial. A BLØDI program for Fig. 1(b) might begin with the input:

```
A        INP        B/1
```

The input block is designated A, its function is denoted by INP, and its output goes to input 1 of block B. The code for the subtractor, block B, is

```
B        SUB        C.
```

The quantizer has two output levels and one input decision (transition) level (see Fig. 1(a)). These must be listed in Column 3 beginning with the lowest output level and progressing alternately through all the input and output levels. For the quantizing characteristic of Fig. 1

```
C        QNT        — 1, 0, + 1,   E.
```

Block E is an integrator, or accumulator, with a decay constant, a, and is coded

```
E        ACC        a, D, F.
```

Block F is a delay line with 1-sample delay,

```
F        DEL        1, B/2,
```

and finally the output is specified.

```
D        OUT
```

These six instructions constitute the input program. From it, the computer prepares a longer numerical program including all the necessary detailed operations and data shuffling. Compilers provide the "open sesame" to computer usage. They adapt a general-purpose computer to particular problem classes. Where large numbers of similar problems are to be solved, a compiler is practically a necessity. Through BLØDI, researchers have been able to program their own experiments in such diverse fields as speech synthesis, speech analysis, vocoder transmission, stereophony, artificial reverberation, architectural acoustics, and audition. However, other programming languages, such as Fortran, [6] have proved useful in speech work, and others will be called for as the scope expands.

Speech processing involves a large amount of data. For instance, ten seconds of speech waveform may be represented by 10^5 samples or about 10^6 binary digits. All of these must stream through the machine during the solution to a problem. This situation demands a fast computer with considerable memory if the processing is to be finished expeditiously.

The IBM-7090 is such a machine, and it was used in obtaining the results reported in the remainder of this paper. The running time for a problem depends, of course, on its complexity. Typically, it takes between 10 and 50 times as long to process a speech sample as it does to speak it. Such time scales are economically feasible. In our opinion, time scales of 1000 to 1 or greater are not.

Despite their potential for speech research, computers are hard taskmasters. They require that we examine our wants with unaccustomed precision. Yet by encouraging precise thinking, they have led us to a deeper understanding of speech processes. Now we turn to some results.

An Objective of Speech Research

A long standing aim of speech research has been to reduce its complex time waveform to a number of simpler acoustic features. One level of reduction relies upon recognizing perceptually-significant features related to the speech-production mechanism. Temporal variations of these features are closely synchronized with the phonetic units of spoken language. Thus they provide the springboard for a further reduction; namely, to the linguistic level. It is upon these transformations that reduced bandwidth voice transmission and man-machine communication by voice hinge. The remainder of this paper will demonstrate the relevance of computer methods to these goals. Traditionally speech processing begins with a measurement of the spectral magnitude. This reduction is justified on two counts. First, the speech mechanism is constrained by its slow gestures to produce a succession of acoustic quasi-steady-states. These are produced when a wide-band excitation wave is filtered through the slowly-varying vocal tract. Thus, there are intervals over which time-invariant spectral analysis is meaningful. Second, the auditory mechanism does not perceive modest phase displacements between speech components, and so only the magnitude of the spectrum is of prime concern. Interestingly, even this ancient ritual of speech analysis has come under finer control through computer usage.

Spectral Measurement

Measurement of a spectrum implies "averaging", or analysis, over some interval of time. Strict Fourier analysis considers an infinite interval and achieves an exact spectral determination. In speech, finite intervals only are available, and thus there will be uncertainties in the spectral estimates. These arise from several causes:

1. For an interval Δt long, Fourier analysis cannot resolve spectral features closer than $\Delta f = 1/\Delta t$ cps apart, nor can an analysis band Δf wide resolve temporal features closer than $\Delta t = 1/\Delta f$ seconds apart.
2. In sounds with a random noise-like structure, the spectral estimate at any frequency is subject to an uncertainty [7] whose standard deviation is proportional to $1/\sqrt{\Delta f \Delta t}$. For 250 cps resolution and a 50 millisecond window, the *root-mean-square* error of measurement is about 1 db.

3. In sounds, such as vowels, with a quasi-periodic waveform, the relation between the duration of the quasi-period and the analyzing interval, or "window", must be taken into account to avoid introduction of spurious spectral features.

The first two factors are basic limitations on resolution and accuracy. Computer analysis, through choice of program parameters, permits precise control of these factors. This feature facilitates analysis pointed either at steady-states or dynamic variations in the spectrum.

The third factor pertains to sounds such as vowels and nasals. Ideally, they can be considered the result of a periodic pulse sequence exciting a time-invariant network, as shown in Fig. 2. Their idealized spectrum consists, then, of harmonics whose amplitudes are determined by $v(j\omega)$, the network response. For representation, we would like to separate $v(j\omega)$, the spectral envelope, from the harmonic fine-structure. However, it is intrinsically impossible to obtain directly the continuous curve $v(j\omega)$ because of the repetitive nature of the waveform. [8] On the other hand, by choosing the analyzing aperture equal to a period, it is possible to measure points *on the envelope*, points spaced at the reciprocal of the period length. A fixed aperture, such as provided by a filterbank or a scanning spectrograph analyzer, does not in general meet this condition. These instruments resolve the fine-structure or lose spectral definition depending upon the relation of voice pitch to spectral-window duration.* The spectral features so introduced are difficult in some cases to separate from actual variations in $v(j\omega)$. With a pitch-synchronous analysis, this problem is avoided. In connected speech, where sounds are only quasi-periodic, pitch synchrony preserves period-to-period temporal resolution as well. We have used this technique in many experiments to control spectral accuracy.

Initially, the speech time-waveform is segmented into pitch periods (by visual inspection). These are then expanded by computer into a Fourier series. The coefficients specify the desired spectral points. Several pitch-synchronous spectra are shown in Fig. 3. Their shape is not simple; many maxima and minima appear. Coupled with the acoustic theory of speech production, computers have aided in deciphering these forms.

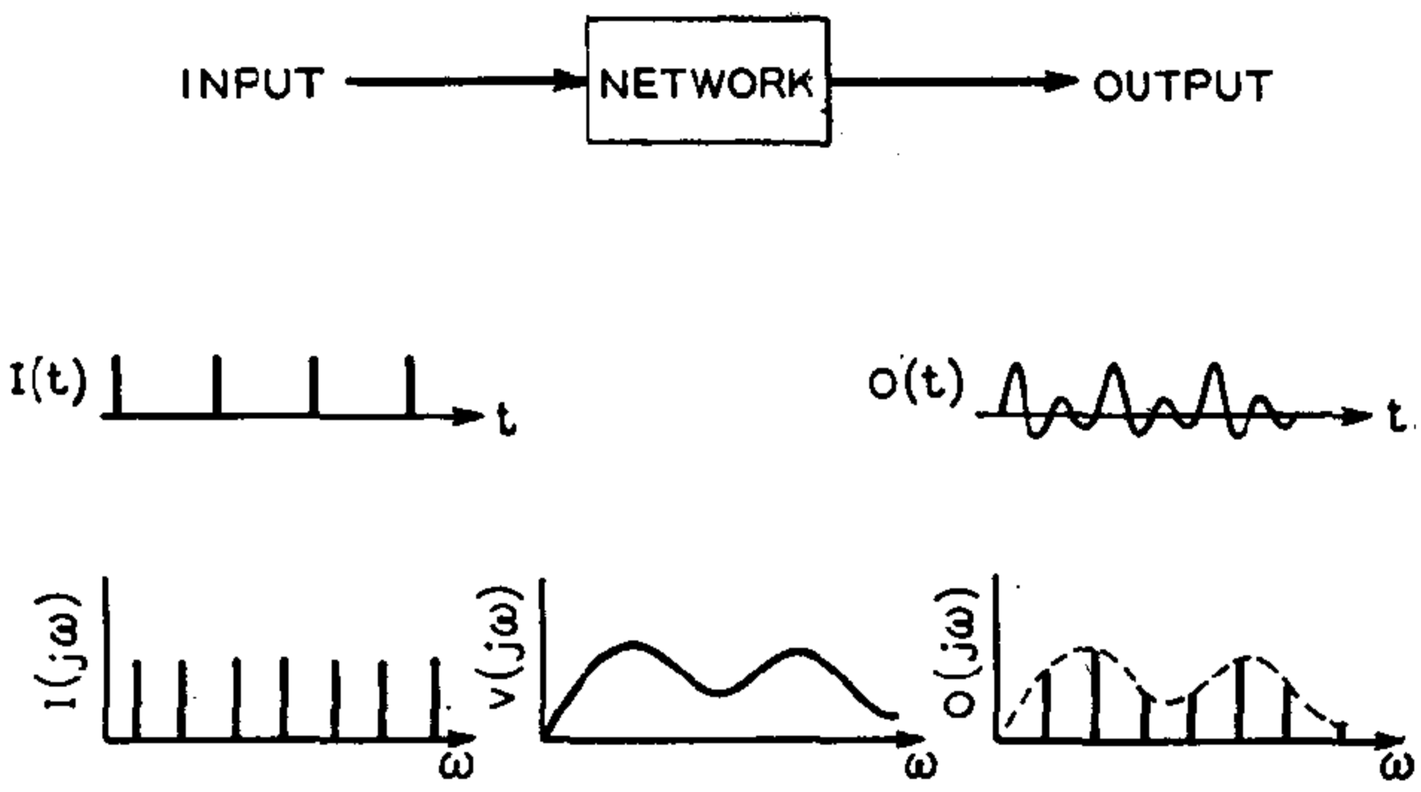


Fig. 2. The spectrum of a voiced sound can be considered to be an envelope enclosing a harmonic fine-structure.

* The equivalent of a pitch-synchronous analysis can be obtained by computation on the output of a filterbank, only, however, with knowledge of pitch and only if there are at least as many filters as harmonics.

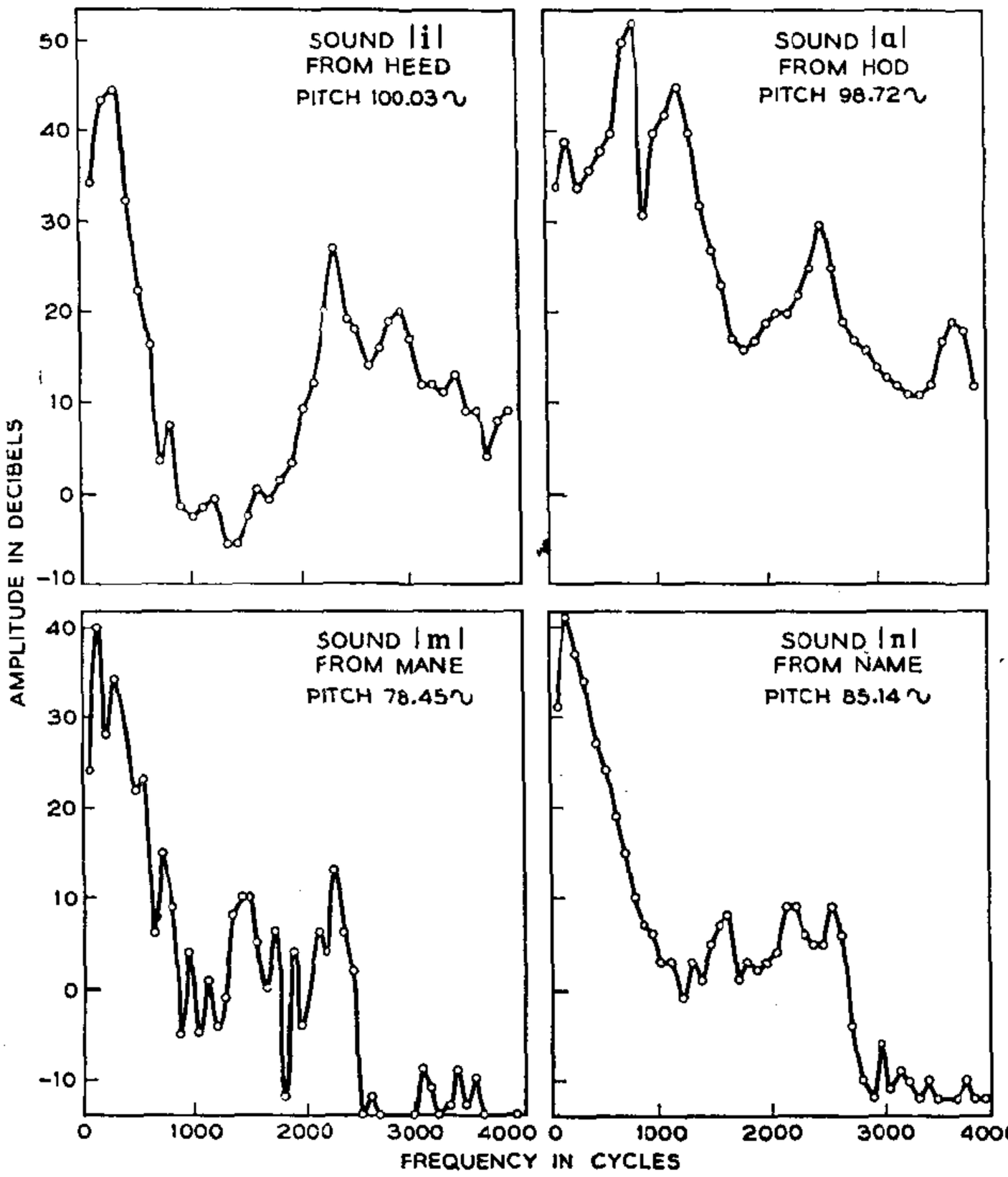


Fig. 3. Pitch synchronous spectra of vowels and nasals show many maxima and minima.

Spectral Representation

The acoustic theory of the vocal tract predicts the speech spectrum. Basically, an energy source is assumed to excite an appropriately shaped tract. Because the source and tract are thought not to interact strongly, the output spectrum is the product of the excitation spectrum and the vocal tract response taken separately. The latter has been derived for various sounds considering only longitudinal modes of vibration in the tract. This simplification leads to a one-dimensional model which specifies the cross-sectional area of the tract as a function of position from glottis to lips. Its response can be specified by poles and zeros in the complex frequency plane. For vowels, the vocal tract is driven by the vocal cords and acts as pipe of nonuniform cross-section, which has poles but no zeros (see Fig. 4). [9] In the case of some consonants, the tract introduces zeros as well as poles. In fricatives, the zeros arise from the cavities behind the source. [9] In nasals, zeros are introduced by the oral cavity which parallels the nasal path. [9] This body of theory, together with cross-sectional areas and excitation spectrum, leads to a predicted spectrum for any given sound. In reconciling predicted with measured spectra, the computer is an apt tool. In essence, it can search for a pole and zero pattern most favorable for approximating or matching the measured data.

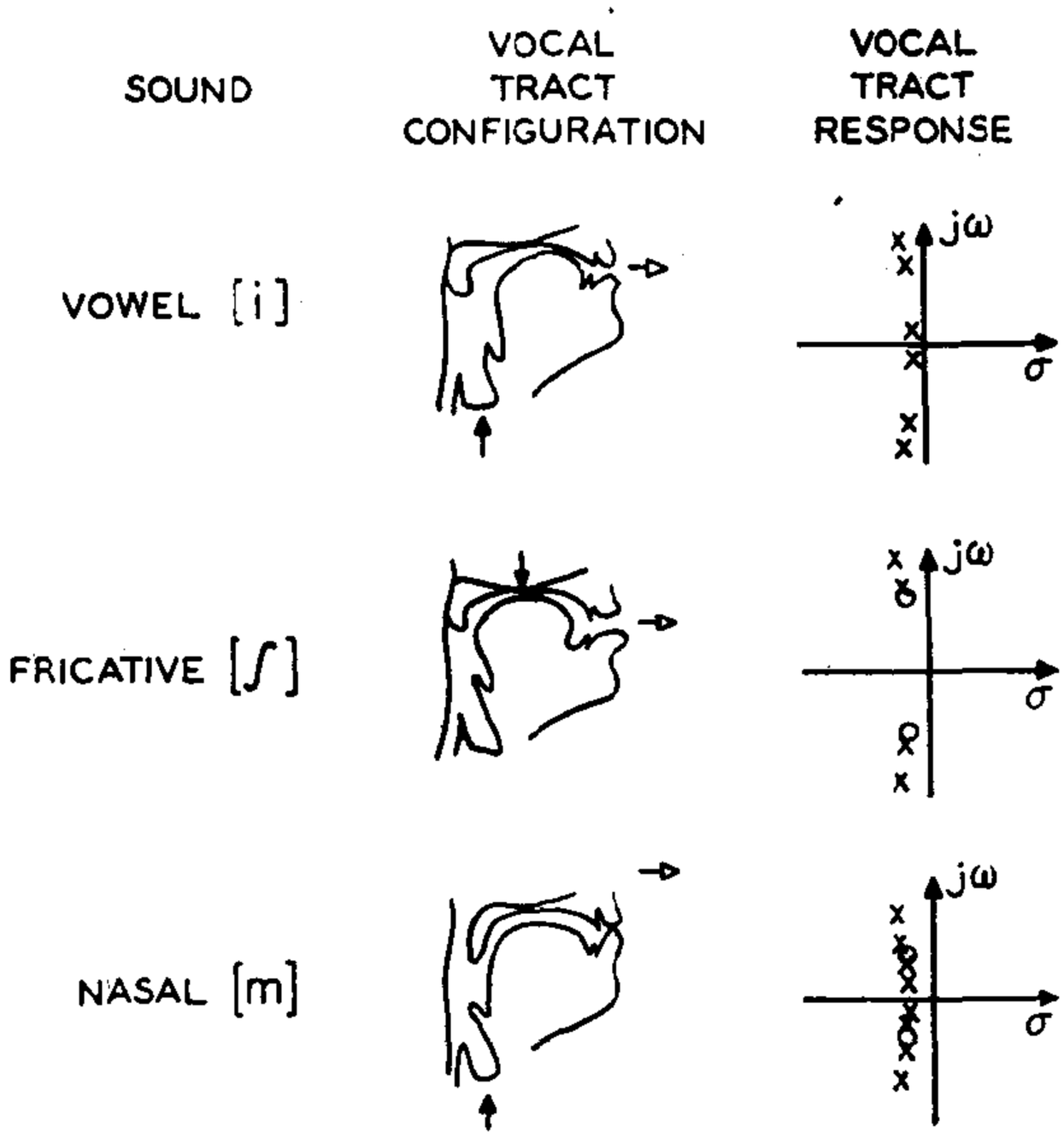


Fig. 4. The canoncial spectral representations of various sound classes can be derived in terms of poles and zeros from the appropriate vocal tract configurations (adapted from Bell, et al; Ref. 14).

Spectral Matching

Spectral matching has a long history. [8, 10–15] Early attempts utilized the experimenter's intuition to determine pole-zero positions and degree of fit. Computers provide the means for examining easily a large number of matching spectra using a quantitative measure of fit, while keeping human effort to a minimum.

In the case of vowels, the matching spectrum (excluding the source) should include three or four poles plus factors for the lip radiation impedance and the higher-frequency vocal tract poles. Specifically, we use the standard expression,

$$V(s) = K_1 \left(\frac{s}{s + s_r}\right) \left[\frac{1}{\frac{4}{\pi} \prod_{n=1} (s - s_n)(s - s_n^*)}\right] e^{K_2 s^2}$$

where s is the complex frequency variable, s_n are pole positions, K_1 is an amplitude constant and K_2 is a constant of the higher-order pole correction. In terms of the previous discussion, $V = K_1$ (radiation effect) (lowest four vocal tract poles) (higher-order pole correction).

To obtain the final expression these terms must be multiplied by the excitation spectrum. Certain reasonable assumptions about the shape of the glottal wave show that it has only zeros within the frequency region of interest. [8] Such a factor may be written

$$G(s) = \prod_i \pi (s - s_i)(s - s_i^*)$$

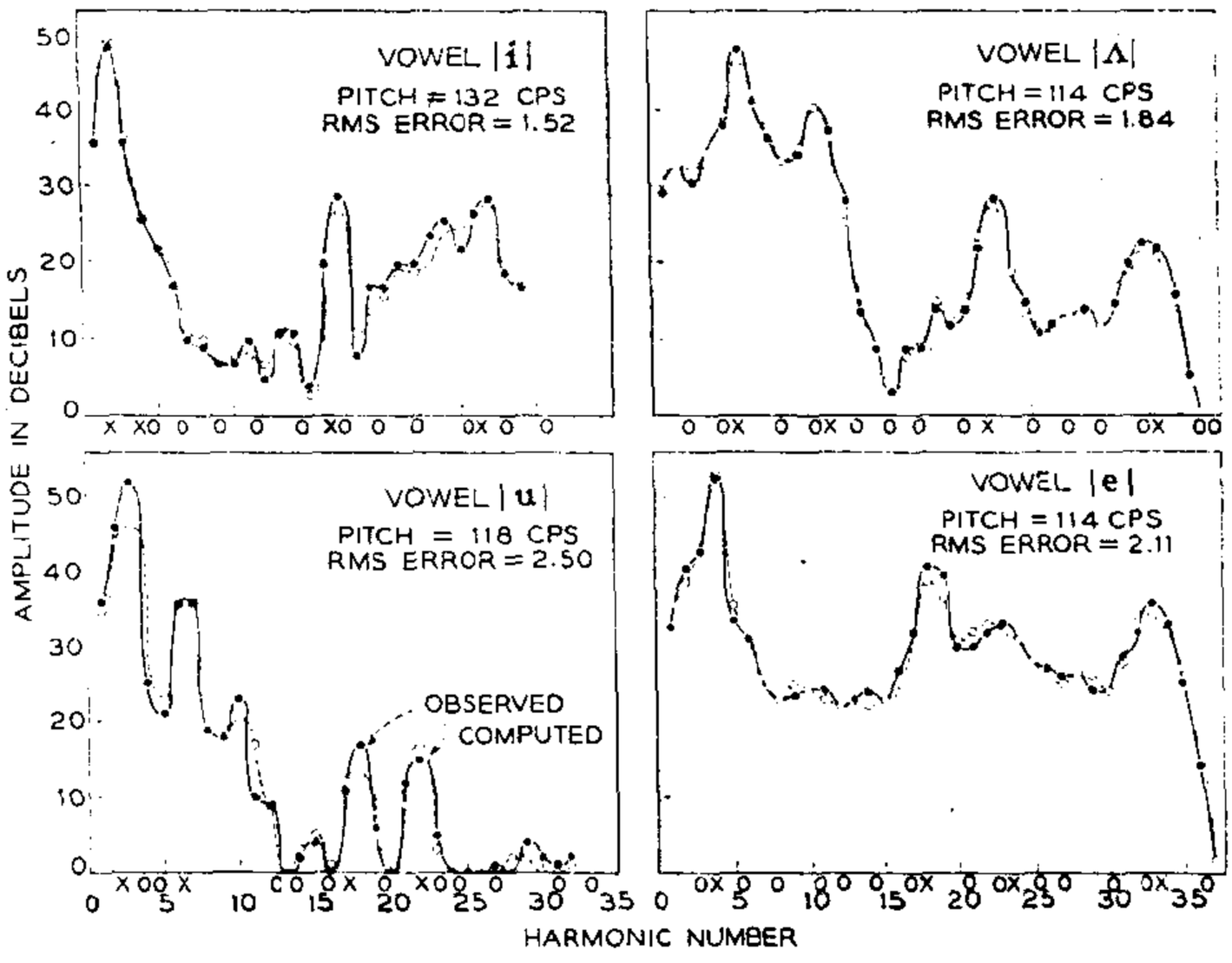


Fig. 5. Using both vocal tract poles and glottal zeros as well as radiation and higher-order pole corrections, the spectrum of vowels can be closely fitted.

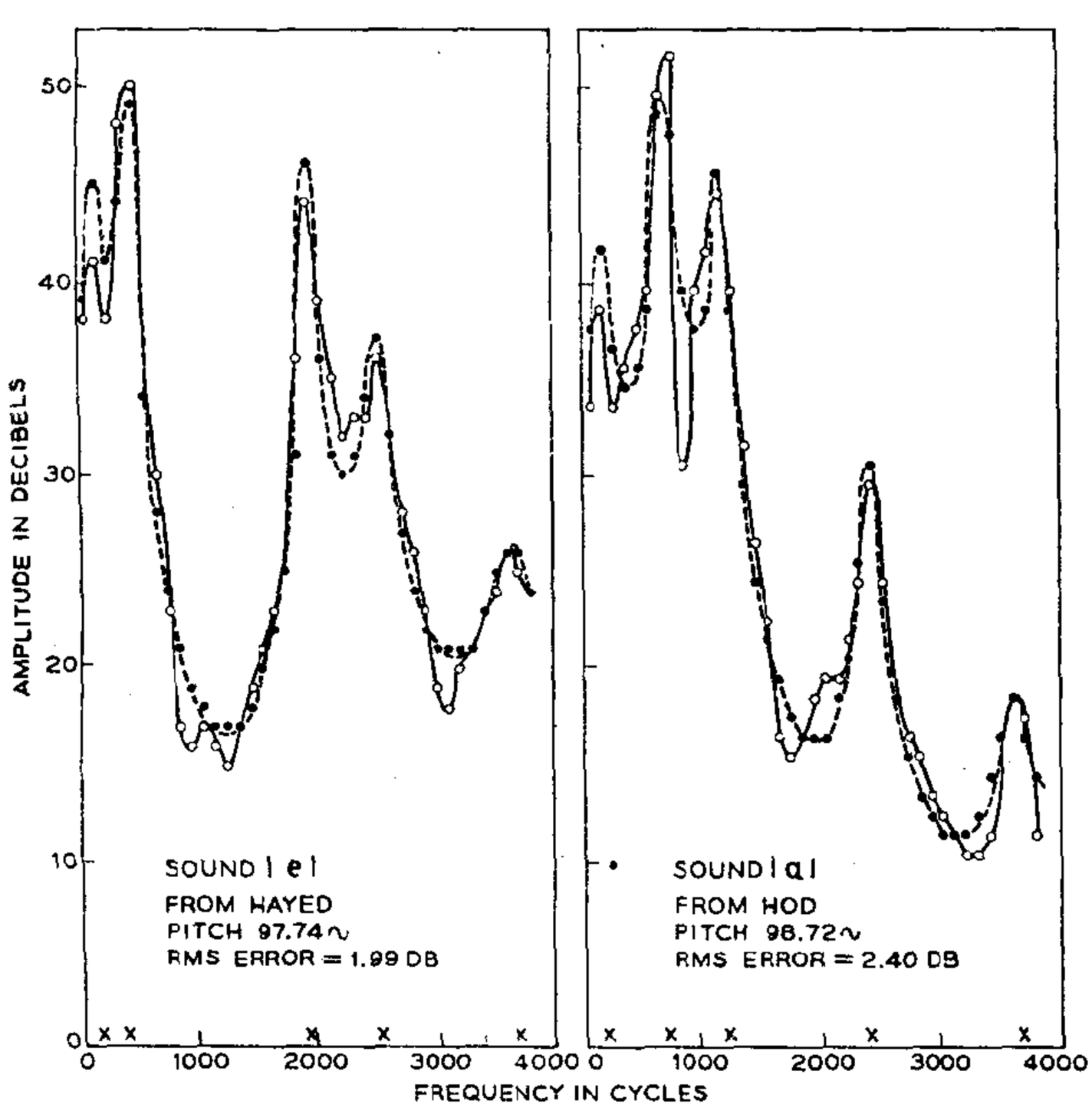


Fig. 6. Ignoring glottal zeros, vocal tract poles provide a good, but not detailed spectral match to vowel spectra.

where the s_i are zero positions. The matching spectrum is then

$$P(s) = V(s) \cdot G(s).$$

In our first vowel-matching studies, the initial pole and zero positions were estimated by eye from the pitch-synchronous spectra. The values of $P(s)$ were computed for $s = jm\omega_0$, where ω_0 is the fundamental frequency of the measured spectrum and m is the harmonic number. These values were then compared with the measured spectrum. The root-mean-square error in decibels is

$$\text{Error} = \left[\frac{1}{M}\sum_{m=1}^{M}[\log|P(jm\omega_0)| - \log|P_d(jm\omega_0)|]^2\right]^{1/2}$$

where P_d is the measured spectrum and M is the total number of harmonics in the spectrum. After the initial pole and zero positions were selected, both their frequency and damping were varied according to the method of successive approximations to minimize the error.

Typical results are shown in Fig. 5. The root-mean-square error in these fits is less than 2 decibels. Both the pole and zero positions are evident from the spectral shape in many places. However, as one expects, zeros produce less effect than poles. Thus, the matching procedure can be simplified by ignoring the glottal zeros, and fitting only the poles and correction terms. Results are shown in Fig. 6. Here the gross effect of the excitation spectrum was included as single pole at 150 cps with a damping of 30 cps. Note the absence of spectral detail in the matching spectrum. Here the rms error is still typically about 2 db.

The poles of the vocal tract spectrum can be divided out of the measured spectrum leaving the glottal spectrum. Examples are shown in Fig. 7. Their irregular shape is notable as is their conventional falling characteristic.

Since these spectra were obtained by Fourier analysis, they can be reconverted to a time waveform by Fourier synthesis. Glottal waves so obtained are shown in Fig. 8 along with the corresponding speech waveforms. The largest oscillations occur during the closed portion of the glottal cycle, presumably reflecting the epoch of maximum energy input and because subglottal damping is removed. The relative timing of these intervals reveals a transport delay, corresponding to the time required for the sound to travel from the glottis to the lips. The glottal waveform shows a more rapid closure than opening. This characteristic seems typical of many talkers.

A similar matching procedure has been applied to nasals. The theory predicts four poles and a zero to be significant in the vocal tract spectrum below about 3000 cps. Radiation, excitation, and higher-order pole correction were used also. The initial positions and values were picked by eye and adjusted by computer. Matches for four cases are shown in Fig. 9 and yield rms errors of from 2 to 4 db. Again, the complexity of the glottal spectrum gives rise to many minor maxima and minima.

These results, as well as those of Fant at RIT, Stockholm, and Stevens et al, at MIT, show that the acoustic theory of speech production permits us to account in detail for spectral patterns. In capitalizing on this fact to extend our quantitative knowledge of speech parameters beyond the vowels, computers will be the key.

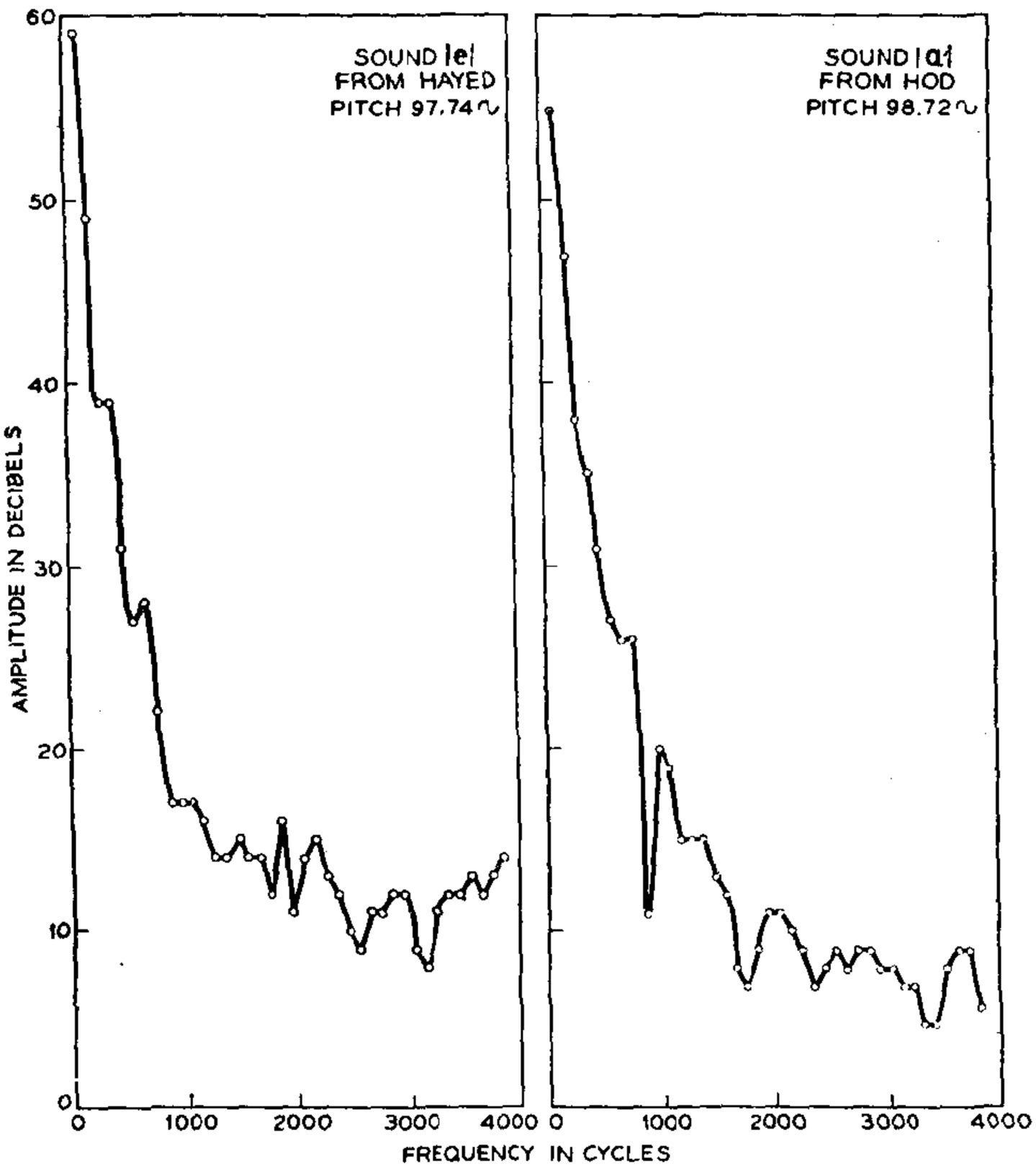


Fig. 7. If the vocal tract poles are removed from measured vowel spectra, the glottal spectrum remains. Note the irregularities.

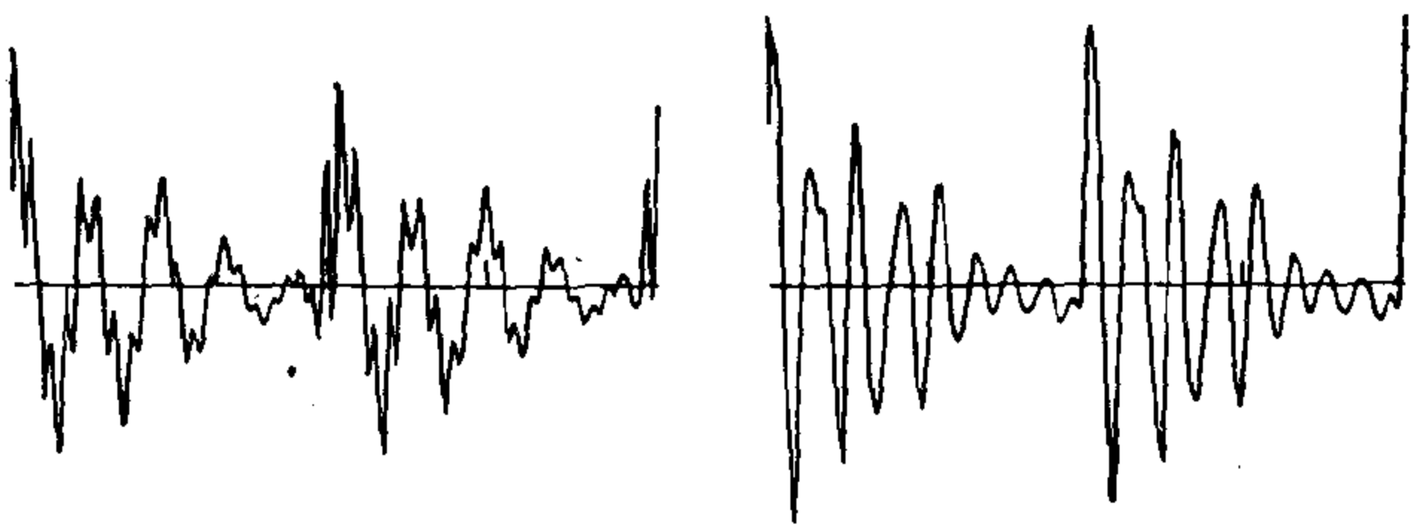


Fig. 8. Vocal cord waves can be obtained from the glottal spectra. The corresponding speech waves are shown for comparison.

Automatic Spectral Matching

In the spectral fitting described above, the computer was used merely to adjust the investigator's choice of initial pole-zero position. For speech coding and recognition, an entirely automatic procedure is needed.

The multi-dimensional nature of the matching problem dictates fundamentally different strategies for initial choice and adjustment. One reasonably successful choice strategy for vowels relies upon the spectral maxima commonly produced by poles. It proceeds by evaluating the fit (error) over all possible ways in which poles can fall on the several maxima found in the spectrum. The best fit, as indicated by minimum error, is taken as the initial guess, and is then adjusted for a more precise fit. The sensitivity of the method can be judged

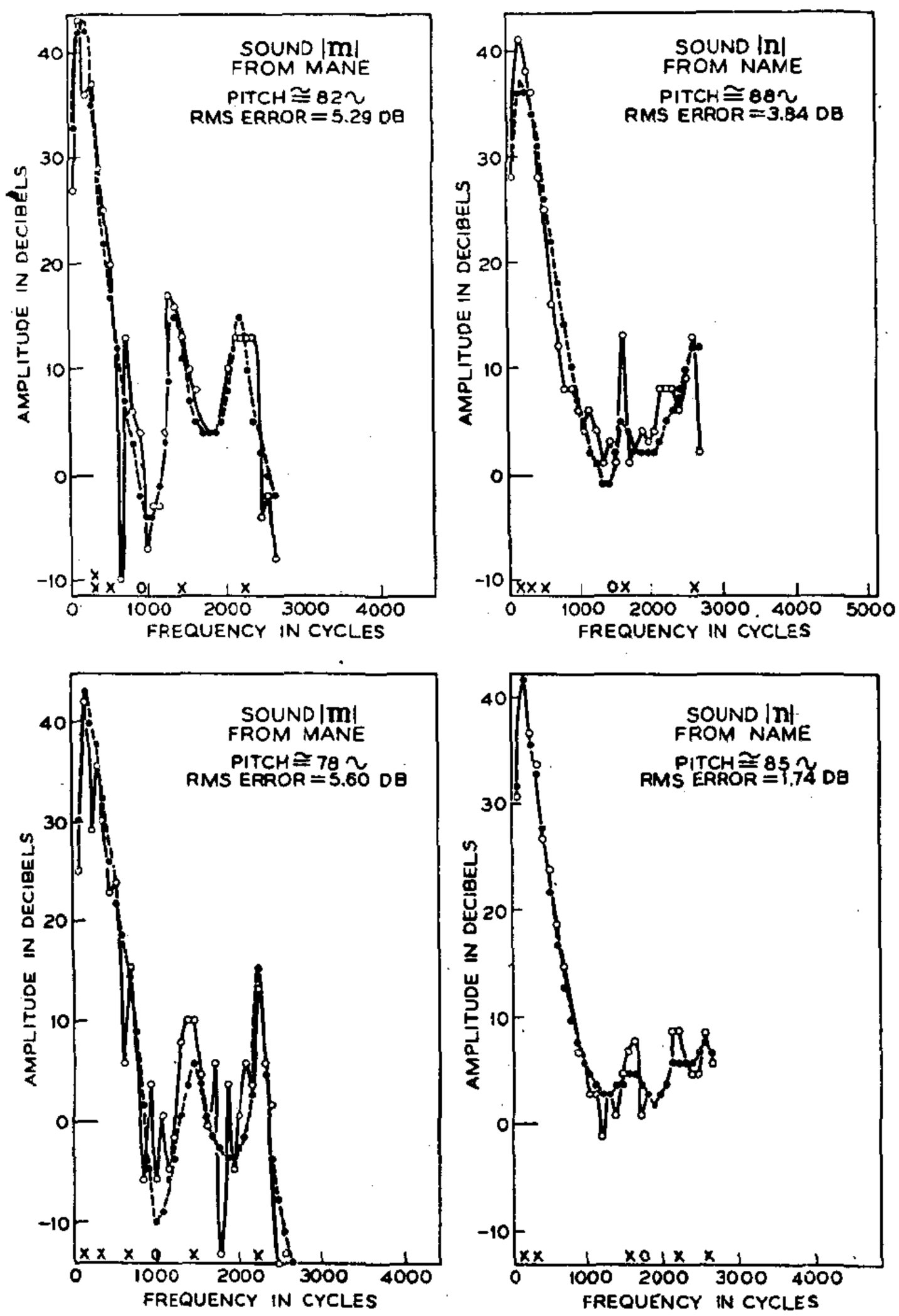


Fig. 9. Nasal spectra can be matched using 4-poles and one zero as the vocal tract response.

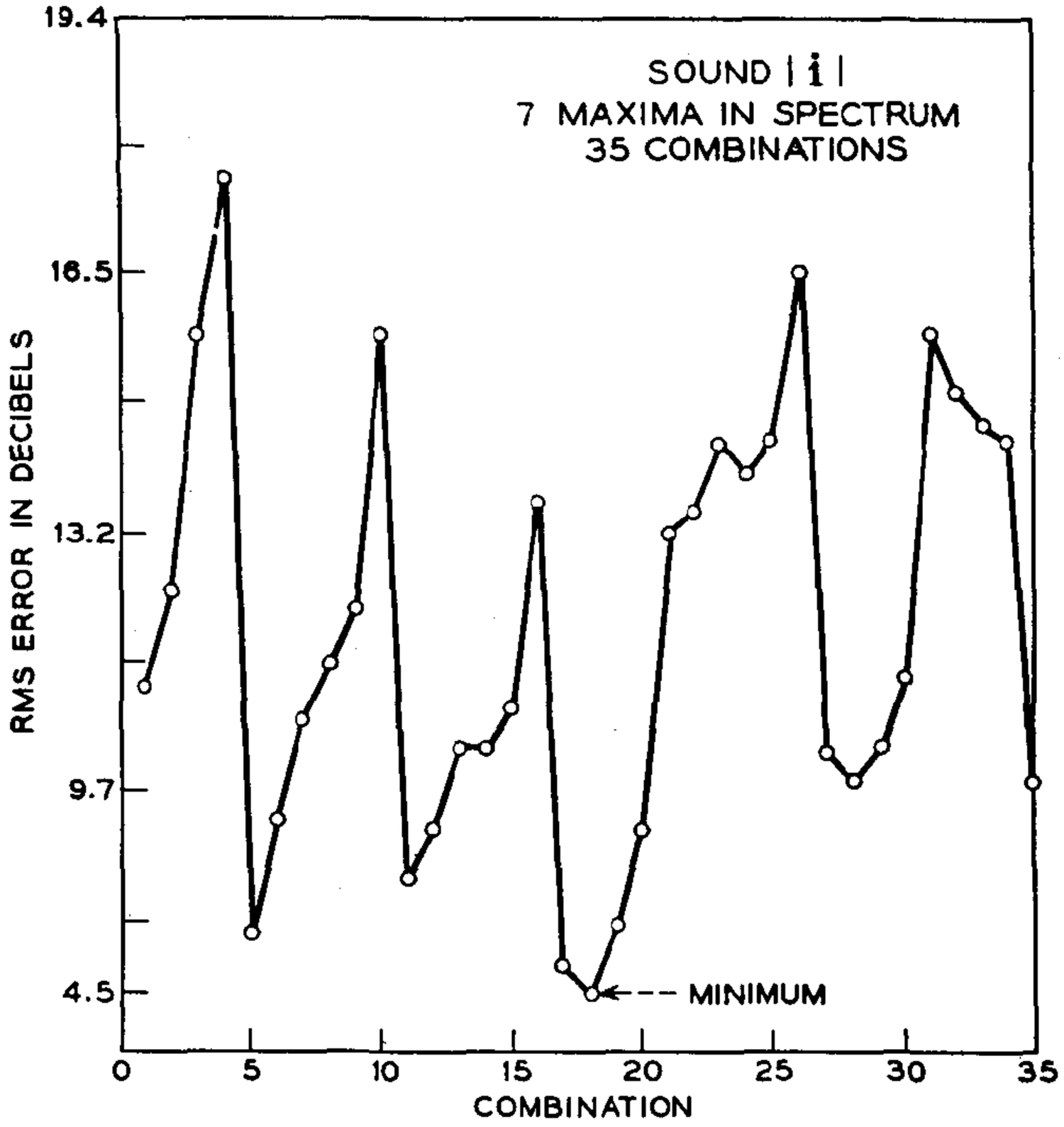


Fig. 10. Testing 4-poles in 35 position configurations yields an error distribution which clearly indicates the correct choice for spectral matching.

by examining the errors over the ensemble of possible positions, as shown in Fig. 10. Here 4-poles were used to fit 7-maxima; thus there were 35 choices to evaluate (in this case, two or more poles were not allowed to fall at the same maximum). As can be seen, the correct choice yields substantially less error than its nearest competitor. Yet, this strategy does not always give the correct answer. Difficulties may arise from several causes. For instance, vowels may be partially nasalized, or the glottal spectrum may depart markedly from the assumed form.

A similar strategy, with the addition of a zero to be tested at spectral minima, has been applied to fricatives. [16, 17] Two poles, a zero, and a radiation factor were used. Typical matches, shown in Fig. 11, give a good fit on the average, but many spectral details remain unaccounted for. These may arise from close-spaced pole-zero pairs contributed by cavities behind the fricative source. Also, the excitation spectrum has been assumed flat in these matches, and may actually contain significant detail.

In searching for an initial choice, the computer examines many positions for the zero. Yet, even in its most favorable position it has little effect on the spectral shape. Thus, additional matches were undertaken using two poles only. These are shown in Fig. 12 for the same sounds as in Fig. 11. As can be seen from the error figures, the zero does not materially

contribute to the fit. Thus, a simplification of the vocal-tract model for fricatives has been established.

These examples illustrate how computers can facilitate automatic representation of spectra. However, automatic analysis is aimed ultimately not at isolated spectra but at connected speech. The techniques outlined above begin with a canonical pole-zero pattern for the sound in question, then search and fit. In applying them to connected speech, either an overall canonical form or segmentation before fitting is required. Any overall form encompassing all sound classes will contain many parameters and lead to a lengthy search. So, segmentation has been the standard approach. Current methods do not fully utilize knowledge of speech production. [18, 19, 20] Their effectiveness may well be increased by noting the constrained shapes predicted by vocal tract models for various sound classes.

Our experiments with segmentation and spectral representation indicate that any automatic-technique is likely to err occasionally. Our models are, of course, incomplete, and uncertainties in measured spectra from noise and reverberation increase the probability of error. Errors can be reduced by utilizing sequential constraints. Early efforts along this line [21, 22] were based upon continuity of pole, or formant, movements, and showed that

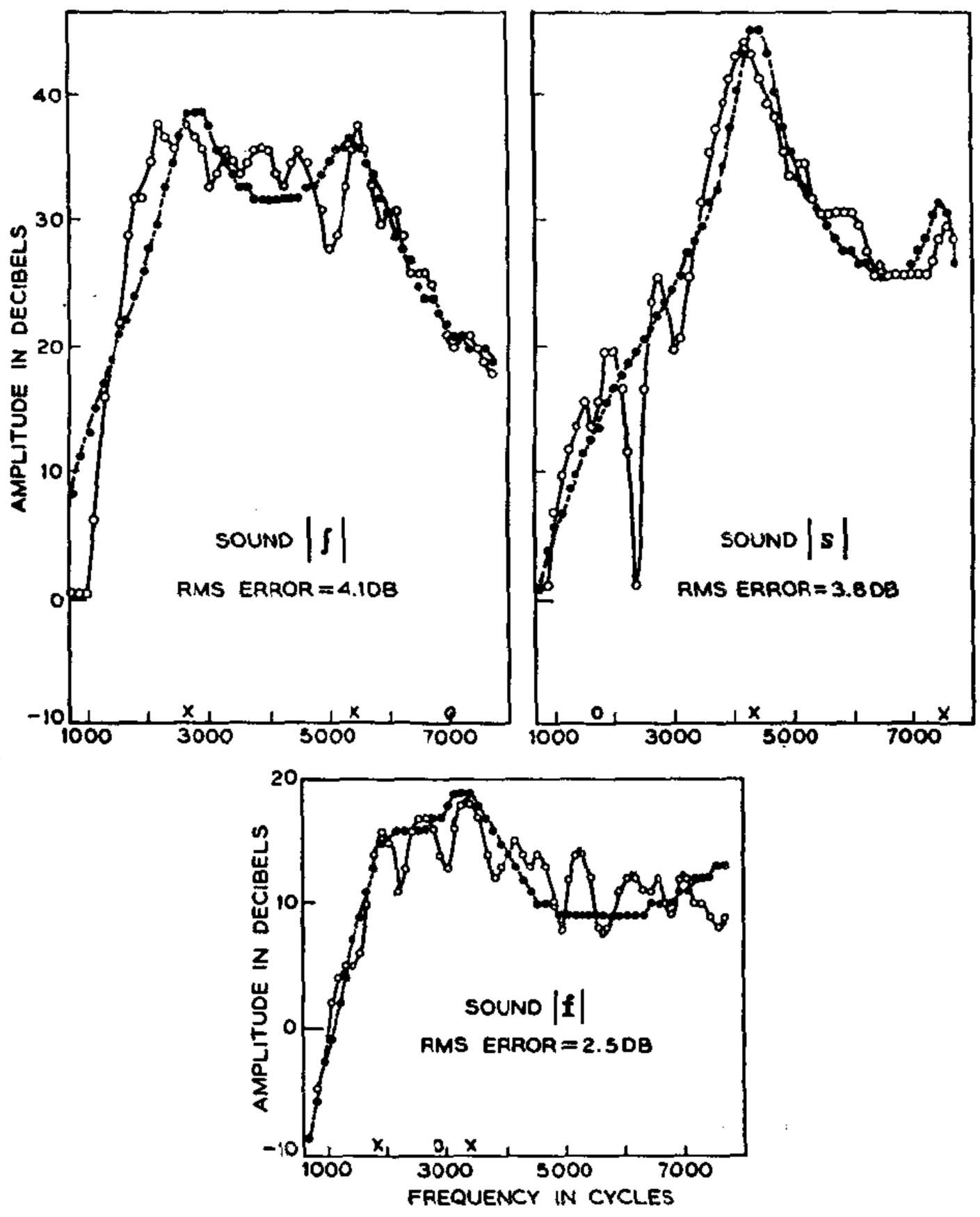
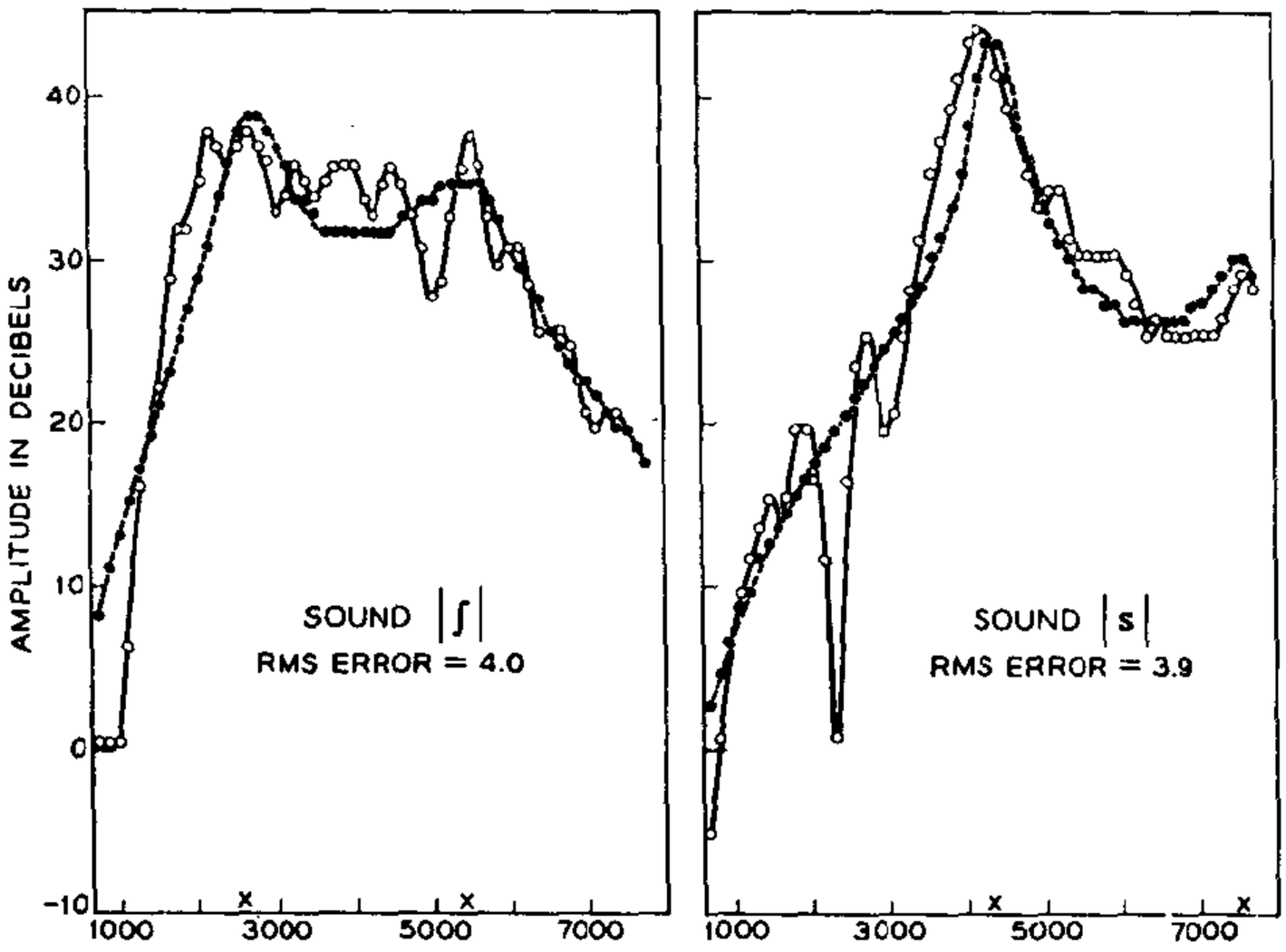


Fig. 11 .Fricative spectra can be matched with two poles and one zero.



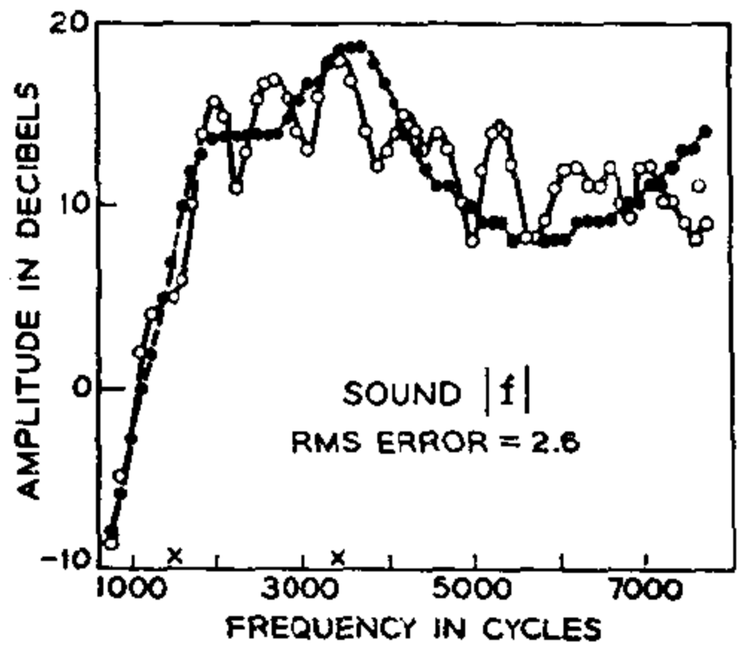


Fig. 12. With only two poles, fricative spectra can be approximated as closely as with two poles and one zero.

errors could be so corrected. However, "continuity" should be related to vocal tract dynamics if the full benefits are to be achieved. Here, again, computers are playing a central role.

Speech Dynamics, Synthesis, and Recognition

In the acoustic structure of speech, the quasi-steady-state aspect is, perhaps, less than half. There are, in addition, two fundamentally different dynamic aspects. First, there are transitions from one spectral shape to the next. These provide cues in the perception of some sounds which could not be recognized from any steady-state segment. [23] Experiments, particularly by the Haskins Laboratory of New York City, have provided much understanding of these cues. [24] Second, the acoustic pattern of any speech sound is strongly influenced by its neighbors. Both transitions and "intersound" influences reflect constraints in vocal tract dynamics. In addition, language structure provides limitations upon the succession of acoustic events.

A recent speech-synthesis study by J. L. Kelly and L. J. Gerstman [25] in which these factors are clearly separated, illustrates the power of computers in studying such complexities. Their basic synthesizer, Fig. 13, is a computer-simulated terminal vocal-tract analog with three resonances controllable both in frequency and damping. Excitation from hiss and buzz sources is controlled in amplitude and frequency. All told, nine control signals are required to drive the simulated machine.

In similar machines, such as Lawrence's PAT, [26] these signals are taken from slides or records prepared for the purpose. In Kelly and Gerstman's machine, they are derived automatically from an input sequence of phonetic symbols, each carrying a duration and pitch value. This process is shown in outline form in Fig. 14. The central feature is a stored table which for each input symbol contains a list of eight control values (pitch excepted). These represent steady-state acoustic parameters, or targets, for the control signals. The pitch value accompanying each symbol completes the target specification. An input symbol sequence, therefore, determines the control signals at a sequence of temporal points.

Two duration values are supplied with each input symbol. One value gives the duration of

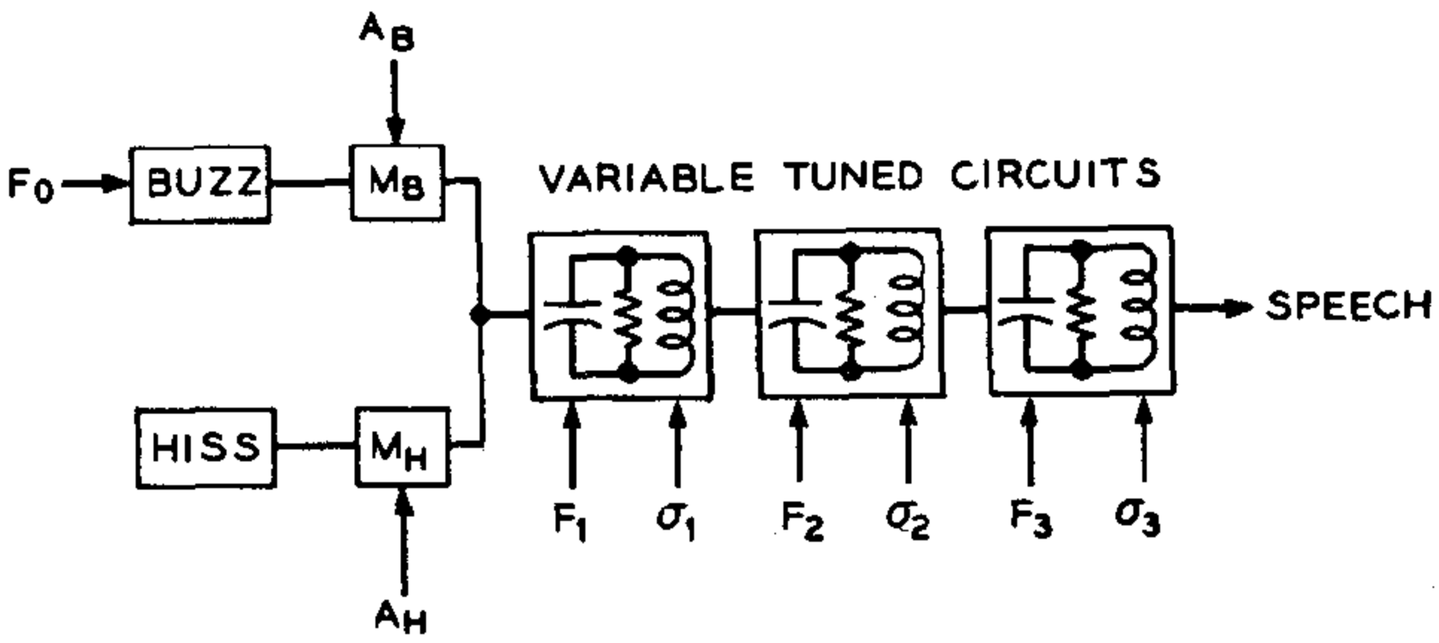


Fig. 13. Nine control signals are required to control the terminal vocal-tract analog synthesizer used in the Kelly-Gerstman talking machine.

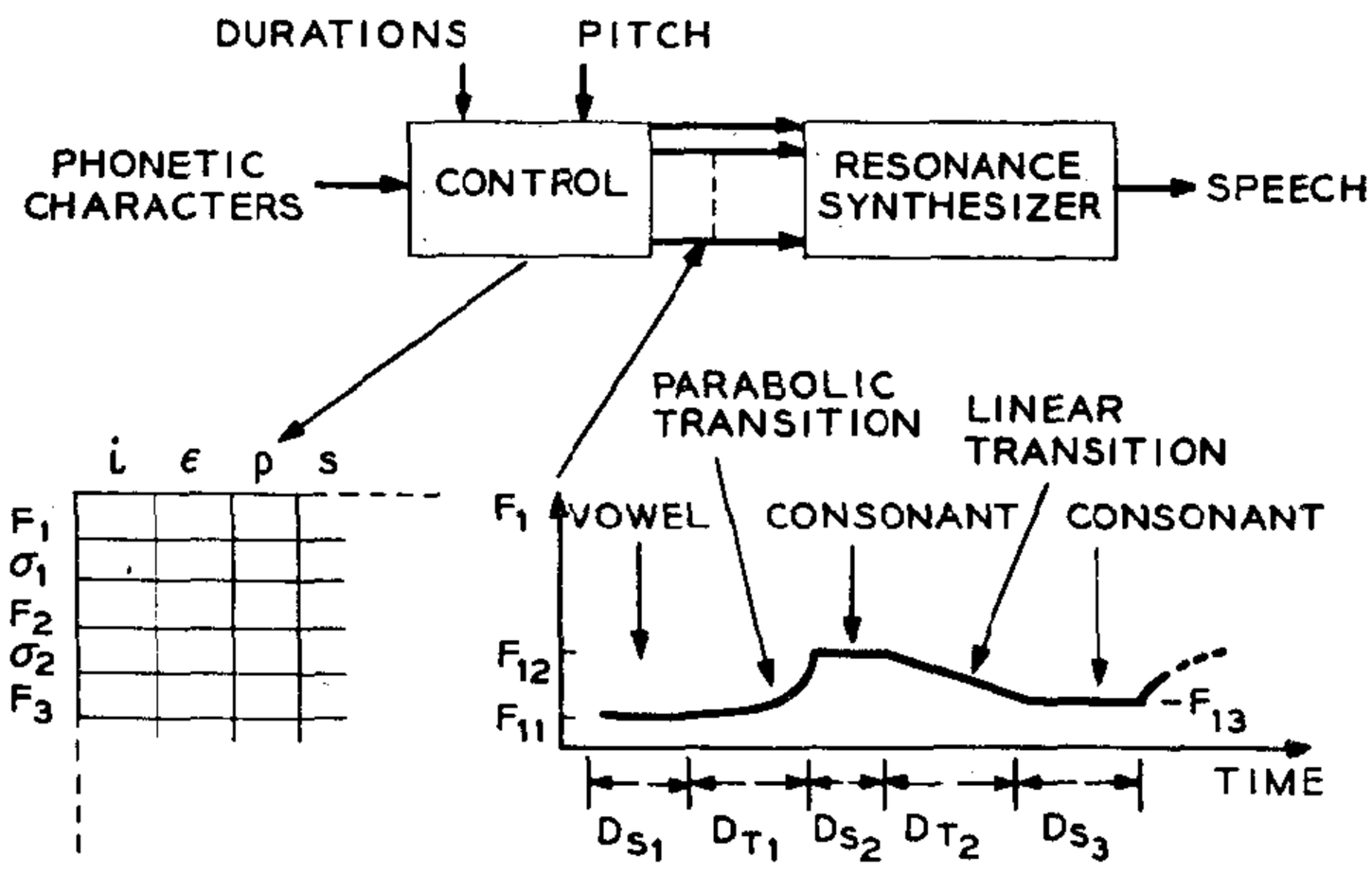


Fig. 14. The Kelly-Gerstman talking machine synthesizes continous control signals from a discrete symbolic input.

each steady-state, while the second specifies the transition duration between steady-states. The transitions themselves are interpolated according to rule. Parabolic curves are used between vowel and consonant, while linear transitions serve between vowel and vowel, or consonant and consonant. This algorithm synthesizes continuous control signals from the discrete symbolic input.

Speech from this synthesizer consists only of steady-states and transitions. Any effect of linguistic structure on the sounds produced is limited to the selection of input symbols, durations, and pitches. Also missing is the interaction of articulatory positions in connected speech. The main effects here are abbreviated formant excursions and temporal coalescing of the usual acoustic patterns. [27] Some articulatory interaction can be included by requiring the steady-state values to depend not only on the immediate input symbol, but also on the ones preceding and following. Such an expanded machine is presently on the programming boards. [28] Temporal and linguistic influences are more difficult to include. They involve structure underlying the input symbol sequences and their relation to prosodic features such as stress and intonation. The latter are attributes related to spectral configuration, duration, intensity, and pitch.

The simplified speech from the Kelly-Gerstman machine raises an interesting question. Is it possible to recover its input symbols by automatic analysis of the output? This query might be divided into two parts; first, can the control signals be recovered from the acoustic output; and second, can the symbolic input be recovered from the control signals? With the aid of a detailed knowledge of the generator, the latter at least can be accomplished exactly.

The control signals consist of parabolic and linear segments, joined smoothly at their end points. These points are the steady-state vowel and consonant values. They can be located easily since they always coincide with positions of zero-slope. Once they are available, a simple look-up procedure will recover the symbolic input exactly (apart from a few inconsequential ambiguities).

Deducing the control signals from the acoustic output is slightly more difficult. Two source functions, and three poles are known to constitute the class of spectra produced by the machine. Our methods for spectral measurements and automatic matching would be infallible here, except for measurement uncertainties. These would not be large on the average. Knowing that all control signals must have a stereotyped form, it seems clear that a close approximation to them could be inferred.

If noise were added to the speech, or if it were otherwise distorted, analysis would become much more difficult. It would still take the same form, but with increasing emphasis on decision theory and detection methods to hold errors to a minimum. Knowledge of syntactic and semantic constraints in the symbolic input would be of aid, too.

Compared to artificial speech, natural speech presents a fearsome spectre when we contemplate template analysis. The complex and fluctuating structure of the genuine article contrasts sharply with the simplicity of the Kelly-Gerstman version. Then, too, there are talker-to-talker variations to complicate the picture. Despite the disparities, the Kelly-Gerstman model can be of significant utility in representing natural speech, if aided by a human operator. Begin with a spectrogram of natural speech; a sentence, "Feed the computer lies", is shown in the upper half of Fig. 15. Knowing the words, an experienced experimenter

(L. J. Gerstman, in this case) segmented the utterance as shown, assigned a phonetic symbol to each pair of segments (one steady, one transition), and measured the pitch frequency and durations. With this data, the Kelly-Gerstman machine generated the speech shown below the original in Fig. 15. The match is surprisingly good, both objectively and subjectively. Thus, given the Kelly-Gerstman machine, the input symbols are a concise representation of the original utterance. Greater accuracy and automatic methods will come not from a different, but from refinements of the same, approach. These refinements will be aimed largely at automating some of the linguistic knowledge which Gerstman exercised in analyzing the speech sample of Fig. 15.

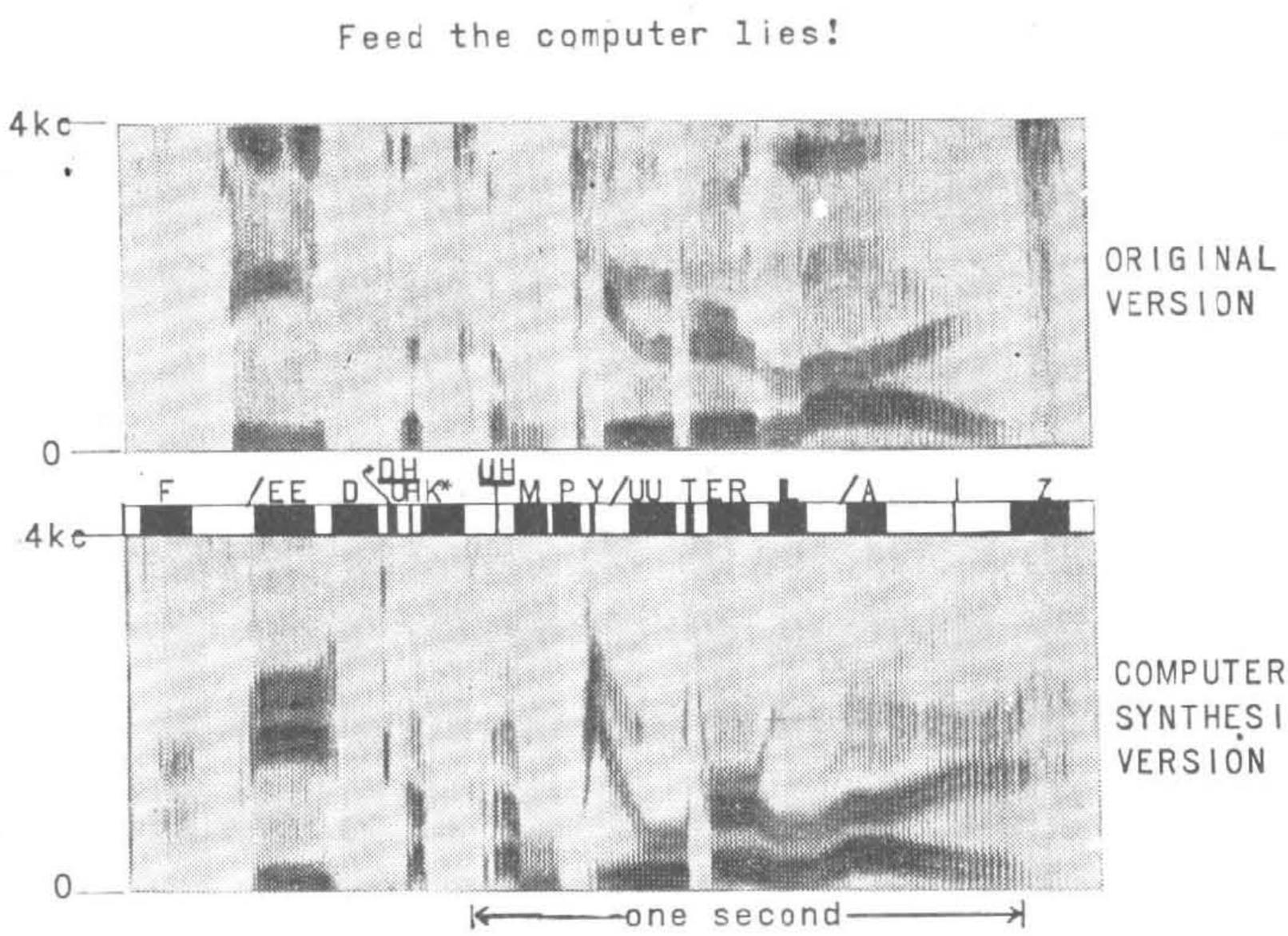


Fig. 15. The human utterance "Feed the computer lies" can be matched surprisingly accurately by the Kelly-Gerstman machine. The necessary analysis consists of measuring, by eye, the durations and pitch of the various phonetic divisions, each of which is assigned a symbol by the experimenter. Compare original (above) with computer synthesis (below).

With this simplified example of speech processing, we catch a glimpse of future speech processors, machines incorporating vocal tract and linguistic knowledge as well as sophisticated statistical decision procedures. The results so far achieved using computer methods are projecting us well beyond the acoustic level in speech analysis. From here forward, we will require linguistic and, perhaps, even semantic factors to uncover the relation between the acoustic event and its linguistic counterpart. Computers through their memory and logical capability permit us with relatively little technological effort to investigate this relation experimentally. The bottleneck no longer lies in technology, but in finding the right questions. It is up to us to ask them.

REFERENCES

1. C. G. M. Fant, "Modern instruments and methods for acoustic studies of speech", Royal Inst. of Tech., Stockholm, Sweden, Rept. No. 8, June 11, 1957.
2. E. E. David, Jr., M. V. Mathews, and H. S. McDonald, "A high-speed data translator for computer simulation of speech and television devices", 1959 IRE Western Joint Computer Conv. Record, 354–357.

3. E. E. David, Jr. "Digital simulation in research on human communication", Proc. Inst. Radio Engrs., *49*. 319–329 (1961).
4. J. L. Kelly, Jr., Carol Lochbaum, and V. A. Vyssotsky, "A block diagram compiler", Bell System Tech. J., *40*, 669–676 (1961).
5. F. de Jager, "Delta modulation, a method of p.c.m. transmission using the 1-unit code", Philips Res. Report 7, 442–466 (1952).
6. IBM Reference Manual, 709/7090 Fortran Programming System, International Business Machines, Corp., New York, 1961.
7. R. R. Blackman and J. W. Tukey, "The measurement of power spectra from the point of view of communications engineering", Part I, Bell System Tech. J, *37*, 185–282 (1958); Part II, Bell System Tech. J., *37*, 485–569 (1958).
8. M. V. Mathews, Joan E. Miller, and E. E. David, Jr., "Pitch synchronous analysis of voiced sounds", J. Acoust. Soc. Am. *33*, 179–186 (1961).
9. C. G. M. Fant, *The Acoustic Theory of Speech Production*, Mouton and Co., The Hague, (1960).
10. J. C. Steinberg, "Application of sound measuring instruments to the study of phonetic problems", J. Acoust. Soc. Am. *6*, 16 (1934).
11. D. Lewis, "Vocal resonance", J. Acoust. Soc. Am. *8*, 91 (1936).
12. R. L. Miller, "Nature of the vocal cord wave", J. Acoust. Soc. Am. *31*, 667 (1959).
13. W. Lawrence, Personal Correspondence, Signals Research and Deve. Est., Christchurch, England (1959).
14. S. Inomata, "Computational method for speech recognition", Bull. Electro-Tech. Lab. (Tokyo) *24*, 597–612 (1960).
15. C. G. Bell, H. Fujusaki, J. M. Heinz, K. N. Stevens, and A. S. House, "Reduction of speech spectra of analysis-by-synthesis techniques", J. Acoust. Soc. Am. *33*, 1725–1736 (1961).
16. M. V. Mathews and Joan E. Miller, "Spectral analysis of fricative consonants", J. Acoust. Soc. Am. *34*, 743 (A) (1962).
17. J. M. Heinz and K. N. Stevens, "On the properties of voiceless fricative consonants", J. Acoust. Soc. Am. *33*, 589–596 (1961).
18. E. E. David, Jr., "Artificial auditory recognition in telephony", IBM J. Research Develop. *2*, 294–309 (1958).
19. J. W. Forgie and C. D. Forgie, "Segmentation scheme for use of a speech recognition computer program", J. Acoust. Soc. Am. *32*, 1516 (A) (1960).
20. P. N. Scholtz and R. Bakis, "Spoken digit recognition using vowel-consonant segmentation", J. Acoust. Soc. Am. *34*, 1 5 (1962).
21. V. A. Vyssotsky, E. E. David, Jr., and M. V. Mathews, "Recognition of voicing, voice pitch and formant frequencies with a digital computer", Proc. III Int. Cong. Acoust. *1*, 229–231, Elsevier Publis. Co., Amsterdam (1961).
22. G. W. Hughes, "The recognition of speech by machine", RLE Tech. Rept. 395, Mass. Inst. Tech., Cambridge, Mass., May (1961).
23. R. K. Potter, G. A. Kopp, and H. C. Green, *Visible Speech*, D. van Nostrand, New York (1947).
24. A. M. Liberman, P. C. Delattre, F. S. Cooper, and L. J. Gerstman, "The role of consonant-vowel transitions in the perception of the stop and nasal consonants", Psychological Monographs, No. 379, *1*, 68 (1954); and F. S. Cooper, P. Delattre, A. M. Liberman, J. M. Borst, and L. J. Gerstman, "Some experiments on the perception of synthetic speech sounds", J. Acoust. Soc. Am., *24*, 597–606 (1952).
25. J. L. Kelly, Jr. and L. J. Gerstman, "An artifical talker driven from a phonetic input", J. Acoust. Soc. Am. *33*, 835 (A) (1961).

26. W. Lawrence, "The synthesis of speech from signals which have a low information rate", Chap. 34, *Communication Theory*, Butterworth Sci. Pub., London (1953).
27. J. Swaffield, J. N. Holmes, and J. N. Shearme, Personal Communication, Joint Speech Research Unit, Eastcote, England; and J. N. Holmes and J. N. Shearme, Proc. IV Int. Cong. Phonetic Science (forthcoming).
28. J. L. Kelly, Jr. and Carol C. Lochbaum, "Speech Synthesis", Proc. IV Int. Cong. Acoust., Copenhagen (1962) (forthcoming).

Importance de diverses conditions expérimentales dans l'étude des effets ultra-soniques en milieu liquide

par A. DOGNON

Professeur á la Faculté de Médecine de Paris

Il est bien connu que les résultats obtenus dans l'étude des effets chimiques ou biologiques des ultra-sons en milieu liquide sont souvent divergents et même contradictoires, non seulement pour des expérimentateurs différents, mais aussi pour le même, lors d'expériences successives apparemment identiques. Nous avons attaqué l'étude des diverses conditions qui peuvent expliquer ces divergences: Il en existe en grand nombre, certaines inattendues ou même paradoxales. Nous ne nous flattons pas de les connaître toutes, car l'expérimentation à cet égard est extrêmement longue, la découverte de chaque nouvelle cause d'erreur forçant à revenir sur les essais antérieurs.

Les fréquences utilisées ont varié de 576 à 960 K Hz, avec des puissances modérées, ne dépassant pas 2 watts/cm^2. Les effets étudiés ont été soit les effets thermiques (sonde thermo-électrique), soit les effets chimiques (oxydation de l'iodure de potassium), soit surtout les effets lytiques (hémolyse en suspension diluée de globules rouges), en raison de leur sensibilité et de la précision de leur mesure. Ces deux derniers sont, comme on le sait, exclusivement liés à la cavitation.

Le dispositif théorique idéal pour une telle étude a été indiqué depuis longtemps par *Giacomini*. Il s'agit d'un vase d'expérience à parois non absorbantes placé sur le trajet du faisceau en milieu indéfini, conditions dont il est possible de s'approcher beaucoup (fig. 1).

Le malheur est que, dans ces circonstances expérimentales parfaites, on n'observe pratique-

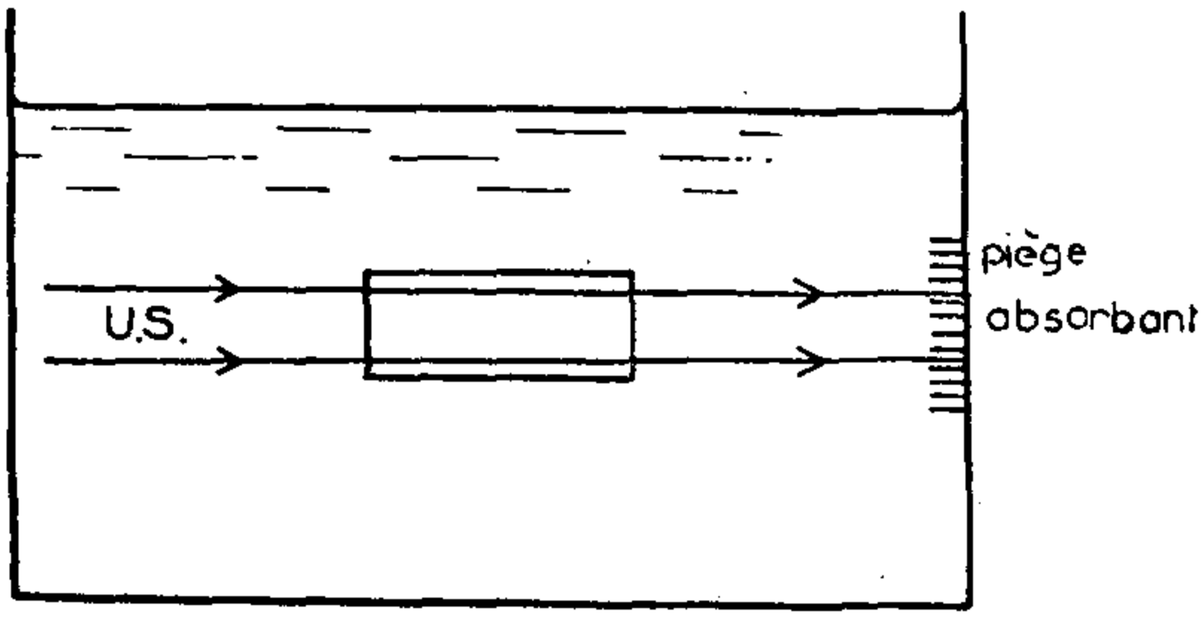
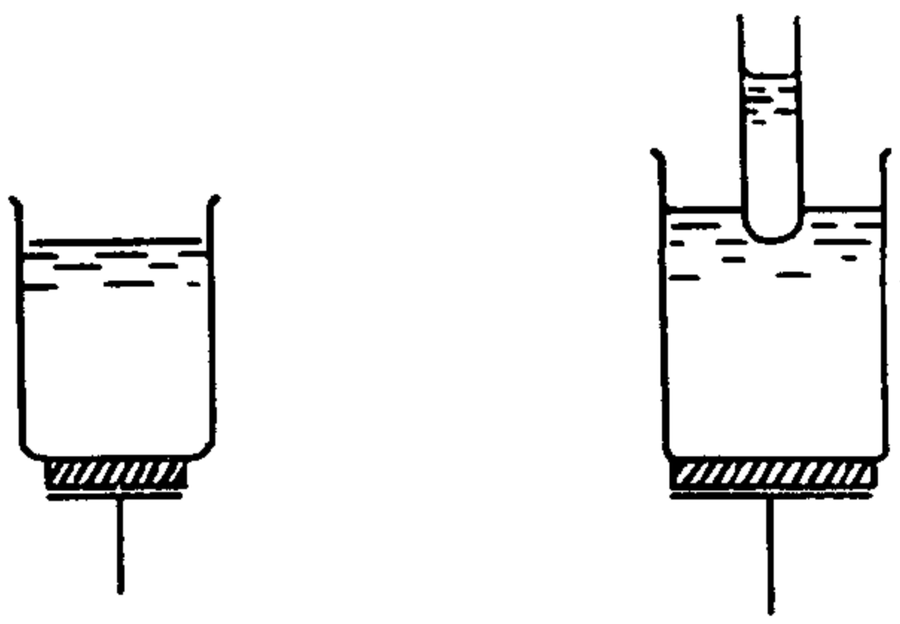


Fig. 1. Dispositif théorique de Giacomini.



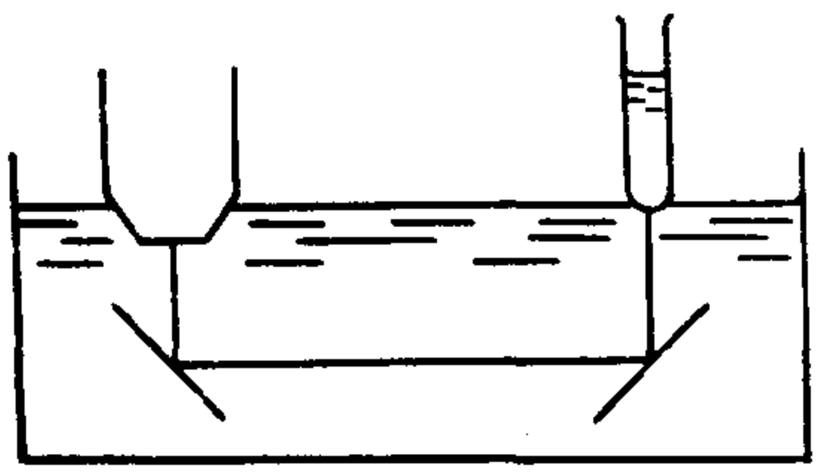


Fig. 2. Différents types schématiques de dispositions expérimentales.

ment rien, ce qui montre a priori que les effets lytiques ou chimiques, que l'on observe très facilement avec les dispositifs usuels (fig. 2), sont liés précisément à l'« imperfection » de ceux-ci et par conséquent aussi à un grand nombre de détails difficiles à prévoir.
Nous avons utilisés tout d'abord un premier dispositif représenté par la figure 3. Le vase d'expérience est un tube d'acétate de cellulose à paroi mince, de 3 cm. de diamètre environ. Il peut être soit dans la position A, qui est celle des conditions usuelles de travail, soit dans la position B, c'est-à-dire immergé et dans des conditions très voisines de celles du vase idéal de *Giacomini*. En A, on constate que l'action hémolysante est intense et débute pour une puissance de quelques dixièmes de watts/cm^2 à l'émetteur, comme la cavitation elle-même. En B, il n'y a aucun effet, même pour des puissances dix fois plus grandes, même encore si, en substituant un miroir concave au miroir plan, on forme dans le tube d'expérience le foyer du faisceau.

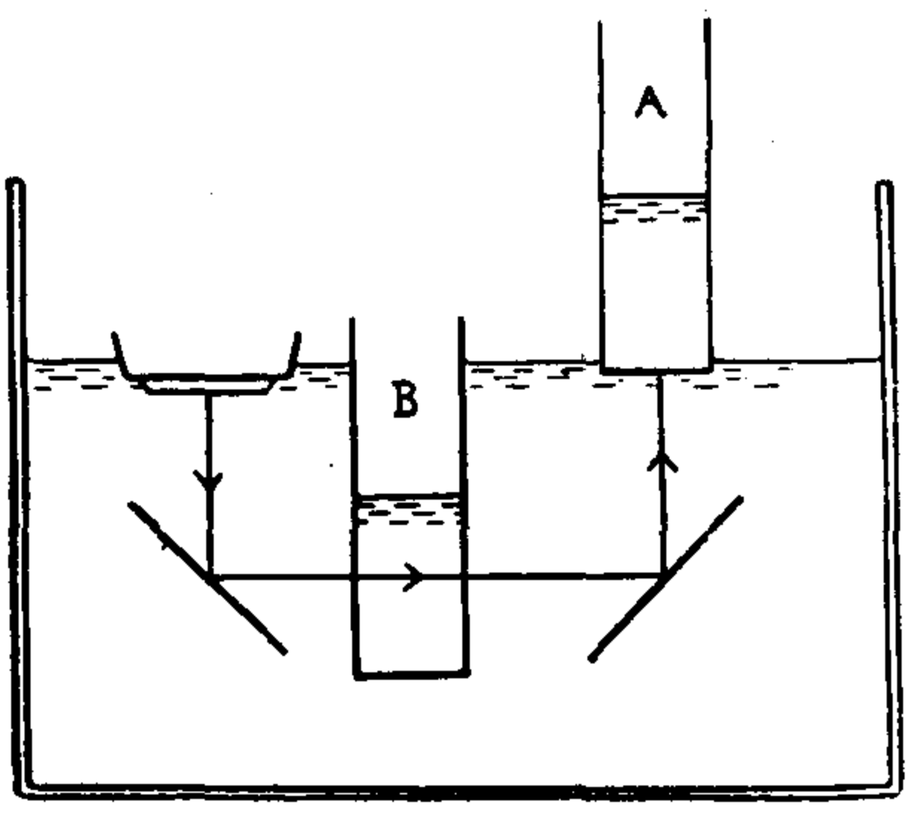


Fig. 3. L'action hémolytique ou chimique a seulement lieu dans la position A du tube d'expérience (en acétate de cellulose).

Dans l'idée que l'absence d'action est liée à l'absence d'ondes stationnaires, on accole un réflecteur à la face postérieure du tube. On observe alors une légère hémolyse pour une puissance d'environ 2 watts/cm^2. Donc, si la présence d'ondes stationnaires est un facteur important, ce n'est pas le seul. A priori, on peut songer à l'intervention de deux autres causes: 1^0) En A, l'environnement étant aérien, et par conséquent la réflection totale, le milieu d'expérience est enfermé dans une sorte de « sphère » d'intégration, alors qu'en B une très petite fraction de l'énergie totale est seule retenue. 2^0) En B, la face de sortie, constituée par la paroi du tube, est rigide. En A, c'est la surface *libre* du liqude. L'expérience va montrer que ces deux facteurs sont importants.

Rôle de l'environnement:

Utilisons le tube d'expérience dans la position A, mais après l'avoir entouré d'un large manchon rempli d'eau (fig. 4). On constate alors: 1^0) que l'énergie dans le tube, évaluée par l'élévation de température en un temps assez court (30 secondes) pour que le système puisse être considéré comme adiabatique, est diminuée de 50%. 2^0) Que les effets chimiques ou lytiques sont également fortement diminués. Quant au taux de cette diminution, elle dépend de la puissance. Si celle-ci est assez voisine du seuil de cavitation, le rôle du manchon d'eau peut être d'abaisser cette puissance au-dessous du seuil, et par conséquent de supprimer l'un et l'autre des effets précédents. Il est très vraisemblable qu'il s'agit là simplement de la répartition dans un volume plus grand de l'énergie qui traverse le tube d'expérience.

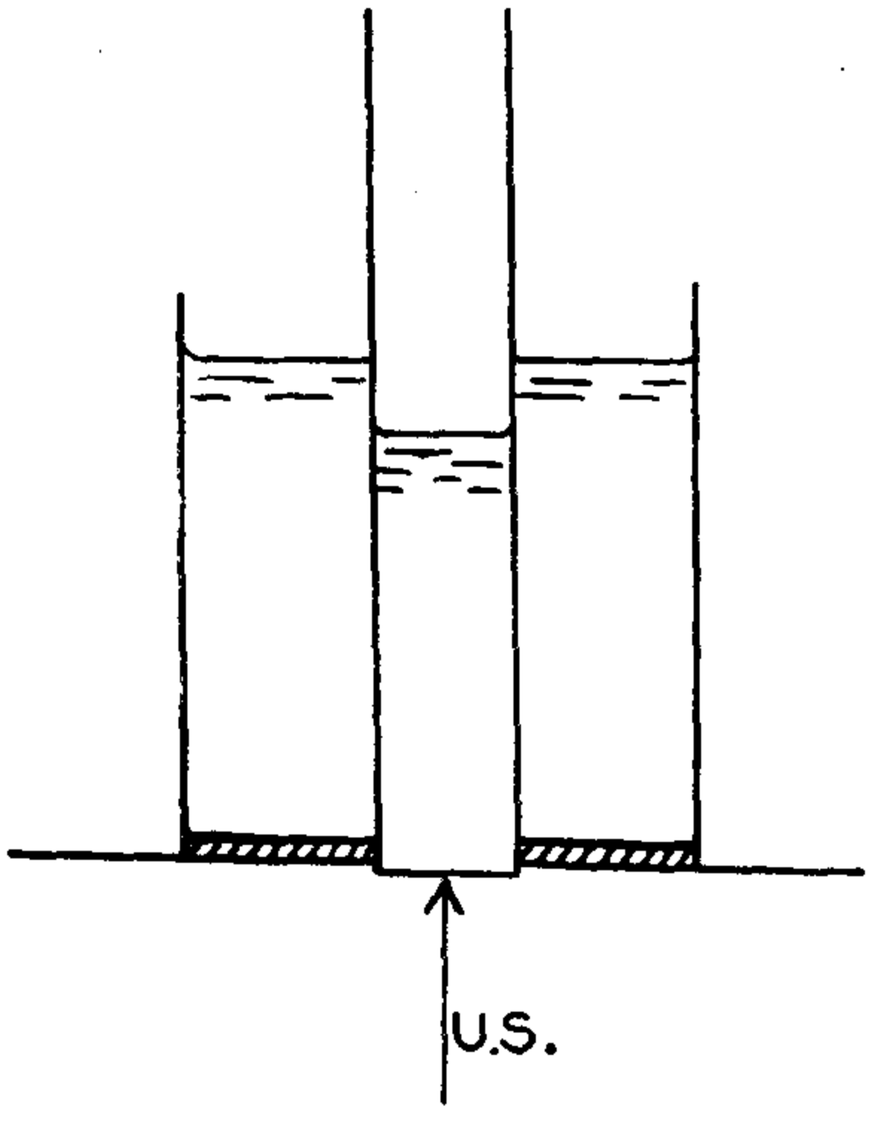


Fig. 4.

Rôle de la face de sortie:

Que se passe-t-il maintenant lorsque, le tube d'expérience précédent étant placé dans les meilleures conditions, on substitue à la surface libre une surface rigide, absorbante ou

réfléchissante ? On remarquera que c'est là le cas soit d'un vase clos entièrement rempli de liquide, soit d'un vase ordinaire à la surface duquel plonge l'émetteur.
La substitution d'un interface solide-liquide à l'interface liquide-air, a toujours pour résultat l'affaiblissement considérable de l'effet lytique ou chimique, qui peut être totalement bloqué pour des puissances aussi élevées que 2 watts/cm. Mais l'aspect du liquide est différent suivant qu'on remplace la surface libre par la section d'un cylindre *absorbant* (paraffine ou picéine), ou par celle d'un cylindre *réflecteur*. Dans le premier cas (ondes progressives), le liquide paraît « gelé ». Il n'y a aucun mouvement de convection et les quelques bulles présentes restent tout à fait immobiles. Dans le second, on observe un beau système d'ondes stationnaires rendu visible par la concentration des globules rouges aux noeuds de pression où ils restent inaltérés. La cavitation est également bloquée et, par conséquent, les réactions chimiques sont aussi absentes.
Les mesures thermiques pratiquées avec un couple revêtu d'une très petite couche de matière absorbante, montrent que, avec le cylindre absorbant, la puissance est environ moitié de celle qui existe dans le cas du cylindre réflecteur ou de la surface libre. On ne trouve pas là de quoi expliquer un blocage des effets tel qu'il ne se produit encore rien pour des puissances dix fois supérieures au seuil. En fait, il s'agit d'un blocage de la cavitation, ou du moins de la cavitation *efficace*. Dans le cas d'ondes stationnaires, on peut imaginer le mécanisme suivant: Pour que la bulle, née dans les plans ventraux de pression, atteigne le diamètre de résonance qui conditionne son efficacité, il faut qu'elle soit nourrie par le gaz diffusant des régions voisines. Si les plans générateurs des cavités sont en position parfaitement fixes dans le liquide, la diffusion ne peut assurer la nourriture de la bulle jusqu'au diamètre de résonance. Tout se passe comme si on avait diminué la concentration du gaz, ce qui, on le sait, diminue toujours, ou supprime, les effets liés à la cavitation.
Si cette hypothèse est exacte, on doit faire réapparaître l'efficacité en agitant le plan réflecteur. Ceci a été réalisé en montant le cylindre sur une tige vibrante (à 100 périodes/sec.), de sorte que la position des plans stationnaires change rapidement au sein du liquide. Dans ces conditions, on fait réapparaître totalement, et la cavitation visible, et les effets qui l'accompagnent. Par contre, comme on devait s'y attendre, aucune modification ne se produit avec le cylindre absorbant. Il faut chercher une autre cause à l'inefficacité des ondes progressives. Est-ce simplement la différence qui existe dans la localisation de l'énergie? La différence des amplitudes est tout à fait insuffisante pour en fournir l'explication.

Rôle de la hauteur du liquide

Dans un tube d'expérience semblable au précédent (parois minces de plastique), lorsque la surface est libre et la puissance acoustique assez faible pour ne pas déformer sensiblement celle-ci en créant l'amorce d'un geyser, le thermo-couple picéiné permet facilement, à 960 K Hz, de mettre en évidence un *effet interférométrique*. Lorsqu'on fait varier très progressivement la hauteur du liquide, par exemple par adjonction de gouttes successives (fig. 5), la puissance acoustique passe par une série de maxima et de minima très marqués distants de $\lambda/4$. On peut vérifier que l'action hémolytique suit les mêmes variations: Il suffit pour cela de mesurer la puissance, pour une hauteur donnée de liquide contenant des globules rouges à grande dilution, mais avec une puissance inférieure au seuil de cavitation,

puis d'élever cette puissance dans un rapport constant. On observe alors des variations du taux d'hémolyse parallèles à celles de la puissance. En s'arrangeant pour que la puissance maximale soit peu supérieure au seuil, on produit ou non l'hémolyse pour une variation de $\lambda/4$ de la hauteur du liquide. Bien entendu, cet effet ne se produirait pas avec des puissances élevées qui amèneraient une déformation importante de la surface et la disparition de toute régularité dans les plans stationnaires. Il est beaucoup plus difficile à observer aussi à fréquence plus basse (576 K Hz) vraisemblablement par un effet perturbateur des bulles de cavitation, qui sont produites beaucoup plus facilement. On y arrive cependant, au moins en ce qui concerne les mesures de puissance, en bloquant la cavitation par l'emploi d'une suspension suffisamment concentrée de particules (p. ex. des levures à 5%).

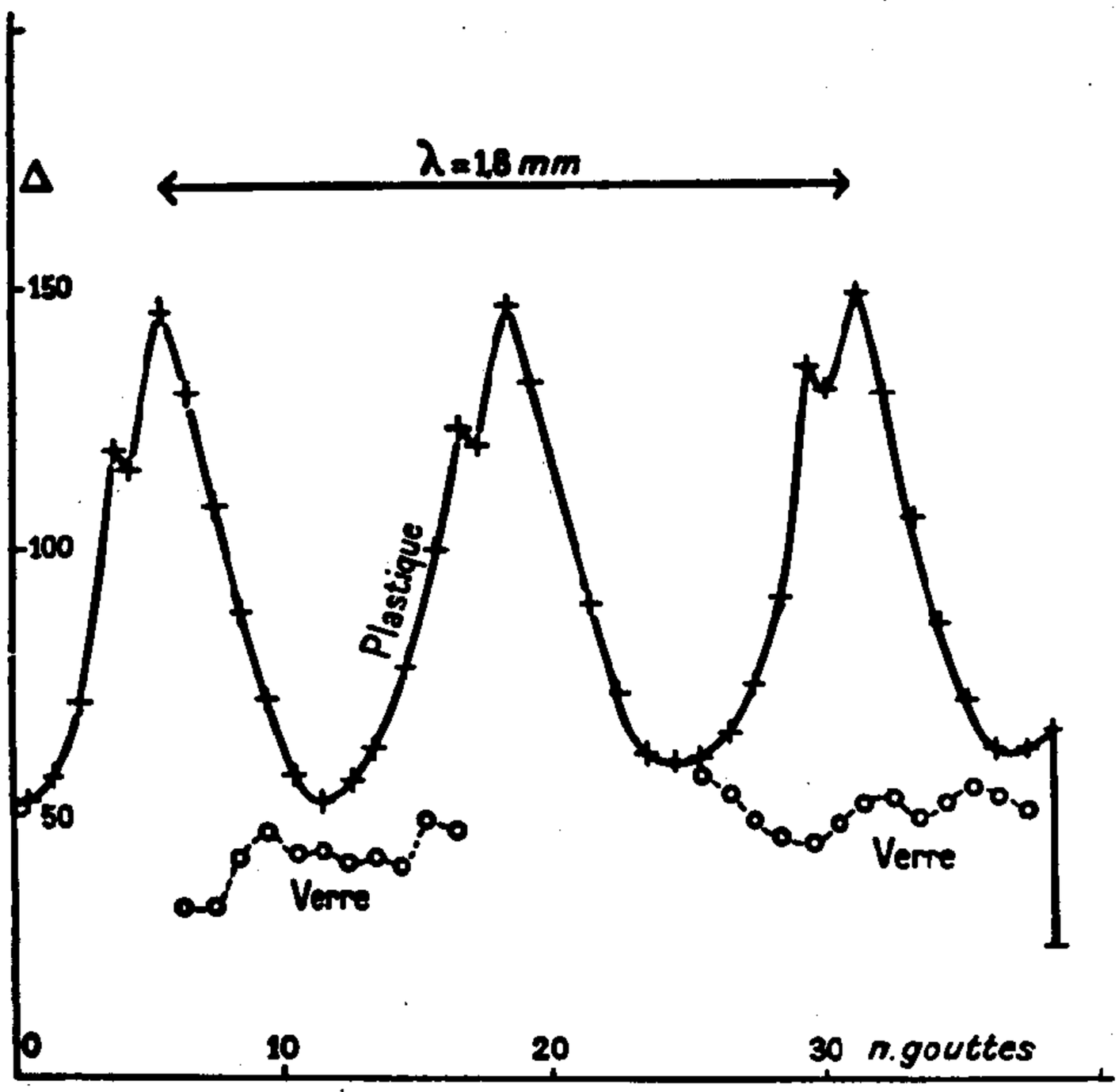


Fig. 5. Effet interferométrique: variation périodique de puissance (et d'action) obtenue par faibles variations de hauteur (adjonction de gouttes). Pas d'effet en tube de verre. N = 960 KHZ.

Une constatation importante sur laquelle on va revenir, est que cet effet interférométrique ne s'observe pas, même à la fréquence la plus favorable, si l'on remplace le tube d'acétate de cellulose par un tube de verre du type tube à essai. La fig. 5 montre les variations faibles et irrégulières que l'on observe dans ce cas. Ceci nous amène à étudier le rôle de la nature des parois.

Rôle de la nature des parois:

Dans l'expérimentation usuelle on utilise fréquemment des tubes à essai, éventuellement à parois épaisses, dont le fond entre en contact avec le bain vibrant, et on observe régulièrement que le rendement est nettement meilleur ainsi que si l'on fait pénétrer le faisceau à

travers une membrane mince. Ce résultat est paradoxal, et même, si l'on veut bien songer qu'une paroi de verre de cette épaisseur réfléchit près de 80% de l'énergie, et plus encore par suite de sa convexité, on devrait s'étonner d'observer encore des actions intenses dans de telles conditions. Autre paradoxe, dans un tube de verre à fond plat, les actions sont toujours plus faibles. La mesure au thermo-couple permet effectivement d'observer, dans l'axe d'un tube de verre à fond convexe, des puissances supérieures à celles qui existeraient en l'absence du tube, c'est-à-dire au même point d'un milieu transmetteur de même nature que le liquide contenu dans le tube. Un tel résultat ne peut s'expliquer que par un régime particulier de vibration des parois qui constituent une sorte de résonateur. Effectivement, lorsque le tube contient une suspension, on voit les particules s'accumuler en une colonne axiale, ce qui ne se produit jamais dans les autres types de vase.

La participation des parois se manifeste très nettement lorsqu'on emploie un vase d'expérience du type « tube à pied », commode dans certains cas (fig. 6). Dans le modèle *a*, le tube de verre est soudé à un pied, également de verre, qui transmet la vibration. La vibration est intense, et les effets très marqués. Dans le type *b*, le raccordement du pied au tube se fait par l'intermédiaire d'une bague de caoutchouc. Dans ce cas, le rendement du dispositif est beaucoup moins bon.

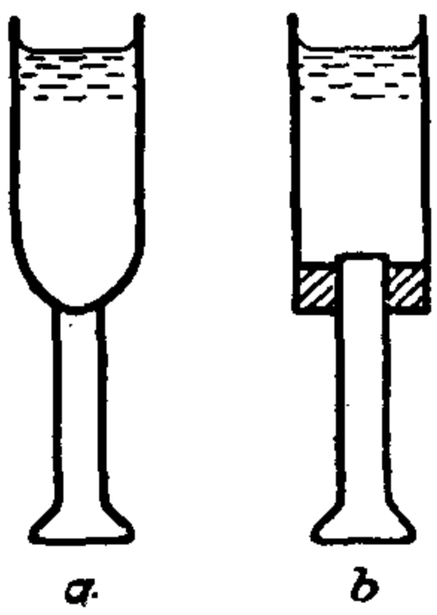


Fig. 6. Importance de la vibration des parois: seul le tube *a* (pied conducteur soudé) permet une action intense.

En résumé, le rendement le meilleur est obtenu avec un tube de verre dont l'épaisseur peut être de plusieurs millimètres, et dont le fond, convexe, entre juste en contact avec le bain vibrant. Un tel dispositif est à la fois focalisateur et intégrateur. La quasi-totalité de l'énergie qui y pénètre y demeure et est finalement absorbée. Mais on voit aussi qu'il serait tout à fait incorrect, dans ces conditions, de chercher à évaluer l'énergie à l'intérieur du tube par un procédé qui suppose un faisceau dirigé, comme une mesure de pression de radiation ou un piézomètre.

Lorsque le liquide en expérience est contenu dans un tube de verre, on constate que l'introduction d'un matériel absorbant, même de volume réduit, comme par exemple un petit agitateur en matière plastique, entraîne une diminution notable de l'efficacité de la vibration. Ceci se comprend si l'on assimile ce problème à celui d'un intégrateur de lumière dans lequel on introduit un corps absorbant. Si le coefficient d'absorption du corps introduit est beaucoup plus élevé que celui de l'eau, ce qui est le cas pour les cires, les plastiques, le caoutchouc etc...., le partage massique de l'énergie se faisant dans le rapport des absorptions, il suffit de petites quantités de matière pour diminuer beaucoup l'énergie disponible dans le liquide. Cet effet est nettement mis en évidence par la figure 7, qui représente le

taux d'hémolyse en fonction du temps, d'une part dans un tube de verre (traits pleins), d'autre part dans le même tube intérieurement revêtu d'une chemise d'acétate de cellulose de 0,25 mm d'épaisseur (traits pointillés). Dans ces conditions, le taux d'hémolyse peut être réduit au $^1/_5$ de sa valeur, et il en est de même de l'énergie mesurée au thermo-couple picéiné dans l'axe du tube.

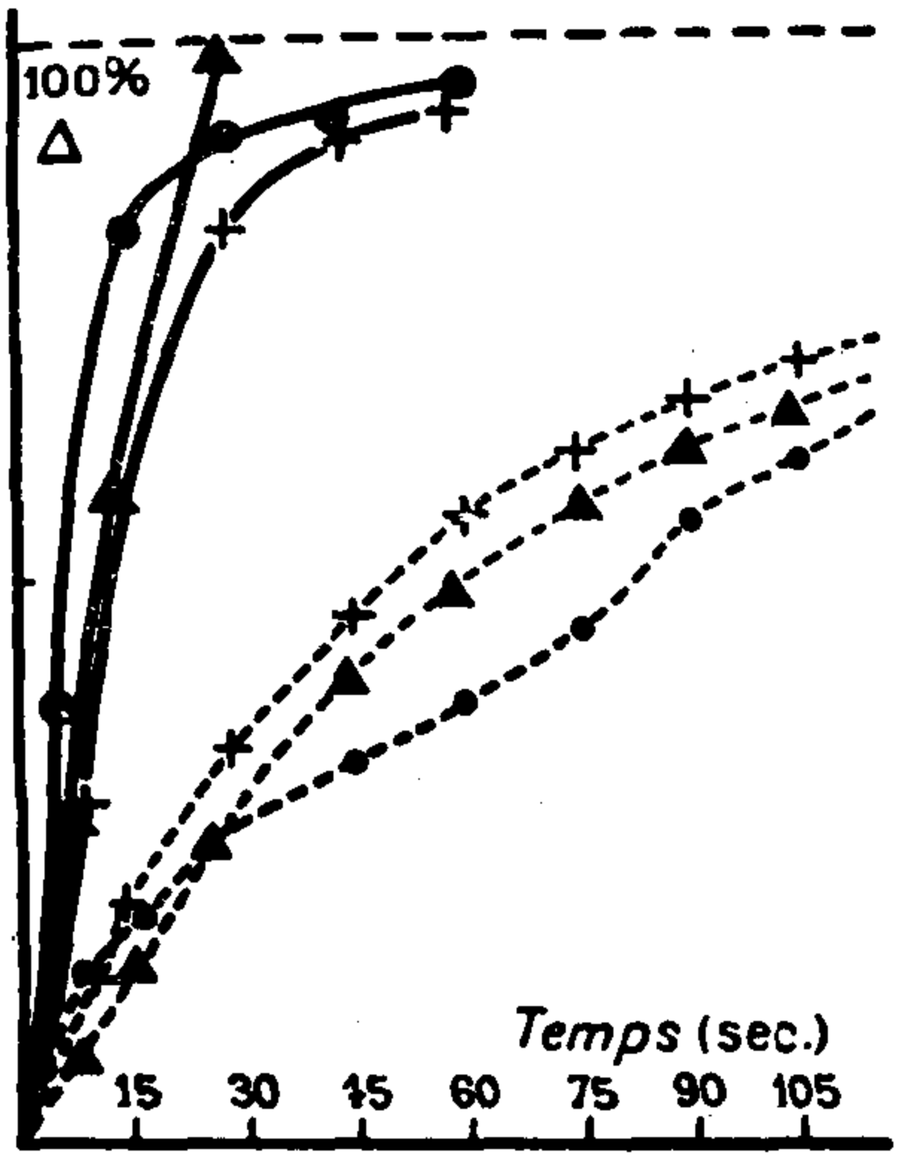


Fig. 7. Hémolyse en function du temps dans un tube de verre (traits pleins) et dans le même tube intérieurement garni d'une chemise de plastique mince (traits pointillés).

On peut illustrer les diverses particularités qui viennent d'être énoncées, par des combinaisons expérimentales diverses dans lesquelles la disposition et la nature des vases d'expérience permettent, avec la même énergie incidente, d'obtenir ou non l'action ultra-sonique, dans des conditions au premier abord paradoxales. La fig. 9 en donne un exemple: A gauche, un tube d'acétate de cellulose contient une suspension d'hématies dans laquelle plonge le fond d'un tube de verre contenant une suspension identique. La vibration arrive de bas en haut, et par conséquent la suspension inférieure, qui sert de milieu de transmission, semble soumise à une intensité vibratoire plus élevée. Cependant, l'action hémolytique y reste nulle, alors qu'elle est très marquée dans le tube supérieur. A droite est réalisée la disposition inverse, et le résultat est également inversé. Le paradoxe apparent réalisé par le dispositif de gauche s'explique par une double raison: D'abord parce que la surface libre du liquide inférieur est remplacée par la surface rigide et réflectrice du tube de verre, créant, comme on l'a vu, des ondes stationnaires parfaitement stables, conditions dans lesquelles ne se produit pas de cavitation efficace. Ensuite parce que les parois de plastique sont en elles-mêmes peu favorables pour les raisons que nous avons exposées. Au contraire, toutes les conditions favorables sont réalisées pour le tube supérieur (parois de verre, surface libre). Dans le dispositif de droite, l'absence d'effet dans le tube supérieur s'explique par les propriétés défavorables de sa paroi. A première vue, on pourrait penser qu'il ne devrait pas y avoir d'action dans le tube inférieur dont la surface n'est pas libre, mais

remplacée par la membrane d'acétate de cellulose qui forme le fond du tube supérieur. En fait, cette membrane mince n'étant pratiquement ni absorbante ni réfléchissante ne compte pas, et la surface effective est la surface libre du liquide supérieur. D'ailleurs, le phénomène de blocage par introduction d'un cylindre absorbant ou réflecteur, tel que nous l'avons décrit plus haut, s'observe moins nettement dans le cas d'un tube de verre dans lequel, comme on l'a vu, le régime de vibration est différent.

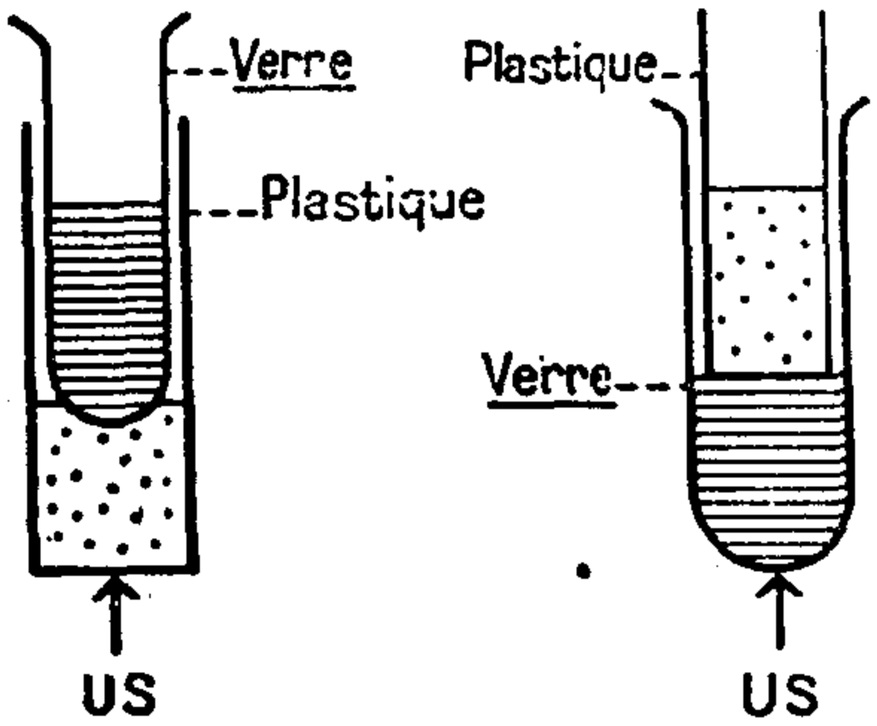


Fig. 8. A énergie égale on peut faire varier l'action hémolytique suivant la disposition expérimentale. L'hémolyse est représentée par les hachures.

Causes de divergence dans l'étude de certains facteurs extérieurs

Toutes les considérations qui précèdent, concernent la géométrie ou la nature du vase d'expérience. Il en est d'autres tout à fait différentes qui sont susceptibles de causer une grande dispersion des résultats lorsque on étudie l'action de facteurs externes tels que température ou pression. En effet, il ne faut jamais oublier que la plupart des effets étudiés

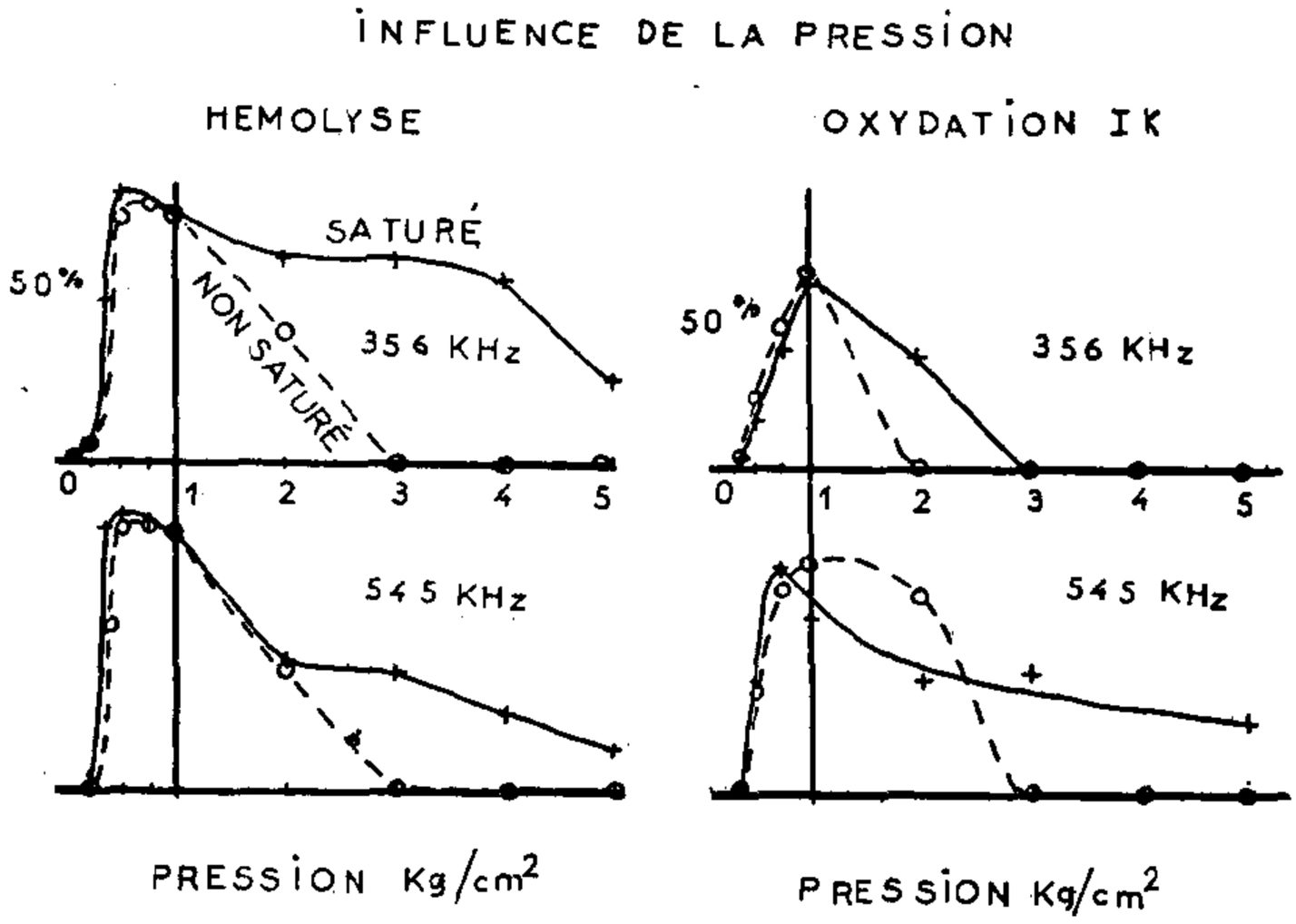


Fig. 9. Influence de la pression sur l'action hémolytique et l'action chimique des U-S, suivant que l'équilibre des gaz est réalisé (traits pleins) ou non (pointillés).

étant dûs à la cavitation, la « tendance au dégagement » du gaz dissous est un des facteurs essentiels.
Pour une température et une pression données, cette « tendance au dégagement » est nulle à l'équilibre, c'est-à-dire à la saturation. Elle est positive si la pression s'abaisse (ou que la température s'élève), négative dans le cas contraire. C'est évidemment seulement dans les conditions de l'équilibre des gaz qu'une étude de l'action des ultrasons aura une signification précise.
Pour illustrer cette considération, nous donnerons dans les fig. 9 et 10 nos résultats concernant, pour des fréquences variant de 356 à 3000 K Hz, l'effet hémolytique et l'effet oxydant, en fonction de la pression. Dans chaque cas, l'effet a été étudié soit dans les conditions d'équilibre des gaz (agitation préalable prolongée du liquide en expérience sous la pression choisie), soit en l'absence d'équilibre (l'action des ultra-sons a lieu dès la mise sous pression).

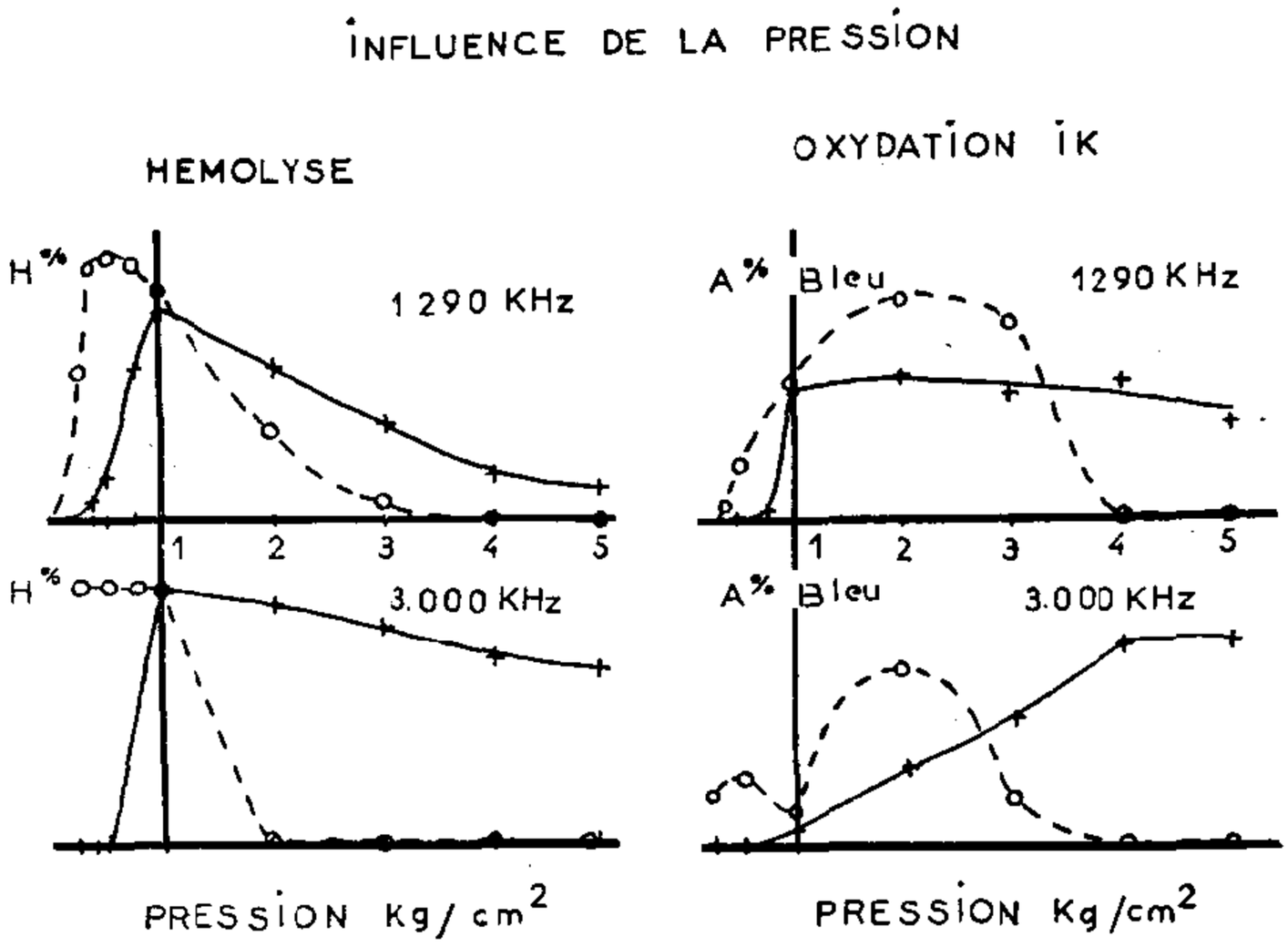


Fig. 10. Même expérience pour des frequences plus élevés.

Nous ne discuterons pas en détail ces résultats, et nous nous bornerons à remarquer l'énorme différence qui existe entre les courbes en traits pleins (saturation) et les courbes en pointillé (absence de saturation). Dans tous les cas, et surtout pour la fréquence la plus élevée qui est aussi la moins « cavitante », la saturation a pour effet de prolonger considérablement vers les pressions élevées l'action ultrasonique, qui, en l'absence de saturation, est toujours totalement supprimée pour une pression de 3 Kg/cm². A 3000 K Hz, et à la plus forte pression étudiée (5 Kg/cm), l'effet est peu inférieur (action hémolytique) ou très supérieure (action oxydante) à celle qu'on observe à la pression atmosphérique.
Parmi les observations qui font l'objet de cet exposé, certaines ont un intérêt pratique, d'autres possèdent surtout un intérêt théorique. Par exemple, constater que l'effet hémolytique ou chimique peut varier du simple au double quand la hauteur du liquide en expérience varie de $\lambda/4$ est sans importance pratique si l'on opère avec des puissances

élevées ou avec des vases de verre, puisque le phénomène est masqué dans ces conditions par l'agitation superficielle ou le mode particulier de vibration. Par contre, l'effet bloquant qui résulte du remplacement de la surface libre par un solide absorbant ou réflecteur possède un double intérêt: Théorique, car le mécanisme du phénomène mérite d'être éclairci, pratique car il proscrit l'emploi de certains dispositifs expérimentaux. Ce double aspect apparaît aussi dans les résultats qui concernent l'influence de la pression. De toute façon, nous espérons avoir contribué à faciliter la tâche des expérimentateurs en expliquant certains échecs et l'incohérence apparente de nombreux résultats.

International Standardization Work in the Field of Acoustics

WILLI FURRER

Berne, Switzerland

The beginning of international standardization in the field of acoustics in its present form dates back to 1953, the same year the 1st International Congress on Acoustics was held. The Dutch National Committee of the International Electrotechnical Commission (IEC) took the initiative to create the Technical Committee 29 (electro-acoustics), meeting for the first time in The Hague in June 1953. Shortly afterwards, i.e., in October of the same year and following an invitation extended by the British Standards Institution, the Technical Committee 43 (acoustics) of the International Organization for Standardization (ISO) convened in London for the first time.

There exists, however, a forerunner of the Committee on acoustics: in 1937, in Paris a group of acousticians from various nations gathered under the auspices of the ISA (International Federation of the National Standardizing Associations). Already then the discussions were mainly about reference levels, loudness function, sound level meters, standard musical pitch, etc. Also in the field of building acoustics the Acoustics Group of the Physical Society, London, had already since 1948 done important preparatory work.

In many cases it is difficult or even arbitrary to draw a distinct line between "acoustics" and "electro-acoustics". In order to avoid duplication of work among the ISO/TC43 and the IEC/TC29, a so called steering committee was established, consisting of the presidents and the secretaries of both committees. It is their responsibility to distribute all the arising problems among the two committees. Up to now this arrangement has worked smoothly, showing excellent results.

In principle the working methods of both the ISO and the IEC Committee are quite similar: A working group is formed to study a specific problem, the chairman and the members of which are designated by the Plenary Committee. Such a group will only exist until its task is concluded. There are, however, special cases, when permanent sub-committees are formed–e.g. the Sub-Committee 29-A "sound-recording"–who in their turn will split up into working groups for their studies.

Without going into minor details, a standardization process can be described as follows: A working group will present a first proposition that is distributed to all the National Committees. Once their comments and critics have been received, the working group will hold one or more meetings in order to discuss these contributions. After having been amended, the document has to be accepted by the Plenary Committee, and then the General Secre-

tariat will put it to the vote of the National Committees. Once this draft has been passed by a certain qualified majority of National Committees, finally a new, definitive standard in the form of an "ISO Recommendation" or an "IEC Publication" is born. It has to be noted that there exist some slight differences in the procedures of ISO and IEC.
All the hitherto published standards as well as some further important draft standards are listed below:

ISO – TC 43

Recommendation R 16 — STANDARD TUNING FREQUENCY
1st Edition 1955

The standard tuning frequency is the frequency for the note A in the treble stave and shall be 440 Hz. This frequency shall be observed within an accuracy of $\pm$ 0.5 Hz when tuning musical instruments.

Draft Recommendation — PREFERRED FREQUENCIES FOR ACOUSTICAL MEASUREMENTS

1000 Hz has been selected as the primary frequency for all series of preferred frequencies for acoustical measurements. For most measurements in the audio range, it is convenient to space the frequencies by octaves or fractions of an octave ($^1/_2$ octave or $^1/_3$ octave). The preferred geometric centre frequencies of octave filter pass bands are then: 63, 125, 250, 500, *1000*, 2000, 4000, 8000 Hz, etc.

WORKING GROUP 1

Draft Recommendation in preparation — NORMAL THRESHOLD OF HEARING BY EARPHONE LISTENING AND ITS APPLICATION TO THE SPECIFICATION OF A STANDARD REFERENCE LEVEL FOR THE CALIBRATION OF PURE-TONE AUDIOMETERS

This recommendation specifies a standard reference zero for the scale of hearing level applicable to pure-tone audiometers, which, it is hoped, will help promote agreement and uniformity in the expression of hearing level measurements throughout the world. Experimental work has been carried out under the auspices of the Committee, at five Standardizing Laboratories in France, Germany, UK, USA and USSR.
Studies for a reference threshold for bone-conduction audiometry have also started.

WORKING GROUP 2 — *Loudness scale*

Recommendation R 131
1st Edition 1959

EXPRESSION OF THE PHYSICAL AND SUBJECTIVE MAGNITUDES OF SOUND OR NOISE

This recommendation defines the sound pressure level, based on a reference pressure of 2.10^{-4} dyn/cm^2. The "phon" is defined as a dimensionless unit expressing the loudness level of a given sound or noise; a function relating the loudness in "sones" to the loudness levels in phons is specified.

WORKING GROUP 3

Draft Recommendation

MEASUREMENTS OF ABSORPTION COEFFICIENTS IN A REVERBERATION ROOM

This recommendation describes how a reverberation room should be used to measure sound absorption coefficients of acoustical materials, furniture, persons, space absorbers, etc. The general principle is that the specimen is introduced into the room and the absorption added is computed from measurements of the reverberation time of the room before and after the introduction of the specimen.

Contents:
Explanation of terms and principle of measurements
Measurement arrangements
Statement of the result

WORKING GROUP 4 — *Airborne and impact sound transmission in buildings*

Recommendation R 140
1st Edition 1960

FIELD AND LABORATORY MEASUREMENTS OF AIRBORNE AND IMPACT SOUND TRANSMISSION

This recommendation defines methods of measuring the airborne sound insulation of walls, and the airborne and impact sound insulation of floors, both in the field and in the laboratory, e.g., in buildings, dwellings, etc. The way in which the airborne and impact sound fields are generated, the frequency range of measurements and the characteristics of the necessary filters are described. Definitions are also given of the quantity measured in each case, and of the method of normalizing the results to make them comparable.

Contents:
Definitions: Levels, level difference, normalized level difference, sound reduction index, normalized impact sound level, etc.
Airborne sound transmission: Field measurements
id. Laboratory measurements
Impact sound transmission: Field measurements
id. Laboratory measurements

WORKING GROUP 5 *Methods of assessment of loudness by objective analysis*

Draft Recommendation in preparation

Most noises are measured objectively in terms of spectrum analysis in octave or $^1/_3$ octave bands. It is often desirable to obtain a single figure that corresponds to the loudness level of such noises. Two different procedures are being studied. The first one, worked out by *S. S. Stevens*, is particularly applicable to octave band analyses. A second one, indicated by *E. Zwicker*, is based on $^1/_3$ octave band analyses and is claimed to be applicable to strong-lined spectra or irregular spectra for which octave band analyses are not appropriate. The first procedure is purely analytical, whereas the second one requires a graphical evaluation.

Several National Committees still express some doubts about the applicability and the usefulness of one or the other procedure.

WORKING GROUP 7 *Traffic Noise*

Draft Recommendation METHODS OF MEASUREMENT OF NOISE EMITTED BY VEHICLES

This recommendation describes methods of determining the noise emitted by motor vehicles, intended as far as possible to meet the requirements of simplicity consistent with reproducibility of results and realism in the operating conditions of the vehicle.

Contents:
Measuring equipment
A sound level meter, measuring dB(A) is used.
Acoustical environment
Measurements with vehicles in motion
Appendix: Measurements with stationary vehicles

WORKING GROUP 8 *Industrial and Residential Noise*

Recommendation in preparation PROPOSAL FOR NOISE RATING NUMBERS WITH RESPECT TO CONSERVATION OF HEARING, SPEECH COMMUNICATION AND ANNOYANCE

The scope of this proposal is to recommend ratings for noises with respect to the risk of permanent ear damage, interference with speech communication and with respect to annoyance. Noise rating numbers are based on noise curves about which still some discussions are going on.

WORKING GROUP 9 *Measurement of Machinery Noise*

The work concerned with this subject has not yet advanced sufficiently to show representative results. The same remark applies to the problem of *Aircraft Noise*.

WORKING GROUP 11 *Equal-loudness Contours*

Recommendation R 226
1st Edition 1961

NORMAL EQUAL LOUDNESS CONTOURS FOR PURE TONES AND NORMAL THRESHOLD OF HEARING UNDER FREE FIELD LISTENING CONDITIONS

This recommendation specifies, for binaural listening in a free progressive plane wave, for the frequency range 20 to 15'000 Hz:

a. The normal relations existing between sound pressure level and frequency for pure tones of equal loudness.
b. Values for the normal threshold of hearing. (Normal binaural minimum audible field).

Recommendation in preparation

EXTENSIONS TO PURE-TONE EQUAL LOUDNESS CONTOURS

A proposal for one extension, namely, the relation between diffuse and free-field contours is under study.

IEC – TC 29

SUB-COMMITTEE 29-A *Sound Recording*

Publication 94
1st Edition 1957
2nd Editions 1962

RECOMMENDATIONS FOR MAGNETIC TAPE RECORDING AND REPRODUCING SYSTEMS: DIMENSIONS AND CHARACTERISTICS

These recommendations apply to non-perforated magnetic tape used for sound recording in both professional and domestic applications and to the equipment used for recording and reproducing it.

Contents:

Mechanical requirements: Tape speed, tape winding, sound track, spools
Electrical requirements: Recording characteristics, reproducing characteristics, standard re-play characteristics.
Magnetic tapes: Dimensions, strength, flammability.
Leaders and Labels for recorded tapes
Programme identification
Appendix: Methods of measuring the magnetization of a tape.

Publication 98
1st Edition 1958
Amendments 1959

RECOMMENDATION FOR LATERAL-CUT COMMERCIAL AND TRANSCRIPTION DISK RECORDINGS

These recommendations include the most important dimensional features and the recording and reproducing characteristics, that are necessary to secure interchangeability of recordings.

Contents:
Definitions
Types of disks: Groove dimensions, speed, diameter, direction of rotation, direction of recording, thickness, diameter and eccentricity of centre hole, finishing groove, etc.
Electrical recording and reproducing characteristics: Tolerances, stylus tip radius, etc.

Publication 98-1
1st Edition 1959

RECOMMENDATIONS FOR STEREOPHONIC COMMERCIAL DISK RECORDS

These recommendations, as a supplement to Publication 98, apply to stereophonic commercial disk records. They give the most important dimensional features and the recording and reproducing characteristics that are necessary to secure interchangeability of records.

Contents:
Same as in Publication 98, with additional recommendations for the stereophonic groove and the reproducing stylus tip radius.

Final Draft
approved for publication

RECOMMENDATIONS FOR MAGNETIC SOUND RECORDING ON 16 mm AND 35 mm FILM

Due to the fact, that the text is not in full agreement with that of a recommendation of ISO/TC36 (Cinematography), this publication is issued as a "Report". It endorses the Recommendations 264 and 265 of the International Radio Consultative Committee (CCIR), relating to recording standards for the international exchange of television programmes. This decision received the agreement of ISO/TC36.

WORKING GROUP 2 — *Audio-apparatus*

Publication 89
1st Edition 1957

RECOMMENDATIONS FOR THE CHARACTERISTICS OF AUDIO-APPARATUS TO BE SPECIFIED FOR APPLICATION PURPOSES

The purpose of these recommendations is to facilitate the determination of the quality of acoustical apparatus, the comparison of such types of apparatus and their proper practical application, by listing the characteristics which are useful for their specification. The publication is confined to a qualitative description of the different characteristics and does not attempt to specify performance.

Contents:
Definition of general terms: Levels, amplification, impedances, distortion, scales, etc.
Sound system amplifiers: Source and load impedances, power, minimum and overload input voltages, frequency response, noise, distortion, etc.
Microphones: Reference axis, free field, pressure and close talking sensitivities, directional pattern, inherent noise level.

WORKING GROUP 5 — *Loudspeakers*

Publication 124
1st Edition 1960

RECOMMENDATIONS FOR THE RATED IMPEDANCES AND DIMENSIONS OF LOUDSPEAKERS

These recommendations refer to simple moving-coil (dynamic) loudspeakers of the direct radiator type. Dimensional recommendations are limited to loudspeakers with circular cones, excluding cones of elliptical section.

Contents:
Recommended values for the impedance
Mounting dimensions

Publication in preparation

RECOMMENDED METHODS OF MEASUREMENT FOR LOUDSPEAKERS

The object of these recommendations is to specify, on the simplest possible basis, practical and uniform methods of measuring certain characteristics of loudspeakers, so that discussions between suppliers, users and testing authorities may be based on clearly expressed and reproducible results.

WORKING GROUP 6 — *Hearing Aids*

Publication 90
1st Edition 1957

RECOMMENDATIONS FOR THE DIMENSIONS OF POLARIZED PLUGS FOR HEARING AIDS

Contents:
Dimensions and tolerances

Publication 118
1st Edition 1959

RECOMMENDED METHODS FOR MEASUREMENTS OF THE ELECTRO-ACOUSTICAL CHARACTERISTICS OF HEARING AIDS

The purpose of these recommendations is to describe practicable and reproducible methods of determining certain physical performance characteristics of air-conduction hearing aids using electronic amplification and acoustically coupled to the eardrum by means of ear inserts.

Contents:
Definition and explanation of terms
Test equipment
Test procedure

Publication 126
1st Edition 1961

I.E.C. REFERENCE COUPLER FOR THE MEASUREMENT OF HEARING AIDS USING EARPHONES COUPLED TO THE EAR BY MEANS OF EAR INSERTS

The purpose of this publication is to recommend a coupler for loading the earphone with a specified acoustic impedance when determining the physical performance characteristics, in the frequence range 200-5000 Hz of air-conduction hearing aids using earphones coupled to the ear by means of ear inserts. The coupler described is a development of an earlier 2 cm^3 coupler.

Contents:
Construction and design data

Final Draft
approved for publication

RECOMMENDATIONS FOR PURE TONE AUDIOMETERS FOR GENERAL DIAGNOSTIC PURPOSES

The audiometer covered by this recommendation is a device designed for general diagnostic use and to determine the hearing threshold levels of individuals by monaural air-conduction earphone listening and by bone-conduction.

Final Draft
approved for publication

RECOMMENDATIONS FOR PURE TONE SCREENING AUDIOMETERS

"Screening" is the process of dividing a group of individuals into two groups according to whether they do or do not have hearing threshold levels greater than certain minimum values at one or more specified frequencies.

WORKING GROUP 7 — *Ultrasonics*

Final Draft
approved for publication

RECOMMENDATIONS FOR TESTING AND CALIBRATION OF ULTRASONIC THERAPEUTIC EQUIPMENT

Contents:
Acoustic output characteristics
Safety and control features
Tolerances
Calibration procedures
Technique of measurements
Bibliography

WORKING GROUP 8 — *Sound Level Meters*

Publication 123
1st Edition 1961

RECOMMENDATIONS FOR SOUND LEVEL METERS

These recommendations apply to sound level meters for general purposes, with only moderate precision requirements. They do not apply to apparatus for measuring impulsive sounds.

Contents:
General technical characteristics
Microphone characteristics
Characteristics of the indicating instrument
Amplifier characteristics
Calibration and checking of the sound level meter

Publication in preparation

RECOMMENDATIONS FOR PRECISION SOUND LEVEL METERS

These recommendations apply to a high precision sound level meter for laboratory use or accurate measurements in which stable, high fidelity and high quality apparatus is required.

Publication in preparation

SPECIFICATION FOR OCTAVE, $^1/_2$ OCTAVE AND $^1/_3$ OCTAVE BAND FILTERS INTENDED FOR THE ANALYSIS OF SOUNDS AND VIBRATIONS

WORKING GROUP 9 — *Shock and Vibration*

Final Draft approved for publication

RECOMMENDATIONS ON METHODS FOR SPECIFYING THE CHARACTERISTICS OF VIBRATION PICK-UPS FOR SHOCK AND VIBRATION MEASUREMENTS

Equipment for the measurement of shock and vibration has found widespread use in recent years because of the increasing importance of the dynamic effects in apparatus having a large power-to-weight ratio and also because of the urgent need for information leading to a better understanding of the effects of vibration and shock to which delicate equipment is subjected.

Publication in preparation

SPECIFICATIONS FOR AUXILIARY EQUIPMENT FOR SHOCK AND VIBRATION MEASUREMENTS

In many instances it is necessary to insert auxiliary equipment between the vibration or shock detector and the indicator or recorder in the form of amplifiers, carrier systems, filters, etc.

Already this summary enumeration permits us to recognize the important amount of information, contained in these standards. It is unavoidable that various figures and methods had to be fixed in an arbitrary manner, in spite of exposing them to criticism. We must, however, consider that they are the benefits of international collaboration on a broad basis, i.e., of a teamwork, whose achievements deserve our implicit trust.

Before closing, we must stress the point that these standards are not only of importance to producers and users of acoustical apparatus and methods, but also to the scientist, who can very often base his research work on the terminology, the definitions, the units and the methods.

Biological Effects of Noise

ETIENNE GRANDJEAN

Department of Hygiene and Work Physiology, Swiss Federal Institute of Technology, Zurich.

Definition

Noise is by definition annoying sound. In fact, sound waves have not only the function of acoustical information but also side effects which are, under certain conditions, perceived as disturbing, wakening or annoying noise.
The physiological aspects of these side effects shall be the main object of this paper; other effects like hearing losses or interference with speech communication may be dropped since they are already well-known.

Auditory mechanisms in the brain

Sound energy is converted into nerve impulses of bioelectrical nature after reaching the receptor organ in the internal ear. The physiological organization of the auditory pathways consists of two distinct systems. One is a "specific projection" system leading directly to the cortical auditory area and concerned with integration and perception of the acoustic signals. The other one is a non specific projection system branching off from the specific one in the "activating" reticular formation and spreading diffusely into different functional systems of the brain. This second projection is concerned with arousal of the different sensory, motor and vegetative functions.
The reticular formation is somehow a central alerting or activating system which enables the whole organism to react in an adequate way to the given outworld situation. The stimulation of the reticular formation by noise or other sensory stimuli will arouse the animal or human being in order to enable it to focus its attention to the external acoustical information. Therefore, the ability of paying attention depends from the reticular formation.
In addition, two other important functions which are regulated most probably by the reticular formation must be pointed out here. Animals and human beings have the capacity of increasing their attention to a special kind of sensory information and of reducing simultaneously irrelevant sensory informations. The reticular formation has therefore the rôle of a discriminative input filter.
Another very important function is the habituation. If a noise stimulus in repeated over and over, the reticular formation will gradually lose its responsiveness to the noise stimuli. This capacity of reducing the alerting response is called habituation.

In the process of habituation, cortex and reticular formation are able to extinguish selectively some stimuli which prove to be irrelevant. This is a possible explanation for the capacity of animals and human beings of sleeping peacefully in the presence of many potentially alerting noise stimuli. It explains also why a mother can ignore all sorts of noise but is wakened up by any small noise of her child: through habituation she learns to distinguish between irrelevant traffic noise and the significant sound manifestation of her child.

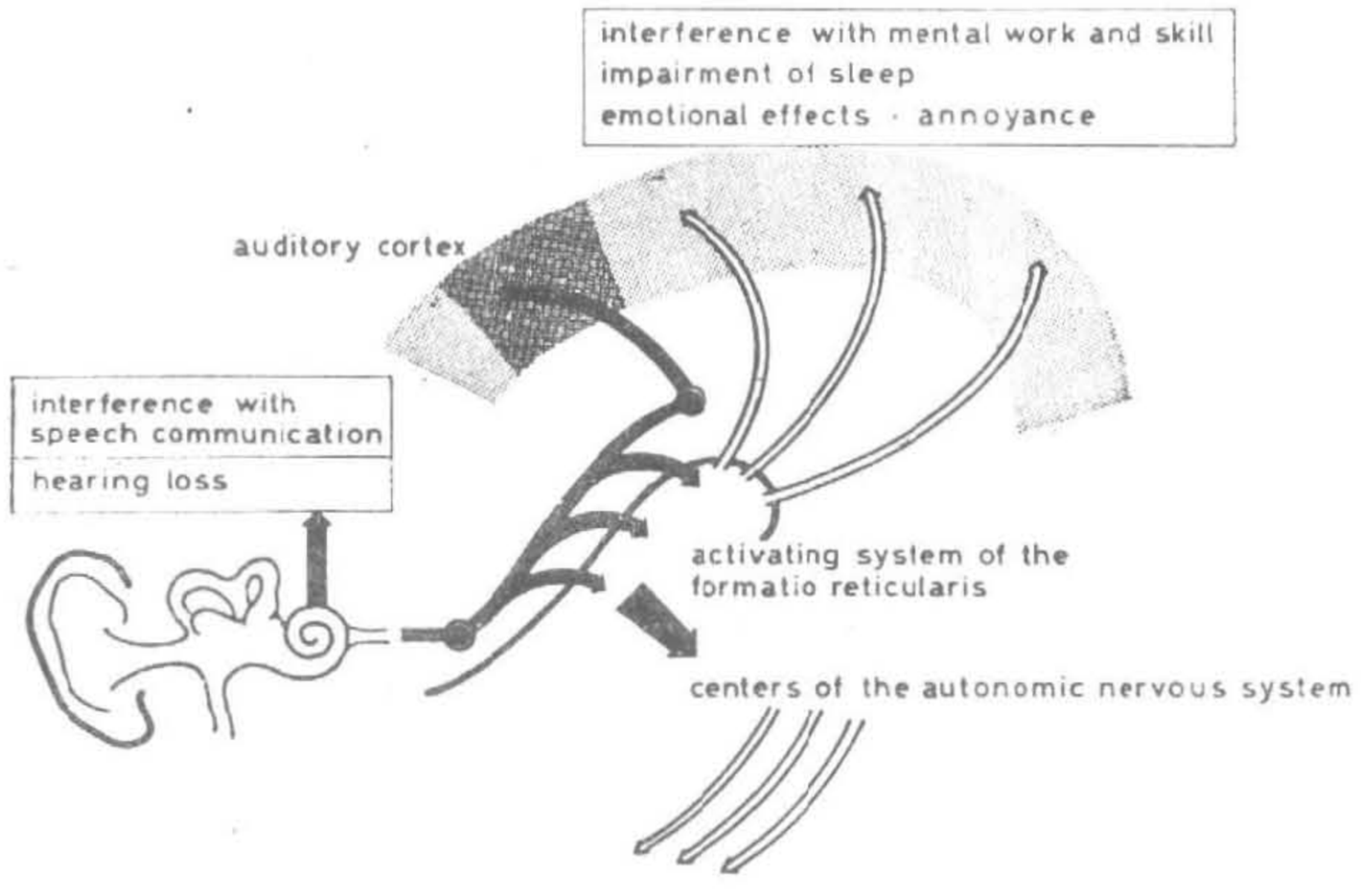


Figure 1, Effects of noise in relation to auditory specific and non specific pathways in the brain.

Figure 1 summarizes the most important biological effects of noise:

1. Noise may affect the internal ear and produce hearing losses.
2. Noise may interfere with speech communication.
3. There are side effects based on the stimulation of the reticular activating system by noise. Such stimulations are interfering with mental and skill work by distraction, are producing impairment of sleep, subjective emotional effects, and vegetative responses.

The following two important physiological facts have to be pointed out once more:

4. Noise is not the only factor stimulating the reticular activating system; all sorts of afferent and emotional stimuli may produce the same alerting responses.
5. The phenomenon of habituation prevents human beings and animals from reacting undistinctly and every time to any kind of noise by all the mentioned side effects.

Interference with mental and skill work

It is known by experience that thinking and creative mental work is more difficult in a noisy than in a quiet environment. The literature contains many investigations dealing with the effects of noise on skill tasks, on mental work and physical capacity of man under laboratory conditions as well as in factories. Some authors were not able to find quantitative effects of noise, while others observed under laboratory conditions a decrease of skill and vigilance, and in factories a decrease of productivity.

Whatsoever the observations of all the authors are, in one point they all agree: noise pro-

duces always subjective feelings of annoyance when it interferes with skill or mental work. Careful analysis of the literature led KRYTER 1950 to the conclusion that continuous and expected noise had no influence on psychomotor functions. It is known to-day that the negative results are partly due to the fact that noise affects vigilance primarily if it is irregular and unexpected, or if vigilance requirements are high and of long duration. This has been demonstrated in an elegant manner by BROADBENT who found a marked decrease of mental capacity.

From the physiological point of view, a decrease of mental capacities can indeed be expected if the noise situation does not allow the process of habituation.

Man is certainly able to do mental or skill work in a noisy environment. However, for this respect he has to increase his attention in order to ignore the noise. Certainly, many of the contradictory results reported up to now are due to such changing levels of human attention.

As an opposite effect of noise, increases of productivity have been observed also in monotonous situations. Noise can improve performance by its wakening effect if environment and work by themselves have no stimulating effects on the reticular formation. This may explain the success of music under monotonous working conditions.

Impairment of sleep

One of the most important noise effect in everyday life is the impairment of sleep. Unfortunately, this is just the kind of noise effects which scientific investigation failed to study until now. Nevertheless, it can be stated that a deep and undisturbed sleep is necessary for a healthy life. During sleep, muscles, brain and various organs show a reduced activity in order to allow the restitution of energy and the assimilation of nutritive elements. The shortening or frequent interruptions of sleep interfere with recovery and reduce well-being and resistance. During sleep, we close the eyes in order to protect the sleeping brain from light stimuli. On the other hand, we cannot close the ears; the only preventive mechanism is the process of habituation which has been already mentioned above. The only important investigation on sleeping disturbances by noise has been done by STEINICKE with 343 subjects. Between 2 and 7 o'clock a.m., noise stimuli with frequencies ranging from 60 to 5000 cps were given during 3 minutes to the sleeping subjects. The results showed prominent individual differences: some 10–20% woke up with 30–35 phones only, while others were still asleep with 70 phones. This confirms the everyday experience that the sleep of some sensitive individuals is impaired by the least noise, while others are able to sleep in presence of high traffic noise, passing trains, airplanes or even firing guns. STEINICKE concluded from these experiments that noise should not exceed 35 phones in bedrcoms during night. Such research should be intensified in order to study problems like habituation to different noise qualities and quantities.

Noise effects on the autonomic nervous system

For long time, physiologists have been interested in vegetative reactions to noise. CANNON found already 1929, that noise inhibited digestive functions. He then compared noise effects

with fear, pain and rage responses. He presented evidence of all such responses belonging to a common physiological emergency system. Many physiologists have been studying the various vegetative effects of noise stimulation since CANNON reported his results. Specially intensive research along this line has been done by LEHMANN and his collaborators since World War II.

All the effects of noise on autonomic nervous system which have ever been observed by physiologists can be summarized as follows:

Under proper laboratory conditions, noise stimuli produce

1. vasoconstriction in the skin, increase of blood pressure and heart rate, expressing changes in the regulation of blood circulation;
2. reduction of salivary and gastric secretion and decrease of peristaltic movements, which can be considered generally as symptoms of a slowing of digestive functions;
3. increase of metabolic functions.

All these responses occur after all kind of afferent or emotional stimulation of the reticular formation and are transitory and not noise specific. Furthermore, they occur also during sleep and are therefore not related to subjective annoyance. The responses of the autonomic nervous system to noise are of a limited value for the evaluation of noise effects, since they are not connected to sleep impairment and to annoyance.

Nevertheless, it is interesting to remember the meaning and the importance of audition in animals: there, hearing is less important for communication than for the perception of danger. With respect to preservation of life, this represents an essential physiological mechanism, since noise as an alerting signal not only arouses the attention but changes also the regulation of the whole organism by adapting circulatory and other functions to flight or to fight. This was certainly the original meaning of the reaction of the autonomic nervous system to noise.

Annoyance. Emotional effects of noise

The most delicate effects of noise are with no doubt the emotional ones. According to general experience, particular sounds may produce in some people pleasant, in other people unpleasant feelings. In addition, the same sound may be perceived by the same subject sometimes as pleasant and sometimes as unpleasant. As it is well-known, the extent of annoyance depends from the quality as well as from the quantity of noise and in addition from subjective conditions varying individually.

Individual personal attitudes toward noise are most important. Motorcyclists, workers, musicians are not annoyed by the noise of their own activity. However, outsiders and neighbors will usually be more annoyed the more they have negative attitudes toward the particular noise or toward the noise producing people.

Another important factor is the kind of experience man has had before with noise. Noise which often produced fear or disturbed sleep or other activities may become increasingly more annoying. That means that sometimes a process opposite to habituation can occur which may be called a hypersensitivity to noise.

As mentioned before, the reticular formation can amplify as well as reduce the intensity of a noise stimulus reaching brain. It is interesting to state that also the emotional effects can be gradually amplified to a hypersentivity or in the contrary be gradually reduced by a process corresponding to habituation. It is reasonable to assume that the reticular formation which is also depending from emotional processes is the focus of these psychological phenomena. To-day, the majority of hygienists and social experts agree that the annoying effect of noise is the most common one and has therefore the greatest importance for the well-being of people living in cities and other noisy surroundings. For this reason, annoyance is the main problems of authorities, health administrations and police dealing with noise complaints.

Noise and health

There is full agreement in considering hearing losses as a serious health injury. There is less agreement about the evaluation of the importance of the side effects (impairment of sleep, annoyance, vegetative responses and interferences with speech communication and mental and skill performance) as a health hazard.

By definition of the World Health Organization, health is a state of complete physical, mental, and social well-being. This definition is wide enough to include not only hearing losses, but also sleep impairment and subjective annoyance. Controversial is still the evaluation of the noise effects on the autonomic nervous system.

Daily repeated exposures of animals like rats to intense noise produce a chronic state of high blood pressure together with other pathological changes. These are due obviously not only to a modification of vegetative regulation, but also to disturbances of endocrine regulation, i.e. of the production of some hormones.

JANSEN, a collaborator to LEHMANN, investigated the health condition of nearly 1000 workers. A group of 669 workers was exposed to noise between 90 and 120 phones at their work places, while another group of 336 workers was exposed to noises below 90 phones. A higher incidence of clinical symptoms like circulatory disturbances, heart rate irregularities and others, was found in the group with high exposure. The interviews of these 1000 workers revealed also characteristic differences between the two groups with respect to their social behaviour. Previous social backgrounds and complaints about living or salary were more or less the same in both groups. However, the frequency of social conflicts at home and at the factory was higher in the group with high exposure.

LEHMANN concluded from the results of JANSEN that the exposure to noise over many years may lead finally to chronic pathological symptoms of the autonomic nervous system. Nevertheless, the results of JANSEN and LEHMANN cannot be considered as conclusive. The two groups of workers do not seem to be comparable; professional conditions and other uncontrolled factors like tobacco consumption, age, etc., could have interfered and could be the reason, as well as noise, of the observed clinical and psychosocial differences. Therefore, the question whether noise of everyday life does produce pathological changes of the autonomic nervous system in human beings or not, remains open, and needs more research.

Conclusion

From a general physiological point of view, frequently repeated wakening and alerting stimuli obviously must interfere with recovery processes not only during sleep but also during day rest.

It is quite clear that some noise stimuli do no harm to man; the alerting reactions as well as the responses of the autonomic nervous system are physiological and normal phenomena. It is only the excess of noise stimuli which interferes with well-being and therefore with health. It is reasonable to assume that these interferences with recovery processes are the main reasons of subjective annoyance. From this point of view, annoyance must be considered as a biological protective mechanism inducing man to avoid noise and to secure recovery processes. The biological meaning of annoyance is therefore quite comparable to other feelings of uncomfort like hunger, fatigue, cold or warm sensations. All these feelings have the importance of life protecting mechanisms.

Annoyance due to noise disturbing rest is an important health hazard from the biological point of view. It is therefore reasonable to consider annoyance as a biological standard for noise abatement.

The difficulty of distinguishing between justified and non justified annoyance complaints remains as a great practical problem. Such individual differences are often caused by personal resentment, financial interests or other purely egoistic motivations. Careful psychosociological field investigations could be expected to contribute to a more complete understanding of these delicate problems.

Noise is increasing gradually in civilized countries. Simultaneously, the frequency of complaints about noise is increasing. It must be concluded that excessive noise exceeds more and more the limits of adaptation–or habituation. Noise is added to other stresses of modern life like haste and restlessness, increasing tobacco consumption, sedentary life with inadequate nutrition and in some places increasing air pollution. Noise is therefore one of the important problems of modern hygiene, and its abatement is a well justified postulate of our days.

Microwave Ultrasonics and Studies of the Solid State

by E. H. JACOBSEN

Department of Physics and Astronomy, University of Rochester, Rochester, New York.

Introduction

To a first approximation we can describe the elastic properties of a solid in terms of an array of point masses joined together by massless ideal springs, the point masses and springs respectively representing atoms bound together by electrical forces. It is the property of such a system to support a vast spectrum of possible elastic waves, the maximum frequency and corresponding minimum wavelength being determined by the mass, binding force, and spacing. For a real solid the elastic wave spectrum (which is responsible for the specific heat capacity) extends from virtually zero to approximately 10^{14} cps, with corresponding wavelengths ranging from macroscopic sample dimensions to angstroms, which is the scale of interatomic separation. Up to frequencies of nearly 10^{12} cps the wave velocity is independent of frequency and wavelength—on the basis of our simple model. At still higher frequencies the velocity is no longer a constant as the wavelength is now short enough to "see" the discontinous point-like structure. This is the phenomenon of dispersion and will always occur for wavelengths comparable in scale to the dimensions of the microstructure. Fig. 1 illustrates this behavior for a simple linear chain of point masses. Unfortunately, or perhaps not, nature is not so simple. For one thing the springs are not ideal, i.e., the restoring force is not quite proportional to the displacement, and this state of affairs causes the energy in one wave of well defined frequency to flow into all of the others having different frequencies. Thus a single wave, once excited, will decay in some characteristic lifetime τ through this irreversible degradation of energy. Generally, the higher the frequency the shorter will be τ because the density of elastic wave modes increases rapidly with frequency and so provides more paths for the degradation to occur. Moreover, the atoms are not simple point masses. They possess an outer cloud of electrons which, in addition to enduing all real solids with imperfect springs, have a life of their own and are able to absorb and emit vibrational energy in a great variety of ways. Certain of these electronic processes manifest themselves as anomalous dispersion, refraction, and rotary polarization of the elastic wave and are often connected with the phenomenon of resonance. Still other processes occur involving various collective excitations which are more difficult to describe and show up in strange ways, electrical superconductivity being an example. There are in fact a great many ways in which the elastic wave system can interact with itself

and with atomic electrons, and such interactions become more numerous and important the higher the frequency.

In contrast to very low frequencies, where the phenomena of anamolous dispersion, etc., reflect the interaction of elastic waves with simple mechanical systems obeying classical laws, the situation at high frequencies requires the use of quantum theory since atoms and electrons usually behave quantum mechanically. That is, they absorb definite units of energy $\hbar\omega$—in this case sound quanta or *phonons*, which are the counterpart to light quanta or *photons*. Thus in part of what follows we shall have to work with quantum equations of motion instead of the more familiar equations of classical mechanics appropriate to most of the problems which crop up in the range of low frequencies. Very often there is no clean break between "low" and "high" frequencies and between "classical" and "quantum" systems. Nonetheless, we can usually say that for frequencies in the audible and megacycle range the quantum treatment, though perfectly valid, is unnecessary and even awkward, and the classical mechanical description suffices. At kilomegacycle (microwave) frequencies and higher the quantum-like behavior starts to appear and so the use of quantum theory often becomes necessary—although not always—sometimes the classical equations are still adequate; it depends on the problem.

In what follows we shall illustrate a limited number of these ideas as they pertain to elastic wave propagation in solids at microwave frequencies. As for their application to studies of the solid state, implied by the title, we shall touch on it only lightly as it is still in an early stage and depends critically on further development of the experimental techniques and theoretical principles discussed below.

Generation of Elastic Waves in Solids

We have at our disposal basically two methods of generating coherent monochromatic elastic waves in and beyond the microwave frequency range. The first is the simple extension of electromechanical transducers into the high frequency region for converting electromagnetic radiation into ultrasonic waves of the same frequency. The second is the use of various excited electronic systems which are arranged so as to emit sound quanta (phonons) by stimulated emission. In principle both methods of sound generation can be used over the entire elastic wave spectrum, which in solids extends from zero to approximately 10^{14} cycles per second. However, as we shall see, current practical limitations favor the use of electrical-mechanical transducers at frequencies below roughly 10^{11} cps, whereas quantum methods such as stimulated emission appear preferable for use above 10^{11} cps.

We shall begin with a brief discussion of the electrical-mechanical transducer of which the piezoelectric crystal is a common example. For such a medium the stress-strain equations are in matrix form

$$T_i = c_{ij}\, S_j - e_{ik}\, E_k \tag{1}$$

where T_i, S_j, and E_k are respectively the components of stress, strain, and electric field, and c_{ij} and e_{ik} are the elastic and piezoelectric constants. They are respectively fourth and third rank tensors which is obscured by the matrix notation. As an illustration we consider compressional wave propagation in quartz along the x-direction so that equation (1) becomes

$$T_x = C_{xx}\, \nabla_x U - e_{xx}\, E_x \tag{2}$$

From (2) we can derive immediately the wave equation (3)

$$\varrho\left(\frac{\partial^2 U}{\partial t^2}\right) = \nabla x \cdot T_x, \tag{3}$$

which becomes

$$\nabla_x^2 U - (1/v^2)\left(\frac{\partial^2 U}{\partial t^2}\right) = \nabla x\,(d_{xx}\,E_x)\,, \tag{4}$$

with the substitutions $v^2 = c_{xx}/p$ and $d_{xx} = e_{xx}/c_{xx}$. For an unbounded one-dimensional medium the inhomogeneous wave equation (4) has the solution

$$U\,(x,t) = \int G\,(\,x - x^1 \,|\, \omega)\,\nabla x,\,(d_{xx}\,E_x)\,dx^1, \tag{5}$$

where the Green's function is

$$G\,(x - x_0 \,|\, \omega) = (i/zk)\;exp\;i\;[\omega t \mp k\,(x - x_0)],\left\{\begin{matrix}(-);\,x > x_0\\(+);\,x < x_0\end{matrix}\right\}, \tag{6}$$

assuming the electric driving field E has the time variation $e^{i\omega t}$. We apply solution (5) to the problem illustrated in Fig. (1) which comprises an infinite medium along the x-direction with an electrical boundary at X_0. To the right of X_0 d_{xx} is finite, and zero to the left. Hence the gradient of the product $(d_{xx}\,E_x)$ is the delta function $\delta\,(x - x_0)\,[dxx\,E_x]x > x_0$ which by (5) yields the solution (6).

$$U\,(x,t) = \frac{i}{2k}\,(d_{xx}\,E_0)\;x > x_0\;exp\;i\;[\omega t \mp k\,(x - x_0)],\left\{\begin{matrix}(-)\;;\,x > x_0\\(+)\;;\,x < x_0\end{matrix}\right\} \tag{6}$$

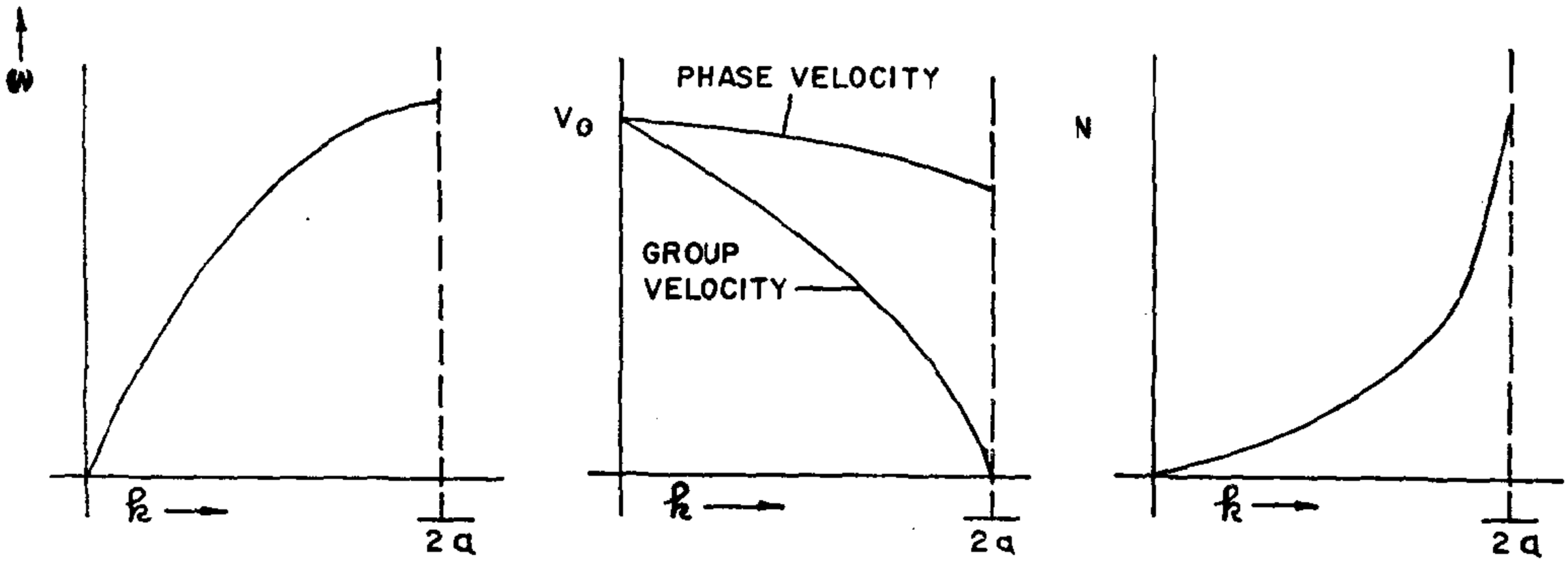


Fig. 1. Properties of wave propagation on a linear chain of point masses separated by a distance d.

(a) Dispersion in angular frequency ω *vs* wave vector $k = \frac{2\pi}{\lambda}$.

(b) Behavior phase velocity $v\phi$ and group velocity $vg = \frac{\delta\omega}{\delta k}$ as a function of k.

(c) Density of allowed elastic wave modes *vs* k.

Eq. (5) shows that $u(x, t) \approx 0$ unless the piezoelectric stress ($d_{xx} E_x$) changes rapidly on the scale the sound wave length λ. The essentially discontinuous character of ($d_{xx} E_x$) pictured in Fig. 1a applies directly to the surface of a piezoelectric crystal except that X_0 then defines an elastic as well as an electrical discontinuity. The effect of an elastic discontinuity at X_0 simply multiplies eq. (6) by 2 and removes the solution for $X < X_0$. This can be easily seen by imagining a boundary at $X_0 - L$, which provides an image source at $X_0 - 2L$, and letting $L \to 0$.

At frequencies of a few kilomegacycles the sound wave length is typically of order 10^{-4} cm which shows that for high frequencies the only important source is expected to be a free surface because of the difficulty of establishing E-fields which vary significantly on the scale of a few thousand Angstroms.

From solution (6) it is easy to calculate the amount of electromagnetic energy transformed into sound. The sonic energy density $W = \frac{1}{2} C_{xx} (\nabla_x U)^2$ multiplied by the sound velocity V is the flow of sound energy per unit area (energy flux S) away from the surface X_0, i.e.,

$$S = \tfrac{1}{2} C_{xx} d_{xx}^2 E_x^2 V. \tag{7}$$

Neglecting material losses, nothing prevents our converting virtually all of the electromagnetic energy into a sound wave. The practical difficulty of so doing depends on the size of the piezoelectric constant d_{xx}. If d_{xx} is large only a modest value of electric field need be provided. If d_{xx} is small, as with quartz, a large E-field is required which implies the use of a high Q resonant cavity. For a piezoelectric rod of cross-section A, placed in a re-entrant cavity the ratio of sound power to electromagnetic power is by (7) simply [1] (see Fig. 1b)

$$\frac{P_{sound}}{P_{r.f.}} = \frac{4\pi C_{xx} d_{xx}^2 AVQ}{\varepsilon \omega V^1}, \tag{8}$$

where ε is the dielectric constant of the rod and V^1 the effective cavity volume. For quartz at 10,000 Mc the practically obtainable parameters in (8) limit this ratio to somewhat less than 10^{-3}. As a result of increased efforts in Materials Science research we may confidently expect the emergence of piezoelectrics possessing significantly larger electro-mechanical coupling constants. Ferroelectric transducers hold some promise in this regard except for their propensity to undergo various phase transformations between room temperature and a few degrees Kelvin, the latter being the required temperature domain for all experiments in ultransonics above a few thousand megacycles.

The foregoing analysis, aside from minor simplifications, applies directly to any transducer which couples elastic strain to *electric* fields E. As we would naturally expect there also exist transducers which couple elastic strain to magnetic fields H. An example of such a magnetoelastic transducer, proposed by Kittel [2] and demonstrated by various workers [3-6], is the ferromagnetic medium. For completeness we discuss briefly the ferromagnetic transducer for the generation of *shear* waves. If we assume the propagation of shear waves along z and parallel to a D.C. magnetic field it is possible to derive an inhomogeneous wave equation [2] which is the counter part of (4)

$$\nabla_z^2 U\pm - \frac{1}{V_t^2} \frac{\partial^2 U\pm}{\partial^2 t} = \nabla_z \left(\frac{\lambda M\pm}{M_s} \right), \tag{9}$$

where $U^{\pm} = Ux \pm iUy$, M_s is the saturation magnetization making a small precession angle φ relative to z and having components M_x and M_y, $M^{\pm} = M_x \pm i\,M_y$, and λ = magnetostriction constant. By the methods outlined above the solution to (9) is

$$U^{\pm}(z, t) = \int G(z - z^1 \mid \omega)\, \nabla_z, \left(\frac{\lambda M^{\pm}}{Ms}\right) dz^{\cdot}, \tag{10}$$

where G is defined as in (6), and the expression for the sonic energy flux S becomes [3]

$$S = \tfrac{1}{2}\varrho\, V_t^3\, \lambda^2\, \varphi^2, \tag{11}$$

with $\varphi \cong M^{\pm}/Ms$ and where φ is proportional to the driving field. Depending on the magneto-elastic coupling coefficients and magnetic field orientatio. compressional waves can also be generated [5]. In any case a steep gradient of magneto-elastic stress must be provided, and as with the piezoelectric transducer, a free surface inherently provides the required discontinuity. The magneto-acoustical generation of sound described above is different from that excited by spin waves as discussed by various authors. However, as pointed out by Bömmel and Dransfeld, [3] there is little distinction between these two processes if the sample is thinner than an acoustical wavelength. The derivation of eq. (11) assumes an infinite sample so that no standing waves are set up. The efficiency of magneto-elastic sound generation depends on the coupling coefficients and appears to be somewhat greater than for quartz.

The previous discussion neglects the fact that wave propagation in a transducing medium consists of a set of coupled waves. For a piezoelectric crystal a polarization and elastic strain field propagate concomitantly so that eq. (1) must be supplemented by another linear equation connecting elastic strain and electric polarization. Thus a complete discription requires the set of simultaneous equations (in tensor form)

$$\begin{aligned} T_{ij} &= C_{ijkl}\, S_{kl} - e_{ikl}\, E_l \\ D_i &= e_{ijk}\, S_{jk} + \varepsilon_{ij}\, E_j \end{aligned} \tag{12}$$

which shows that the wave energy is distributed between the elastic strains and electric polarization. The fractional distribution will depend on the relative magnitude of the piezoelectric and elastic constants. For quartz, the electric polarization field contains of order 10^{-2} of the elastic strain field energy, whereas in ferroelectrics the polarization energy will be appreciably higher. A more complete analysis [7, 8] of eqs. (12) shows that eq. (1) is a good approximation provided the piezoelectric coupling is small, in which case the electric polarization effects act as a small perturbation on the elastic constants of order $e^2/c\varepsilon$ for a perfect insulator. Strictly speaking we should label the elastic constants in (12) as those taken at constant electric field, and ε, the dielectric permittinity, at constant strain.

Similar ideas apply to elastic wave fields in ferromagnetic media in which case eq. (9) is replaced by the set of equations (13) which take into account the coupling between elastic shear and magnetic moment [2].

$$\begin{aligned} &(\omega \pm \omega^1)\,(Mx \pm i\,My) = \pm\, i\gamma bk\,(Ux \pm i\,Uy) \\ &(bk/Ms)\,(Mx \pm i\,My) = -\,i\,(\omega^2\varrho - k^2K)\,(Ux \pm iUy), \end{aligned} \tag{13}$$

where γ is the gyromagnetic ratio, K the elastic shear constant, and b the magnetoelastic coupling constant for shear waves. However, for the ferromagnetic medium, additional complications arise having to do with spin waves and dispersion effects when the elastic wave frequency approaches that of ferromagnetic resonance. We shall not discuss these phenomena further in the context of ferromagnetic systems, but will return to a closely related matter in connection with elastic wave propagation in paramagnetic crystals, the nature of which is analogous to that encountered in a ferromagnet. Before going on to discuss the paramagnetic case, we digress briefly to consider the nature of wave propagation in bounded and infinite media.

Propagation of Elastic Waves

Neglecting losses, the equations of motion for a simple elastic medium, by which we mean the absence of electrical and magnetic polarization effects, are of the form [9]

$$\varrho \frac{\partial^2 Ui}{\partial t^2} = \sum_{j=1}^{3} \frac{\partial T_{ij}}{\partial x_j}, \; i = 1, 2, 3, \tag{14}$$

and posses the plane wave solutions

$$U_j = A_j \, e^{i\,(\omega t - k \,.\, r)}. \tag{15}$$

Insertion of the solution (15) in eq. (14) yields a dispersion relation between phase velocity and wavelength in the usual way by equating the determinant of the coefficients to zero. The wave vector k defines the direction of propagation, i.e., k is normal to the surface of constant phase. However, the direction in which energy flows will not be parallel to k in general. The energy flux vector S defines the direction and rate of energy transport normal to a unit cross-section, and can be defined by means of the continuity equation:

$$\text{(rate of energy production/unit volume)} = \frac{\partial}{\partial t}\text{(energy density)} + \text{div (energy flux)} \tag{16}$$

Though we omit the details [9, 10], we can obtain from (16) an expression for the components of S for a wave propagating along an axis n:

$$S_i^{(r)} = \frac{\omega^2}{2 V_r} C_{ijkn} A_j^{(r)} A_k^{(r)}, \tag{17}$$

where r denotes the mode of propagation (i.e. one of the three possible directions of particle motion), ω the frequency, and $A_j^{(r)}$ the component of particle motion along j for mode r. The c_{ijkn} are the elastic constants in tensor form relative to the particular coordinate system used. Eqs. (15) and (17) show that S and k are not parallel unless k is parallel to certain crystal symmetry axes. The results of a more detailed analysis [10] reveal that for compressional waves S deviates from k unless k is parallel to an axis of twofold or higher rotational symmetry, or normal to a reflection plane or normal to an axis of sixfold symmetry. For transverse waves the same result obtains if the propagation direction is a two-

fold, fourfold, or sixfold symmetry axis, or normal to a plane of reflection symmetry. Such axes are termed pure mode axes. Unless k is directed along such an axis, we may expect in general a deviation of energy flux from the propagation direction, two examples of which follow.

The first concerns propagation of transverse waves along an axis of threefold symmetry such as the (111) direction of a cubic crystal or the (001) direction of a trigonal crystal. As these directions are degenerate, there being no preferred directions of particle vibration for a pure transverse wave, the deviation of energy flux manifests itself in the form of internal conical refraction—a phenomenon which seems to have made its debut in optics [11].

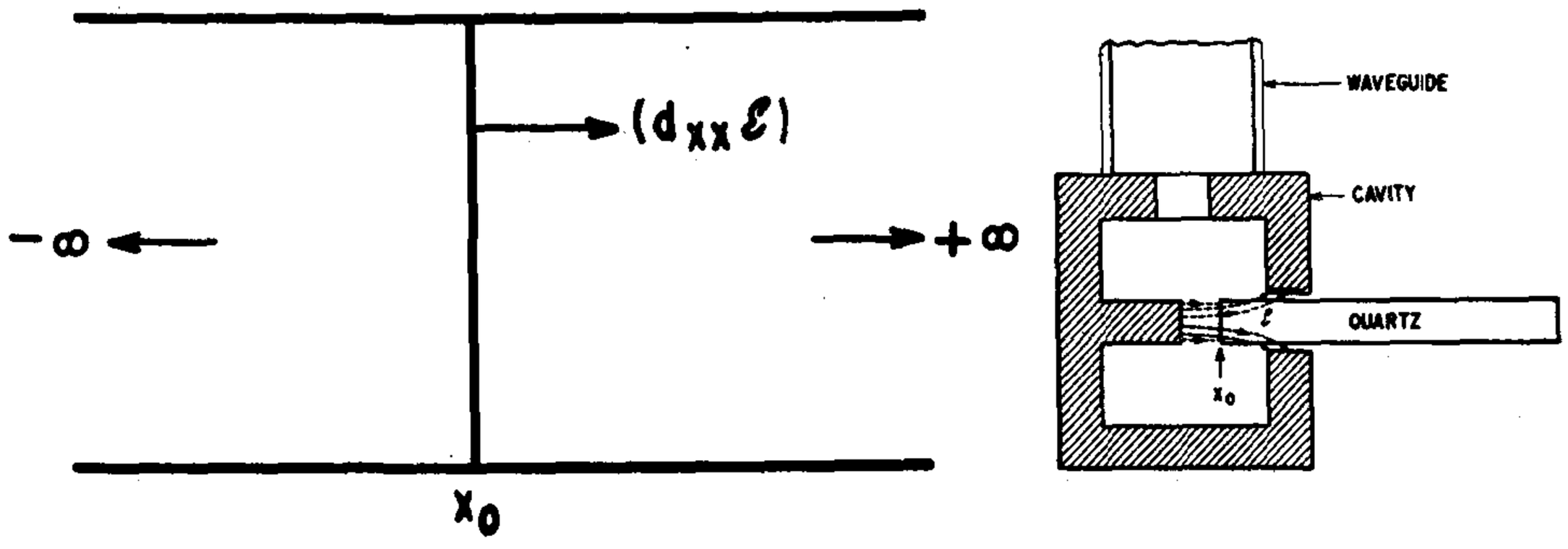


Fig. 2.
(a) Discontinuity in the piezoelectric stress $d_{xx} E_x$ at x_0 is wave source.
(b) Coupling of microwave electric fields to a quartz rod by means of a re-entrant resonant cavity. The free surface at x_0 is the major sound wave source.

The geometry of the cone of internal conical refraction is sketched in Fig. 2; as the direction of particle motion is rotated through an angle φ, the energy flux vector rotates through 2φ in the opposite direction. Questions of superposition arise in this case, of which we omit discussion here, but which are delt with at length by Waterman, and we refer the reader to his excellent paper [10]. The second type of energy flux deviation occurs when the direction of propagation k is misoriented relative to a pure mode axis. For example, if we assume propagation of a longitudinal wave nearly along the 100 direction of a cubic crystal eqs. (17) predict a result like that given in Fig. 3a. Furthermore, one can show from these equations that such a beam striking a stress-free plane surface with the angle of beam incidence equal to the deviation angle of the energy flux, and with surfaces of constant phase parallel to the reflecting surface, is reflected back on itself. This peculiar law of reflection is illustrated in Fig. 3b. Thus for a quasi-infinite medium elastic wave propagation is described adequately by the simple plane wave solutions (15) and the energy flux vector (17). In contrast to the infinite domain, however, most experimental arrangements employ samples with cross section comparable to that of the transducer and of finite length. Consequently very complicated elastic wave fields will exist owing to the presence of boundaries, and the simple plane wave solutions (15) are no longer adequate. Under these circumstances, elastic disturbances propagate via guided wave modes similar to those en-

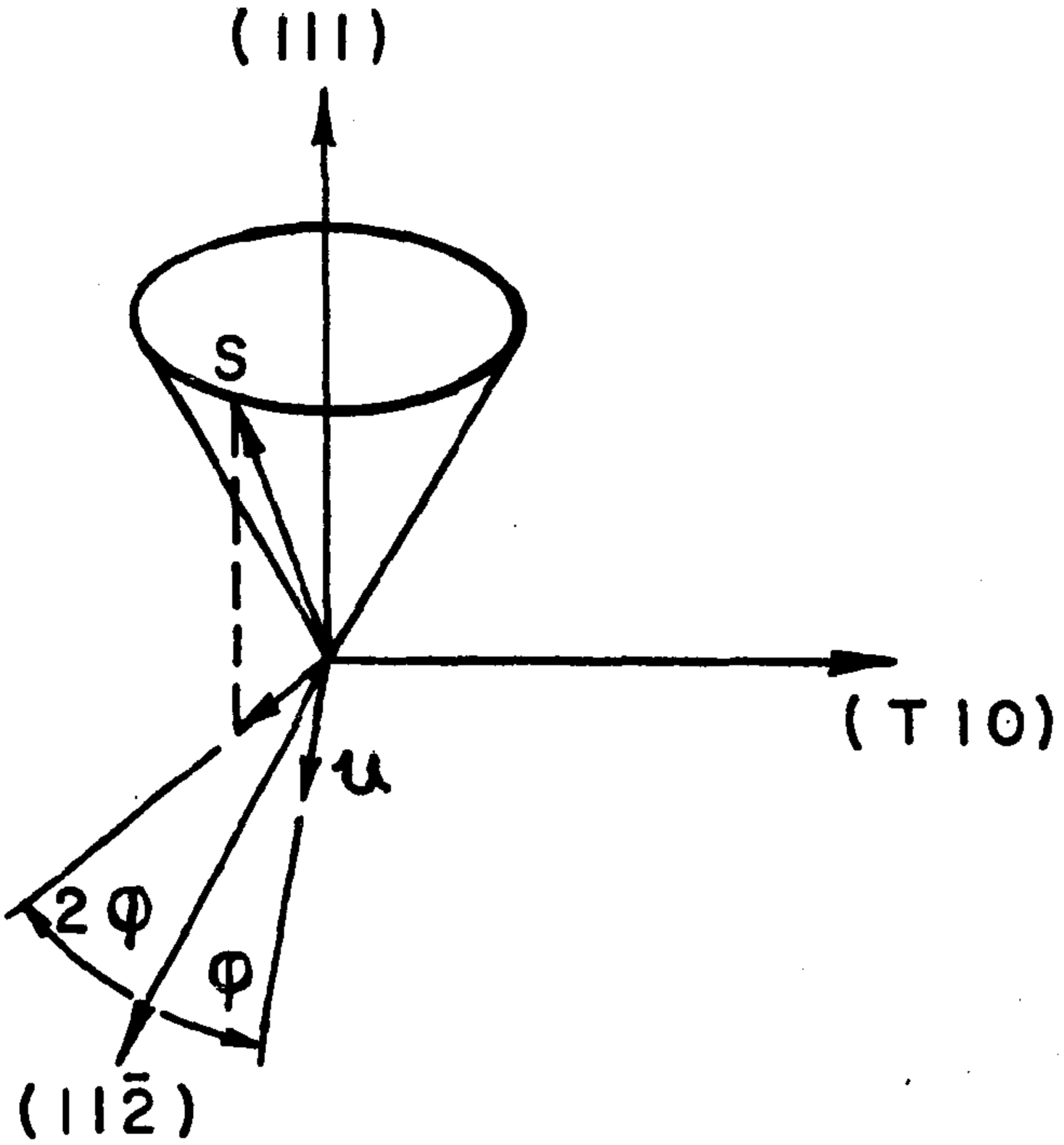


Fig. 3. Internal cone of refraction generated by energy flux vector S as transverse displacement vector rotates through angle π.

countered in the transmission of electromagnetic energy through dielectric or metalic pipes. Analysis of even the simplest case of guided elastic waves in solids is quite involved, and we shall merely quote the salient features.

Basically, pure transverse or pure longitudinal waves are not allowed modes of propagation in a bounded solid—even an isotropic one. Imagine a length of rod cut from a cubic crystal with axis parallel to the (100) direction, and ends flat and perpendicular to this axis. If now one end is caused to vibrate as a piston source parallel to the axis, the ensuing pure longitudinal plane wave front so generated will interact with the walls and eventually take on a more complicated spacial variation as a function of radial distance from the center. Analysis shows that under such circumstances a plane wave front is a superposition of a number of guided wave modes each of which comprise a mixture of longitudinal and transverse particle displacement and each traveling with a slightly different phase velocity. Fortunately in practice a plane wave so launched will remain more or less in tact over a usable distance, depending on the elastic properties, crystal perfection, dimensions, and shape of sample. Torsional waves, in contrast to quasi-compressional or shear waves, can be propagated as plane waves in an isotropic bounded medium. The lowest mode in this case is a plane wave and is nondispersive, and its excitation requieres that the displacement amplitude be proportional to the radius. The nondispersive character of this mode is made use of in delay lines when undistorted pulse propagation is required. For an extensive treat-

ment of elastic wave propagation in bounded media we refer the reader to the work of Redwood [12].

Our discussion up to this point has neglected the attenuation arising from anharmonic forces which produce phonon-phonon scattering. Experiments [13, 14] show that at microwave frequencies the attentuation is unexpectedly low at temperatures in the helium and hydrogen range, and rises steeply at about 20°*K*, slowly leveling off above 77°*K*. Fig. 4 illustrates a typical behavior for compressional waves in quartz. Bömmel and Bransfeld [13] have presented a simple phenomenological theory in which the microwave sound perturbs the thermal phonon distribution and so looses energy irreversably. A refined version of this theory has been developed by Woodruff and Erhenreich [15]. To date our knowledge of phonon-phonon loss mechanisms is very incomplete and refined experimental data and theoretical analysis are needed.

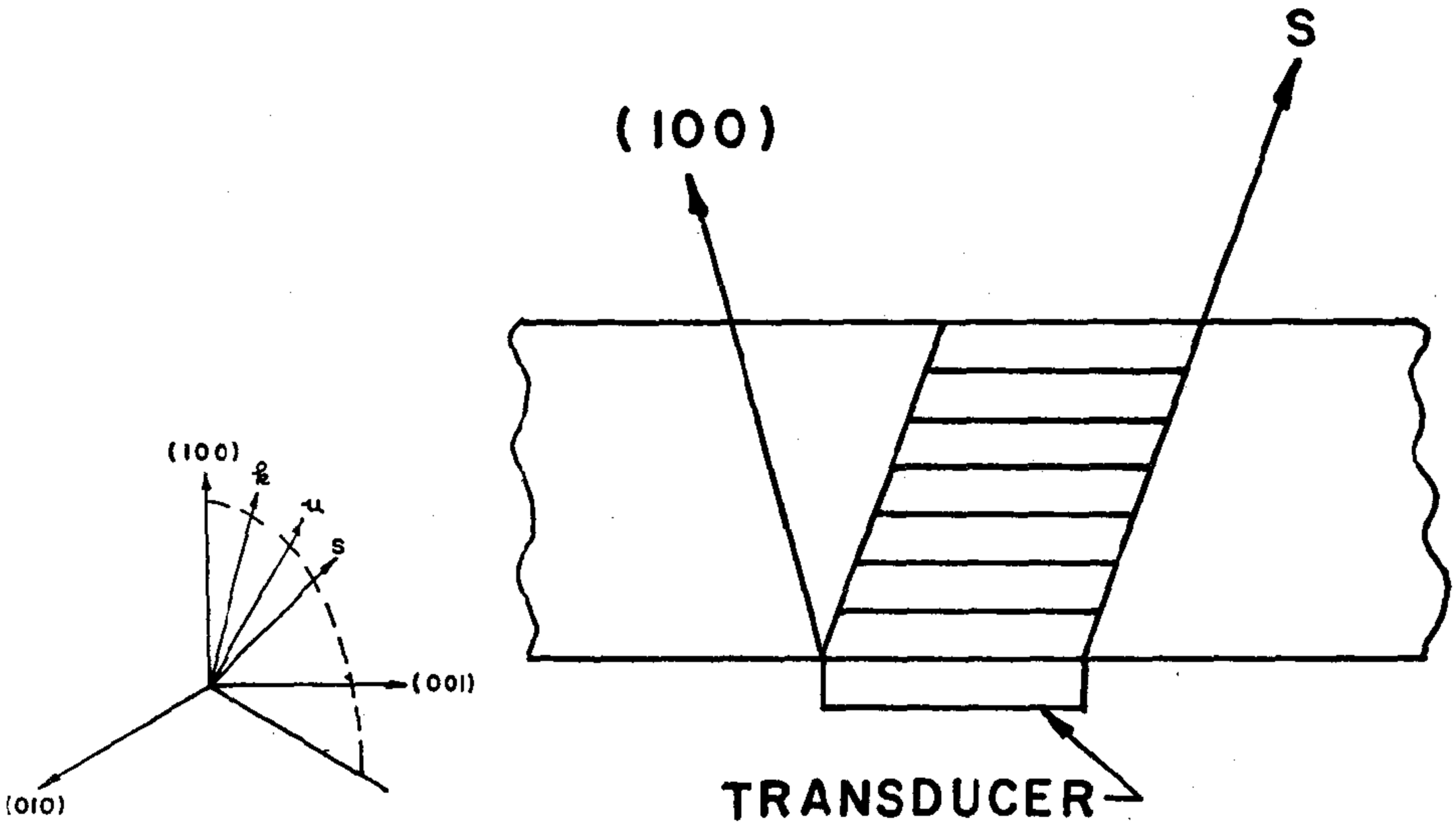


Fig. 4.
(a) Deviation of energy flux vector *S*, displacement *u*, and wave vector *k* caused by small misorientation relative to a pure mode axis.
(b) Reflection of beam when angle of incidence to free surface equals deviation angle of the energy flux. See text.

Spin Systems and Ultrasonic Dispersion

For kilomegacycle elastic waves we should expect to induce transitions between the zeeman levels of paramagnetic ions when the phonon energy $\hbar\omega$ is equal to the level separation. This is an example where the quantum aspects of elastic wave propagation begin to show up. Such transitions have in fact been observed in a limited number of systems [16-18]. In any case the spin-phonon coupling leads to a variety of spin-lattice relaxation mechanisms, and, in some cases, to behavior such as rotary polarization and anomalous dispersion of elastic waves in a manner analogous to that of light. Such behavior occurs quite readily for

electrons comprising the outer shells of atoms, in which case the electron feels elastic strains through its electrical environment. When the spin-phonon coupling is strong, i.e., when the energy levels are strongly influenced by elastic strains via the crystalline electric fields, dispersion effects show up directly in pulse echo experiments, as first demonstrated by Shiren [17].

To describe this situation we shall start with a model [19] analogous to that of the harmonic oscillator used in elementary treatments of electromagnetic dispersion. Our model is the spin system $S = \frac{1}{2}$. We proceed by writing down the Hamiltonian of the complete system (lattice + interaction + spin) and derive from it a set of coupled linear equations of motion, the simultaneous solution of which yields a dispersion relation between wave velocity and frequency. For simplicity we consider first the case of longitudinal waves propagating along the x-direction of a point mass lattice having one atom per unit cell. The Hamiltonian is then

$$\mathfrak{H} = \sum_n \left\{ \frac{P_n^2}{2m} + \frac{K}{2}(U_n - U_{n+1})^2 + \hbar\varepsilon\,(U_{n+1} - U_{n-1})\,S_x^n + g\beta H S_z^n \right\}. \qquad (18)$$

U_n and P_n are respectively the displacement and momentum of the n'^{th} atom along x. K is the restoring force between nearest neighbors and H the d.c. magnetic field along the z-direction. S_x^n and S_z^n are the x and z components of the spin for the unpaired electron on atom n, and ε is the coupling constant (which we treat phenomenologically) between the strain at position n and the spin components S_x^n.

In order to determine the behavior of the system we must obtain the quantum or Heisenberg equations of motion for the quantities P_n, U_n, and S^n. We do this by means of the so-called commutation relations which give the time rate of change of any quantity A according to the rule $i\hbar\dot{A} = [A\mathfrak{H} - \mathfrak{H}A] \equiv [A, \mathfrak{H}]$. (e.g. see *Quantum Mechanics* by L. I. Schiff, McGraw-Hill, Chapter 6) Thus

$$\begin{aligned}
\dot{P}_n &= \frac{1}{i\hbar}[P_n, \mathfrak{H}] = K(U_{n+1} + U_{n-1} - 2U_n) + \hbar\varepsilon\,(S_x^{n+1} - S_x^{n-1}),\\
\dot{U}_n &= \frac{1}{i\hbar}[U_n, \mathfrak{H}] = \frac{P_n}{m}\\
\dot{S}_x^n &= \frac{1}{i\hbar}[S_x^n, \mathfrak{H}] = -g\frac{\beta H}{\hbar}S_y^n,\\
\dot{S}_y^n &= \frac{1}{i\hbar}[S_y^n, \mathfrak{H}] = -\varepsilon\,(U_{n+1} - U_{n-1})\,S_z^n + \frac{g\beta H}{\hbar}S_x^n,
\end{aligned} \qquad (19)$$

where $(S_x^n, S_y^n) = S_z^n$, etc. $\beta = eh/2mc$ so that the S_i do not contain h. We omit for the time being the effects of spin-spin interaction, and we assume $\dot{S}_z \approx 0$.

The eqs. (19) can be re-arranged to give the set of coupled linear equations

$$\begin{aligned}
m\ddot{U}_n &= K(U_{n+1} + U_{n-1} - 2U_n) + \hbar\varepsilon\,(S_x^{n+1} - S_x^{n-1})\\
\ddot{S}_x^n &= \omega_0\varepsilon\,(U_{n+1} - U_{n-1})\,S_z^n - \omega_0^2 S_x^n.
\end{aligned} \qquad (20)$$

Assuming a traveling wave solution of the form $e^{i(\omega t - kna)}$ for U_n and S_x^n, where "na" labels the position along x of the n^{th} atom, these equations become

$$\begin{aligned} -\omega^2 m U_n &= 2K(\cos ka - 1)\, U_n - 2i\hbar\varepsilon (\sin ka)\, S_x^n \\ -\omega^2 S_x{}^n &= -2i\omega_0 \varepsilon (\sin ka)\, S_z{}^n U_n - \omega_0^2 S_x^n . \end{aligned} \tag{21}$$

where $k = 2\pi/\lambda$ and $\hbar\omega_0 = g\beta H$. Under the assumption that S_z is constant in space and time, the determinant formed from (21) leads to the dispersion relation (22).

$$\left(m\omega^2 - 4K \sin^2 \frac{ka}{2}\right)(\omega^2 - \omega_0^2) + 4\varepsilon^2 g\beta H \langle S_z \rangle \sin^2 ka = 0 \tag{22}$$

$\langle S_z \rangle$ is the average value of S_z per atom. At microwave frequencies the sonic wavelength is large compared to the atomic spacing "a", and so we can take the long wave limit by replacing sin ka by ka in (22).

With the definition $V_0^2 \equiv Ka^2/m$, the square of the phase velocity in zero magnetic field, we obtain finally the relation for sonic index of refraction as a function of frequency:

$$\left(\frac{V_0}{V}\right)^2 = \frac{1}{1 + \dfrac{4\varepsilon^2 g\beta H \langle S_z \rangle / K}{\omega_0^2 - \omega^2}} = 1 - \frac{\dfrac{4\varepsilon^2 g\beta H \langle S_z \rangle}{K}}{\omega_0^2 + \dfrac{4\varepsilon^2}{K} g\beta H \langle S_z \rangle - \omega^2} . \tag{23}$$

It is interesting to compare the above equation with that obtained for optical dispersion near the resonant frequency ω_0 of a dipole harmonic oscillator.

$$\left(\frac{c}{V}\right)^2 = 1 + \frac{4\pi Ne^2/m}{\omega_0^2 - \omega^2} \tag{24}$$

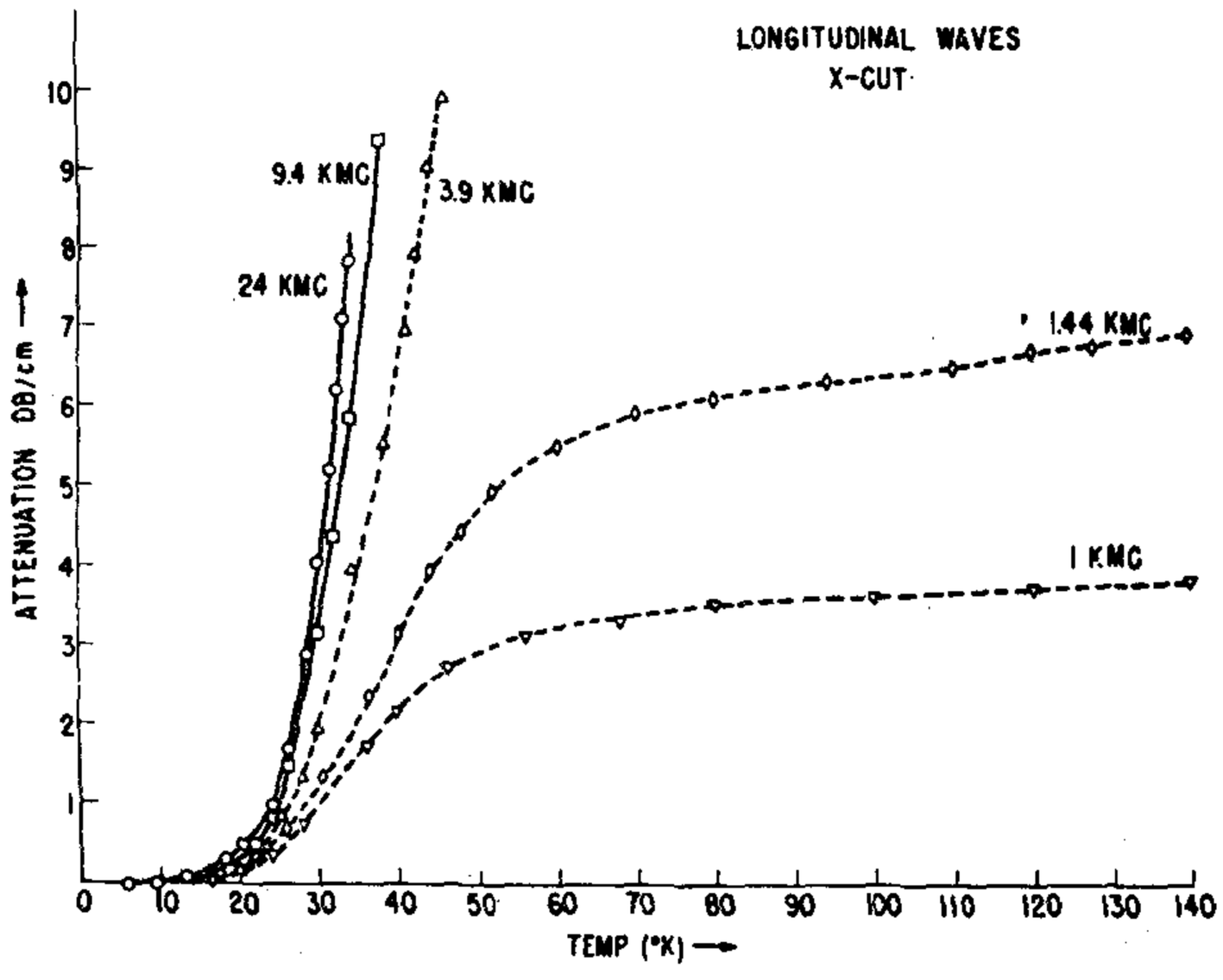


Fig. 5. Attenuation of longitudinal elastic waves in quartz as a function of temperature. Dotted lines are taken from the work of Bömmel and Dransfeld (ref. 13), solid lines the work of the author (ref. 14).

Both equations (23) and (24) take the same form when $\langle S_z \rangle$ is negative, i.e., when there exists a normal population. The small difference between the two expressions comes from the fact that the spins are coupled to the elastic *strain* rather than to the *amplitude* of atomic displacement. A plot of $(V_0/V)^2$ appears in Fig. 5a for negative values of $\langle S_z \rangle$. A positive value of $\langle S_z \rangle$ corresponds to an inverted population, and we shall return to this matter presently. An alternative description of a dispersive medium is given by the relation $\omega = f(k)$, directly obtainable from (22), and appears in Fig. 5b. These graphs illustrate the following points: We are in reality dealing with coupled systems; the disturbance which propagates is a mixture of sound and transverse spin waves. For small coupling ε, most of

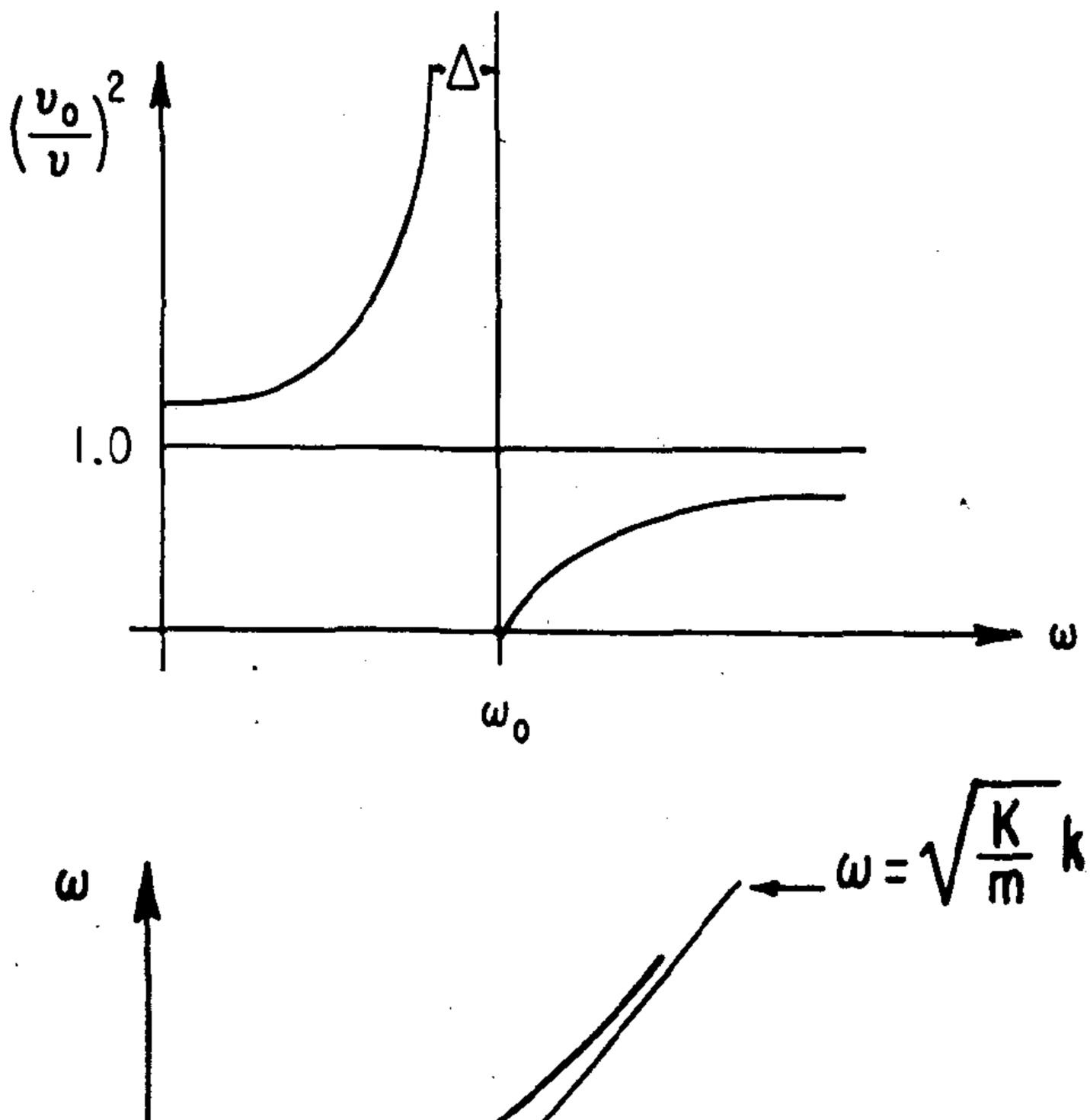


Fig. 6 a, b. Dispersion curves for spin = ½ system. Δ represents a stop band. Taken from the work of Jacobsen and Stevens, reference 19. Curves are for normal population, $\langle S_z \rangle$ negative.

the wave energy is contained in the elastic strain field and so propagates as a nearly pure sound wave. As ε increases, and particularly near resonance, more energy is propagated in the companion spin wave with the result that the disturbance is no longer purely sonic and is propagated at a reduced velocity. In our model, which assumes no losses, there exists near resonance a stop band within which k is imaginary so that normal wave propagation is not possible over certain frequencies.

In nature any resonant system will have a finite Q-factor, or line width, which implies a finite loss for wave propagation in the resonant medium. In our model of $S = \frac{1}{2}$ it is convenient to treat this line width phenomenologically in terms of a spin-spin or transverse relaxation time τ. We do so in the usual way by adding the quantities S_x/τ and S_y/τ to the left hand side of the Bloch equations for S_x and S_y in (19). The modified equations of motion, which now include damping, lead to the more general dispersion law (25).

$$\left(\frac{V_0}{V}\right)^2 = 1 - \frac{\dfrac{4\varepsilon^2 g\beta H \langle S_z\rangle}{K}}{\left(\omega_0^2 + \dfrac{1}{\tau^2} + \dfrac{4\varepsilon^2 g\beta H}{K}\langle S_z\rangle\right) - \omega^2 + \dfrac{2i\omega}{\tau}} \tag{25}$$

For convenience, let us assume $\left| 4\varepsilon^2 g\beta H \langle S_z\rangle \tau/2\omega K \right| << 1$ and take the square root of (25). Then

$$\left(\frac{V_0}{V}\right) = (\alpha + i\beta) \simeq 1 - \frac{2\varepsilon^2 g\beta H \langle S_z\rangle}{K}\left\{\frac{\Omega_0^2 - \omega^2 - \dfrac{2i\omega}{\tau}}{(\Omega_0^2 - \omega^2)^2 + \left(\dfrac{2\omega}{\tau}\right)^2}\right\}, \tag{26}$$

where $\Omega_0^2 \equiv \left(\omega_0^2 + 1/\tau^2 + \dfrac{4\varepsilon^2 g\beta H \langle S_z\rangle}{K}\right)$.

The real and imaginary parts are plotted against ω in Fig. 6a, b. As before when $\omega \to \Omega_0$ the sound wave experiences anomalous dispersion in the vicinity of resonance, but now with a concomitant rise in attenuation. Moreover, with finite τ a stop band no longer exists; wave propagation at frequencies within the resonance band width is possible although the physical meaning of group velocity is not clear if the dispersion is pronounced. However, an energy velocity can always be defined. The problem of wave propagation at frequencies within the region of anomalous dispersion, particularly as it applies to the propagation of pulses, is an interesting and delicate matter which we shall not take up here, but which is treated at some length in references (20, 21, 22). We merely point out, however, that a pulse of sound traversing a medium under the conditions of anomalous dispersion will be distorted owing to the variation of α, and β as portrayed by the graphs in Fig. 6.

The case of an inverted population, $\langle S_z\rangle$ positive, merits further comment. This situation corresponds to the dotted lines in Figs. 6. We note first that a literal interpretation of eq. (26) for positive $\langle S_z\rangle$ suggests that as $\omega \to \Omega_0$ the measured velocity of pulses will increase because of the resulting "inverted" dispersion. However, the analysis of Sommerfeld [20] and Brillouin [21] indicate that, to the contrary, it is not possible to transmit signals at velocities greater than V_0 ($\omega \to \infty$). Hence for an elastic continuum, where V_0 is independent of

frequency, we should not observe an *increase* in the velocity of pulses with an inverted population. The only experimentally distinguishable effect of dispersion between the cases of normal and inverted population appears to be the manner in which a pulse is *distorted* which in turn depends on the variation of phase velocity with frequency. This behavior seems to have been observed in early steady-state experiments with light traversing a resonant vapor, the level populations of which could be altered [23].

The second consequence of inverted population is negative absorption (amplification) by stimulated emission as represented by the dotted line in Fig. 6b. Amplification of sound pulses by an inverted population in ruby has been demonstrated in accordance with the behavior implied by eq. (25). [18] It is to be noted that the product $\varepsilon^2\langle S_z\rangle$ must not be too large when amplifying pulses if distortion caused by large dispersion is to be avoided, a restriction which can be removed for *cw* narrow band amplification. However, in the latter case, some means is needed to eliminate reflected waves within the resonant medium so as to avert the buildup of self-sustaining standing waves, i.e., feed back must be eliminated as with any amplifier in order to suppress self-sustained oscillations. In contrast to amplification, the production of self-sustained oscillation, if desired, should be easily realizable from an inverted spin system under proper conditions of reflection or feedback. Such an oscillator, at microwave sound frequencies, would involve essentially the same features encountered in the optical maser since the wavelengths are comparable. In particular we would expect to observe a series of modes excited within the natural spin-resonance line width, the spacing of which depends upon the ratio of sonic wavelength to crystal dimensions.

The previous discussion of dispersion applies equally to the case of transverse waves, the details depending of course on the form of the coupling constant ε. But in addition to the phenomenon of dispersion there exists for transverse waves the possibility of rotary polarization, and we shall complete this section by describing such behavior for transverse waves interacting with the $S = \frac{1}{2}$ system. We start with a simplified total Hamiltonian similar to (18) of the form

$$\mathfrak{H} = \sum_n \left\{ \frac{1}{2m}(P_x^2 + P_y^2) + \frac{C}{2}(U_x^n - U_x^{n+1})^2 + \frac{C}{2}(U_y^n - U_y^{n+1})^2 + \hbar\delta\,[S_y^n (U_y^{n+1} - U_y^{n-1}) + S_x^n (U_x^{n+1} - U_x^{n-1})] + g\beta H_z S_z^n \right\}, \tag{27}$$

and from it derive the equations of motion for the operators P, U, and S. With the definitions $S_n^{\pm} = S_x^n \pm iS_y^n$ and $U_n^{\pm} = U_x^n \pm iU_y^n$, and assuming plane wave solutions of the form $e^{i(\omega t - kz)}$ for transverse elastic waves and spin waves propagating along z we arrive finally at the set of simultaneous equations of motion (28).

$$\begin{aligned} &[m\omega^2 + 2C(\cos ka - 1)]\, U_n^{\pm} = \hbar\delta\,[2i \sin ka]\, S_n^{\pm} \\ &[2\delta\,\langle S_z\rangle\, i \sin ka]\, U_n^{\pm} = \left[-\frac{g\beta H}{\hbar} \pm \omega\right] S_n^{\pm} \end{aligned} \tag{28}$$

where the distance along z is specified by na, and $\omega_0 \equiv \frac{g\beta H}{\hbar}$.

Taking the long wave limit eq. (28) leads to the dispersion law (29).

$$[m\omega^2 - ca^2k_{\mp}^2]\,[\omega_0 \mp \omega] - [4\delta^2\hbar\langle S_z\rangle]\,a^2k_{\mp}^2 = 0 \tag{29}$$

For $\omega \neq \omega_0$ eq. (29) becomes

$$k_{\mp}^2 = k_0^2\left(\frac{1}{1 + \dfrac{4\delta^2\hbar\langle S_z\rangle/C}{\omega_0 \mp \omega}}\right), \tag{30}$$

where $k_0^2 \equiv \dfrac{m\omega^2}{a^2C}$. This equation shows that there are two senses for transverse waves propagating along the z-direction. Thus any given transverse wave propagating along z can be decomposed into a right and left handed circularily polarized wave, with wave vectors $k+$ and $k-$ respectively. Hence the net plane of transverse polarization is seen to rotate by an angle $\theta = \frac{1}{2}(k+ - k-)$ per unit distance along z, the amount of rotation depending on the magnetudes of the coupling constant δ and $\langle S_z\rangle$, and on the frequency ω. Rotary polarization has been theoretically predicted by Kittel [2] and experimentally observed by Matthews and LeCraw [6] in ferromagnetic systems to which Hamiltonians of the form (18) and (27) apply realistically. To date neither rotary polarization nor anomalous dispersion has been observed for paramagnetic atoms coupled to the phonon field through the *g-tensor*. These phenomena, though possible in principle, are expected to be weak for such coupling in contrast that which occurs through the crystalline field "D" and "E" terms. Before going on to discuss the latter type of coupling, it is perhaps well to summarize the assumptions underlying our spin $S = \frac{1}{2}$ model.

In our analysis we have assumed that at some definite time the spins are in known states of S_z (designated by $\langle S_z\rangle$, the average value of S_z per unit volume) and that a lattice wave is being propagated. We have then derived a set of linear coupled equations of motion on the assumption that the spins do not change their states significantly. Off resonance this approximation is reasonable. But as we approach resonance the interaction between spins and lattice motion increases, so that the disturbance is then a combination of spin and lattice waves. Moreover, if the coupling constant is large the interaction will be correspondingly large and the above approximation of constant S_z fails. When this happens the equations of motion become very complicated owing to the nonlinear terms introduced by the time variation of S_z. It is possible to show, though we shall not do so here, that for time varying S_z and near resonance the elastic wave is severely attenuated, much of its energy being converted to harmonics and to sum and difference frequencies. The interpretation of τ as purely a spin-spin relaxation time and the validity of the dispersion equations hold only so long as the conditions of small coupling and constant S_z are met. When these conditions are met, it is then natural to think of the attenuation of sound on resonance as a process by which a phonon $\hbar\omega_0$ is broken down into many smaller units and given to the dipolar interactions, eventually being fed back to the lattice over the entire spectrum since each mode is slightly coupled to the spins.

As noted earlier our model of $S = \frac{1}{2}$ coupled to the lattice through the g-tensor is idealized and, though illustrative in the essential ideas of dispersion, usually provides only a weak spin-lattice interaction. (But see reference (24).) Included in the Hamiltonian of a typical

spin system are the terms arising from the action of the crystal field. As it turns out, these crystal field terms are usually much more sensitive to lattice distortion than are the components of the g-tensor and so provide the main coupling between sound field and spins. We conclude this section with a discussion of the latter.

An example of the above is F_e^{++} in MgO, with effective spin $S = 1$. Shiren [22] has observed the elastic wave dispersion for this system by pulse-echo methods, and Jacobsen and Stevens [19] have given an elementary theory of the results, a summary of which follows.

The complete Hamiltonian for the propagation of a longitudinal wave along the (100) direction is

$$\mathfrak{H} = \sum_n \left\{ \frac{P_n^2}{2m} + \frac{K}{2}(U_n - U_{n-1})^2 + g\beta H_z S_z^n + \frac{D\hbar}{4}\left(S_z^{n2} + S_x^{n2} + S_x^n S_z^n + S_z^n S_x^n - \frac{4}{3}\right)(U_{n+1} - U_{n-1}) \right\}, \tag{31}$$

where P_n and U_n are parallel to (100), and subscripts x and z refer to a coordinate system rotated 45° to (100). By methods similar to those employed for the model of $S = \frac{1}{2}$ we can derive an expression analogous to eq. (29), i.e.,

$$\left(\frac{V_0}{V}\right)^2 = \frac{1}{1 + \left(\frac{D}{4}\right)^2 \frac{g\beta H\langle S_z\rangle}{K}\left\{\frac{1}{\omega_0^2 - \omega^2} + \frac{1}{4\omega_0^2 - \omega^2}\right\}}, \tag{32}$$

where ω_0 and $2\omega_0$ refer to the resonance frequencies corresponding to transitions $\Delta S_z = 1$ and 2 respectively. Eq. (32) is ploted in Fig. 7a, b. These curves reveal the familiar phenomenon of anomalous dispersion resulting from the interaction of sound waves with resonant spin systems, and similar in almost all detail to that encountered in the optical case. Because of the relative slowness of sound and the ease of changing level populations at microwave frequencies, it is possible to study the dispersion curves in some detail. The above relations were derived with the assumption of no loss or level broadening. We know from our earlier study of $S = \frac{1}{2}$ system that the general behavior is not greatly modified when a loss term or relaxation time is introduced, the main effect being to eliminate the stop bands and provide either absorption or amplification for waves propagating near the resonant frequencies.

Two interesting consequences of ultrasonic dispersion emerge from the above discussion. First, such experiments afford an alternate determination of the various spin-lattice coupling constants. Secondly, application of these ideas appear relevant to the generation and detection of monochromatic coherent elastic waves at thermal phonon frequencies. We have already called attention to the implications of eq. (26) when $\langle S_z\rangle$ is positive, i.e., generation of phonons by stimulated emission. To this process we can add parametric [25] and traveling wave amplification [26], which, though differing in details and point of view, are nonetheless fundamentally very close to the maser principle in that electrons are caused to emit sound quanta coherently. All of these methods will undoubtly play a role in extending our present investigations into the upper limits of the elastic wave spectrum of solids. We shall not, however, delve further into these matters here, except in connection with a novel quantum detection scheme due to Shiren [27]. Before doing so, one final comment

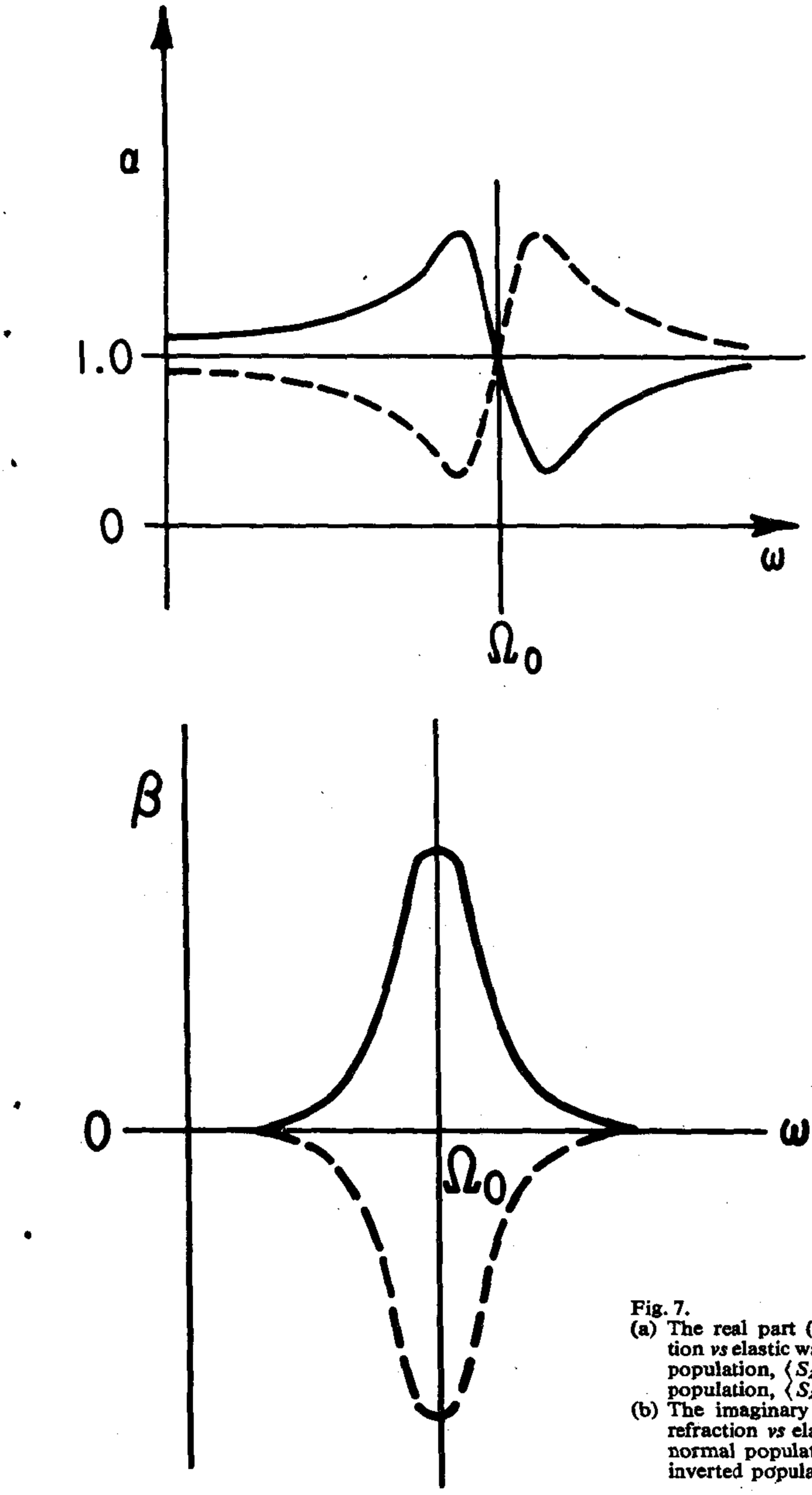


Fig. 7.
(a) The real part (dispersion) of the sonic index of refraction *vs* elastic wave frequency ω. Solid line denotes normal population, $\langle S_z \rangle$ negative. Dotted line denotes inverted population, $\langle S_z \rangle$ positive.
(b) The imaginary part (absorption) of the sonic index of refraction *vs* elastic wave frequency ω. Solid line denotes normal population, $\langle S_z \rangle$ negative. Dotted line denotes inverted population, $\langle S_z \rangle$ positive. See text.

seems in order relating to the derivation of the equations of motion and resulting dispersion relations.

We have carried out our analysis of spin-phonon dispersion in terms of the point mass model of a crystal lattice. This is basically correct and necessary when dealing with elastic waves of very short wavelength. However, for microwave frequencies, where the corresponding elastic wave is still very long relative to atomic spacings, we could have used a continuum representation from the outset. To do so we have merely to write the Hamiltonian in the continuum form so that eq. (18) becomes

$$H = \int_v \mathfrak{H}\,(p, u, s)\,dv = \int_v \left\{\frac{p^2}{2\varrho} + \frac{K^1}{2}\left(\frac{\partial u}{\partial x}\right)^2 + \varepsilon^1\left(\frac{\partial u}{\partial x}\right) S_x + g\beta\, H\, S_z\right\} dv\,, \quad (33)$$

where $\mathfrak{H}$ is the Hamiltonian density, and p, u, s are continuous functions of x. The appropriate commutation relations are then of the form

$$\begin{aligned} &[Ui\,(x, t), Pj\,(x^1, t)] = i\hbar\,\delta\,(x - x^1)\,\delta ij\,, \\ &[Si\,(x, t), Sj\,(x^1, t)] = i\,\hbar\,\delta\,(x - x^1)\,S_k\,(x, t)\,. \end{aligned} \quad (34)$$

Or, alternatively, we could have treated the various operators appearing in the Hamiltonian as classical variables and derived the equations of motion from Hamilton's canonical equations for a continuous medium following Kittel's treatment of phonon-magnon interaction in a ferromagnetic system [2].

Detection of Very High Frequency Elastic Waves

In principle electrical-mechanical transducers will function as detectors of very high frequency sound. However, for frequencies above the microwave range (and sonic wavelengths less than $\sim$500 angstroms) the use of transducers appears impractical. The trouble comes about mainly through the smallness in wavelength of both the electromagnetic and sonic radiation. The small electromagnetic wavelength implies a small resonant cavity in order to obtain sufficient driving field intensity. The small sonic wavelength implies near perfection in quality and alignment of sample and transducer in order that all regions of the surface vibrate in phase and so provide maximum sensitivity. These conditions become increasingly difficult to satisfy above, say, 100 Kmc.

A way around this difficulty is to detect sound quanta incoherently, analogous to photo electric detection of light quanta. Unfortunately no such convenient "photo-electric" device yet exists for sound waves, and very little work has been done along this line to date. However, a preliminary step has been taken in this direction by N. S. Shiren [27] who demonstrated a novel two-quantum detector for microwave sound. The scheme makes use of the paramagnetic spin levels of $F_e{}^{++}$ in $Mg0$ outlined in Fig. 8. The method consists of observing the absorption of one radio frequency quantum which accompanies the simultaneous absorption of one sound quantum. Because the r.f. and sound quanta are of unequal energy, but their sum is arranged so that $\hbar\omega, -\hbar\omega_2 = 2\hbar\omega_0$, no transition occurs except when sound field and microwave field appear simultaneously, whereupon the transition

$\Delta S_z = 2$ proceeds as shown. The probability of the two quantum process, derived from second order perturbation theory, contains terms of the form

$$\left| \frac{A_1 A_2}{(\omega_1 - \omega_0)(\omega_1 + \omega_2 - 2\omega_0)} \right|^2 \tag{35}$$

where A_1 and A_2 denote the respective amplitudes of the elastic and radio frequency fields. The double quantum character is evident from the resonance denominator. The transition

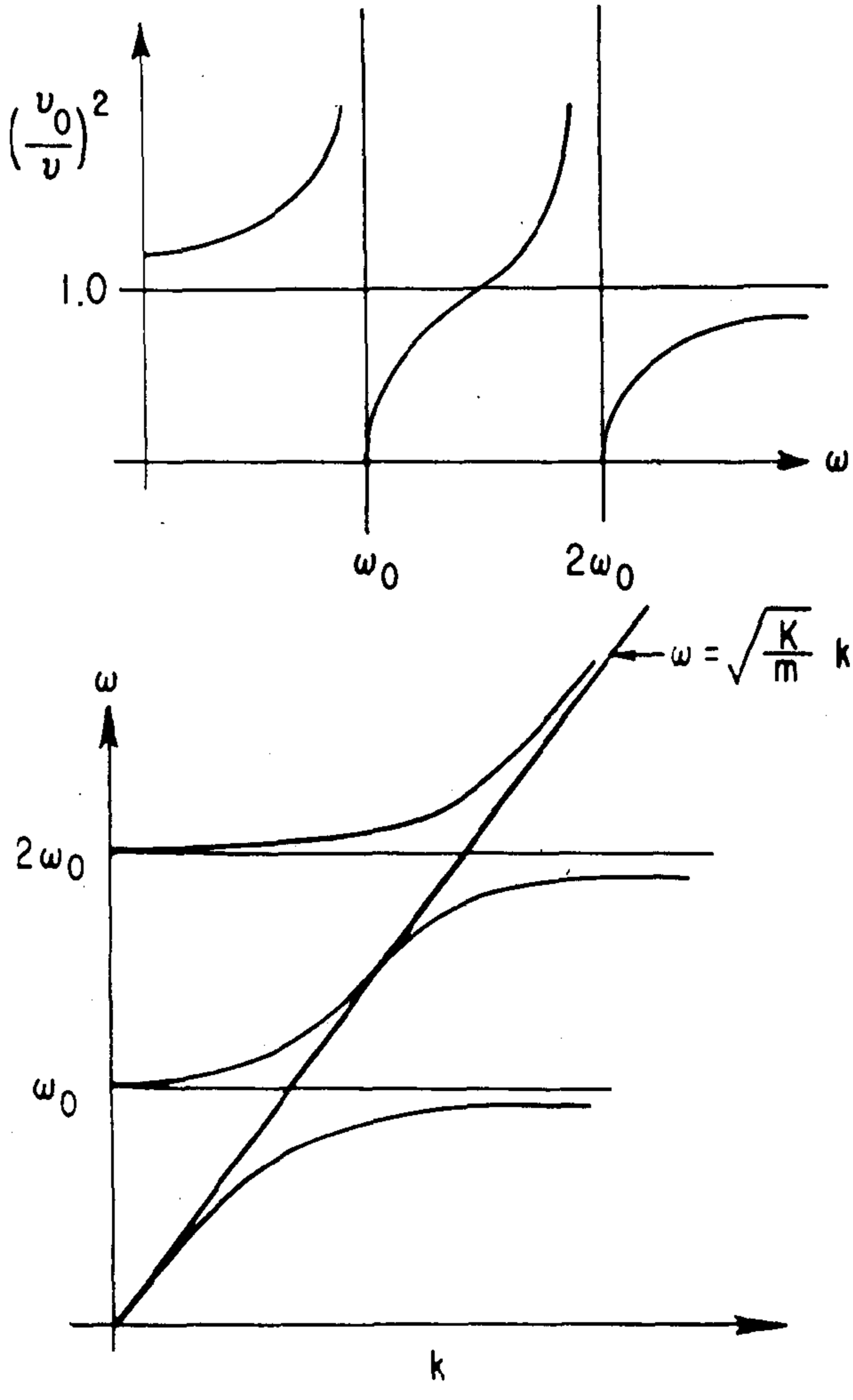


Fig. 8.
(a) Sonic index of refraction as a function of elastic wave frequency for spin = 1 system. Two resonances appear, one for $\Delta S_z = 1$ and the other for $\Delta S_z = 2$. See text.
(b) Dispersion relation between elastic wave frequency ω and wave vector k. See text.

is inherently fast and can follow the steep leading and trailing edges of well defined sound pulses. The use of paramagnetic resonance as a detector for microwave phonons suffers mainly from a slow recovery time: when the sound pulse ends, the resonance signal will return to its previous value in a time T_1 and so lengthen the trailing edge considerably, whereas the characteristic time for a double quantum transition can be shown to be of order $T_2 \sim 10^{-8}$ sec. Furthermore, the double quantum scheme could be used to detect thermal frequency phonons using microwave frequency r.f. fields with paramagnetic ions possessing appropriate level spacings. Also, as suggested by Shiren, conversion gain can be achieved by the inverse process—high frequency r.f. fields can be used to detect relatively low frequency sound waves.

However, an obvious drawback of the double quantum method is the reduced sensitivity arising because the process is second order. Even so the realizable sensitivity may exceed that obtained from a poorly oriented transducer or an imperfect sample which distorts the wave front—the latter occurring frequently in practice. Moreover, it is possible to improve the double quantum sensitivity by increasing the amplitude of the r.f. field since the product $A_1 A_2$ appears in eq. (35). Refined variants of quantum detection of sound are needed and will doubtlessly come into being in the future. Such detection methods have received little attention to date, but will be increasingly important for experiments in the ultramicrowave region and beyond.

We conclude this section with a brief account of the role of x-ray and neutron diffraction in the detection of thermal elastic waves in solids. [28] So far this method has provided most of what we know about the elastic spectrum for frequencies above roughly 10^{11} cps.

Fundamental to the interaction of electromagnetic and matter waves with elastic waves is Bragg scattering. Consider first the scattering of electromagnetic waves by sound. The sound wave, represented by, $\exp i\,(\omega_1 t - k_1 \cdot r)$, acts as a moving grating which diffracts the incoming electromagnetic wave, represented by $\exp i\,(\omega_0 t - k_0 \cdot r)$. If the interaction between the two waves is small, we can show that the scattered wave is represented by $\exp i\,[(\omega_0 \pm \omega_1)t - (k_0 \pm k_1)\cdot r]$, the (+) sign signifying that one photon of energy $\hbar\omega_0$ and momentum $\hbar k_0$ is destroyed along with one phonon of energy $\hbar\omega_1$ and momentum $\hbar k_1$ to create a new photon of energy $\hbar\,(\omega_0 + \omega_1)$ and momentum $\hbar\,(k_0 + k_1)$, and the (—) sign signifying the inverse process. These processes express the conservation of energy and momentum which leads to the Bragg law

$$\lambda = 2\Lambda \sin\theta \tag{36}$$

where Λ is the sound wavelength and λ that of the incident electromagnetic wave, illustrated in Fig. 9. The above refers to a first order scattering process. If the interaction between the waves or the amplitude of the sound wave is large we should have to replace eq. (36) by (36a)

$$n\lambda = 2\Lambda \sin\theta, \tag{36a}$$

realizing that a perfectly sinusoidal disturbance can produce higher order reflections involving more than one photon and phonon per scattering event. Our discussion so far refers to a continuum and omits the possibility of Bragg reflections coming from discrete planes of atoms. When the atomic crystalline structure is taken into account the construction of Fig.

9 is modified in that k_1 must be replaced by $k_1 + K_n$, where K_n is a reciprocal lattice vector. Eq. (36) must then be written

$$n\lambda = \frac{2\pi}{|k_1 + K_n|} \sin\theta\,, \tag{36b}$$

where n now refers only to the order of the *crystalline* reflection; the reflection from the elastic wave is still first order since we are assuming its amplitude is small. Fig. 10 illustrates the meaning of eq. (36b) for a compressional wave propagating normal to planes of atoms;

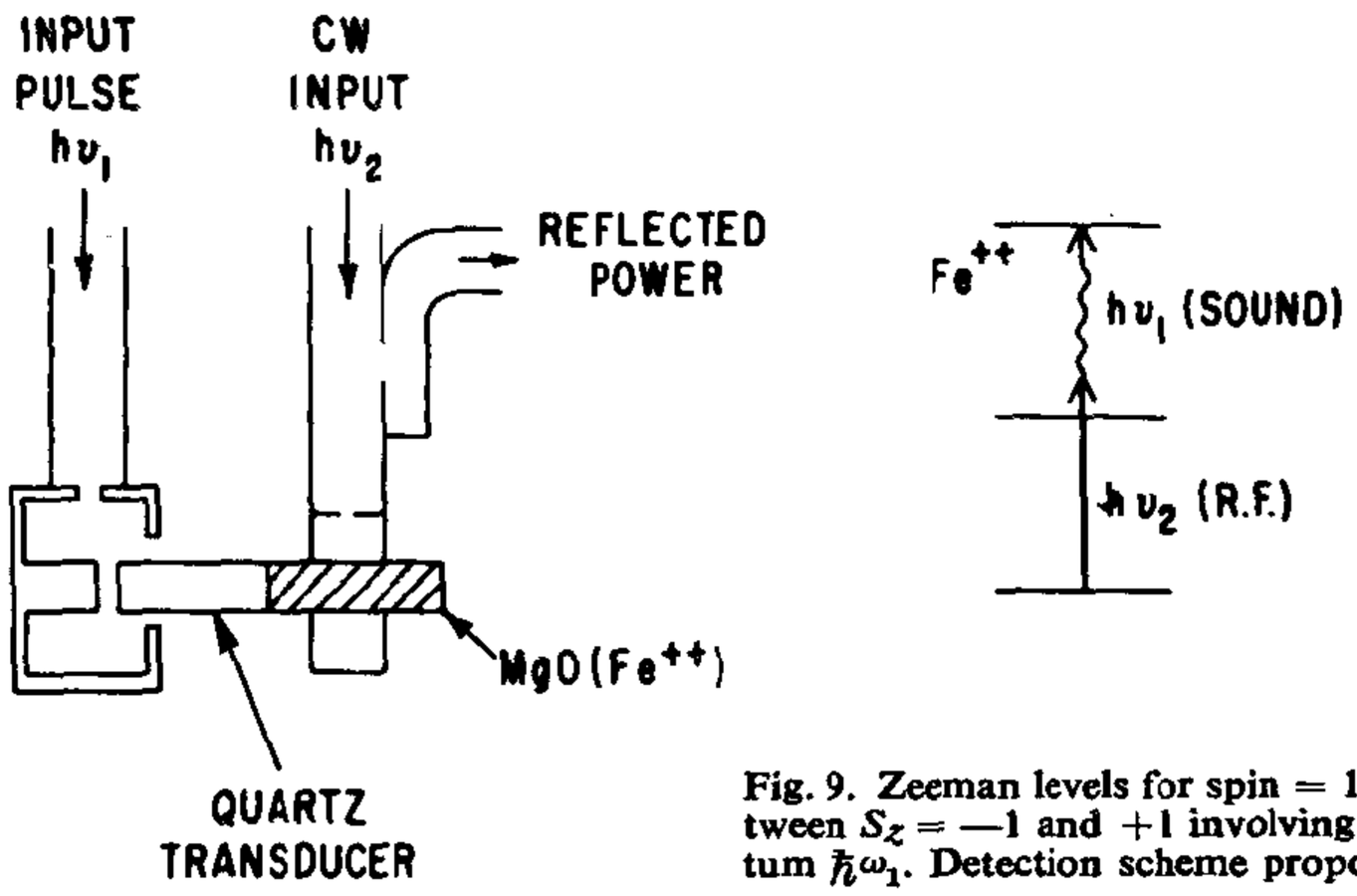


Fig. 9. Zeeman levels for spin = 1 system, and double quantum transition between $S_z = -1$ and $+1$ involving one sound quantum $\hbar\omega_2$ and one r.f. quantum $\hbar\omega_1$. Detection scheme proposed by Shiren, see reference (27).

the intensity pattern exhibits side bands about each crystalline Bragg reflection as shown. Eqs. (36) indicate that we shall best observe Bragg reflections from elastic waves when their wavelength and that of the incident wave are of comparable magnetude. Therefore with typical x-ray wavelenghts $\sim 1A$ we can see the presence of elastic waves of between approximately 1 and 20 A wavelength. But this is an important region since it covers the highest density of possible elastic wave modes and extends over a frequency range ofroughly 500 Kmc to cut-off $\sim 10^{13}$ cps. Thus, if we can generate monochomatic phonons in this frequency range we shall be able to detect them unambiguously.

All of the previous discussion pertains equally to neutron diffraction. However, in this case an important practical difference arises. Because of the large mass, a neutron suitable for diffraction experiments will have a kinetic energy comparable to that of the lattice thermal phonons. Consequently, Bragg reflection from a thermal wave produces a significant change in the energy of the scattered neutron beam which can be measured accurately by time-of-flight methods. Hence neutron diffraction experiments measure the frequency of thermal elastic waves directly. In fact one can also determine their life time as it is manifested in the breadth of the energy spectrum at each scattering angle θ. With x-rays, on the other hand, this energy shift is normally insignificant, and one can measure only the *intensity* of the scattered wave and from this deduce the amplitude and frequency of the elastic wave. [28] The wavelength in any case is always determined by the angle θ through the Bragg law. (See Fig. 10). The extreme monocnromicity of Mössbauer x-ray sources

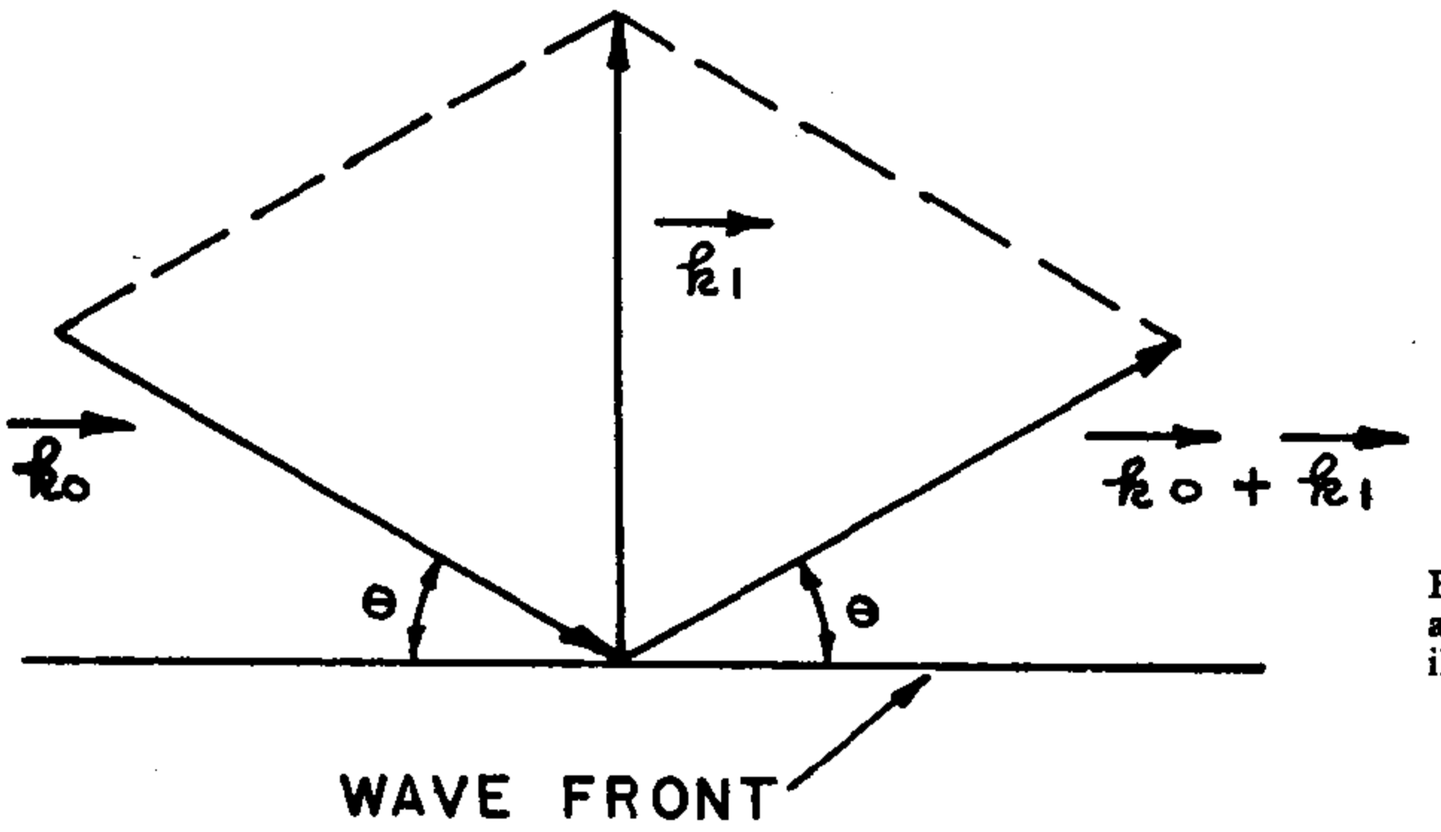


Fig. 10. Diffraction of electromagnetic and matter waves by sound wave grating, illustrating Bragg's Law.

affords, of course, an energy resolution equal to or exceeding that of neutrons. However, the practicability of using Mössbauer emission for thermal elastic wave studies is presently handicapped by low intensity and limitations in measuring frequency shifts by Doppler methods.

During the last decade numerous studies of thermal waves in crystals have been carried out by *x*-ray and neutron diffraction [29]. The main result of such studies has been determination of the linear restoring constants between atoms. Since the atomic spacing is of the same order as the phonon wavelength, such experiments measure directly the dispersion produced by the periodic arrangement of atoms, which in turn yields the various inter atomic force constants. The atomic force constants are tensor quantities which in various linear combinations define the macroscopic elastic constants. It seems likely that future experiments combining diffraction with controlled generation of monochromatic thermal phonons may lead to a more extensive picture of atomic binding and relevent electronic structure in solids, as such experiments would afford an opportunity to observe higher order phonon process.

Conclusion

Our excursions thus far into the high frequency phonon region have been very modest. The highest lattice frequencies in solids lie between 10^{13} and 10^{14} cps, whereas the highest elastic

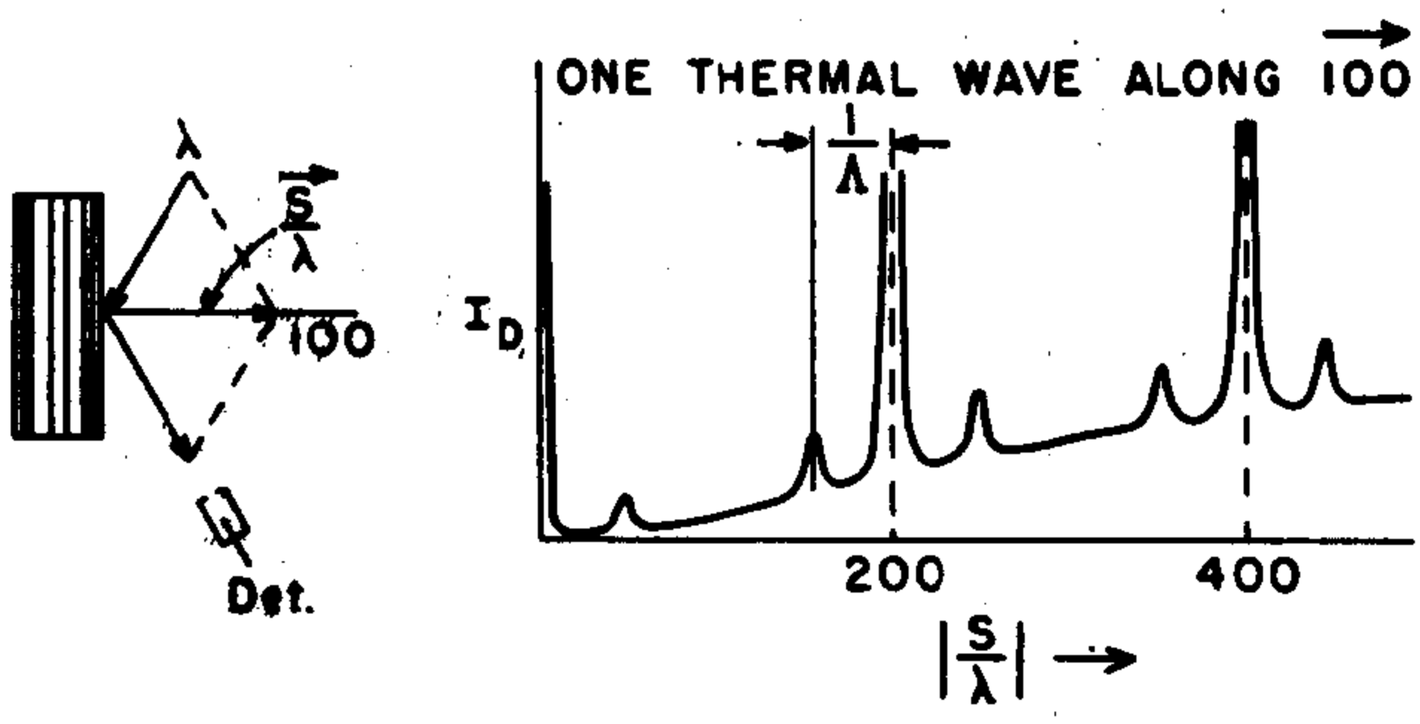


Fig. 11. Diffraction of x-rays from a crystal lattice on which is superimposed a compressional sound wave. λ is the incident x-ray wavelenght, Λ that of the sound wave. Similar ideas apply to thermal neutron diffraction.

wave frequency generated by laboratory methods falls some three orders of magnetude lower. Nonetheless, even at these lower frequencies there are doubtless hitherto unexplored and perhaps unknown phenomena which will bear fruitfull investigation. Moreover, those areas which have received current attention such as, for example, phonon-phonon scattering, spin-phonon interaction, and traveling wave amplification are still in an elementary stage and will likely reward improved analytical and experimental techniques. Furthermore, electron-phonon interaction as related to electrical and thermal conductivity, and magneto-plasma phenomena have received little attention to date. There exists also the possibility of parametric interaction between spins and phonons for amplification, detection, and harmonic generation of elastic waves. These techniques in combination with the phonon maser principle, besides possessing interest in their own right, may provide coherent sources of thermal phonon radiation and so make possible a detailed study of inter atomic forces, energy transport, and a variety of collective excitations in complex systems.

REFERENCES

1. K. N. Baranskii, Doklady Akad. Nauk S.S.S.R. *114*, 517 (1957); Soviet Research in Physics (English Translation) *2*, 237 (1957); H. E. Bömmel and K. Dransfeld, Phys. Rev. Letters *1*, 234 (1958); E. H. Jacobsen, Phys. Rev. Letters *2*, 249 (1959).
2. C. Kittel, Phys. Rev. *110*, 836 (1958); See also K. B. Vlasov and B. Kh. Ishmukhametov, J. Explt. Theoret. Phys. (U.S.S.R.) *37*, 725 (1959); (English translation: Soviet Phys. – JETP *37*, 531 (1960)).
3. H. E. Bömmel and K. Dransfeld, Phys. Rev. Letters *3*, 83 (1959).
4. M. Pomerantz, Phys. Rev. Letters *7*, 312 (1961).
5. M. F. Lewis, T. G. Phillips and H. M. Rosenberg, Physics Letters *1*, 198 (1962).
6. H. Matthews and R. C. LeCraw, Phys. Rev. Letters *8*, 397 (1962).
7. J. J. Kyame, J. Acoust. Soc. Am. *21*, 159 (1949); *26*, 990 (1954).
8. A. R. Hutson and D. L. White, J. Applied Physics *33*, 40 (1962).
9. A. E. H. Love, *The Matematical Theory of Elasticity*, 1944, 4th Edition, Cambridge University Press.
10. P. C. Waterman, Physical Review *113*, 1240 (1959).
11. Max Born and Emil Wolf, *Principles of Optics*, Pergamon Press (1959), p. 685; W. A. Wooster, *Crystal Physics*, Cambridge University Press (1938), p. 148.
12. Martin R. Redwood, *Mechanical Waveguides*, Pergamon Press (1960).
13. H. E. Bömmel and K. Dransfeld, Phys. Rev. Letters *2*, 298 (1959).
14. E. H. Jacobsen, *Quantum Electronics* (edited by C. H. Townes), Columbia University Press (1960) p. 479.
15. T. O. Woodruff and H. Ehrenreich, Phys. Rev. *123*, 1553 (1961).
16. E. H. Jacobsen, N. S. Shiren, and E. B. Tucker, Phys. Rev. Letters *3*, 81 (1959).
17. N. S. Shiren and E. B. Tucker, Phys. Rev. Letters *6*, 168 (1961).
18. E. B. Tucker, Phys. Rev. Letters *6*, 183, 547 (1961).
19. E. H. Jacobsen and K. W. H. Stevens, Physical Review (to be published).
20. A. Sommerfeld, Ann. Physik (4), *44*, 177 (1914).
21. L. Brillouin, *Wave Propagation and Group Velocity*, Academic Press (1960).
22. N. S. Shiren, Physical Review (to be published).
23. Ladenburg, Zeits. fur Phys. *48*, 15 (1928); Kopferman and Ladenburg, *ibid. 48*, 26, 51 (1928); Carst and Ladenburg, *ibid. 48*, 192 (1928).
24. Laura M. Roth, Physical Review *118*, 1534 (1960).

25. J. A. Armstrong, N. Bloembergen, J. Duwing, and P. S. Pershan, Physical Review (to be published); Norman M. Kroll, *ibid* (to be published). These articles discuss parametric amplification of light waves by induced nonlinear electric dipole and higher moments in an atomic system; the same general ideas apply to the nonlinear generation of sound waves under certain conditions when S_z varies with time.
26. A. R. Hutson, J. H. McFee, and D. L. White, Phys. Rev. Letters *7*, 237 (1961).
27. N. S. Shiren, Phys. Rev. Letters *6*, 168 (1961).
28. Max Born, Reports on Progress in Physics (Physical Society), *9*, 294 (1942–3); J. Laval, C. R. Acad. Sci. Paris, *207*, 169 (1938); R. W. James, *The Optical Principles of The Diffraction of X-Rays*, G. Bell and Sons Ltd. (1950), Chpt. 5.; John C. Slater, Reviews of Mod. Phys. *30*, 197 (1958).
29. An extensive bibliography covering work on elastic spectra determination in crystals by x-ray and neutron diffraction is contained in *Electrons and Phonons* by J. M. Ziman, (Oxford University Press, 1960). This text, though not dealing with spin-phonon interaction, gives the most complete account to date of electron-phonon interactions pertaining to electrical conductiviy.

Neuere Analogien zwischen akustischen und elektromagnetischen Schwingungen und Wellenfeldern

von Erwin Meyer

Göttingen

1. Einleitung

Dass man das physikalische Verhalten von akustischen und elektromagnetischen Schwingungen und Wellen in vieler Hinsicht miteinander vergleichen kann, ist altbekannt und gehört vielerorts zum Vorlesungs- und Übungsstoff für den Studenten. Die Schwingungsgleichungen ebenso wie die Wellengleichungen und die Leitungstheorie für die Wellenausbreitung sind auf beiden Gebieten ausserordentlich ähnlich oder sogar gleich. Freilich, es gibt auch Unterschiede, vor allem in den Wellenfeldern. Das Schallfeld hat eine skalare Feldgrösse, den Schalldruck, und eine vektorielle Feldgrösse, die Schallschnelle; das elektromagnetische Feld hingegen ist in seinen beiden Feldgrössen, der elektrischen und der magnetischen Feldstärke, vektoriell. Polarisationsvorgänge für Luft- oder Wasserschall sind nicht bekannt. Es gibt aber viele Vorgänge auf beiden Gebieten, bei denen die eben genannten unterscheidenden Merkmale eine untergeordnete Rolle spielen.

Was zunächst das Gebiet der Schwingungen betrifft, so entsprechen die verschiedenen Ausführungsformen der einfachen und gekoppelten Schwingungskreise sowie der Filtersysteme einander sehr genau. Nichtlineare und rheolineare Prozesse sind für Akustik wie Elektrik in gleicher Weise zu behandeln. Ein besonders reizvolles Beispiel sind die parametrische Erregung und Verstärkung von Oberflächenschwingungen einer Wasserschicht; man kann diese Erscheinung als instruktives Demonstrationsexperiment in einer Vorlesung gut vorführen und in Parallele zur parametrischen Erregung oder Verstärkung eines elektrischen Schwingungskreises, z.B. mit einer Kapazitätsdiode setzen.

Auch auf dem Gebiete der Wellenausbreitung gibt es zahlreiche Vergleiche. Besonders schön sind hier die alten Versuche von Winston Kock mit der Bündelung bzw. Fokussierung von elektrischen Zentimeterwellen durch Linsen aus einem »synthetischen« Dielektrikum, bzw. für akustische Wellen, durch Linsen aus einem »synthetisch« dichteren Medium. Im III. Physikalischen Institut in Göttingen, das sich in den letzten Jahren viel mit Analogien befasst hat, sind eine Reihe von neuen, gerade durch Analogiebetrachtungen angeregten Untersuchungen auf akustischem bzw. elektrischem Gebiete gemacht worden. Von einigen soll hier die Rede sein. Sie betreffen folgende drei Gruppen:

a) Die Schallausbreitung in schallgedämpften Rohren mit periodischen Wandstrukturen bei überlagerter starker Gleichluftströmung sowie ihre Analogie zu Vorgängen in Wanderfeldröhren für Mikrowellen. (1)

b) Die Entwicklung von elektromagnetischen reflexionsfreien Absorbern in Analogie zu verschiedenen akustischen Absorbertypen, insbesondere der Fall der Wellenausbreitung parallel zu einer absorbierenden Fläche. (2)

c) Der Bau eines gemeinsamen Hallraumes für elektromagnetische und akustische Wellen. (3)

2. Schallausbreitung in Rohren mit periodischer Wandstruktur bei überlagerter starker Gleichluftströmung und die Analogie zur Wanderfeldröhre

Auf akustischem Gebiet ist es ein interessantes Problem, die Schallausbreitung, insbesondere die Schalldämpfung in Kanälen zu studieren, deren Seitenwände für ein breites oder ein schmales Frequenzband schallabsorbierend ausgekleidet sind, und die gleichzeitig eine starke Luftströmung mit oder gegen den Schall führen. Im Falle des Schmalbandabsorbers, nämlich bei der Verwendung von vielen gleichabgestimmten Helmholtz-Resonatoren in der Kanalwand ergab sich dabei ein bemerkenswertes Ergebnis. Ein Teil der Apparatur zeigt Abb. 1 a. Der Schall wird über einen 100 Watt Lautsprecher in den Kanal eingespeist, dessen Abmessungen 10 cm × 3,3 cm sind. (Abb. 1 b.) Gleichzeitig wird hier der Gleichluftstrom, aus einem 60 kW Radialgebläse kommend, zugeführt, durch grosse Schalldämpfer fast geräuschlos und durch mehrere Siebe fast turbulenzfrei gemacht. Längs des Kanals wird ein Mikrofon bewegt, ein elektrostatisches Sell-Mikrofon, eingebaut in einen Laminar-Profilkörper (Abb. 1 c). Durch sorgfältigen Aufbau war es möglich, ein Signal-Rauschverhältnis von 30–35 dB bei 200 m/s Windgeschwindigkeit und noch höhere Werte bei kleineren Windgeschwindigkeiten zu erzielen.

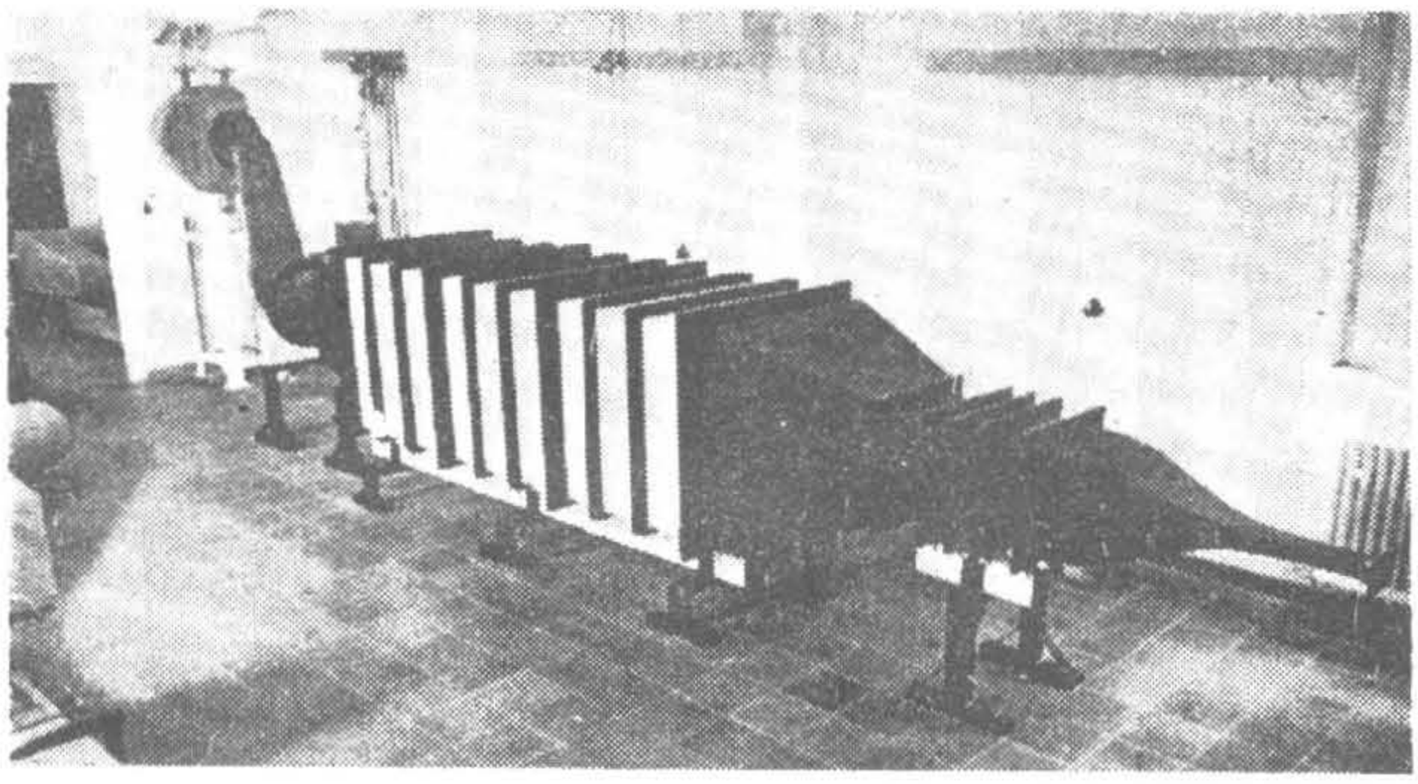


Abb. 1a
Windkanal
(Schalldämpfer und Turbulenzsiebe)

1. Bearbeiter: Fr. Mechel, P. Mertens, W. Schilz
2. Bearbeiter: D. Eisfeld, R. Pottel, H. J. Schmitt, H. Severin, F. Wiekhorst
3. Bearbeiter: H. W. Helberg, S. Vogel, N. Roy

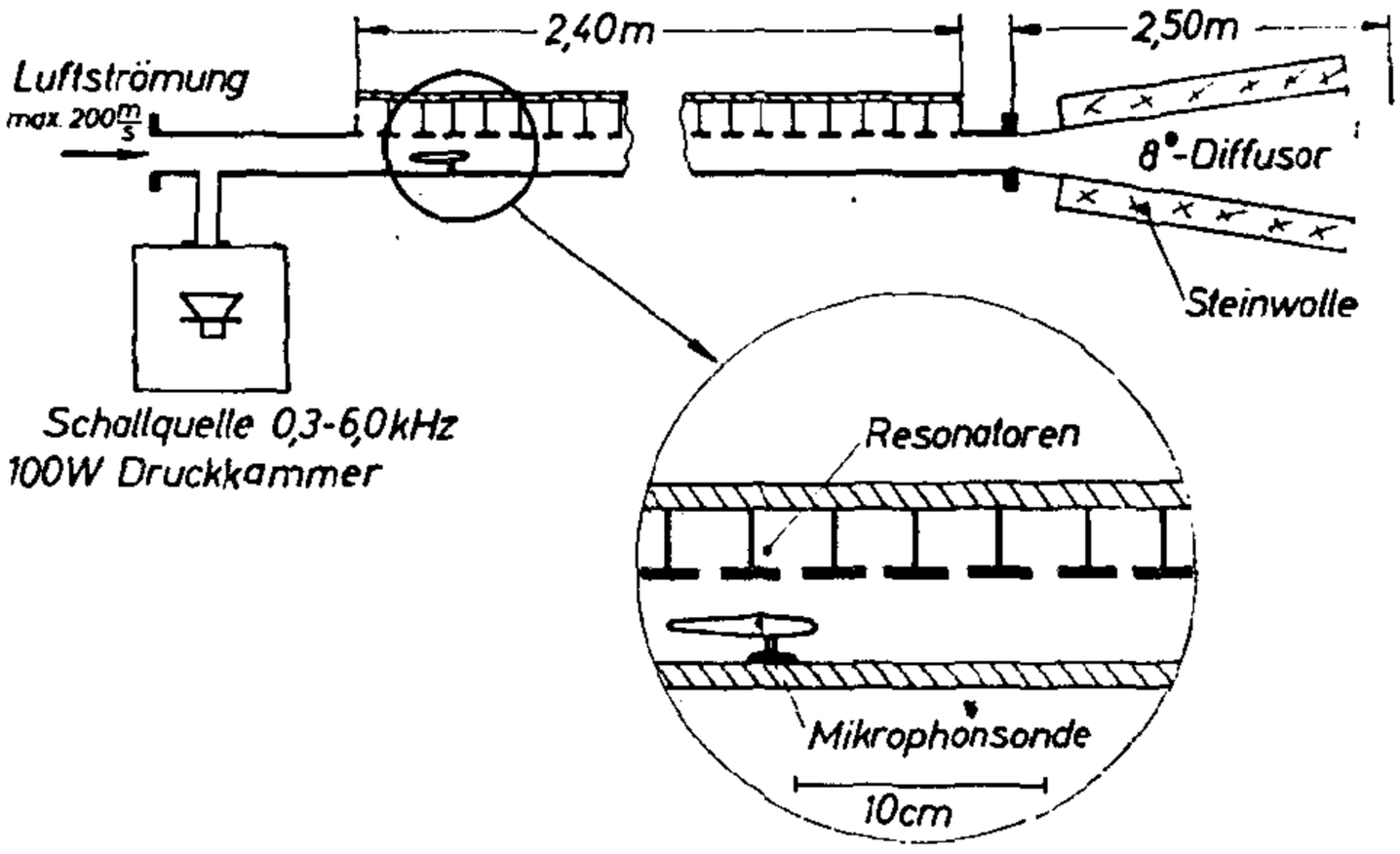


Abb. 1b
Testkanal mit Helmholtzresonatoren

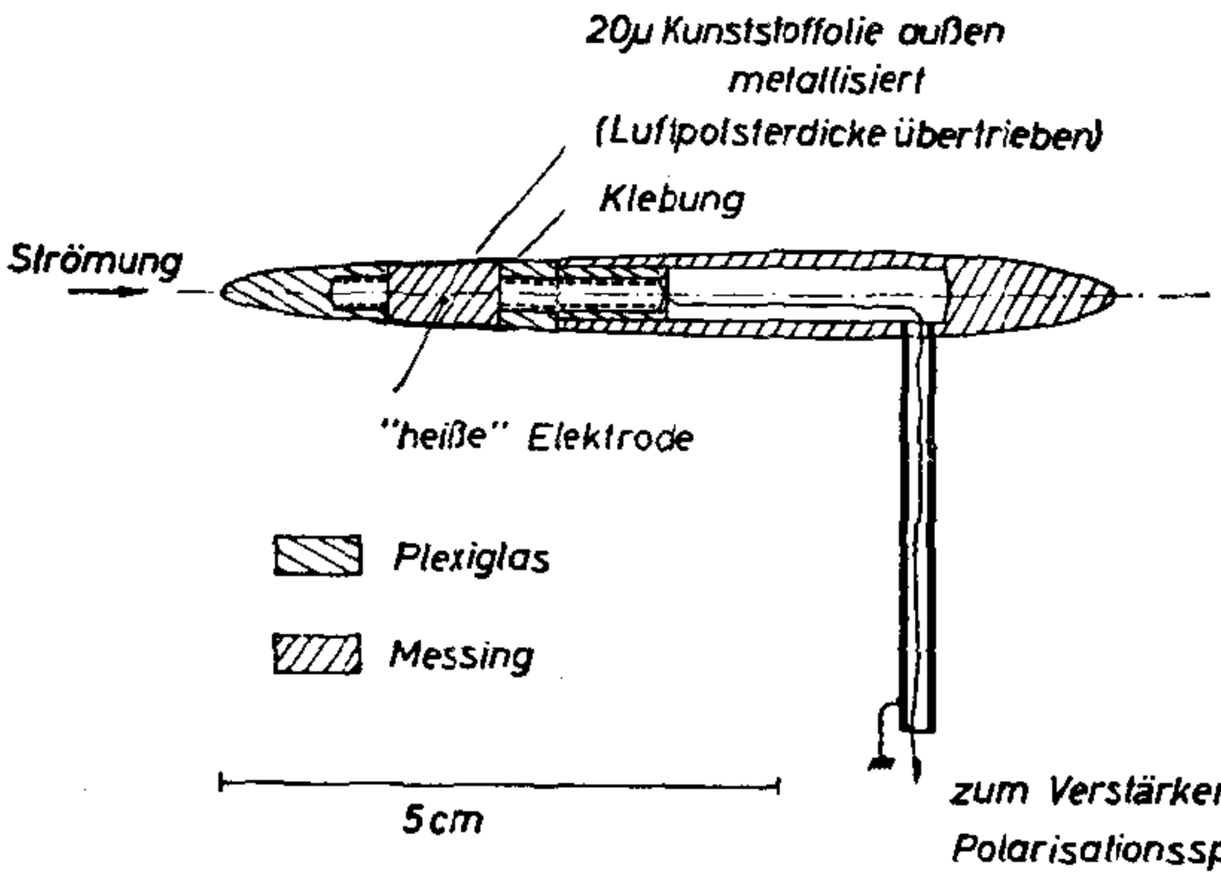


Abb. 1c
Längsschnitt des Sondenmikrophons

Abb. 2 zeigt die für verschiedene Strömungsgeschwindigkeiten v gefundenen Schalldämpfungen in dB/m als Funktion der Frequenz; Schall und Strömung sind gleichgerichtet. Die Resonatoren sind auf rund 1 000 Hz abgestimmt. Auf der Ordinate ist nach oben eine positive Dämpfung, nach unten eine negative Dämpfung, d.h. also eine Verstärkung aufgetragen. Abb. 2 zeigt deutlich, dass es für jede Strömungsgeschwindigkeit einen kleinen Frequenzbereich gibt, in welchem die ohne Gleichströmung recht hohe Schalldämpfung in eine negative Dämpfung, in eine »Verstärkung«, umschlägt.
Noch deutlicher tritt dieser Effekt in Abb. 3 auf, für den der Messquerschnitt des Rohres auf $10 \times 1{,}6$ cm² verringert wurde. Auch hier finden sich Verstärkungen von mehreren dB/m! Wie schon in den ersten Versuchen dieser Art erkannt wurde, hängt die Verstärkung stark von der Ausbildung der Turbulenz an den Resonatoröffnungen ab (runder Querschnitt, rechteckiger Querschnitt, Öffnung mit Gaze überspannt oder nicht etc.)
Der geschilderte Tatbestand zwingt dazu, anzunehmen, dass die für die Verstärkung benötigte Energie aus der Gleichströmung kommt. Ausserdem liegt es nahe, sich an einen

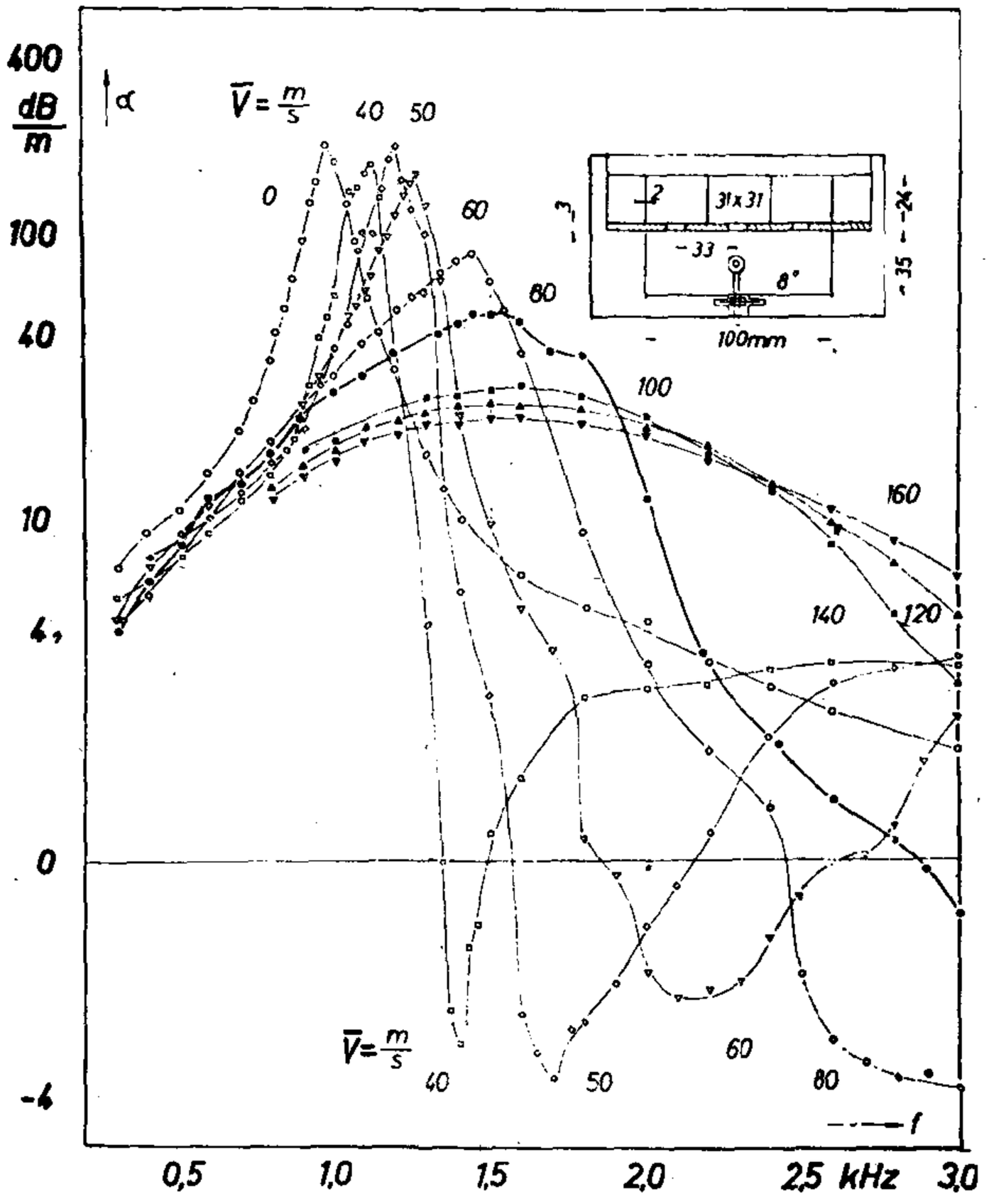


Abb. 2
Dämpfung α in Abhängkeit von der Frequenz f bei einem unbedämpften Lochresonator. Parameter: Mittlere Strömungsgeschwindigkeit $\bar{V}$.

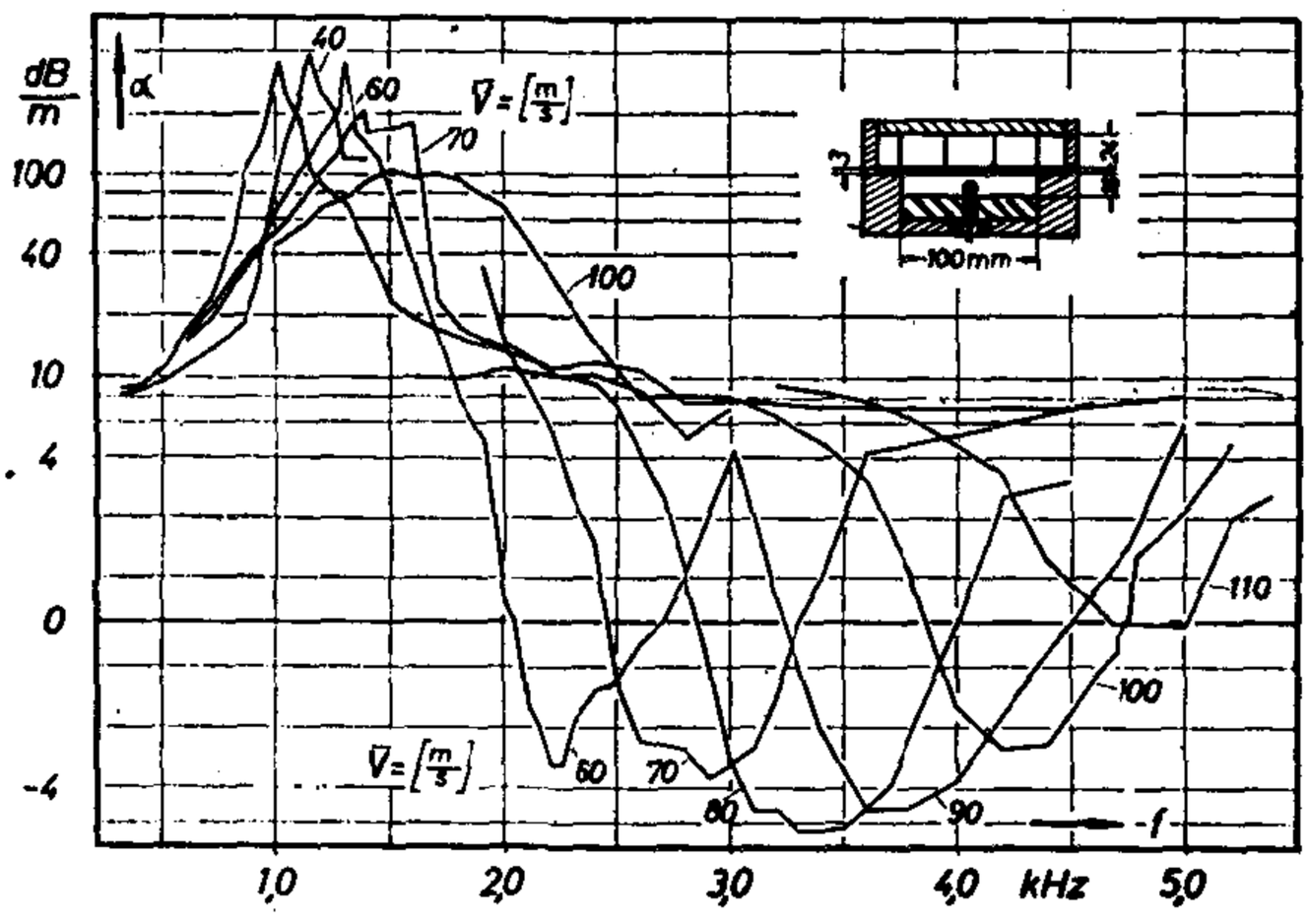


Abb. 3
Dämpfung αi Abhängigkeit von der Frequenz f bei Lochresonatoren mit verringter Kanalhöhe. Parameter: Mittlere Strömungsgeschwindigkeit $\bar{V}$.

analogen Mechanismus auf dem elektromagnetischen Gebiet zu erinnern, nämlich an die sogenannte »Wanderfeldröhre«. Sie hat folgenden Aufbau: Eine elektromagnetische Welle breitet sich auf einem Leiter, z.B. in einem gefältelten Hohlleiter, aus; bei geeigneter Dimensionierung wird ihre Ausbreitungsgeschwindigkeit von Lichtgeschwindigkeit auf etwa die Geschwindigkeit des Elektronenstrahls, der im Zentrum des Hohlleiters fliesst, herabgesetzt. Damit ist die Möglichkeit einer Wechselwirkung zwischen Welle und Elektronenstrahl gegeben. Ob dabei die Wechselwirkung mit der aufgeprägten Signal-Grundwelle oder mit einer ihrer räumlichen Harmonischen eintritt, die durch den periodischen Aufbau des Wellenleiters verursacht werden, ist gleichgültig.

Abb. 4 stellt links den gefältelten Hohlleiter, d.h. einen Hohlleiter mit periodischer Struktur, eine Verzögerungsleitung dar. Rechts ist der eben beschriebene Schallkanal mit den Helmholtz-Resonatoren in der Wandung skizziert. In beiden Fällen sei die Breite des Wellenleiters die gleiche. Während im glatten Hohlleiter bei einer sinusförmigen Einspeisung der Welle die longitudinale elektrische Feldstärke E_x auch nahe der oberen Berandung $y = b$ räumlich einen sinusförmigen Verlauf hat, treten im Hohlleiter mit periodischer Struktur (Wiederholungslänge L) entsprechende Veränderungen auf. Das Gleiche gilt akustisch für die Transversalschnelle v_y an der Stelle $y = b$ längs der x-Achse im Resonator-Kanal. In beiden Fällen folgt für die Phasenkonstante der $n$ten räumlichen Partialwelle

$$\beta_n = \omega/c_{ph,n} = \beta_0 + n \,.\, 2\pi/L,$$

wobei $n = 0, \pm 1, \pm 2 \ldots$ sein kann. Daraus folgt

$$c_{ph,n} = \frac{c_{ph,o}}{1 + n\,\frac{\lambda o}{L}}$$

Eine Bedingung für Verstärkung ist, dass bei einer der räumlichen Teilwellen, die auch "Hartree-Harmonische" genannt werden, die Phasengeschwindigkeit etwa gleich der Elektronenstrahlgeschwindigkeit bzw. der Windgeschwindigkeit ist. Die Analogie geht aber

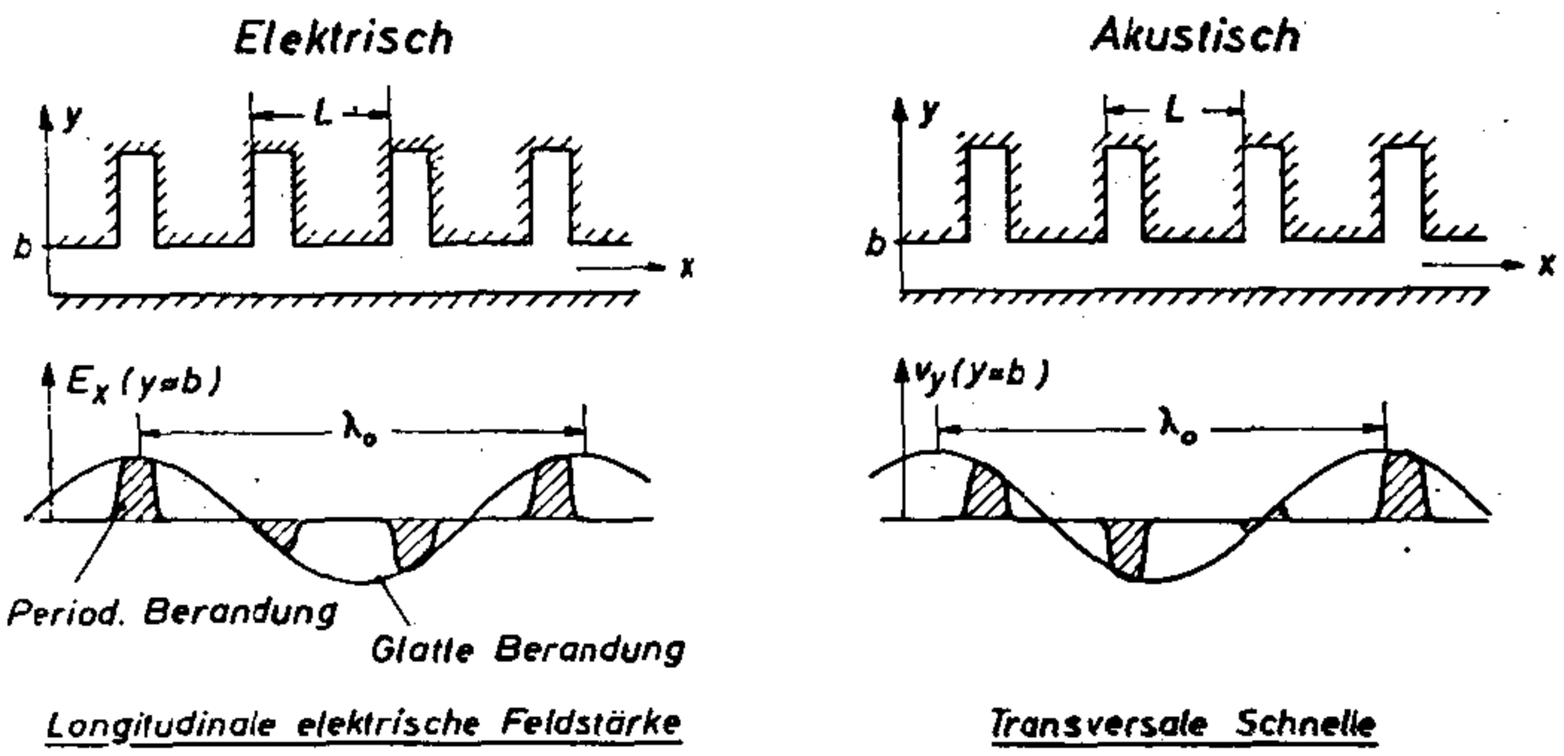


Abb. 4
Räumliche Teilwellen
(Hartree Harmonische)

noch weiter. Im elektrischen Fall bilden sich längs des Elektronenstrahls Raumladungswellen aus, die mit der eingespeisten elektromagnetischen Welle entlang laufen. Im akustischen Fall ist es die Grenzschichtwelle. Ebenso wie diese eine strahlungslose Geschwindigkeitsschwankung darstellt, sind die Raumladungswellen strahlungslose Raumladungsschwankungen. Die Interferenz passiert in beiden Fällen direkt an den Schlitzen der Hohlleiterwandung bzw. an den Resonatoröffnungen in Wandnähe. Raumladungswellen bzw. Grenzschichtwellen sind nur vorhanden, wenn sie einen längeren gemeinsamen Weg mit der eingespeisten Signalwelle gleicher Phasengeschwindigkeit zurücklegen. Dann verstärken sich beide gegenseitig.

Da in unserem Zusammenhang Versuche an der elektrischen Wanderfeldröhre weniger interessieren, seien hier nur für den akustischen Fall einige Messungen zum Beweis des eben Vorgetragenen angeführt. Mit Hilfe eines Hitzdrahtanemometers ist man in der Lage, die Turbulenz einer Luftströmung, d.h. die Geschwindigkeitsschwankungen zu messen. Man kann dies bei allen Signalfrequenzen mit und ohne Schallsignal in der Nähe der Resonatorwandungen tun. Dann zeigt sich, daß es einen Unterschied der Turbulenz mit und ohne Schallsignal nur für die Verstärkungsfrequenz gibt. Diese Differenz $\Delta\sqrt{\overline{v}^2}$ ist für die verschiedenen aufeinanderfolgenden Resonatoren in Abb. 5 aufgetragen. Die Turbulenz wird sozusagen durch die Signalfrequenz und die durch sie bedingte Anregung der Resonatoren gesteuert oder »synchronisiert«.

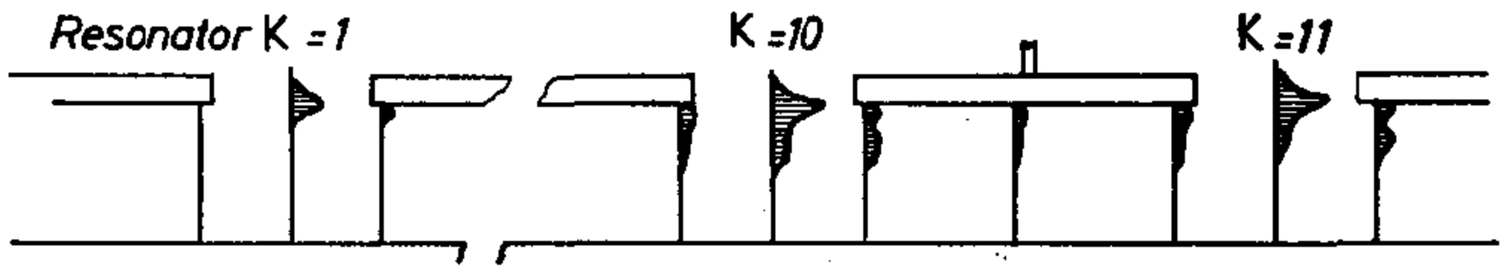


Abb. 5
Pseudoschallverteilung im Kanal

Die Turbulenz, besser die gesteuerte Turbulenz, die auch »Pseudo-Schall« genannt wird, bewegt sich längs des Rohres mit der Strömungsgeschwindigkeit $\bar{V}$, oder genauer gesagt mit $\varkappa\bar{V}$, wobei $\varkappa$ etwa zwischen 0,75 und 1,0 liegt ($\varkappa\,\bar{V}$ = Wirbeltransportgeschwindigkeit), der wirkliche Schall dagegen mit der Schallgeschwindigkeit c. Die Phasendrehung des Schalles längs der Strecke L ist $2\pi\,.\,L\,.f/c$, wobei f die Frequenz ist. Die Phasendrehung des Pseudo-Schalles ist $2\pi\,.\,L\,.\,f/\varkappa\bar{V}$. Wenn eine Verstärkung eintreten soll, muß $2\pi\;\; L\,.\,f(1/c - 1/\varkappa\bar{V}) = 2\pi\,.\,n$ sein ($n = 1, 2, 3$ etc.). Setzt man noch $c = c_0 + \bar{V}$, wobei c_0 die Phasengeschwindigkeit des Schalles im Rohr bei $\bar{V} = 0$ ist, folgt als Funktion der Windgeschwindigkeit für die Verstärkungsfrequenz f der Wert: $f = \varkappa \cdot \frac{n}{L} \frac{\bar{V}(\bar{V} + c_0)}{c_0}$.

Mit verschiedenartig aufgebauten Resonatoren sind die Verstärkungsfrequenzen gemessen worden; es wurde geprüft, ob sie den obengenannten Bedingungen genügen. Abb. 6 mit der Verstärkungsfrequenz $f/\varkappa$ über der Geschwindigkeit $\bar{V}$ zeigt, daß die obige Gesetzmäßigkeit gut erfüllt ist. Die Meßpunkte lassen sich entweder in den Kurvenverlauf $n = 1$ oder $n = 2$ einordnen.

Wenn man anstelle der Resonatoren, d.h. des reaktiven Wandbelags breitbandige Absorber in Form von porösen Platten, d.h. dissipative Absorber, nimmt, ist die Analogie zu den entsprechenden elektrischen Vorgängen nicht mehr so auf der Hand liegend; z.B. anstelle eines Steinwolle-Belages wäre dann ein Graphit-Belag für den elektrischen Fall anzuwen-

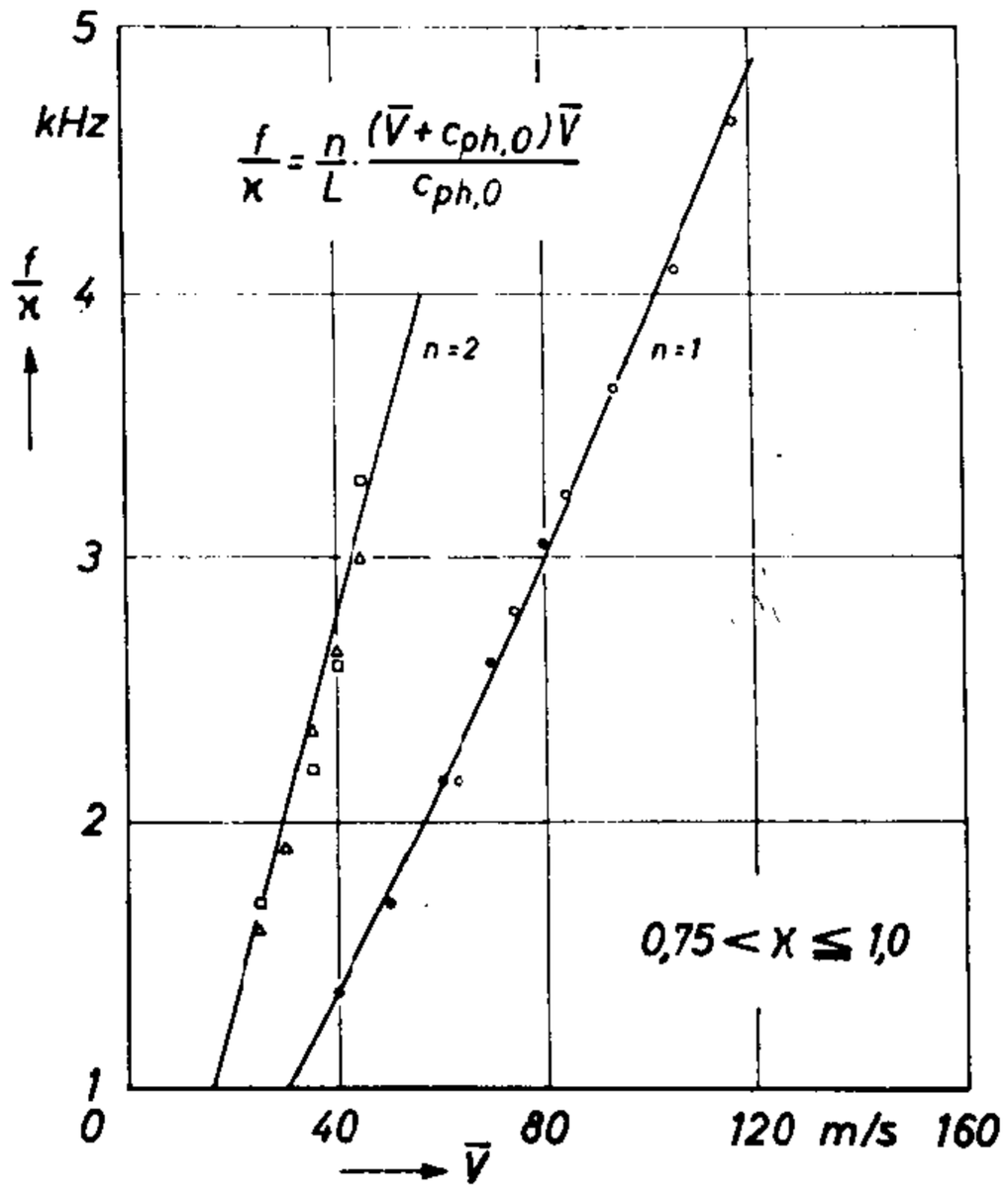


Abb. 6
Dispersionskurven der ersten beiden Hartree Harmonischen und Meßpunkte maximaler Verstärkung.

den. Von Vergleichen dieser Art sei hier abgesehen. Der Vollständigkeit wegen seien hier aber die akustischen Meßergebnisse für Sillanwolle in Abb. 7 und 8 gezeigt; sie stellen die Schalldämpfung in Abhängigkeit von der Frequenz bei einer Schallausbreitung mit und gegen die Strömungsrichtung dar. Der freie Kanalquerschnitt ist in diesem Fall 35 × 100 mm². Im ersten Fall erhält man eine erhebliche Verringerung der Schalldämpfung, im zweiten Fall, Schall- und Luftströmung entgegengesetzt, eine wesentliche Vergrößerung. Im einzelnen sind die Ursachen für diese Effekte inzwischen weitgehend klargestellt. Im Fall der gleichen Richtung von Luftströmung und Schall wird die Wellenlänge vergrößert und damit die Energiedichte verkleinert, was eine verringerte Schalldämpfung bewirkt. Das Umgekehrte tritt im zweiten Fall auf, die Wellenlänge wird verkürzt, und die Energiedichte steigt, was zu erhöhter Absorption führt.

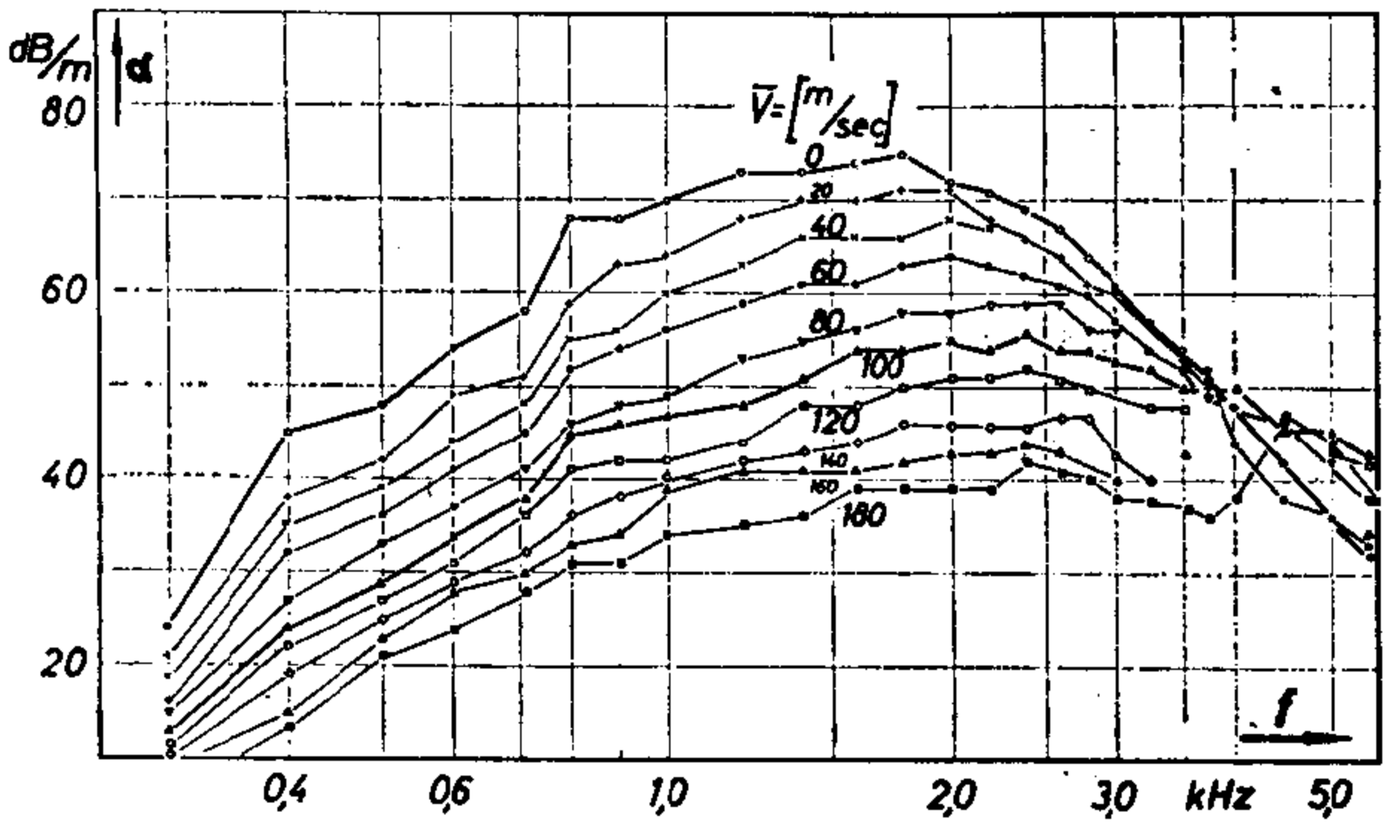


Abb. 7
Dämpfung α in Abhängigkeit von der Frequenz f bei sinem porösen Absorber für Schallausbreitung in Strömungsrichtung. Parameter: Mittlere Strömungsgeschwindigkeit $\bar{V}$.

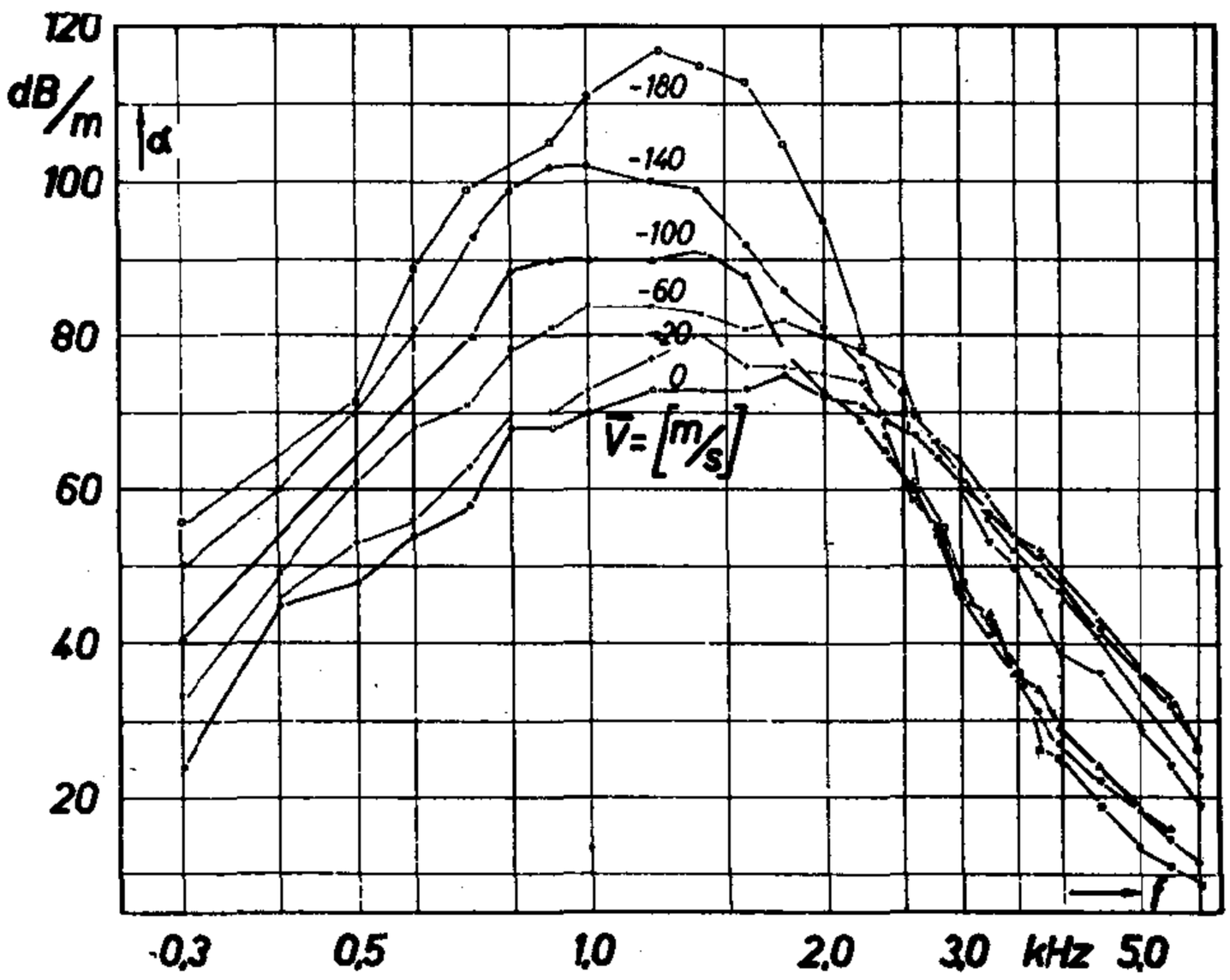


Abb. 8
Dämpfung α in Abhängigkeit von der Frequens f bei einem porösen Absorber für Schallausbreitung gegen die Strömungsrichtung. Parameter: Mittlere Strömungsgeschwindigkeit $\bar{V}$.

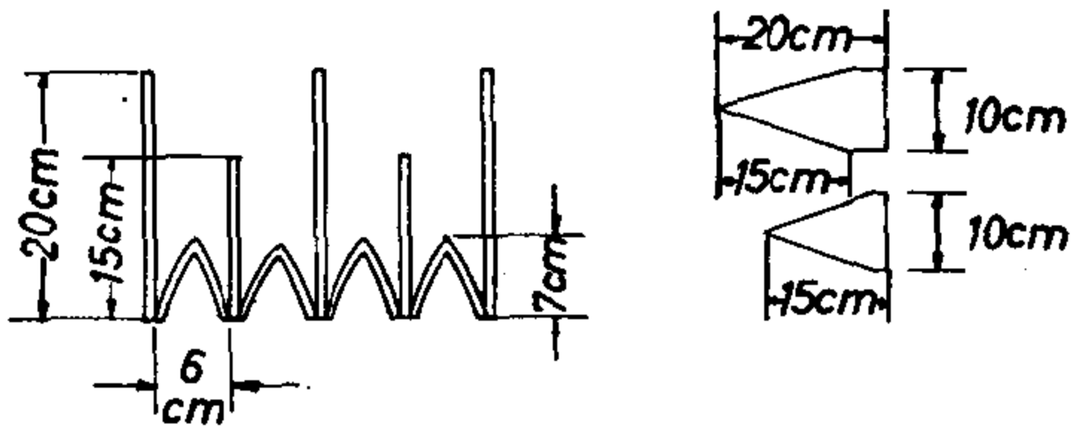
3. Reflexionsfreie akustische und elektromagnetische Absorber

Ein anderes Gebiet, auf dem die Analogie zwischen akustischen und elektrischen Erscheinungen nahe liegt, sind die reflexionsfreien Absorber, d.h. diejenigen Materialien oder Anordnungen, bei denen die auftreffende Wellenenergie möglichst ohne Reflexion völlig absorbiert wird. Für derartige Absorber ist von jeher die Akustik führend gewesen, einfach, weil hier der praktische Bedarf an solchen Anordnungen immer außerordentlich groß war. Es lag deswegen nahe, die elektrische Seite von der akustischen etwas »lernen« zu lassen. Es ist ganz unmöglich, im Rahmen dieses Vortrages die vielen Parallelen einigermaßen erschöpfend darzustellen. Wir wollen nur zwei Gruppen von Absorbern besprechen, bei denen die Analogie besonders ins Auge fällt, und zwar einen *breitbandigen* und einen *schmalbandigen* Absorber. Wir sprechen dabei zunächst nur vom senkrechten Welleneinfall und interessieren uns für die erreichbaren Absorptionsgrade und zugehörigen Schichtdicken.

Für den reflexionsfreien Breitbandabsorber ist die »traditionelle« Form der sogenannte Keilabsorber, d.h. der Typ eines Absorbers, bei dem nach vorn zu die Materialdichte allmählich abnimmt (allmählicher Übergang der Eigenschaften von Luft zu denen des Absorbers). Zur Veranschaulichung sollen die Wasserschall-Breitbandabsorber mit einem Mikrowellen-Absorber ähnlicher Bauart verglichen werden.

Da für Wasserschall der Schalldruck die entscheidende Feldgröße ist, müssen Wasserschallabsorber aus einem gummiartigen Kunststoff bestehen, dessen Verformung große Verluste verursacht. Verluste sind aber nicht bei allseitigem Druck, sondern nur bei Schubdeformation groß, was man durch Hohlräume in dem Absorbermaterial erreicht; z.B. setzt man es aus drei dünnen Folien zusammen, wobei die mittlere Folie entsprechend perforiert ist. Damit bietet sich als einfachste Konstruktion der »Rippenabsorber« an, bei dem nach vorn spitz zulaufende Rippen, eben diese Dreischichtplatten, parallel in einem gewissen

Abstand voneinander angeordnet sind. Eine solche Kombination aus Rippen verschiedener Länge, mit Wasser dazwischen, gibt einen sehr kleinen Reflexionsfaktor in einem großen Frequenzbereich 5–60 kHz (Bild 9).



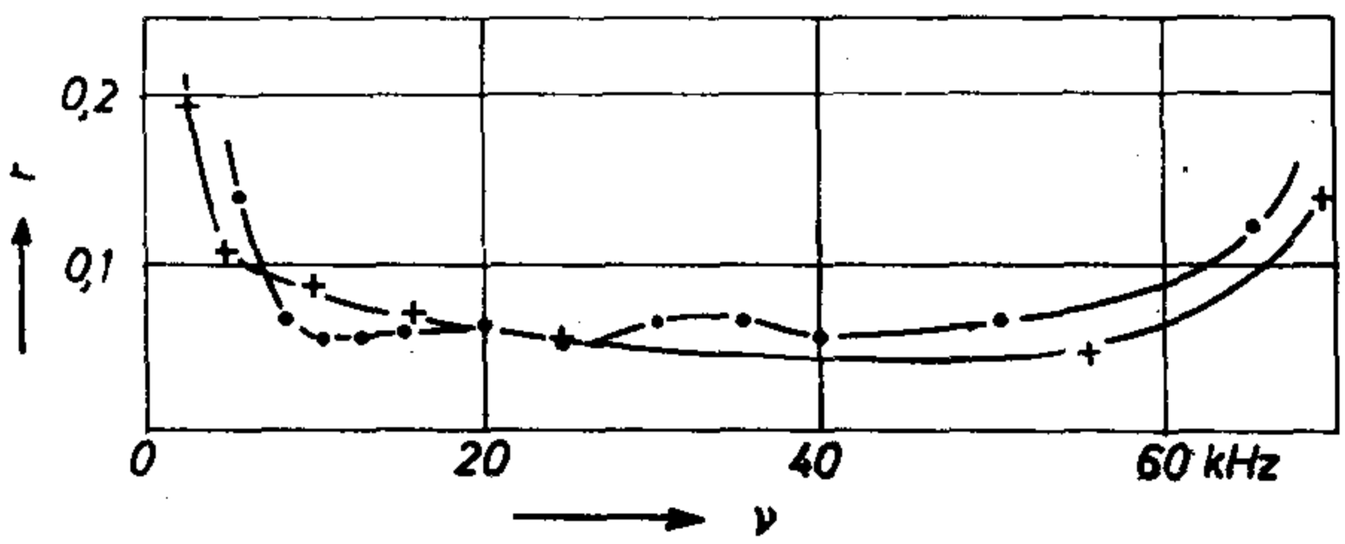


Abb. 9
Aufbau und Reflexionsfaktor eines Wasserschall-Breitbandabsorbers.

Nun zu dem elektrischen Analogon. Die eben beschriebenen Gummifolien sind gegenüber Wasser relativ schallweich. Der Schalldruck der daran vorbeilaufenden Welle wird daher in ihrer Nähe verringert. Im elektrischen Fall treten an die Stelle der Gummifolien Graphitfolien, an denen ebenfalls die elektrische, parallel zu den Rippen liegende Feldstärke mehr oder weniger zusammenbricht. Abb. 10 zeigt schematisch den Aufbau des elektrischen Breitbandabsorbers sowie die Frequenzkurve des auf Amplituden bezogenen Reflexionsfaktors, der in einem großen Wellenlängenbereich unter 5% liegt.

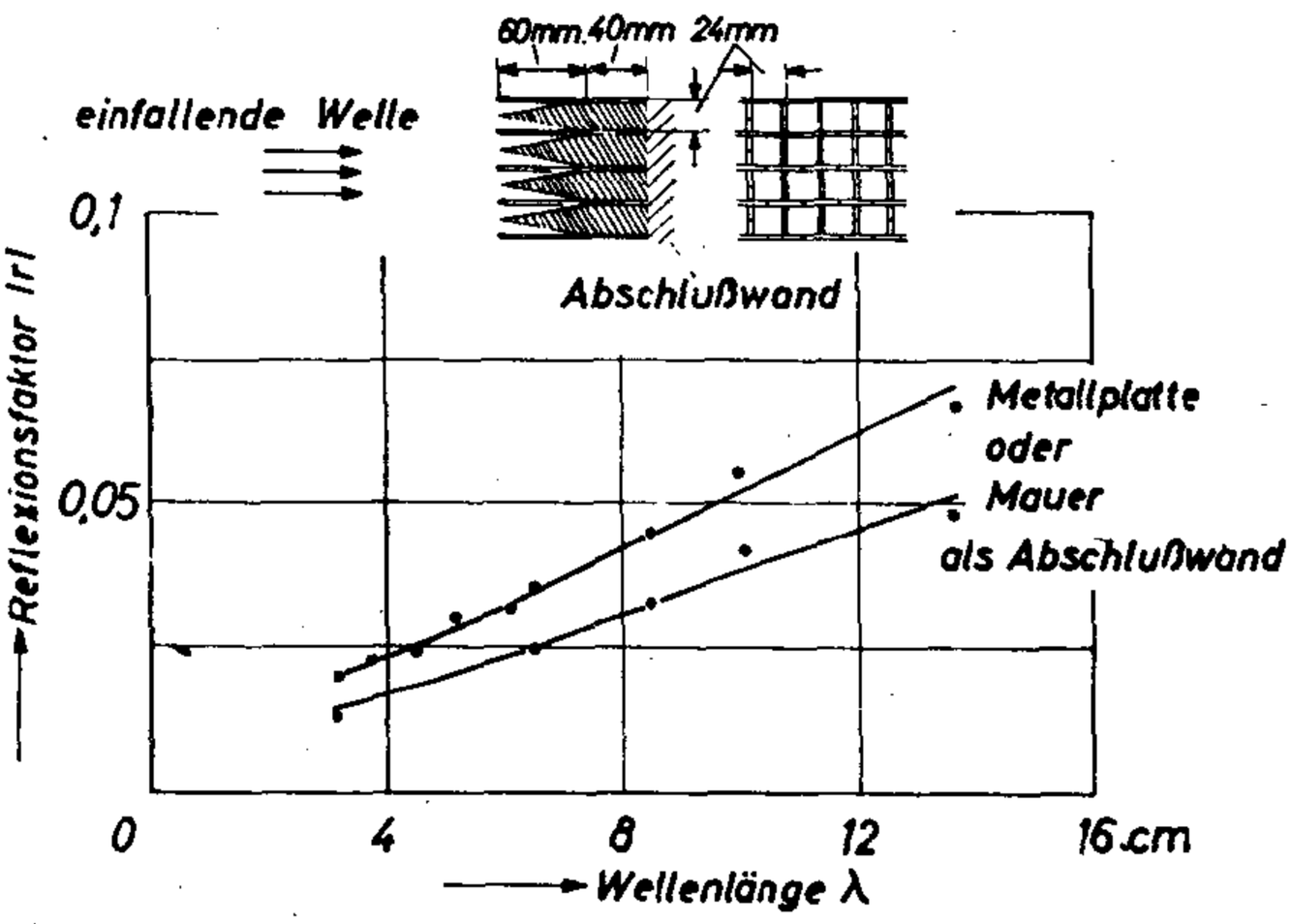


Abb. 10
Aufbau und Reflexionsfaktor eines Rippenabsorbers für elektromagnetische Wellen.

In beiden Fällen ist der wirksame Frequenzumfang groß; der Nachteil ist, daß man wegen der tiefsten Frequenz des Bereiches eine Schichtdicke in der Größenordnung von $\lambda/4$ bis $\lambda/2$ dieser Frequenz benötigt. Daher wird dieser Absorbertyp fast ausschließlich zur Auskleidung von reflektionsfreien Räumen, d.h. von reflexionsfreien Wasserschallbassins, von reflexionsfreien Luftschallräumen und von reflexionsfreien Räumen für elektromagnetische Wellen benutzt.

Kleine Schichtdicken erhält man im akustischen wie im elektrischen Fall nur durch »Resonanzabsorber«. Natürlich wird damit die Bandbreite wesentlich verringert. Wir wollen als Beispiel einen Absorber betrachten, der aus zwei Resonanzkreisen besteht. Für die akustische Seite gehen wir auch hier wieder von der Wasserschalltechnik aus. Der eine

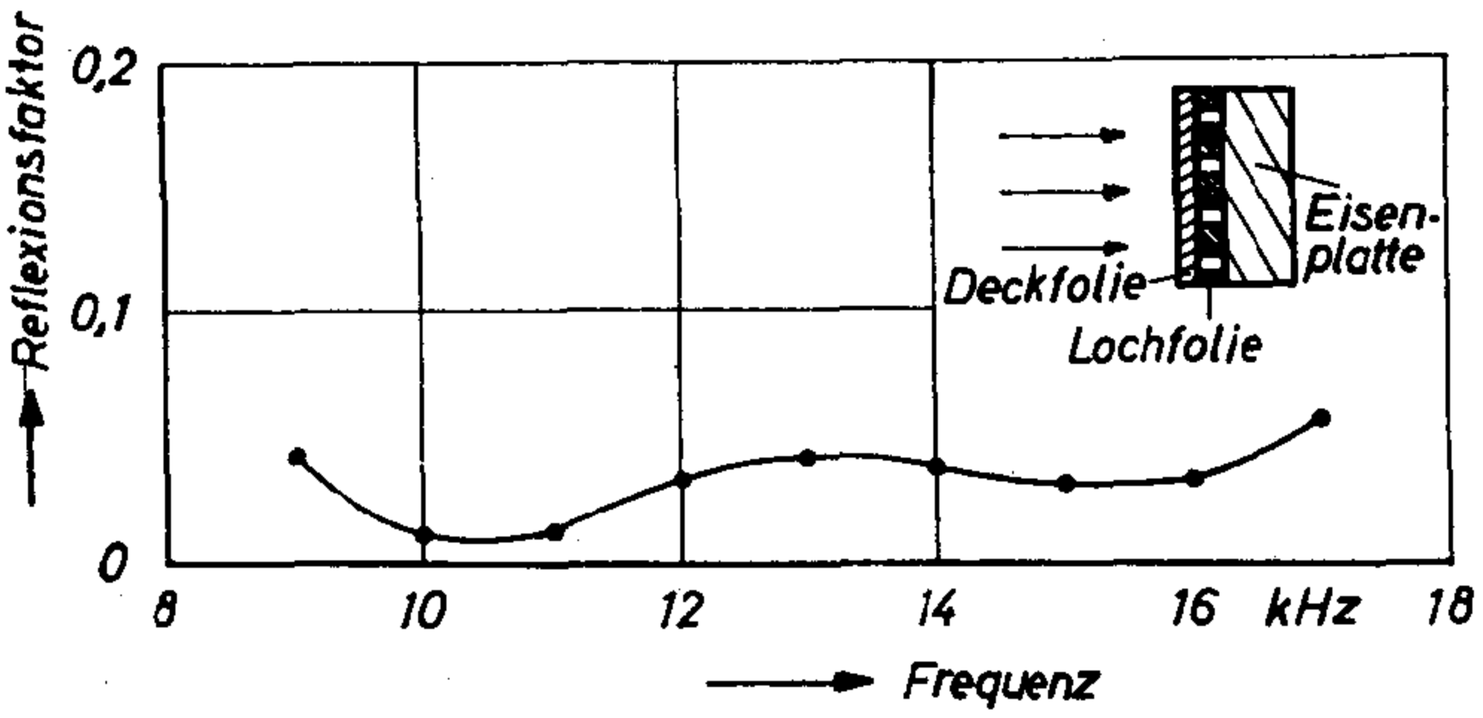


Abb. 11
Reflexionsfaktor eines Zweikreisresonanzabsorbers für Wasserschall.

Kreis wird z.B. aus gleichabgestimmten, d.h. gleich großen Luftblasen gebildet, der zweite besteht aus einem $\lambda/4$-System. Dazu befinden sich die Luftblasen in einem Abstand von $\lambda/4$ ihrer Resonanzfrequenz vor einer schallweichen Wand. Man kann übrigens die Luftblasen auch »einfrieren«, indem man Hohlräume in eine entsprechende Gummischicht einbringt, außerdem ersetzt man den $\lambda/4$ Resonanzkreis durch ein System mit konzentrierter Masse und Federung. Man hat dann die in Abb. 11 (rechts oben) dargestellte Anordnung vor sich. Abb. 11 zeigt zugleich die gemessene Frequenzkurve des Reflexionsfaktors.

Es ist klar, wie die entsprechende Anordnung eines Zweikreisabsorbers im Gebiet der elek-

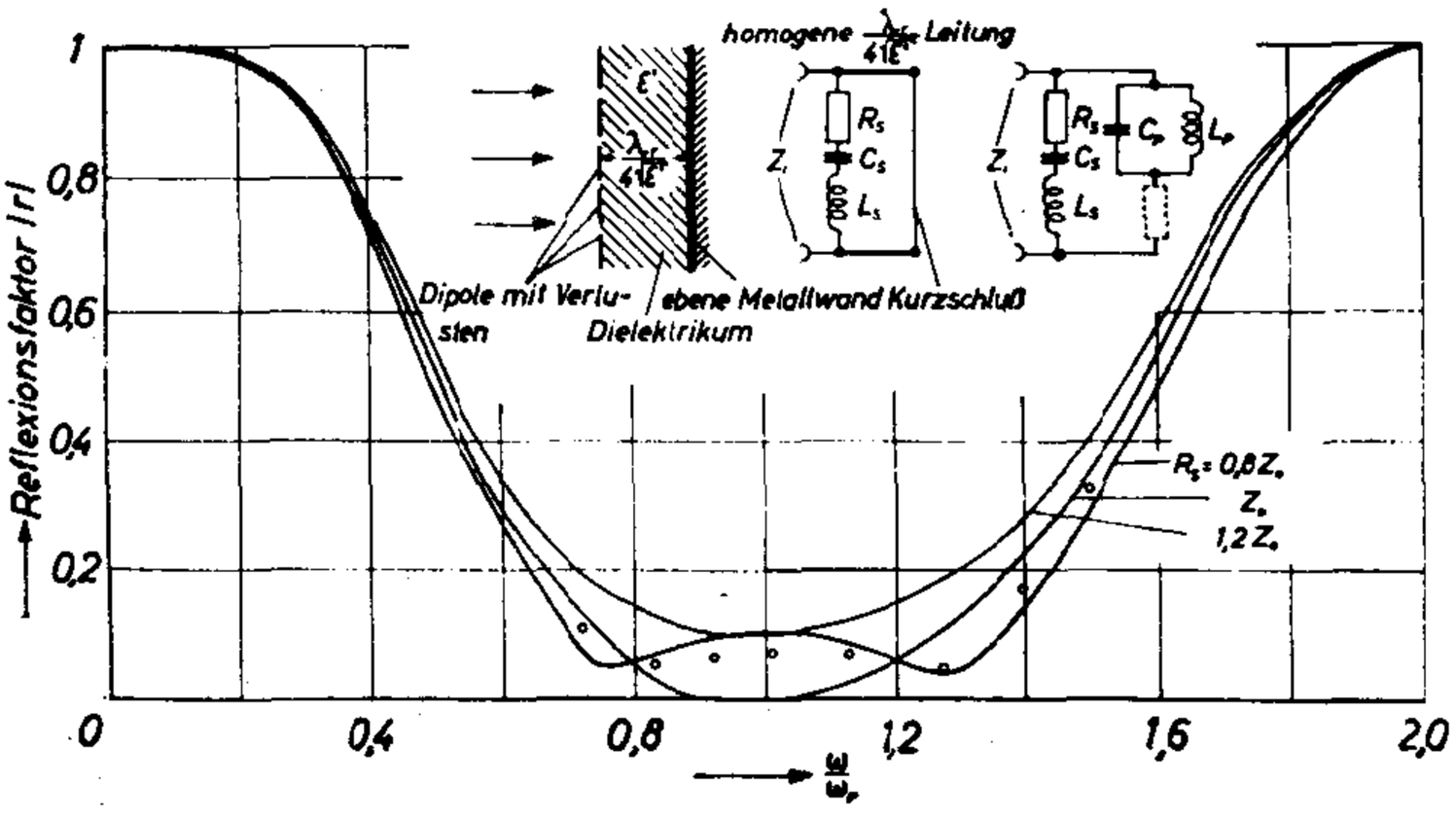


Abb. 12
Aufbau, Ersatzschaltbild und Frequenzkurve eines Zweikreisdipolabsorbers für elektromagnetische Wellen (10 kMHz).

tromagnetischen cm-Wellen aussehen muß, nämlich eine Reihe von elektrischen verlustbehafteten Dipolen (z.B. aus Graphitfolie) — analog zu den Luftblasen — im Abstand $\lambda/4$ vor einer gut leitenden Wand. Das Ersatzschaltschema mit $\frac{\lambda}{4}$ Leitung bzw. mit Parallelresonanzkreis ist das gleiche (Abb. 12) wie bei der akustischen Anordnung. Auch hier ist die Bandbreite des Absorbers und seiner Ersatzschaltungen etwa eine Oktave.
Soweit der akustisch-elektrische Vergleich bei senkrechtem Einfall der Wellen. Auch die Winkelabhängigkeit des Reflexionsfaktors ist in beiden Fällen ähnlich. Auf einen Unterschied muß allerdings hingewiesen werden; er ist bedingt durch die Polarisation der elektrischen Wellen. Der Reflexionsgrad eines elektromagnetischen Absorbers hängt im allgemeinen noch von der Polarisation der einfallenden Welle ab, aber durch geeignete Anordnungen gelingt es, z.B. im Falle des eben genannten Dipolabsorbers durch gekreuzte Dipole oder Kreisscheiben aus Graphitfolie, den Polarisationseinfluß völlig auszuschalten.
Wir gehen jetzt von dem senkrechten Welleneinfall zum streifenden Einfall über. Auch hier gibt es einen interessanten Vergleich in der Struktur des Wellenfeldes bei Ausbreitung parallel zur absorbierenden Fläche.
Beim Luftschall kann man hierzu Untersuchungen in einfacher Weise machen. Man läßt den Schall parallel zu einer schallabsorbierenden Fläche sich ausbreiten und mißt durch Zusatz von leichten Teilchen, Talkumpulver und dgl., die im Schallfeld mitgenommen werden, Richtung und Betrag der Schallschnelle. Bei einer schallharten Fläche als Untergrund müssen die Teilchen sich parallel zu ihr hin und her bewegen und für das Auge gerade Striche ergeben (Abb. 13). Über einem absorbierenden Stoff wird die Phasenfläche der Schnelle (Wellenfront) nach vorn über geneigt, und außer der Schnelle V_x entsteht noch eine senkrechte Schnellekomponente V_z. Die Bewegung der Teilchen ist eine Ellipse (Abb.

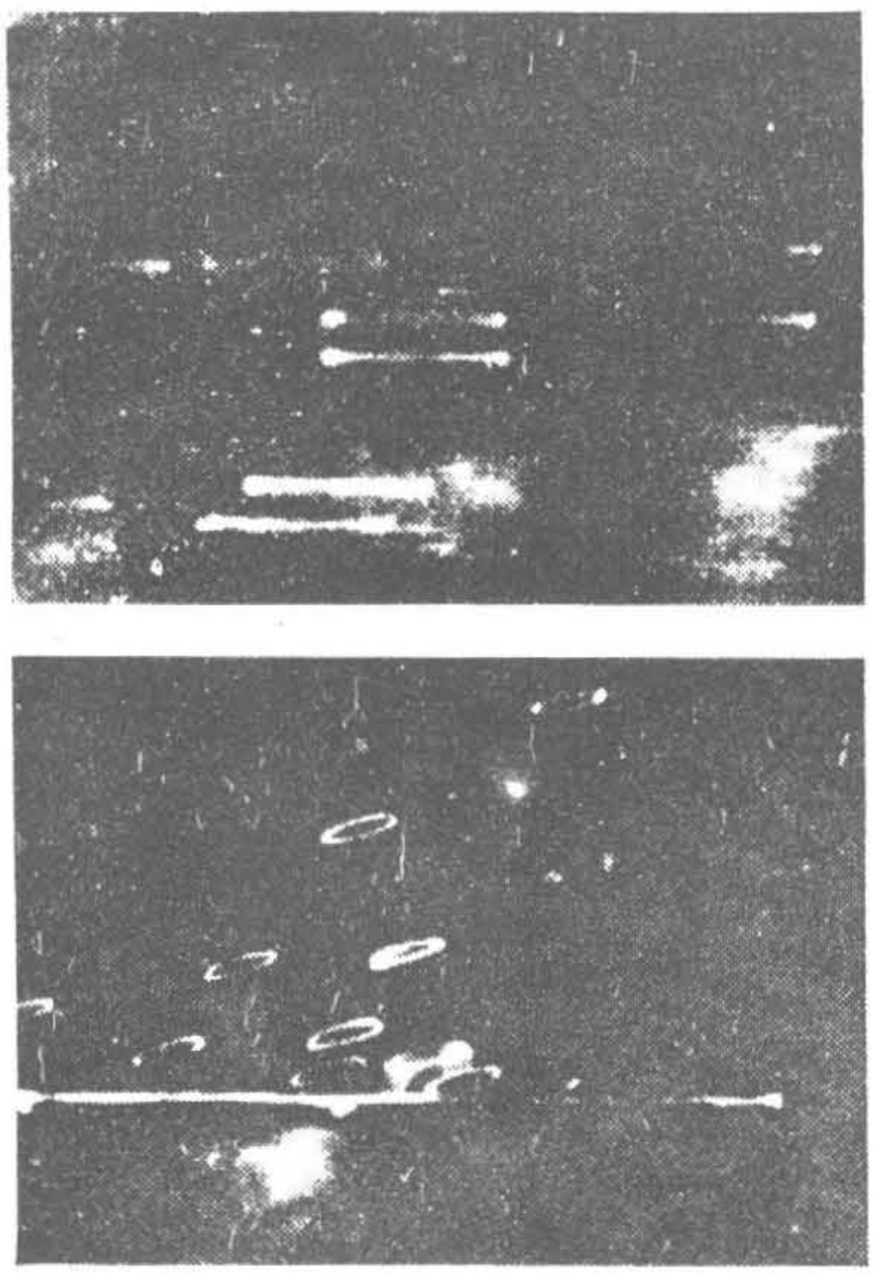


Abb. 13
Teilchenbewegung vor schallharter und schallabsorbierender Wand (ca. 200 Hz).

13). Solche Messungen sind schon vor 10 Jahren ausgeführt worden, und es lag nahe, ähnliche Versuche für die Ausbreitung elektromagnetischer Wellen längs einer elektrisch absorbierenden Schicht zu machen. Dies ist kürzlich geschehen.
Mikrowellen mit einer Wellenlänge von 3 cm, d.h. mit einer Frequenz von 10 kMHz, werden mit einem Trichter über einer Glykolfläche abgestrahlt; Glykol hat eine stark verlustbehaftete Dielektrizitätskonstante ($\varepsilon = 1{,}85 - j\,8{,}6$). Es ist interessant, daß alle Messungen anstatt mit der Frequenz 10 kMHz im Hörfrequenzbereich mit etwa 100 Hz bei hoher Selektivität des Messverstärkers ausgeführt wurden. Dazu wird nämlich über einen rotierenden Phasenschieber (Drehzahl 100 Hz) die Sendefrequenz von 10 kMHz auf 10 kMHz + 100 Hz verschoben und nach dem Empfänger mit der ursprünglichen Frequenz von 10 kMHz gemischt. So entsteht eine Differenzfrequenz von 100 Hz.
Abb. 14 zeigt die erhaltenen Schwingungsellipsen der elektrischen Feldstärke längs einer Linie konstanter Phase, aufgetragen in Abhängigkeit von der Höhe z über der Glykolfläche. Die verschiedene Neigung der Ellipsen, die sich in einer Höhe von etwa einer Wellen-

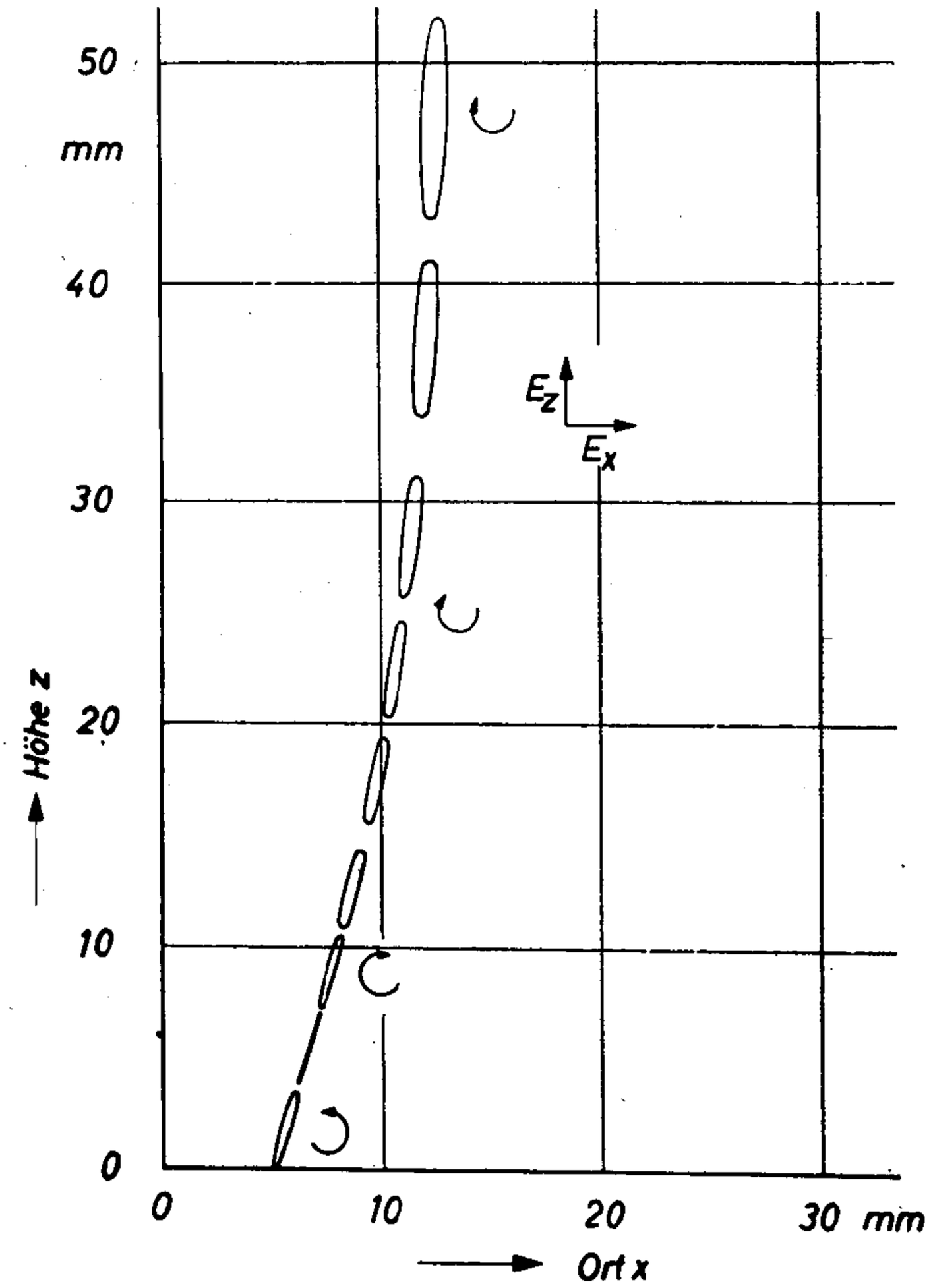


Abb. 14
Schwingungsellipsen der elektrischen Feldstärke längs der Wellenfront einer elektromagnetischen Welle über Glykol (10 kMHz).

länge bereits aufgerichtet haben, ist deutlich. Die longitudinale Komponente E_x wird dann verschwindend klein, während die transversale E_z-Komponente entsprechend dem transversalen Charakter der elektrischen Wellen anwächst. Dicht über der absorbierenden Fläche liegen in Übereinstimmung mit der Theorie die Verhältnisse allerdings etwas anders als beim Schall.

In der Raumakustik ist gut bekannt, daß bei der Ausbreitung von Schall über einer absorbierenden Fläche, z.B. über Publikum, eine erhebliche zusätzliche Schalldämpfung auftritt. Der Schall wird am stärksten unmittelbar über der Absorptionsfläche geschwächt, während einige Wellenlängen darüber die normalen Ausbreitungsbedingungen vorherrschen. Genau das Gleiche trifft für die elektromagnetischen Wellen zu und wurde auch in unserem Modellfall gut beobachtet. Dazu wurde z.B. in einer Entfernung von 3 m vom Sender die elektrische Feldstärke / E_z / (und auch die zugehörige magnetische Feldstärke H_y) in Abhängigkeit von der Höhe z über der Glykolfläche gemessen. Als Vergleich dienten Messungen über einer nicht absorbierenden Metallfläche. Die Ergebnisse zeigt Abb. 15. Die Zunahme der Feldstärke von einem Punkt unmittelbar über der Glykolfläche bis zu einer Höhe von 14 cm darüber beträgt rund 20 dB. Ein solcher Effekt ist für die Metallfläche natürlich nicht vorhanden.

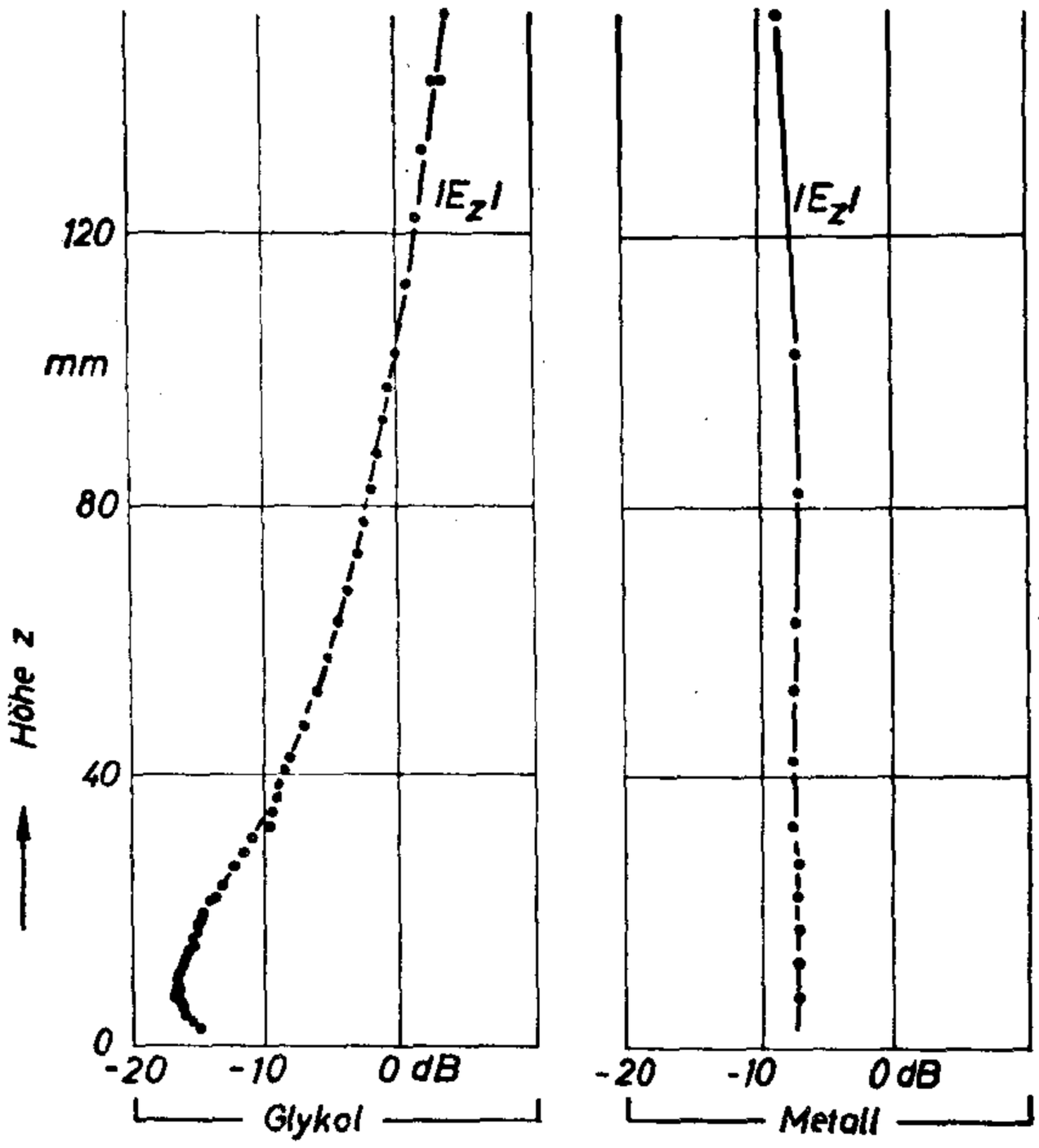


Abb. 15
Zunahme der elektrischen Feldstärke mit wachsender Höhe z über Glykol bzw. Metall (10 kMHz).

4. Hallräume für akustische und elektrische Wellen

Über die Bedeutung des Hallraumes für die akustische oder, besser gesagt, für die raumakustische Meßtechnik an dieser Stelle etwas zu sagen, erübrigt sich. Der Hallraum ist seit mehr als 50 Jahren ein außerordentlich wichtiges akustisches Meßinstrument.

Es ist interessant zu untersuchen, ob das Hallraumverfahren sich auch auf elektromagnetische Wellen übertragen läßt. Derartige Versuche wurden in den letzten Jahren in Göttingen angestellt.

Der Hallraum erlaubt zwei verschiedenartige Messungen durchzuführen. Die Hauptanwendung besteht, nach dem Vorschlag von W. C. Sabine, darin, den Absorptionsgrad von Materialien zu bestimmen, indem man die Nachhallzeit des Hallraumes vor und nach dem Hineinbringen des Prüfmaterials mißt. Die zweite Möglichkeit liegt darin, bei bekannter Gesamtabsorption des Hallraumes die abgestrahlte Leistung eines Schallsenders durch eine Messung der Energiedichte zu bestimmen, analog dem optischen Verfahren der Ulbrichtschen Kugel.

Voraussetzung für solche Messungen ist ein Raum, dessen Eigenabsorption so gering als möglich ist. Im akustischen Falle sind folgende Absorptionsursachen vorhanden: 1) die Zähigkeitsgrenzschicht der Luft dicht vor den Umfassungswänden, 2) die molekulare Absorption der Luft und 3) die Absorption der Umfassungswände (Verluste durch Mitschwingen oder Porosität). Wenn man die Verlustursachen zu Punkt 2) und 3) ganz ausschalten könnte, z.B. durch Verwendung von Stickstoff anstelle von Luft im Raum, und durch Herstellung völlig starrer und unporöser Wände, so bliebe trotzdem noch die erste Verlustursache übrig.

Während man in der Akustik einen Raum durch seine Nachhallzeit charakterisiert, ist es in der Elektrik üblich, für einen Hallraum oder in der »elektrischen« Sprache für einen Hohlraum die »Güte« einzuführen, die die Resonanzüberhöhung jeder Eigenfrequenz angibt. Wenn man im akustischen Fall nur die erste Verlustursache, die Zähigkeitsgrenzschicht, hätte, ist die Güte Q durch den Quotienten Raumvolumen dividiert durch Zähigkeitsschichtvolumen gegeben. Dabei ist das Zähigkeitsschichtvolumen das Produkt aus Raumoberfläche mal Dicke der Zähigkeitsgrenzschicht. Im Falle des Göttinger Hallraumes von 340 m^3 Volumen und rund 300 m^2 Oberfläche führt dies im mittleren Frequenzbereich auf einen Zahlenwert von Q von einigen Tausend. Er hängt natürlich von der Frequenz ab. Um den großen Göttinger Hallraum auch für elektromagnetische Wellen zu benutzen, wurden seine Betonwände mit Kupferfolie bekleidet, deren einzelne Streifen mit Silberlack leitend untereinander verbunden sind. Zur Nachhallmessung wurden kurze Impulse (0,5 μsec) der jeweiligen Frequenz f angewendet. Die Nachhallzeit T (60 dB) betrug für 10 kMHz 410 $\pm$ 15 μsec, daraus ergibt sich eine Güte des Raumes von rund $1{,}7 \cdot 10^6$, da $Q = 0{,}454 \cdot f \cdot T$ ist. Man kann andererseits die Güte leicht ausrechnen; denn die Verluste sind im elektrischen Fall lediglich durch den Skin-Widerstand der Kupferfolie gegeben, die Ausbreitung der Wellen in der Luft ist verlustfrei. Ähnlich wie im akustischen Fall ist die Güte praktisch der Quotient: Raum-Volumen durch »Skin-Volumen«, wobei unter Skin-Volumen das Produkt Raumoberfläche $\times$ Eindringtiefe (Skin-Tiefe) zu verstehen ist. Die Ausrechnung ergibt in guter Übereinstimmung mit dem experimentell gefundenen Wert $1{,}7 \cdot 10^6$.

Ein Bild des Raumes zeigt Abb. 16. Man erkennt die zur Erhöhung der akustischen Diffusität notwendigen Streukörper, wofür Plexiglasschalen verwendet wurden. Derartige Körper sind im elektrischen Fall, insbesondere bei kleinen Wellenlängen nicht erforderlich. Einfacher ist es stattdessen, in der altbekannten schon von W. C. Sabine angegebenen Weise vorzugehen, nämlich einen großen Flügel, hier aus Metall, im Raum rotieren zu

lassen. Er wird elektrisch als »Moden-Mixer« bezeichnet. Seine Wirkungsweise geht aus Abb. 17 a and b hervor. Beide Aufnahmen zeigen für 10 kMHz Nachhallvorgänge in einem kleinen Modellhallraum von rund 6 m^3 Inhalt; dabei ist der eine Nachhallvorgang ohne, der zweite mit rotierendem Flügel aufgenommen. Der Nachhallvorgang ist im zweiten Falle viel gleichmäßiger. Die Nachhallzeit für diesen Modellraum ist übrigens 32 μsec. Da die Nachhallzeit $T = 0{,}185$ V/A ist, erhält man für diesen Modellhallraum ($V = 6$ m^3) eine Gesamtabsorptionsfläche A von rund 230 cm^2 bei einer totalen Oberfläche von $S =$ 23 m^2. Die elektrische Absorptionsfläche beträgt also nur 1% der Gesamtoberfläche. Übri-

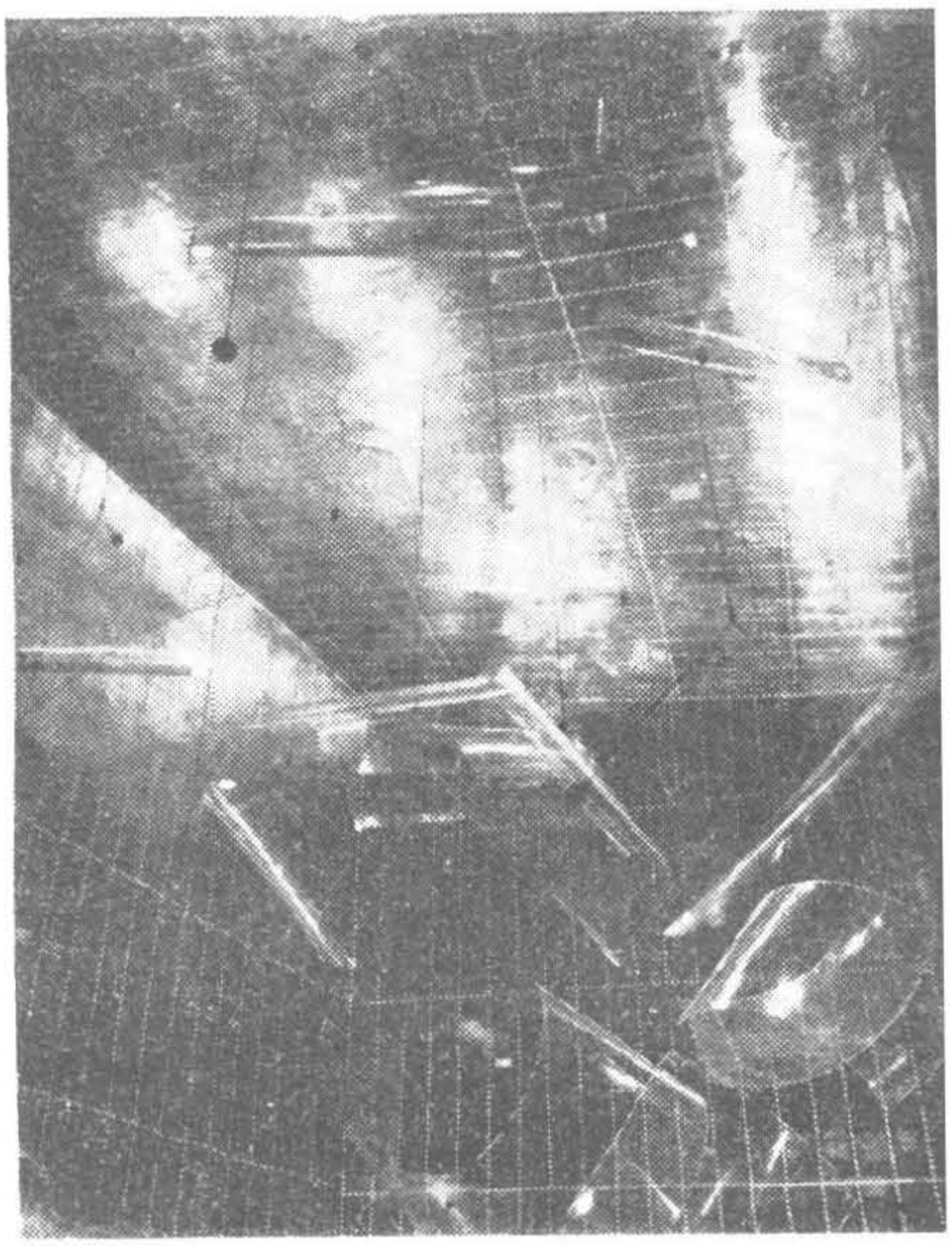
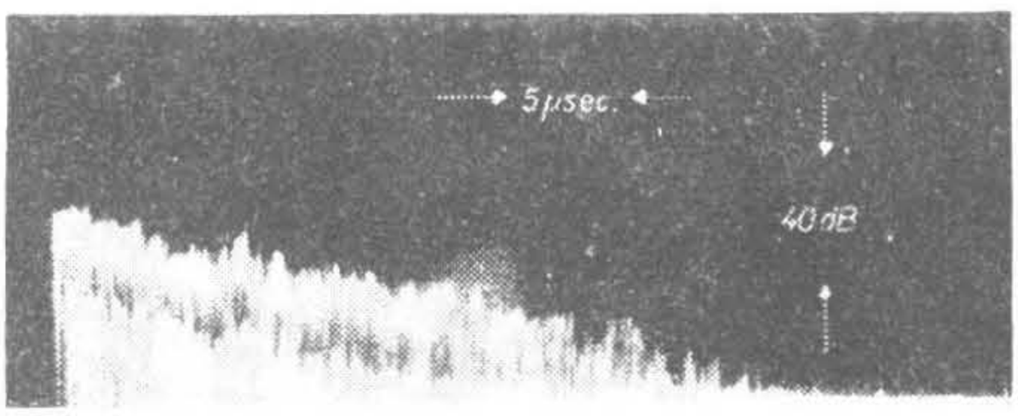


Abb. 16



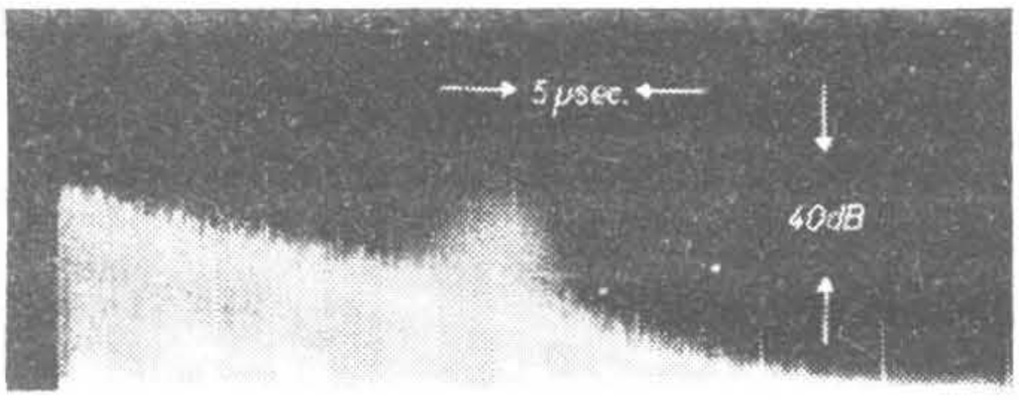


Abb. 17
Nachhallkurven eines Hallraumes für elektromagnet. Wellen (10 kMHz) ohne und mit rotierendem Flügel.

gens sind die entsprechenden Zahlen für den großen Hallraum: $V = 340$ m³, Oberfläche $S = 300$ m² und Gesamtabsorptionsfläche $A = 1530$ cm², d.h. $A \sim 0{,}5 \cdot 10^{-3}\ S$. Im akustischen Fall ist übrigens $A/S = 1{,}4 \cdot 10^{-2}$ bei 1 kHZ. (Tab. 1).

Elektromagnetisch:

$T = (0{,}185\ \frac{\mu sec}{m}) \cdot \frac{V}{A}$ $\qquad Q = 0{,}454 \cdot f\,T$

	V[m³]	S[m²]	A[cm²]	A/S	Q	f[Hz]	T[μsec]
Großer Hallraum	340	300	1530	0,5‰	$1{,}7 \cdot 10^6$	10^{10}	410
Kleiner Hallraum	6	23	230	1‰	$1{,}5 \cdot 10^5$	10^{10}	32

Akustisch:

$T = (0{,}163\ \frac{sec}{m}) \cdot \frac{V}{A}$ $\qquad Q = 0{,}454 \cdot f\,T$

Großer Hallraum	340	300	42600	1,4%	$5{,}9 \cdot 10^3$	10^3	

Tab. 1
Daten Göttinger Hallräume

Die Gesamtabsorptionsfläche kann man übrigens für einen elektrischen Hallraum auch noch in anderer Weise ermitteln, wenn man auf das Sabinesche Original-Experiment mit dem »offenen Fenster« zurückgreift. Der eben erwähnte kleine Modellhallraum enthielt ein Fenster, das durch einen Metallschieber allmählich geöffnet werden konnte und dadurch mehr und mehr Energie nach außen übertrug. Bei konstanter Energiezufuhr wurde die jeweilige Energiedichte E, ausgehend von dem Wert E_0 bei geschlossenem Fenster gemessen. Der Wert $\frac{E_0}{E} = 1$ ist in Abb. 18 über der Fensterfläche aufgetragen. Für den Wert $E_0/E = 2$

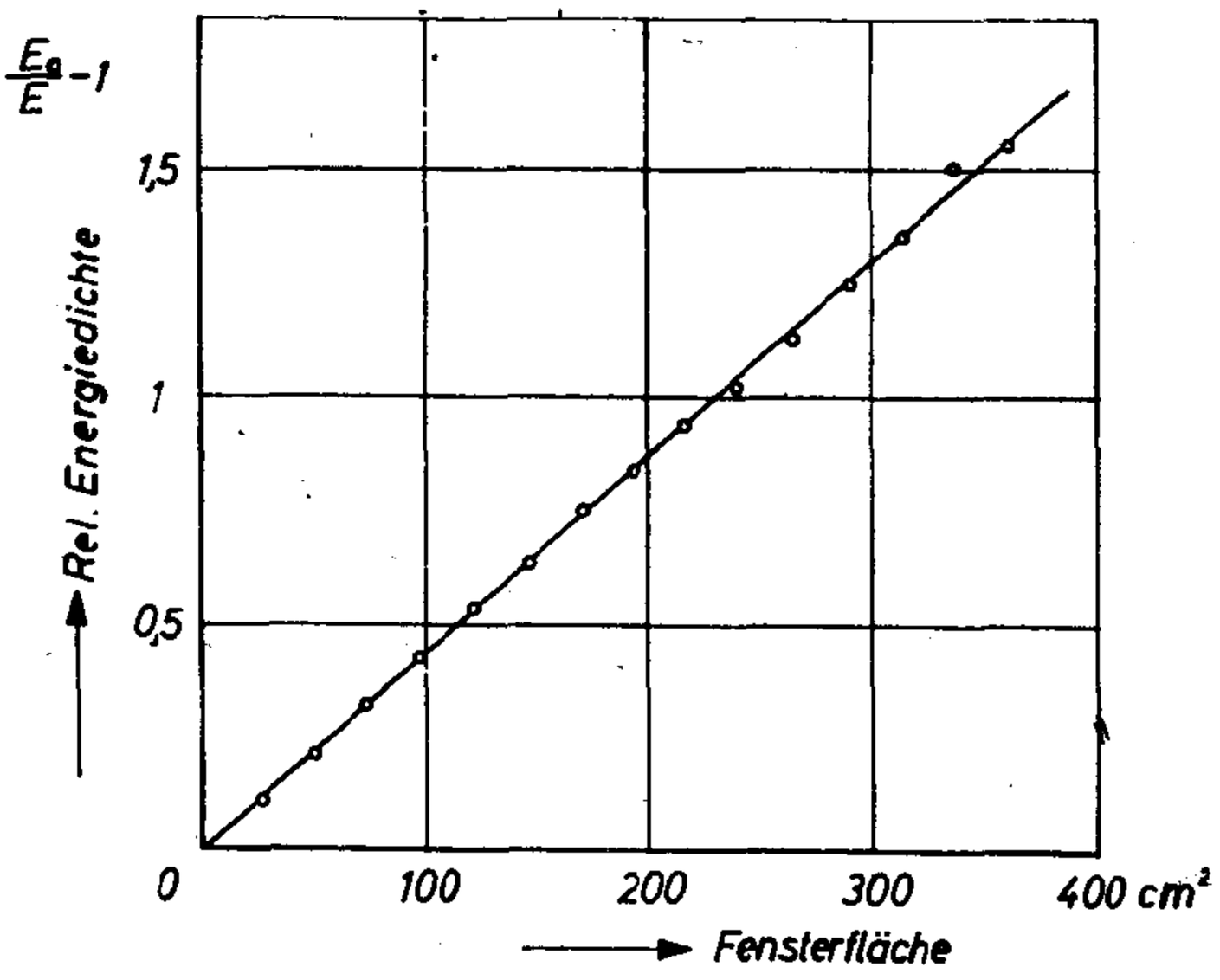


Abb. 18
Gemessene Energiedichte im Hallraum für elektromagnetische Wellen (10 kMHz), in Abhängigkeit von der Fensterfläche.

liest man sofort als Eigenabsorptionsfläche des geschlossenen Modellhallraumes 228 cm² ab, in guter Übereinstimmung mit dem aus den Nachhallversuchen gefundenen Wert.
Diese Übereinstimmung und die sonstigen Messungen für die elektromagnetischen Wellen sind eine schöne Bestätigung der raumakustischen Nachhalltheorie. Den elektrischen Hallraum benutzt man wie den akustischen zur Absorptionsmessung von Materialien oder Materialkombinationen bei diffusem Welleneinfall. Diffus heißt hier allerdings nicht nur gleichmäßige Verteilung aller Einfallsrichtungen, sondern auch gleichmäßige Besetzung aller Polarisationsrichtungen.
Aus einer großen Zahl von Messungen sei hier die Bestimmung der Absorptionsfläche von Personen herausgegriffen. Die Personen wurden im Hallraum einzeln aufgestellt, so daß man den Absorptionswert je Person erhält, übrigens anders als es zur Zeit in der Raumakustik üblich ist (Absorptionsgrad je Publikumsfläche) Abb. 19 zeigt das Resultat (Absorptionsfläche als Funktion der Personenzahl). Über insgesamt 5 Personen gemittelt, war die Absorptionsfläche je Person rund 1 m²; dies gilt für eine Frequenz von 10 kMHz oder eine Wellenlänge von $\lambda = 3$ cm.

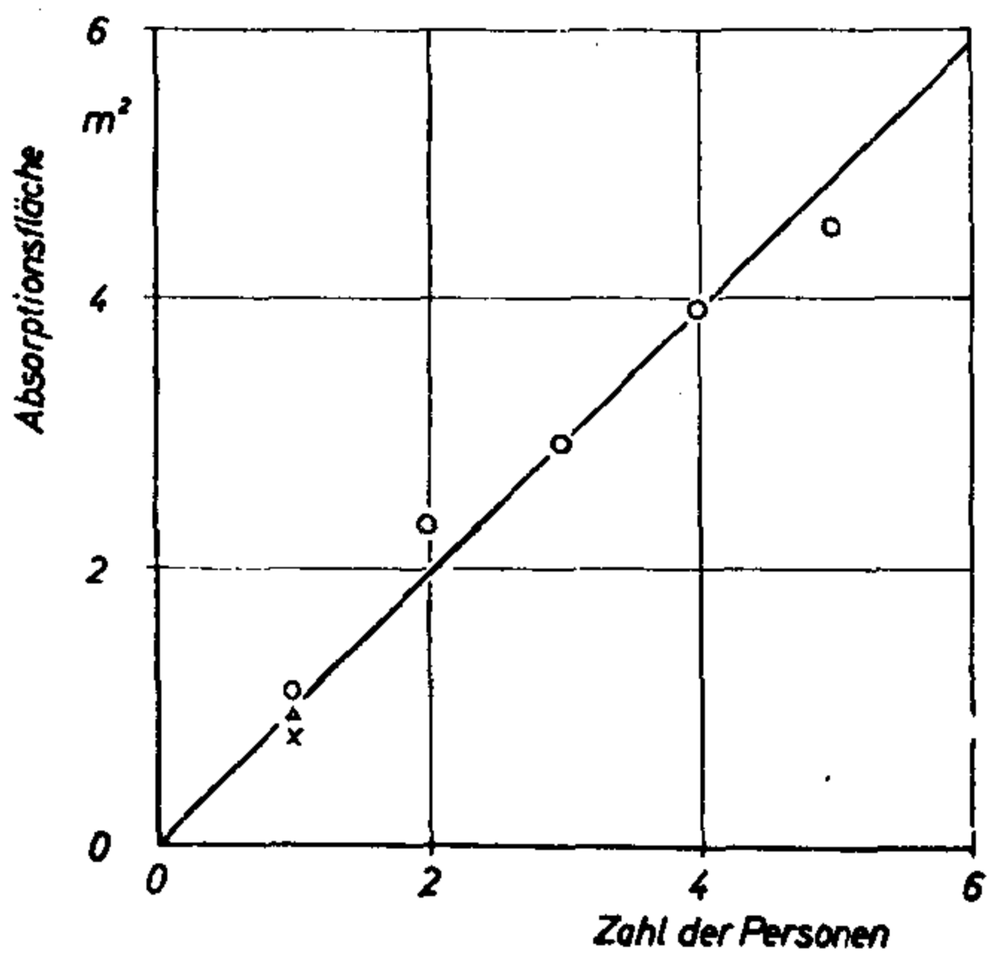


Abb. 19
Absorptionsfläche stehender Personen für elektromagnetische Wellen (10 kMHz), gemessen im Hallraum.

Zum Schluß sei noch eine zweite Anwendung des Hallraums für elektromagnetische Wellen erwähnt, gleichfalls analog zur Anwendung eines akustischen Hallraums. Kennt man nämlich die Gesamtabsorptionsfläche A, so kann man durch eine einzige Messung der Energiedichte an einem beliebigen Punkt des Hallraumes die gesamte abgestrahlte Leistung P_s eines hineingebrachten Senders bestimmen. Elektrisch ausgedrückt, bedeutet dies: $P_s = A/F_e \cdot P_e$.
P_e und F_e sind aufgenommene Leistung und Empfangsfläche der Meßantenne. Beide Größen müssen bekannt oder sonst irgendwie gemessen sein, ähnlich wie das die Schallenergiedichte im Hallraum messende Mikrofon absolut geeicht ist. Auf diese Weise konnten im 10 kMHz-Gebiet die Wirkungsgrade einer großen Reihe von Modellantennen gemessen werden, ohne daß man das sonst übliche, umständliche Meßverfahren der genauen Aus-

messung des Freifeldes der Antennen im reflexionsfreien Raum und der Integration der Meßwerte zu benutzen braucht. Auch hier war die Übereinstimmung mit den im Freifeld zur Kontrolle gemessenen Werten gut.

Soweit der Vergleich akustischer Hallraum — elektrischer Hallraum. Die vorstehenden Ausführungen mögen gezeigt haben, daß man immer wieder mit großem Nutzen die Analogien zwischen elektrischen und akustischen Schwingungen und Wellen zur Lösung von interessanten Problemen und zur gegenseitigen Befruchtung beider Gebiete heranziehen kann.

Recent Advances in the Subjective Measurement of Noise

by D. W. Robinson

The increasing incursions of noise into the privacy of the home, into the business conference, and even into the tranquillity of the remote countryside create a mounting series of problems—social, psychological and hygienic—which transcend the boundaries of acoustics alone. Even when a solution is within the technical grasp of the sound engineer, effective action lies rather with administrators, men of finance, and governments. High cost is usually a concomitant of the amelioration of noise nuisance and must be weighed against advantages which are less tangible. Moreover, the alleviation of noise for one interest has sometimes to be purchased at the expense of some restriction or loss of amenity for others. It is outside my competence to go into these wider questions, but it is salutary for those of us who work in more specialized fields to stand aside occasionally and review our own contribution against this broader backcloth. The special contribution of subjective acoustics is to provide some quantitative guidance on the probable human reaction to noises not yet heard and to assess the reasonableness of reactions to existing ones. In a matter involving human opinion and judgement it is unrealistic to expect criteria possessing the inescapability of a physical law. On the other hand decisions of finance, legislation or control must usually be very clear-cut, which places the subjective acoustician in something of a dilemma. Others will sometimes read more into his results than he would claim for them himself.

In this lecture I confine myself to the aspect of noise implied in the wellknown definition, "sound undesired by the recipient", in other words to noise considered as a nuisance. Amongst the various problems caused by noise, the interference with people's amenities has a claim on an important share of the available research effort, in that it is widespread in occurrence, affects the greatest numbers of people, and is insidious in its growth; also it presents a much less clearly defined area of study and thus a greater challenge to the research worker. In addition to the experimental difficulties of noise measurement and the unavoidable need to attack the problem in a statistical way with large numbers of people, there arises at the outset the fundamental question whether quantitative assessments can usefully be made at all, for it is not at all obvious how to set about measuring the "annoyingness" of a noise, nor on what basis to set a level above which a noise becomes a nuisance.

During the interval between the 3rd and 4th I.C.A.'s much experimental progress has been made, with aircraft noises—present and future—providing one of the main incentives. The science of subjective noise measurement is thus in a state of growth and there is no plateau

at this time on which one can conveniently rest to survey the horizon with detachment. My own view is coloured by the particular work with which I have been associated at the N.P.L., and which forms the basis of this lecture, and I am only too conscious of the inadequacy of my coverage. These are shortcomings which I hope you will understand and forgive.

Loudness

It has been said that there are just three basic dimensions of acoustic sensation [1], pitch, loudness and quality, in the sense that two sounds may be presented which differ from each other with respect to one of these dimensions while being the same with respect to the other two. Certainly loudness has been singled out for the most extended studies; pitch to a lesser extent. Loudness is a quality in rather close correspondence with the physical intensity of sound, and as such it is *sui generis.* Alone among the substantives defining the subjective magnitude of a sound the word "loudness" conjures up no overtones of annoyance, acceptability, or pleasant or unpleasant associations, but describes simply the "size" of the sensation. This makes it a fairly tractable subject of experimental investigation. When we turn to actual noises, their quality, context, relevance for the listener and personal associations enter in. While it remains appropriate to speak of the loudness of noises, we also need somehow to take into account these other, more "psychological" characteristics as well. Broadly speaking, the study of loudness has reached a point where it is possible to calculate the probable judgement of the "normal listener" given the sound pressure level spectrum and some facts about the direction of arrival of the sound wave and its variations with time. I shall qualify this remark later, but in comparison with the still primitive methods of predicting acceptability, annoyance, noisiness, and so on, the theory of loudness must be accounted a highly-developed one.

Nature of Subjective Measurement

Before turning to the methods of subjective investigation, it is appropriate to consider shortly what is meant by subjective measurement, although this is a topic that has been argued at great length by psychophysicists for over a century.

There is no inherent technical difficulty in expressing the physical events which constitute a noise in any desired degree of accuracy and detail. For example, if we have oscillographic records of the waveforms of sound pressure at the two ears of the hearer these must contain all the information of an acoustical character which it can possibly be necessary to know about the noise. I say, advisedly, "of an acoustical character" because I intend to leave aside the complication that people's experience of noise is sometimes influenced by other modes of perception, as well as by associations, fears and neuroses; I leave these problems aside because they belong to other fields than acoustics. An example is the creaking door in the night which can lead to feelings and drastic action that greatly exaggerate the usual response to such small sources of sound energy. Returning to the purely acoustical stimulus, the oscillogram obviously contains far more information—in the technical sense of the word—than the listener can make use of and a more practical physical description of the

sound is one in which a great deal of the redundant information has been suppressed. The purpose in subjective measurement is to establish a correlation between certain qualities of the sound as judged by the hearer and the physical properties of the same sound, so we aim to simplify the physical representation in a way that makes the correlation as straightforward as possible. In the nature of things the approach must be empirical, though experiment can well be supplemented by intuition, by principles already established (for example the idea of the ear as a frequency analyzer), or by models provided by physiological and neurological research into hearing. It is also important to keep an eye on the practical objective, so that the measurement of subjective values can be kept as simple as is consistent with the facts. As research workers we are concerned with the true state of affairs revealed by experiment, but it is not necessary to burden the non-specialist with finer points which can only be demonstrated by elaborate tests. I have in mind two examples. One is the simplified rule relating loudness in sones (the scale in which the number is proportional to the average person's estimate of the magnitude of the loudness sensation) to loudness level in the phon scale. This states that the logarithm of the sone value increases linearly with the phon value by the formula $S = 2^{(P-40)/10}$. Geometrical intervals in the sone scale arise directly out of subjective tests. Phon intervals are basically unrelated to hearing, being simply steps on a logarithmic sound intensity scale for the basic case of the 1000 c/s reference tone. To expect the mathematical equation to be exact, therefore, is to ask that the ear and brain conspire to make loudness exactly proportional to the 0.3 power of sound intensity. Looking at Fig. 1, it may be thought remarkable that a relation which is so nearly true over the immense range of more than 10^8 to 1 of sound intensity is not in fact exact; but experiments by Stevens and myself [2, 3] seem to confirm this. The practical point is that the departure is of no consequence in estimating the loudness of everyday noises, and the simple formula has now been standardized by the I.S.O. [4]. Another example occurs in the latest version by Zwicker [5] of his method of calculating the loudness of sounds of arbitrary

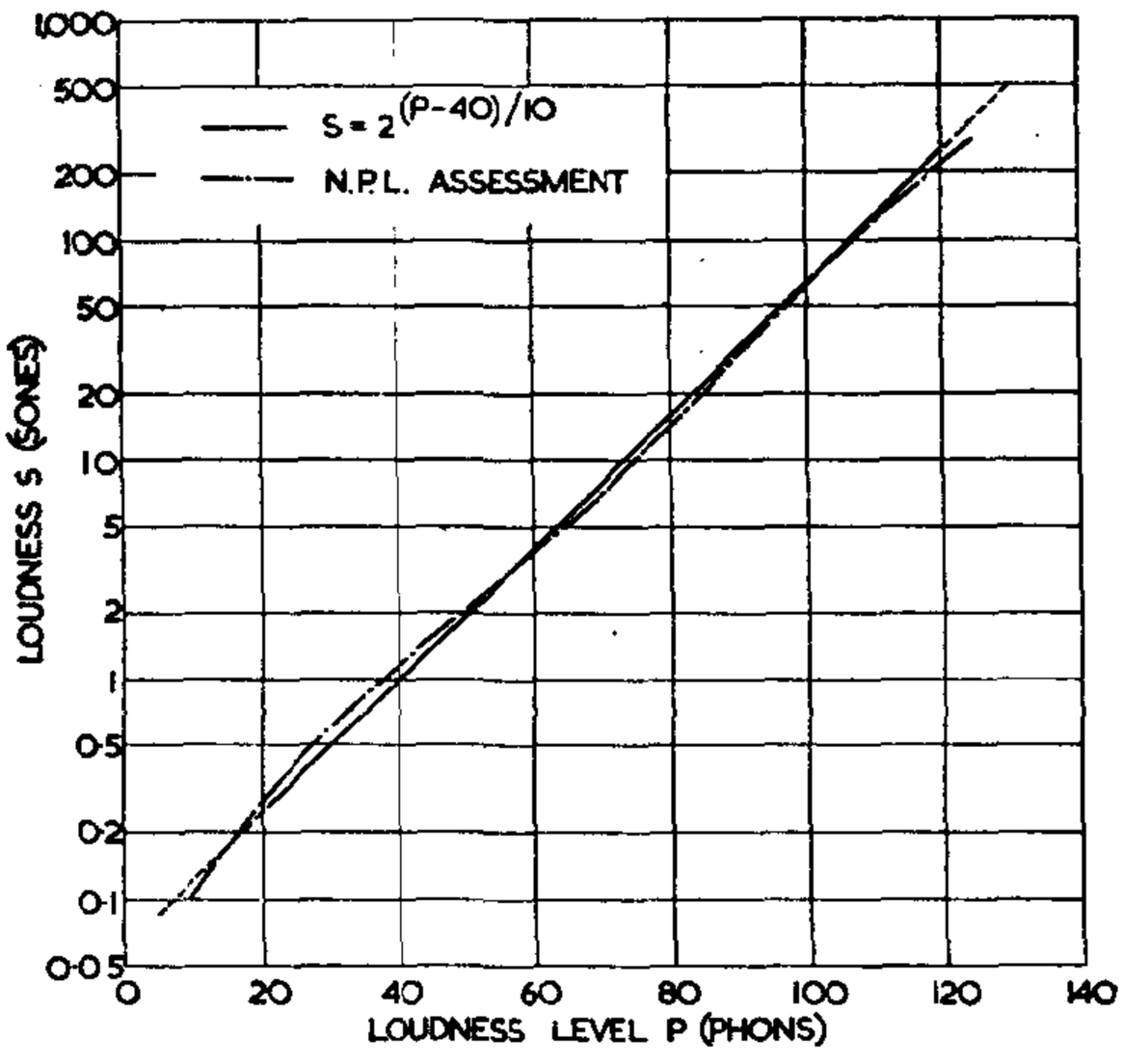


Fig. 1. Relation between loudness and loudness level.

spectra. The method was originally conceived in terms of "critical bands" of frequency [6] which in turn stemmed from experiments on the relation between the loudness and bandwidth of random noises; the bands turned out to be somewhat uneven divisions of the audible spectrum. In the simplified method the bands are approximated by multiples of one-third of an octave which is without appreciable effect on the calculations but makes the procedure directly adaptable to engineering use.

Subjective measurement forms a bridge between the objective and human worlds of experience that is indispensable whenever decisions must be based on meter readings instead of a direct appeal to the opinions of the people concerned, and its spread (not only in acoustics) testifies to the faith that is nowadays placed in a scientific approach to human problems. Is this faith justified by the present state of knowledge and current research in psychoacoustics? Opinions about noise vary from time to time, from one set of circumstances to another, and from person to person. Opinions about some things are so sensitive to extraneous influences, or merely capricious, as to make the use of the word "measurement" almost absurd. It is impossible to visualize, even in principle, a device constructed from sensing elements and a memory in the form of fixed "instructions" that would correctly register in all cases the degree of annoyance that a noise would provoke in the "average" person. Even in the construction of less ambitious meters there are certain disadvantageous factors at work. One of these is that the scale of noticeable differences is much coarser in the subjective world of sound than in the physical. This has important consequences; in the first place, it takes rather large physical reductions of a noise to achieve a modest degree of subjective improvement. Secondly, the imprecision which is inherent in subjective measurement is greatly magnified when translated into physical terms. Thus, to set close limits on a noise in terms of meter readings is arbitrary in the sense that one would have difficulty (or would be taking quite a risk) to put a marginal decision to the test of direct appeal to a group of listeners. At the same time one must face the paradox that it is necessary to set limits sharply in order to avoid endless difficulties of enforcement; if not, the producer of the nuisance would be quite entitled to take advantage of the upper tolerance that is allowed.

These facts are the debit side of the ledger. Against them may be set the fact that subjective measurement is already being successfully applied to specific problems of noise nuisance. Provided one can assume a fairly definite environment and a limited class of noises, reasonably accurate subjective predictions become possible. This is the approach that is being adopted towards the control of motor vehicle noise. Admittedly from the scientific point of view it is unsatisfactory to have to abandon a general solution, but partial solutions are better than none at all.

Which qualities can and which cannot be subjectively measured, and what exactly is meant by "measurement" in this context [7] are questions of a philosophical kind that can best be answered pragmatically. We cannot observe someone else's "sensation" so that propositions like Fechner's Law which relates the "sensation" to the logarithm of the "stimulus" will certainly not appeal to the logical positivist. However, behaviour in the form of a subject's verbal or other response is easily observed, for example when he says of two sounds, "They are equally loud to me". The test of subjective measurability resides essentially in consistency checks. For example, if A and B are equally noisy, and B and C are

equally noisy in the opinion of a listener, then one can verify experimentally whether he also finds *A* and *C* equally noisy. In a similar way, the possibility of measurement on a ratio scale of loudness (the sone scale) depends on the experimental verification of propositions such as this: if *A* is twice as loud as *B*, and *B* is twice as loud as *C*, *A* must be four times as loud as *C*. In practice we exact a further condition of measurability, namely a reasonable consensus of opinion between different "normal" listeners, because we are ordinarily concerned with decisions taken on behalf of communities of people rather than individuals. Some qualities of sounds pass these tests of measurability; loudness is one of them. We are not yet sure whether the same can be said of other terms used to describe noises, especially their unpleasant associations.

Measurable qualities of noise

How many different aspects of human reaction to noise do we need to measure, and how shall we choose them? At first sight we might equate the number of possible scales with the number of adjectives that can qualify the noun "noise". However, the number of *independent* qualities is much fewer. The way to distinguish between the essentially distinct and the "hybrid" qualities is to ask groups of listeners to put a series of sounds in ascending order with respect to each quality and see how many different orders result. Experience leads us to expect at least a slight influence of the test "instructions" on the performance of subjects in experiments like this, which may be due in part to semantic causes; a pair of words may seem synonymous to one person while another would make a distinction. An example, in the English language, consists of the two terms "disturbingness" and "bothersomeness" which have actually been used in tests; I am sorry that both words are so inelegant, but I suggest that it would take an ingenious philologist to say that they were not synonymous. Similar considerations arise in translating from foreigh languages. We have wondered whether this has some bearing, for instance, on differences between motor vehicle noise judgements carried out in terms of "lästigkeit" in Switzerland and of "noisiness" in England [8, 9]. If we had chosen a nearer English equivalent, say "disturbingness", would our results have been in better agreement? It appears to be too soon to answer such questions with confidence, but some light is thrown on the matter by recent work which I refer to later.

The guiding principle should be, for the present, to keep the number of subjective units and scales in practical use to a minimum. A few years ago it was considered enough to have just two, namely the scales of loudness and of loudness level, with the *sone* and *phon* as the respective units. Even then confusion with their physical counterparts, sound pressure level, weighted sound level etc., which are expressed in decibels, was common. Lately, as a result of supposed shortcomings of the loudness scales to compare the noisiness of aeroplanes properly, Kryter and his associates [10, 11] introduced the two new scales of *perceived noise level* and *noisiness*, with the *PNdB* and *noy* as units. It may turn out that these scales measure noise along some other dimension than loudness, but this must finally be decided by careful experiment. Economy of hypothesis is a sound maxim in science and new units ought only to be admitted if the existing ones will not serve. Moreover, it is important to prevent subjective acoustics from falling into disrepute in the eyes of the non-specialist through an

excess of esoteric terminology: in reality it is a subject concerned with ideas which are within the personal experience of everyman.

It would probably be generally accepted today that the minimum need is for two kinds of noise measurement in addition to loudness. One of these would cover the element of "noisiness", "disturbingness" of "annoyingness" of a noise, qualities which go beyond the idea of simple subjective "size" implicit in the "loudness". This idea was in the minds of the originators of the perceived noise level scale. However, the PNdB scale arose out of studies of jet engine noise and it has come to be generally associated with rather intense levels. There is some evidence that very loud sounds are disturbing simply on account of their intensity, and if this is so it explains our recent results [12, 13] that loud noises are judged equally *loud* when their *PNdB* values are nearly the same. It would be interesting to discover whether a larger distinction is made between loudness and noisiness or disturbingness judgements with sounds of lower intensity; so far as I am aware this has not been tested experimentally. Bowsher and I have carried out a pilot experiment [14] on the unpleasantness of noises, using 1300 observers, about which I will only say that it appears possible in principle to erect a scale of unpleasantness and that this would not be simply related to the scales of loudness or noisiness. (Fig. 2). If such a scale were developed, it would appear to fall into the same group as Kryter's.

The other scale would meet a need of a different kind. Brownsey [15] has suggested that a

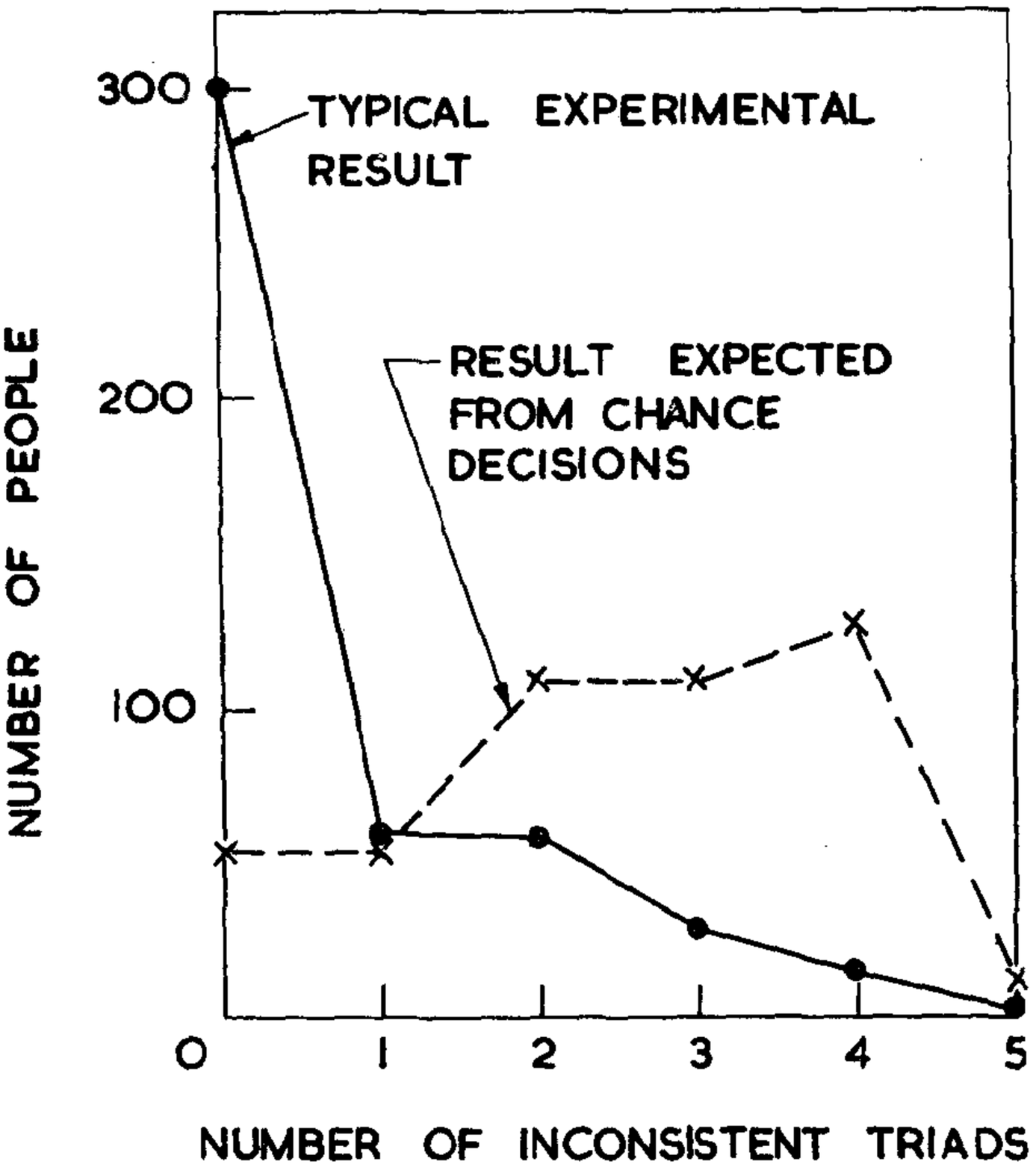


Fig. 2. Consistency check in paired-comparison tests on the unpleasantness of sounds.

distinction should be made between those scales in which the *magnitude* of some subjective quality is measured (like loudness) and those which involve a criterion of *acceptability*. Since it is impossible to observe "acceptability" directly, this may perhaps be interpreted in terms of the number and strength of complaints from people hearing the noise. One method of measuring noises in this sense already exists; its latest form is the draft I.S.O. proposal for Noise Rating Numbers [16].

Experimental methods

In this general lecture it is impracticable to catalogue, much less to comment on, the many variants of experimental method in subjective acoustics. I shall confine myself to some general observations except for one case which, as I believe it to be new, is given in more detail.

Broadly speaking, there is the choice of taking the subject to the scene of the test, or of bringing the experiment to the subject's own habitat.

There are obvious advantages of working in the laboratory; the sounds can be produced under defined acoustical conditions, accurately measured, and presented to the listeners in convenient sequence with no uncertainty as to what they hear. But when the sounds of interest are drawn from real life, the somewhat artificial conditions may influence people's judgements. Experimental investigations on this point seem to be lacking but it is wise to treat results with caution. At the N.P.L. we make use of an anechoic room for detailed studies with one listener at a time, mainly on the loudness of sounds; for subjective experiments with recorded noises in which we want to use a jury of several people at a time, slight variations in the sound field at different positions being tolerated, we make use of more conventional rooms, usually large ones. We prefer to avoid the use of recorded and reproduced sounds altogether when this is feasible, but when recordings are unavoidable we restrict the judgements to the comparison of sounds which are of similar type or which can be assumed to evoke similar associations for the listener. Thus, we have compared in this way the loudness and disturbingness of the noises of helicopters and fixed-wing aircraft, using the method of paired comparisons.

Stereophonic sound to reduce the lack of realism in laboratory tests with reproduced sounds seems an obvious improvement, though here again there is a lack of experimental evidence in support. It may be that the gain is outweighed by the extra complication and restriction of the useful sound field; moreover the addition of stereophony is unlikely to make up for the more serious loss of visual perception. In this connection I may mention our own experience in attempting to rate recorded motor vehicle noise by means of a category scale. Out of doors, in the presence of actual traffic, listeners have no trouble with this task; but with the noises recorded and played back in a sound-absorbent room our judges did not seem to be able to make up their minds at all.

In complete contrast to laboratory tests stands the opinion poll, or social survey [17]. Here the normal environment of the respondents is practically undisturbed by the intervention of the interviewer, so it stands to reason that the results have the best chance of being *valid*. On the other hand it is extremely difficult to make them *precise*. For example, in an opinion poll on the disturbing effect of aircraft noise there are formidable technical and statistical

problems in relating the opinions to actual numbers and levels of noises experienced, except in broad terms. Considerable use of the opinion poll method is being made at the present time in Great Britain but results have not yet been published.

We have often felt the advisability of a compromise between the extremes of the laboratory study and the opinion poll—a middle way whereby some of the advantages of both can be enjoyed. The work on motor vehicle noise which I have just referred to is an example of this approach [8, 18], our method being to gather together a group of people to listen to live noises as they actually occur, mostly out of doors. We have done the same thing with aircraft noise [19]. Certain standard procedures such as the method of constant-stimuli [20] or paired-comparisons, which would be used for preference, are unfortunately ruled out for practical reasons and the most convenient experimental tool in the circumstances appears to be some form of category scaling. In a sense the results of such tests are neither valid nor precise but they are possibly open to less criticism than other methods on one or other of these scores. Reliability (in the sense of repeatability) can easily be tested as part of the experimental design, and turns out to be very satisfactory. Moreover, it is reasonable to argue that the results actually are valid for the particular circumstances of listening, and since it is fairly easy to define these circumstances in their main particulars it becomes possible to compare results from different places or times, or with different classes of noise, directly. Our experience with category scaling in this kind of situation has shown that trends and comparisons are shown up very clearly; on the absolute interpretation of the judgements we are obliged to adopt a more cautious position.

The majority of the experimental investigations reported in the acoustical, as opposed to the psychological, literature are based on subjective methods which may be described as direct, that is the subject makes conscious appraisals about the magnitude or quality under study. What mental processes he undergoes to do this, especially in the more difficult cases like matching the loudness of radically different sounds or deciding whether one sound is less or more than 4 times as loud as another, is really immaterial. No assumptions are involved, not even the question whether the task required of the subject is "possible" or "meaningful", since the desired result emerges straight from the experimental judgements and we agree in advance to bar results which do not stand up to consistency tests. There is a school of opinion which objects to introspective decisions on principle; they would substitute indirect methods based on the observation of behaviour. In a discussion of this point [21] Stevens quotes, as one of their arguments, that introspective judgements are difficult and therefore the decisions are likely to be wrong. On the other hand the point is not really whether the judgement is difficult but whether it is possible, for if not it would cast doubt on the question whether the apparently more "objective" method of observing behaviour were really testing the same thing. In practice it seems best to adapt the psychophysical method to the nature of the investigation, using direct or indirect methods freely according to the need. For example, in a study of the interfering effect of noise on telephone communication [22] the choice would be some form of articulation scoring or the time to get an unpredictable message across—indirect methods—rather than direct appraisal of the degree of "talking difficulty"; the same may be said of studies of the effect of noise on the performance of working tasks, in which it is obviously most relevant to count the mistakes made in the performance of the task itself. This kind of investigation has been extensively

exploited by Broadbent and his associates, both in the field [23] and the laboratory [24, 25]. When the difficulty of direct appraisals reaches the point of refusal on the part of the subject, recourse must be taken to indirect methods. A problem in which we are greatly interested at the present time is to find what relation there is between the frequency of occurrence of occasional noises (like passing aircraft) and the annoyance they produce. Judgement by paired-comparisons is very difficult to arrange with realistic occurrence rates of, say, two an hour ranging to perhaps thirty an hour, because of the long duration of the comparison "stimuli". One might adopt a category scale, but there is a strong probability that people's memory would invalidate their second and subsequent judgements unless the experiment was spread out with many days or even weeks between successive judgements. The alternative of using a group of listeners once only would be extremely inconvenient. The case seems an appropriate one in which to use an indirect observation of behaviour as the index, but up to now we have not found a suitable criterion which is free from the objection that it may test something quite different from annoyance. This seems also to be the drawback of various methods which make use of physiological observations, for example the work of Lehmann [26] on the change of finger pulse volume on exposure to noise. A further objection is that these indirect methods do not hold out much hope of results with confidence limits narrow enough to be of predictive value in practical noise control.

We have concentrated in our own work on various forms of direct psychophysical method in which we have a greater confidence.

When appropriate we prefer the method of paired-comparisons because of the ease of

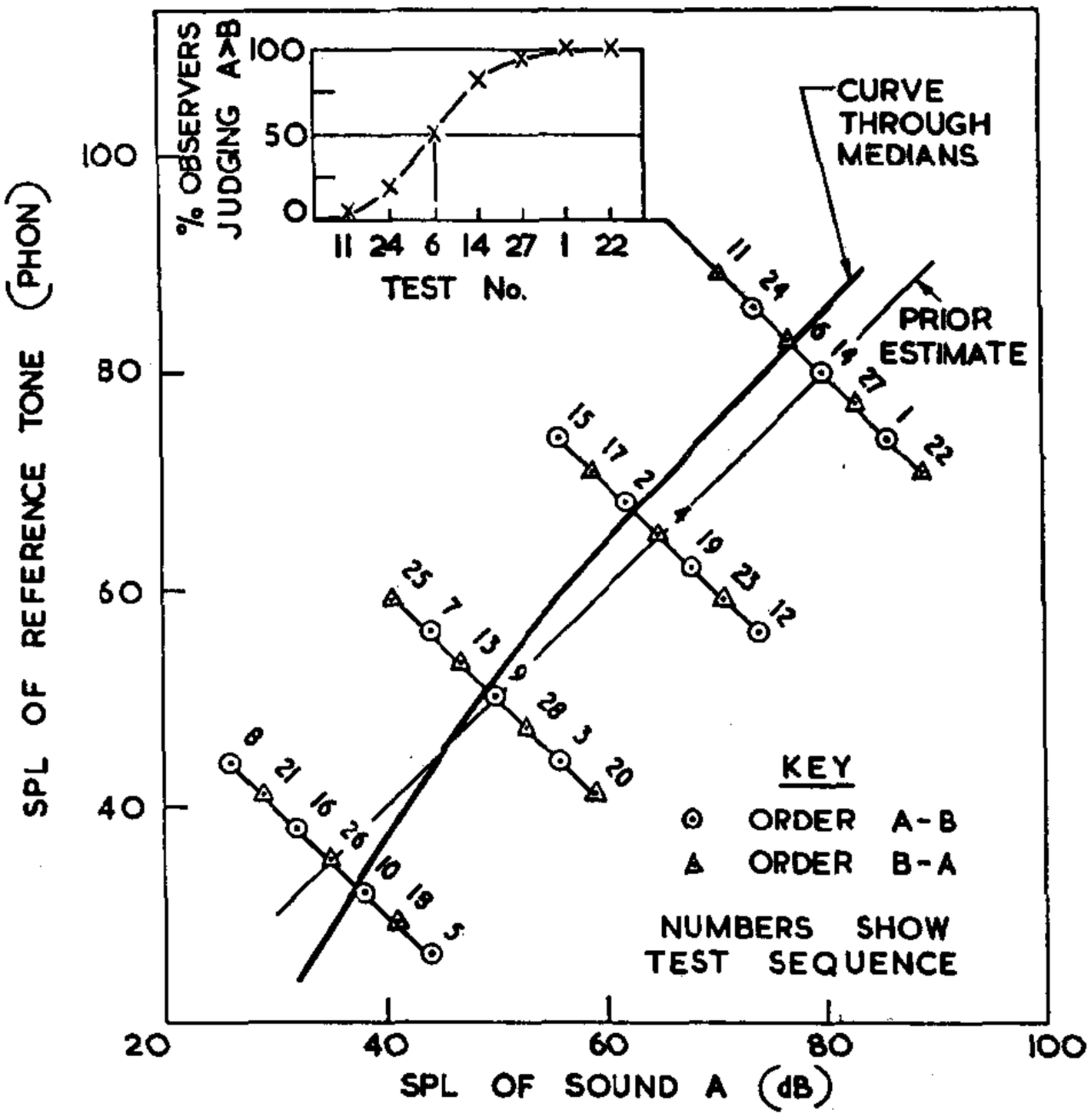


Fig. 3. Schema for a determination of loudness level.

scoring and the convenient built-in system of consistency-triad checks. However, the number of permutations increases rapidly with the number of items for comparison and to get the full advantage of the method a complete permutation is needed. Therefore this method is especially suitable for experiments with large juries (assuming it is difficult to recall them for serial tests) and for comparing a small number of noises. For example, this has been our method of choice in the comparison of aircraft noises in experiments designed to test the reliability of the PNdB and phon scales [12, 13]. It is not appropriate for detailed studies of the loudness of sounds, for example on the relation between the phon and sone scales, or on the relation between loudness level and sound pressure level. In these cases we prefer to work with one observer at a time in a controlled acoustical environment, usually the anechoic room, presenting a large number of sounds for judgement, often at fine intervals of sound pressure etc. For this type of test we prefer the method of constant-stimuli with a randomized pattern of presentation both as regards the order of sounds within pairs and as to level.

An adaptation of this method which we have found very convenient for obtaining a large amount of information at a given test session is illustrated in Fig. 3. A justifiable criticism of the constant-stimulus procedure in its rigorous form, whereby a reference sound is repeated every time a judgement is made, is that it is very time-consuming. This objection can be overcome by the scheme shown in the Fig., which illustrates a fictitious determination of the relation between sound pressure level and loudness level of a sound *A* of given character. The relation is determined over a wide range of absolute levels at a single session of subjective testing lasting only a few minutes. All subjects are presented with the test sounds in the same sequence but alternate ones hear it with the internal order in each pair reversed; this is to cancel out any order effect. To determine the equality relation for sounds *A* and *B* for each individual is not necessary, and the prolonged testing necessary is the main drawback of the standard constant-stimulus method. In the modified procedure, a point representing the group equality judgement is located along each of the four diagonal lines in the pattern, by plotting the cumulative fraction of observers judging *A* louder than *B* and deducing the median as shown in the sketch at the top left hand corner. The success of the method depends on careful preselection of the principal diagonal of the pattern (that is, on an intelligent guess at the answer) and on judicious spacing of the test levels (unanimous judgements yield no information on the location of the median). The method has the advantage that the subject is kept very much "on his toes", not knowing in which order or at what level the next sounds may come. It avoids his making use of false cues; it is efficient in time-saving, and as far as we can see free from biassing effects. The method has been used to study the loudness level of bands of noise, mainly octave bands [27], in connection with the work of I.S.O.

We have not made much use of the method of adjustment, in which partial or total control of the order of presentation and of the level of the comparison stimuli is given over to the subject himself. Of course, this makes the experimenter's task a lot lighter but there seem to be some objections to it on the score of reliability. Stevens [28] has more than once referred to a biassing effect in which the result depends on which sound is adjustable in level; secondly instead of requiring only inequality judgements as in the constant-stimulus method, the subject has to strive, by bracketing, towards a point of equality. It has been

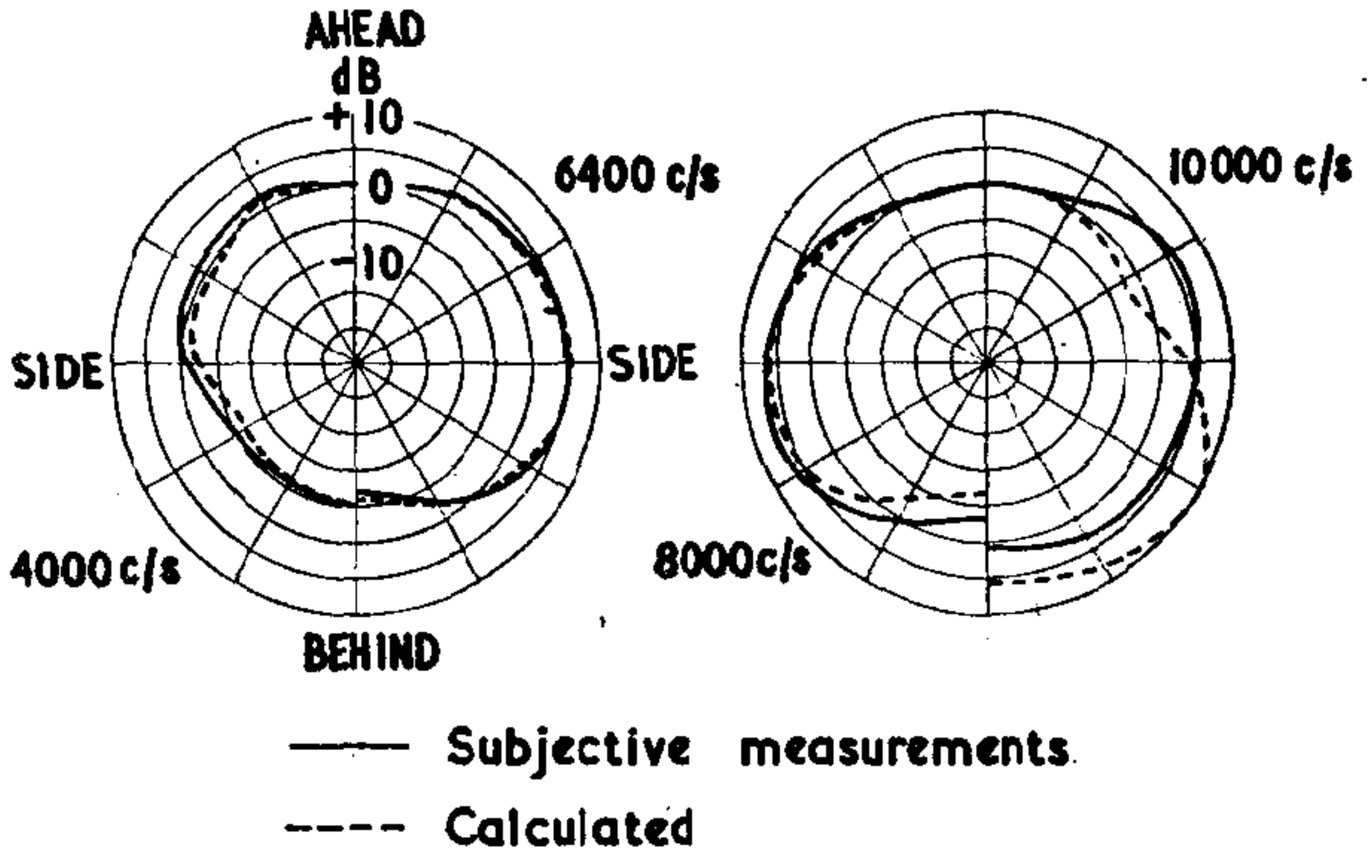


Fig. 4. Comparison of calculated and subjectively measured equal-loudness contours for a source of sound rotating in the horizontal plane.

noticeable in our experience that some observers are dissatisfied with their first approach to the target and seem to try too hard to improve on it, finishing up far away from the "proper" place (it is possible to trace this kind of mistake by means of consistency tests).

The present state of loudness theory

There is considerable experimental evidence that judgements of loudness are little affected by the context in which they are carried out, apart from such specific effects as auditory fatigue. As examples of this I may quote the intercomparison of tones of 33, 200 and 1000 c/s, 1000, 3000 and 10000 c/s, and a number of other triads which comprise a series of

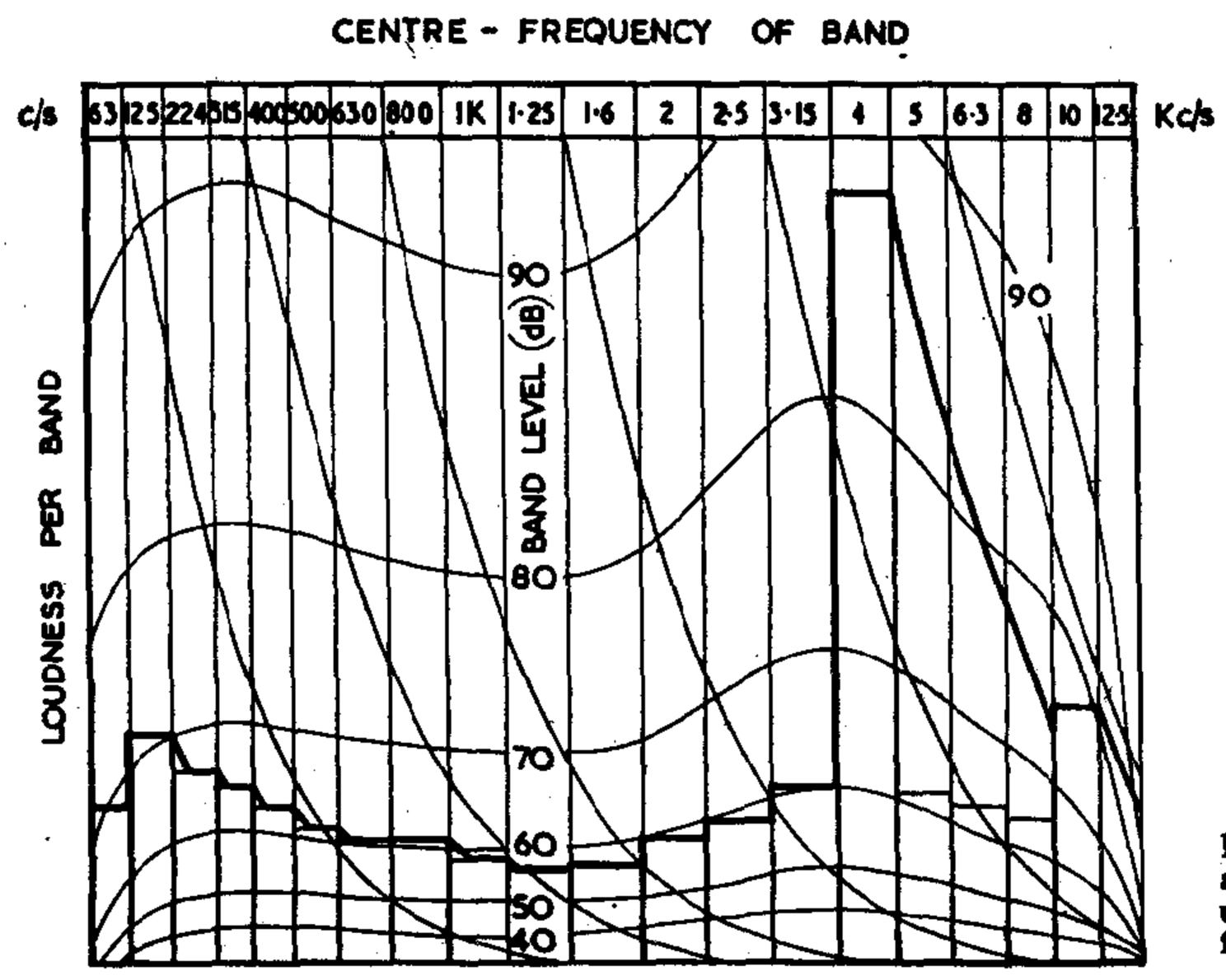


Fig. 5. Diagram illustrating the application of Zwicker's method to the calculation of the loudness of the noise from a multi-stage compressor.

rather severe consistency checks which were satisfied within one or two decibels by various groups of listeners [29]. Similar evidence comes from 2-fold and 10-fold loudness ratio tests; the 10-fold step ought to be 3.3 times as large as the 2-fold step when measured in phons; the experimental values were about 13 and 43, in very good agreement with each other [30]. There is also evidence that loudness is governed solely by the sound pressure received at the listener's ears; this takes two forms. The loudness of a pure-tone from a source in one direction in a free field has been compared with that heard from in front. The sound pressures at the two ears were measured at the same time (they differ because of the head shadow on the remote side and due to diffraction of the sound field); the relative loudnesses were in satisfactory agreement with the values calculated from the sound measurements, after making allowance for the loudness gain due to binaural bearing (Fig. 4). Exactly the same kind of comparison was carried out between a diffuse sound field and a frontally-incident plane wave [32], with the same result, namely that the relative loudness was predictable solely from the changes of sound pressure at the ears in the two sound fields.

For an ideal calculation of loudness, three data concerning the sound are required, the sound pressure level spectrum, the direction of arrival of sound at the listener's position, and the temporal variations if any. With the development by Zwicker [5, 6] of a theory realistically based on the principles of auditory masking and critical bands, the aspect of spectral distribution, at any rate for sounds that are continuous, is well-understood. A practical question remains, whether the complexity of the calculation procedure involved is excessive for some purposes and whether one of the simplified procedures based on octave-band analysis—for example the so-called Stevens Mk. VI method [33]—is not adequate. A recommendation embodying the Stevens method or the Zwicker method or both is at present under consideration by the I.S.O. As regards the accuracy of these procedures some data are presented below in the discussion on perceived noise level; the general conclusions favour Zwicker's method. The principle of the latter is the association, with each of 20 frequency bands, of a partial loudness represented in a diagram as the area of a rectangle and a curvilinear triangle. The triangle takes care of the masking effect on sounds of higher frequency. The total loudness is given by the area included under the stepped diagram so formed (see Fig. 5) and is, of course, less than the sum of the partial loudnesses by an amount depending on the relative levels in adjacent bands. The diagram also shows directly which frequency band makes the major contribution and which bands are totally or partially masked, thus serving as a kind of "subjective spectrum". The labour involved in the graphical method undoubtedly limits the use of Zwicker's method, so we have reduced it, in collaboration with the Mathematics Division of the N.P.L., to a digital computer programme [34], making only slight modifications. In this form it is very convenient for straightforward calculations of the total loudness of sounds.

Directional effects in loudness perception have been rather neglected until lately, and indeed they turn out to be of little practical importance where broad-band noises are concerned. Nevertheless in certain cases the direction of arrival should be taken into account, for example a sound of 8 kc/s is 8 phon louder overhead as compared with directly ahead of the hearer [31]. As between frontally-incident and diffuse sound, there is no more than about 5 phon difference at any frequency. Considerable experimental ingenuity has gone into the determination of this comparison, which is illustrated in Fig. 6. [32, 35, 36, 37].

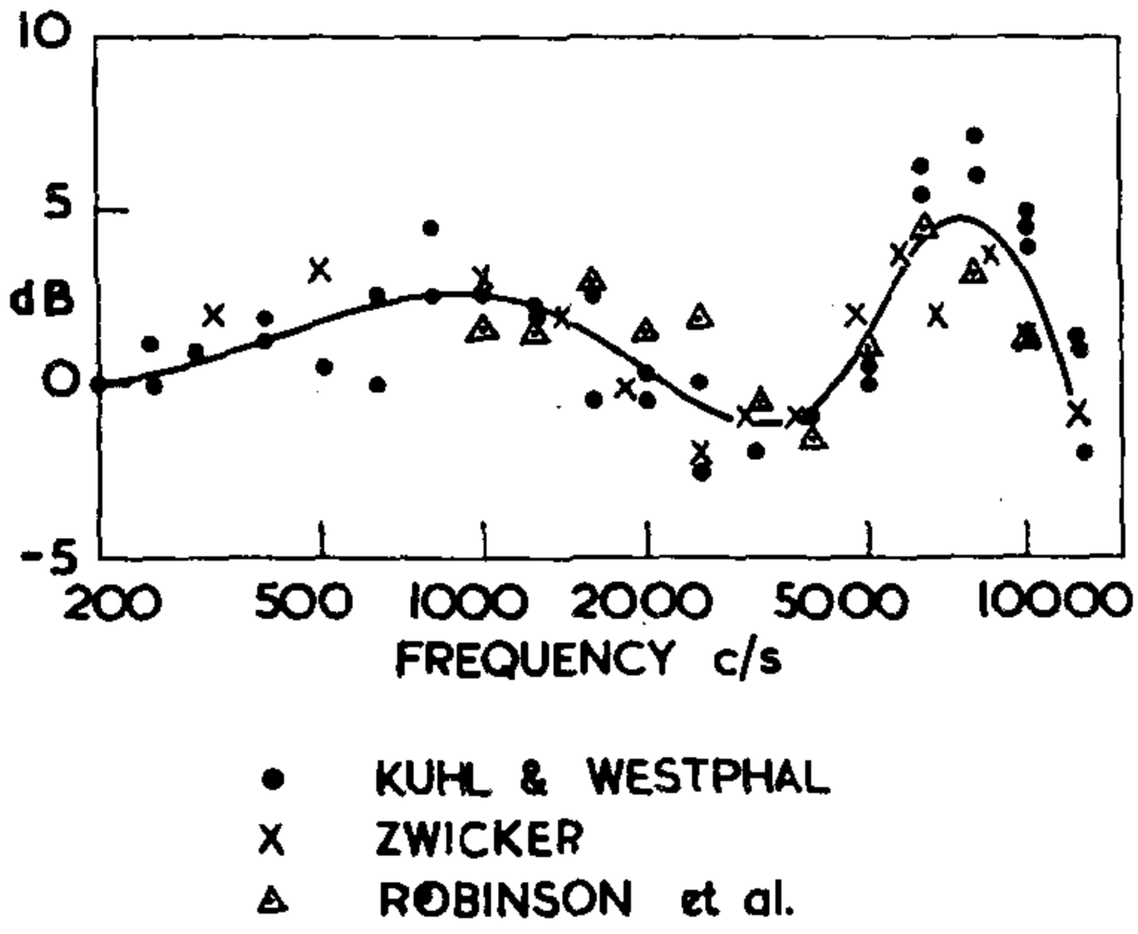


Fig. 6. Difference between sound pressure levels, for equal loudness, of frontally-incident and diffuse sound.

The least satisfactory part of loudness theory concerns time-dependent sounds. Three cases must be distinguished. The first comprises single, transient sounds of short duration on which the experimental data remain inconclusive in spite of the early date of Steudel's original work [38]. The subjective assessment of noises in fields as diverse as sonic bangs, electrical switchgear and drop-forging is hampered by this. There are indications that loudness depends more on the shape of the wave-form than on the spectral distribution, a way of saying that the ear is phase-sensitive to transient sounds, but there do not appear to be any reported attempts to confirm Steudel's conclusion that the loudness is related to the maximum impulse within any 300 μsec period. The second case concerns sounds of limited duration but constant character, such as single bursts of tone of given sound pressure amplitude and frequency, lasting from a few milliseconds up to a fraction of a second. Sounds of this type were studied by Munson some years ago [39], with the main result that full loudness sensation is reached after about $\frac{1}{4}$ sec. For shorter bursts (5–10 ms), Munson's experimental results support the deduction that the loudness level increases about 6 phon on doubling the duration; on the other hand his mathematical model yields loudness in sones initially proportional to time, which is equivalent to a 10 phon increment for double duration. These values both contrast with results extracted from Pollack's work [40] which indicate energy summation, or only a 3 dB increment for double duration up to quite long times of the order 100 ms. The third case consists of quasi-continuous sounds, i.e. trains of short-lived sounds separated by intervals of partial or total inactivity, or modulated sounds or noises. From the physical standpoint such sounds might be treated as continuous by means of a spectrum representation, and at first sight this appears as an attractive extension to the existing theory of loudness for continuous sounds. However, the latter only requires a knowledge of the power spectrum, a representation in which all phase information is discarded. It may well be that a knowledge of the phase spectrum is not a necessary ingredient in the calculation of the loudness of interrupted sounds whose repetition rates are appreciably greater than the reciprocal of the integration time of the ear, although I am not aware that anyone has yet systematically compared such simple cases as the loudness of the pair of waves

$$\left(\sin x + \frac{1}{3}\sin 3x + \frac{1}{5}\sin 5x + \ldots\ldots\right) \text{ and } \left(\cos x + \frac{1}{3}\cos 3x + \frac{1}{5}\cos 5x + \ldots\ldots\right),$$

for example. On the other hand, it seems almost self-evident that the perception of sounds having repetition rates appreciably lower than the reciprocal of the integration time of the ear must depend on the actual shape of the modulation envelope, and to represent this would entail the complex Fourier spectrum. The representation of such sounds in the frequency domain obviously presents difficulties for loudness calculation, and it is more convenient to treat them in terms of duration and repetition rate considered as separate variables. In this way the loudness can be expressed in terms of the loudness of the related (unmodulated) sound having a 100% duty cycle (which, being continuous, can be evaluated from the simple power spectrum) plus a correction which is a function of repetition rate and duty cycle (or "burst-time fraction" as some authors call it). Owing to the numerous possible combinations of parameters (waveform within the pulse, pulse length, repetition rate, departures from true periodicity) there remains scope for much further research. Published results so far available suggest, in brief, that for repetition rates below about 50 c/s and continuing down to about 5 c/s, the loudness tends to be greater than is suggested by the long-term r.m.s. sound pressure alone, by up to 10 phon. Data by Pollack [40] and Niese [41] on this aspect are in fair agreement, though it is doubtful whether a convincing generalized theory for sounds with an arbitrary time variation could yet be written.

Perceived Noise Level

I have already mentioned the origin of this scale, in attempts to reconcile judgements of the noisiness of jet-engined and piston-engined aircraft [10]. The calculation of perceived noise level proceeds analogously to the calculation of loudness level by Stevens' method [33] and differs only in that greater weight is given to the contributions of the higher octave-bands of the spectrum. This implies, among other things, that noisiness possesses an additive property in the same sense as loudness, namely, that the contributions due to different parts of the spectrum can be assumed to sum arithmetically after making an empirical allowance for masking between bands. These similarities throw some doubt on whether noisiness and loudness are entirely distinct qualities. We have carried out two experiments with recorded aircraft noises using altogether more than 2000 observers, half of them judging in terms of "loudness" and half in terms of "disturbingness" [12, 13]. The results are summarized in Fig. 7, showing, firstly, that it made little difference in these comparisons which of the two criteria was in force. The Fig. also compares the effectiveness of five different scales in predicting the loudness and disturbingness matches. The loudness levels calculated by Stevens' and by Zwicker's methods, the perceived noise level in PNdB, the overall sound pressure level and the *A*-weighted sound level [42] are shown for each noise at the level which makes the noise as loud (or disturbing, as the case may be) as the other noises in the same section of the Figure. The most successful scale is that which gives a constant value in each section, and it may be seen from the spread shown at the right hand side of each group of results that Zwicker's method is superior. This is the case whether one looks at the equally

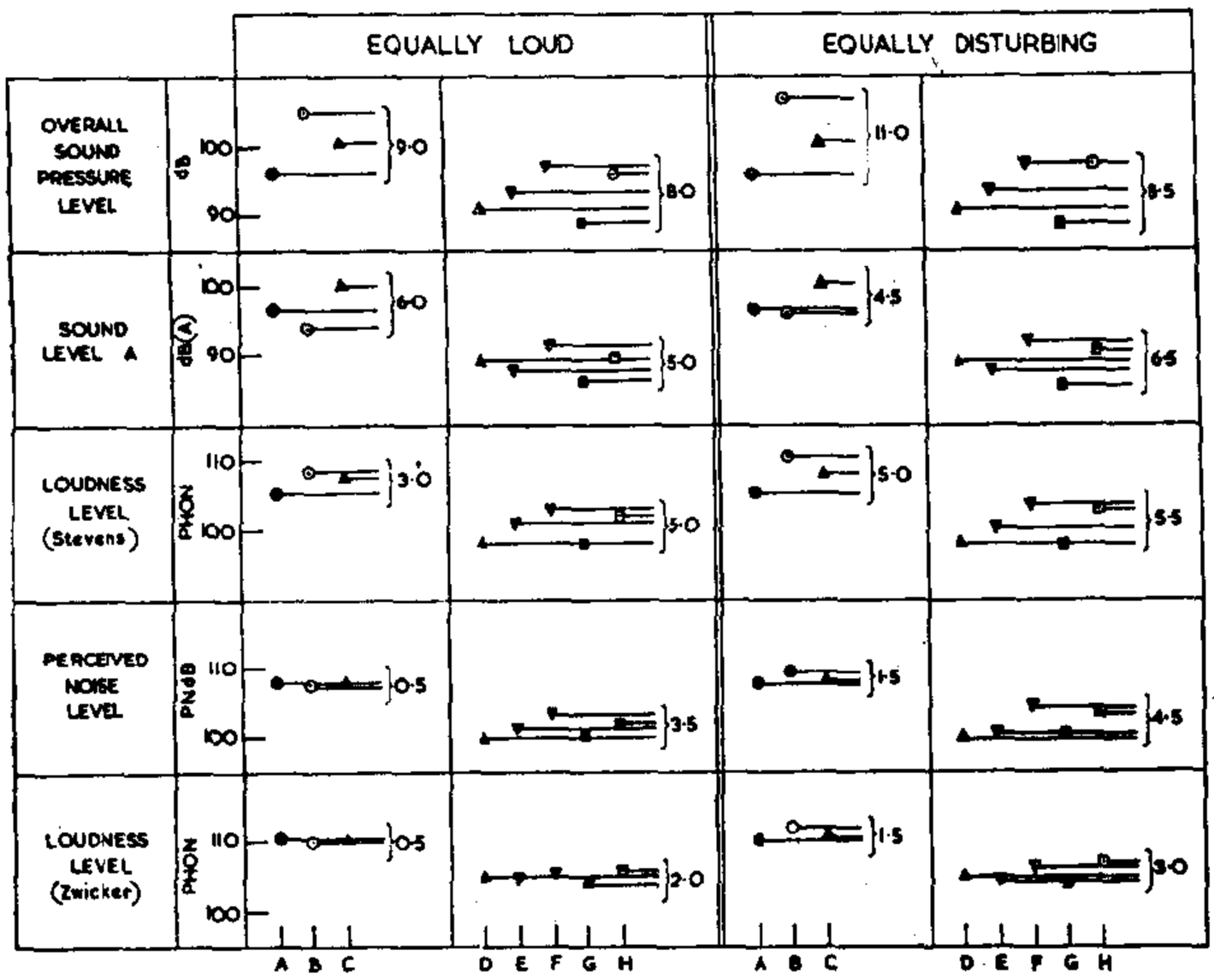


Fig. 7. Results of equal-loudness and equal-disturbingness judgements by 2000 listeners, showing 5 scales of measurement in ascending order of accuracy of prediction of the subjective comparisons.

Group I: A — Boeing 707/120 jet airliner at take-off.
B — Lockheed Super-Constellation at take-off.
C — Avro Vulcan Mk. I (Rolls-Royce Avons), fly-past.

Group II: D — Boeing 707/120 jet airliner at take-off.
E — Bristol helicopter type 171.
F — Westland helicopter type S-55.
G — Fairey Rotodyne at take-off with rotor-tip jets.
H — Fairey Rotodyne, flypast with turboprop engines.

Note: the levels refer to the sound reproduced in the listening room. Groups I and II are separate, unrelated, experiments.

loud or the equally disturbing noise levels, though it might have been expected that perceived noise level was the best predictor in the latter case.

Since these experiments were completed, a paper by Little has appeared [43] proposing a modification to the formula for perceived noise level for so-called "spiked noises", that is noises with a line spectrum superimposed on a random noise spectrum. It is of great current interest in relation to ducted-fan aero engines which emit noise of this character. According to Little an upward correction must be made based on the protrusion of a spectrum line

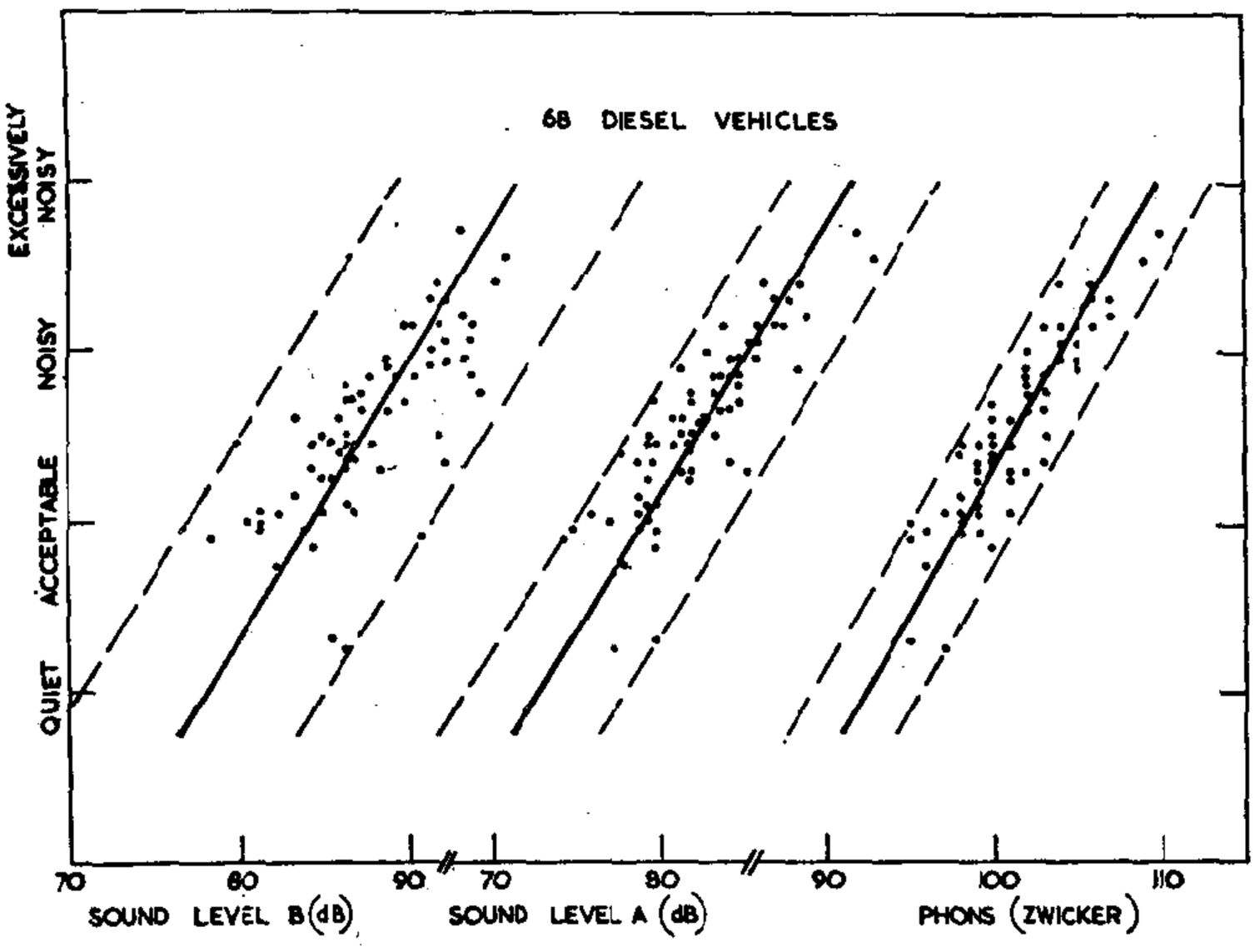


Fig. 8. Roadside judgements of the noisiness of 68 diesel vehicles compared with the maximum values of sound level B, sound level A, and loudness level (Zwicker phons), showing that the correlation depends on the scale of measurement.

above the surrounding random spectrum level reckoned in one-twenty-fourth octave bandwidths. This finding does not appear to be supported by the work of Kryter and Pearsons [44] who remark, however, that pure-tone components probably have a considerable effect on the judged noisiness of an otherwise continuous spectrum. I had hoped to report on our own contribution to this problem, which took the form of a subjective experiment in which several hundred members of the public took part. Unfortunately the analysis of this rather complicated series of comparisons is still going on, but enough has already emerged to suggest that the rather clear-cut indications of Fig. 7 do not apply with equal force to noises with mixed line and random-noise spectra.

Some experiments with category scales

I have set out the arguments which led us to carry out a number of investigations with live noises by means of category scales. The first of these [8] related to motor vehicle noise; a jury of 19 judged the noisiness of individual vehicles passing along a main road on a scale of six points, four named and the other two unspecified extremes. The purpose was to study the correlation between the mean judgement and various objective measures of the maximum noise level to help the work of I.S.O. in selecting a meter to be recommended. Clear indications emerged that *A*-weighted sound level was preferable to *B*-weighted from this standpoint (see Fig. 8), a result confirmed by Weber and Lauber [9] in similar tests. A

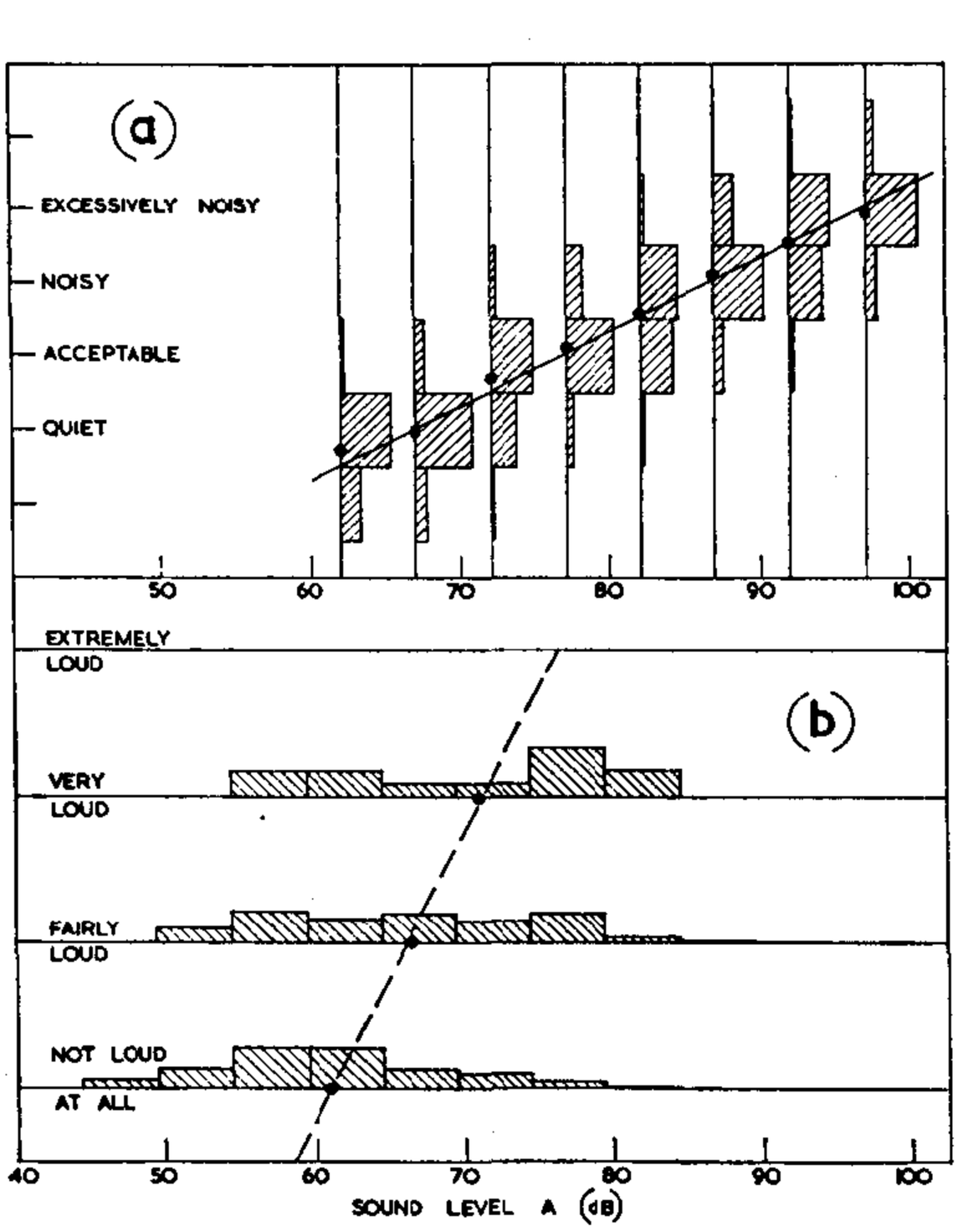


Fig. 9. Comparison of judgements of motor vehicle noise (a) and of the background traffic noise in residential streets (b).

still better correlation exists between the judgements and loudness levels calculated by Zwicker's method, but the latter is not suitable for routine field measurements.

Later we carried out a larger-scale experiment with 57 subjects [18] on the same lines with results that were surprisingly consistent with the first. This is shown by several distinctive features of the regression lines in the correlation diagrams. Firstly, the absolute levels were repeated within 1 or 2 dB for each class of vehicle (cars, motor-cycles and diesels) though with quite a different group of subjects. Secondly the rates of growth of "noisiness" with sound level for the three vehicle classes showed characteristic and repeatable differences from one another. There were also close resemblances in the two investigations as regards the rather small effects of age and sex classification. Such correspondences lend confidence to the category scaling method, at least for the prediction of trends. As already discussed, it is quite another question whether the various levels of noisiness can be read off the correlation diagrams in terms of corresponding sound levels and the latter applied to situations differing from those of the tests.

An example of this is seen when one compares the motor vehicle noisiness judgements with ostensibly similar data obtained by Parkin [45] (not yet published). Parkin's observers were a 10% sample of the respondents in an opinion poll, revisited at their front doors and asked to judge the prevailing street noise on a category scale of four points. It happened that the noise so judged was largely due to motor traffic. His results are shown in histogram form in Fig. 9(b), the abscissa showing the highest value of sound level (A) recorded during the

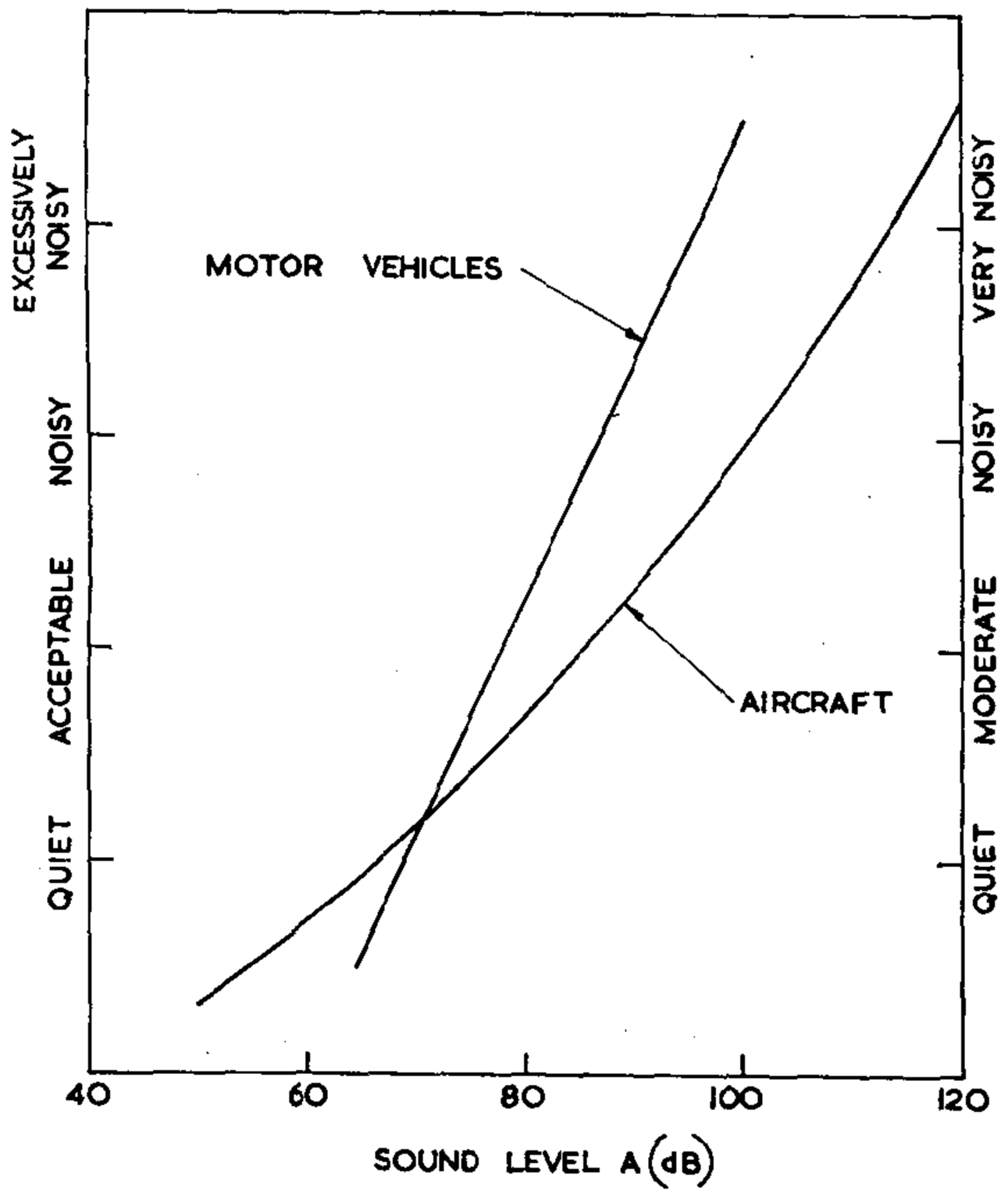


Fig. 10. Comparison of noisiness judgements on motor vehicles and aircraft.

quarter-minute or so of the test. The category scales were unfortunately different, but despite the considerable dispersion no reasonable alignment can be made which reconciles the results. We seem forced to the conclusion, therefore, that people's attitude to the *noisiness of a vehicle* is markedly different (in the more tolerant direction) from their attitude towards the *background noise* in residential streets, even though that is also composed of vehicle noise.

We have extended the category scaling to aircraft noise [19] making use of the annual air display at Farnborough which provides a well-mixed assortment of aeronautical noises. One part of the experiment was modelled closely on the motor vehicle tests so that we could determine whether people's attitudes to these two sorts of noise are comparable. The result is clearly shown in Fig. 10, in which for simplicity a mean line is shown for all motor vehicles; a noise of around 65 dB (*A*), whether from a car or an aeroplane, was generally called "quiet"; but a noise of 90 dB (*A*) was called "excessively noisy" if due to a road vehicle but only "moderate" to "noisy" if from an aircraft. With the same 60 subjects who judged these aircraft noises we also made two other tests. In one of these they were out of doors engaged in a miscellany of occupations, reading, knitting, playing cards, and so on, and they had to judge aircraft noises on a scale of "intrusiveness" and "annoyance" with six points, on a cue from the experimenter after the passage of selected aircraft. In comparison with the results of the "noisiness" judgements which, in contradistinction, engaged

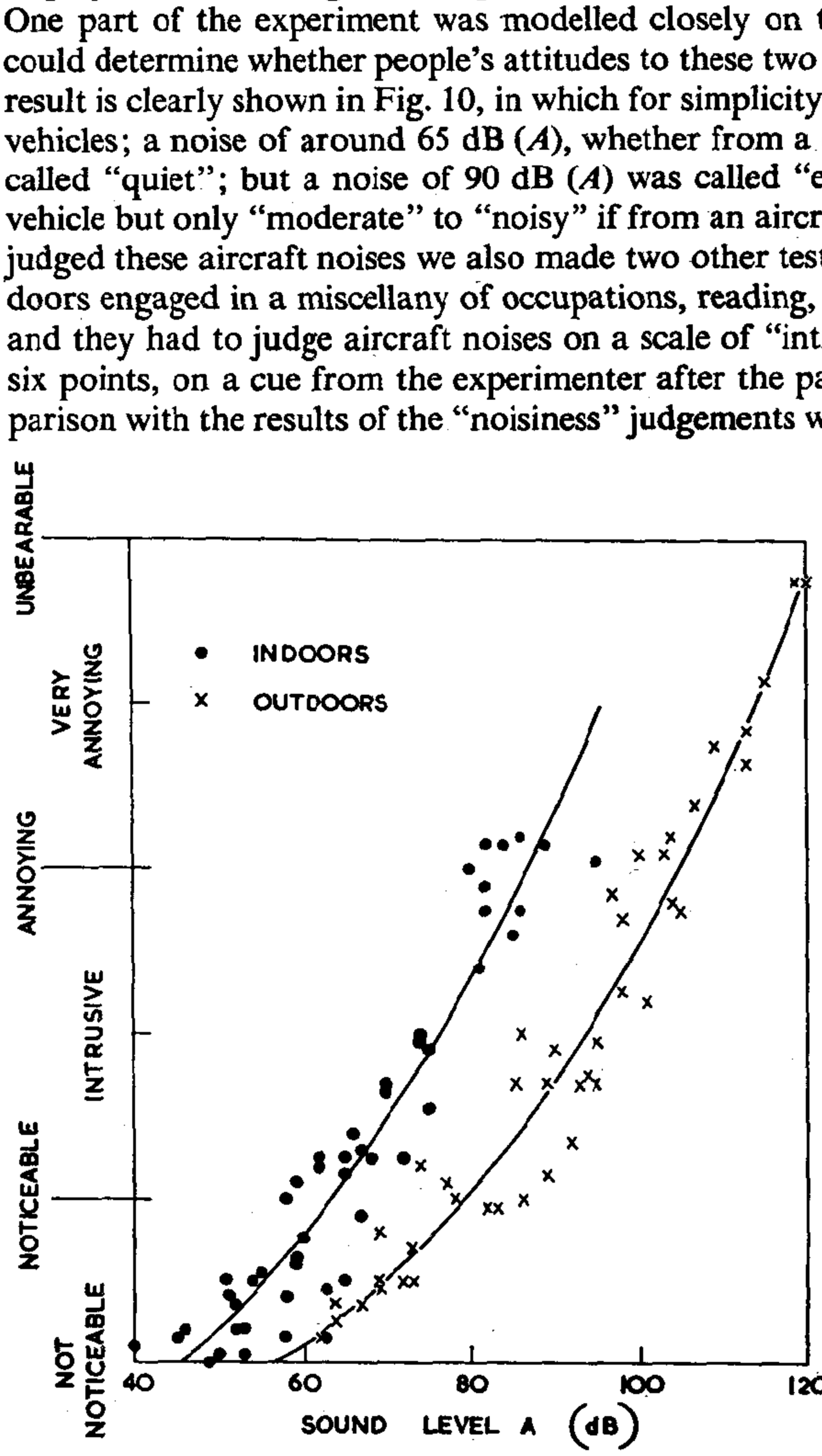


Fig. 11. Comparison of intrusiveness judgements of aircraft noise made out of doors and indoors by the same group of 60 people.

the subjects' whole attention, "noticeable" came above "quiet" but below "moderate in noisiness", "intrusive" corresponded to a level between "moderate" and "noisy", and (perhaps more significant) "very annoying" corresponded almost exactly to "very noisy". There was a third part to our aircraft noise tests, in which the intrusiveness categories were applied to live aircraft noises heard indoors, the "task" (if one may use the term) being to watch a film show. Expressing the results in terms of the sound level received by the listeners, a sound heard indoors was judged much more severely than one of the same sound level outdoors (by about 18 dB (*A*) on the average) as shown on Fig. 11. The interesting feature of this result is that the 18 dB (*A*) difference is much the same as the reduction of sound due to the building used. It might therefore be argued that people recognized the various aircraft as such and judged accordingly; expressed otherwise, when they were indoors they imagined what the source of noise would sound like outside and judged on that. There are other possible interpretations of the results and we are presently carrying out a programme of further tests to distinguish between them. One such test, suggested by the present data, is to see whether the relation between outdoor and indoor judgements is the same for meaningless noise sources. Another is to test whether the shift of judgement from outdoors to indoors is actually connected with people's experience of the sound reduction of buildings, or whether the experimental result suggesting this is a mere coincidence.
Representation of the "intrusiveness" and "noisiness" categories as equal steps of a scale, as in the correlation diagrams (Figs. 10 and 11) is, of course, only a convention. By using

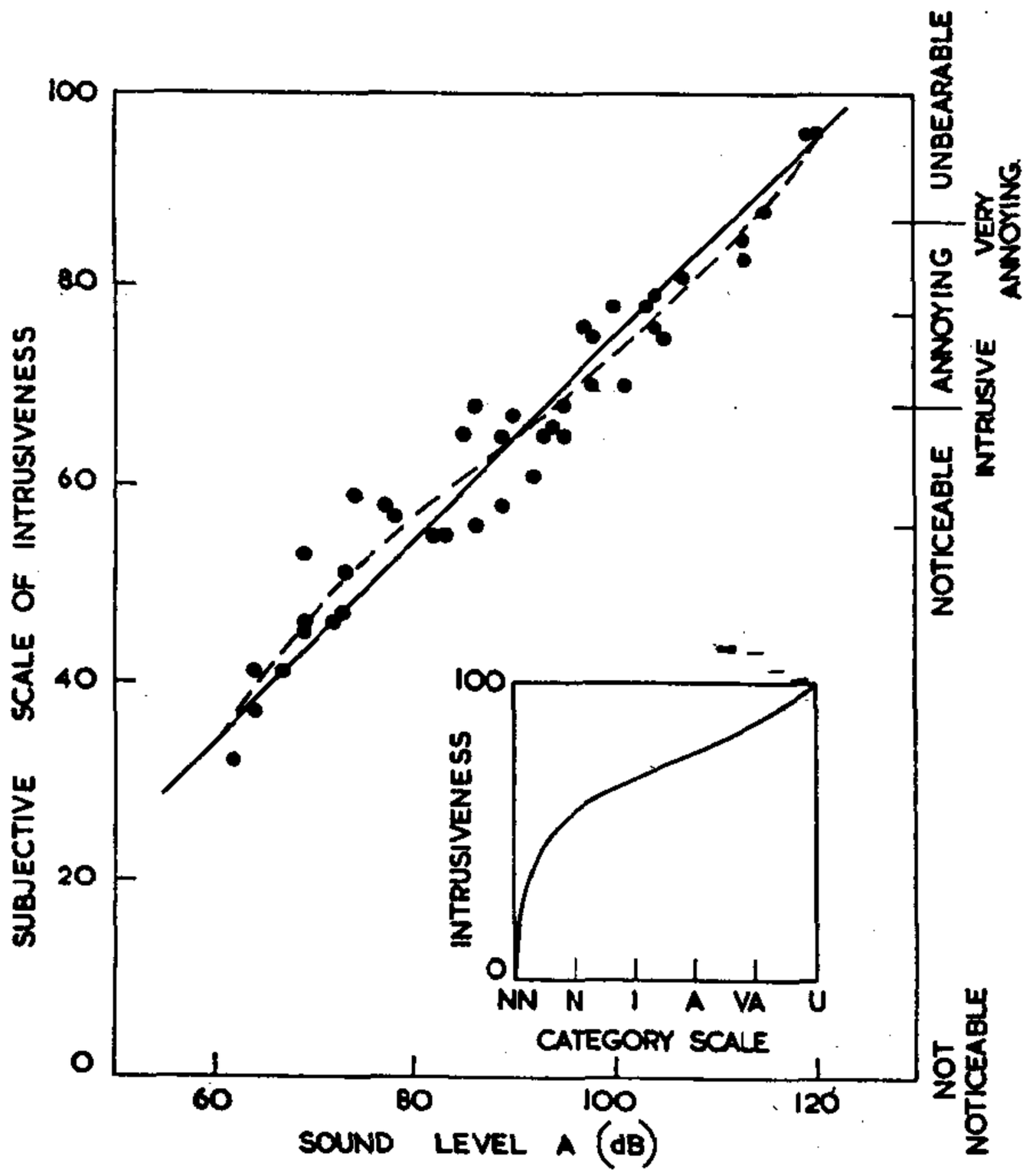


Fig. 12. Judgements of the intrusiveness of aircraft noise transformed into a scale with constant discriminal dispersion.

Thurstone's unfashionable principle of equal discriminal dispersion [46] the ordinate can be transformed to a scale in which equal intervals represent equal subjective steps. This is shown, for the example of the outdoor "intrusiveness" data given on Fig. 11, in the next Figure. The inset diagram on Fig. 12 shows the relevant scale transformation derived from the standard deviations, while the main part of Fig. 12 illustrates how the experimental results appear in the new scale. The dotted curve is the transform of the parabolic regression line of Fig. 11. Instead of being strongly curved it is now more nearly straight; indeed the newly-calculated linear regression line also shown in Fig. 12 fits the data equally well. Thus the subjective magnitude of "intrusiveness" is roughly linear in the sound level and the same is true of "noisiness". This characteristic is also true of loudness measured in equal-ratio steps (log sones or phons), which suggests that the subjective attributes of loudness, intrusiveness and noisiness all grow in a very similar way with the intensity of noise.

Acknowledgements

Much of the work referred to and the results obtained form part of the general research programme of the National Physical Laboratory, and this paper is published by permission of the Director of the Laboratory.

REFERENCES

1. D. B. Fry, Nature, *165*, 950, (1950).
2. D. W. Robinson, Acustica, *7*, 217, (1957).
3. S. S. Stevens, J. Acoust. Soc. Amer., *29*, 603, (1957).
4. International Organization for Standardization, "Expression of the physical and subjective magnitudes of sound or noise", Recommendation ISO/R131, Geneva, 1959.
5. K. E. Zwicker, Acustica (Akust. Beihefte), *10*, 304, (1960).
6. K. E. Zwicker, Acustica (Akust. Beihefte), *8*, 237, (1958).
7. S. S. Stevens, Science, *127*, 383, (1958).
8. D. W. Robinson, W. C. Copeland and A. J. Rennie, Engineer, *211*, 493, (1961).
9. H. Weber and A. Lauber, Generaldirektion PTT (Switzerland), Forschungs-und Versuchsanstalt, Report 22637, 1959.
10. Bolt, Beranek and Newman Inc., "Studies of noise characteristics of the Comet 4 jet airliner and of large conventional propeller-driven airliners", Report for the Port of New York Authority, October, 1958.
11. K. D. Kryter, J. Acoust. Soc. Amer., *31*, 1415, (1959).
12. W. C. Copeland, I. M. Davidson, T. J. Hargest and D. W. Robinson, J. Roy. Aero. Soc., *64*, 33, (1960).
13. D. W. Robinson and J. M. Bowsher, J. Roy. Aero. Soc., *65*, 635, (1961).
14. J. M. Bowsher and D. W. Robinson, Brit. J. Appl. Phys., *13*, 179, (1962).
15. C. M. Brownsey, NPL Conference Proceedings, "The control of noise", p. 257, HMSO, London, 1962. (Contribution to discussion on a paper by D. W. Robinson).
16. C. W. Kosten and G. J. van Os, NPL Conference Proceedings, "The control of noise", p. 373, HMSO, London, 1962.
17. P. N. Borsky, "Human factors underlying community reactions to Air Force noise", W.A.D.D. Technical Report 60–689, 1961.
18. C. H. G. Mills and D. W. Robinson, Engineer, *211*, 1070, (1961).

19. D. W. Robinson, J. M. Bowsher, and W. C. Copeland, Acustica, *13*, (in the press).
20. J. P. Guilford, "Psychometric Methods", Mc-Graw Hill, New York, 1936.
21. S. S. Stevens, "Measurement, psychophysics and utility". Chapter 2 from "Measurement: definitions and theories", editors C. W. Churchman and P. Ratoosh, p. 60, John Wiley & Sons, New York, 1959.
22. D. L. Richards and J. Swaffield, J. Inst. Elect. Engrs., Part B, *106*, 77, (1959).
23. D. E. Broadbent and E. A. Little, Occup. Psychol., *34*, 133, (1960).
24. D. E. Broadbent, Quart. J. Expt. Psychol., *6*, 1, (1954).
25. D. E. Broadbent, Psychol. Rev., *64*, 205, (1957).
26. G. Lehmann, Int. J. Prophyl. Med. u. Sozialhyg., *2*, 1, (1958).
27. D. W. Robinson and L. S. Whittle, "The loudness of octave bands of noise", (to be published).
28. S. S. Stevens, J. Acoust. Soc. Amer., *27*, 815, (1955).
29. D. W. Robinson and R. S. Dadson, Brit. J. Appl. Phys., *7*, 166, (1956).
30. D. W. Robinson, Acustica, *3*, 344, (1953).
31. D. W. Robinson and L. S. Whittle, Acustica, *10*, 74, (1960).
32. D. W. Robinson, L. S. Whittle and J. M. Bowsher, Acustica, *11*, 397, (1961).
33. S. S. Stevens, J. Acoust. Soc. Amer., *33*, 1577, (1961).
34. NPL Conference Proceedings, "The control of noise", p. 407, HMSO, London, 1962.
35. W. Kuhl and W. Westphal, Acustica, *9*, 407, (1959).
36. G. Jahn, Hochfreq. u. Elektroakust., *69*, 75, (1960).
37. K. E. Zwicker, private communication, 1959.
38. U. Steudel, Hochfreq. u. Elektroakust., *41*, 116, (1933).
39. W. A. Munson, J. Acoust. Soc. Amer., *19*, 584, (1947).
40. I. Pollack, J. Acoust. Soc. Amer., *30*, 181, (1958).
41. H. Niese, Hochfreq. u. Elektroakust., *66*, 115, (1958).
42. International Electrotechnical Commission, "Recommendations for sound level meters", Publication 123, Geneva, 1961.
43. J. W. Little, Noise Control, p. 11, May/June, 1961.
44. K. D. Kryter and K. S. Pearsons, Sound, p. 24, March/April, 1962.
45. P. H. Parkin, private communication, 1962.
46. L. L. Thurstone, "The measurement of values", University of Chicago Press, 1959.

Einige physikalische Erscheinungen, die in hochintensiven Ultraschallfeldern entstehen

L. D. Rosenberg

Einleitung

Für die moderne Entwicklung der Schalltechnik ist ein Übergang nicht nur zur hohen Leistungen, sondern auch zu hohen Schallintensitäten kennzeichnend. Wie bekannt, liegen die von unserem Ohr normal wahrzunehmenden Schallintensitäten im Bereich von 10^{-16}–10^{-5} W/cm^2. Die industrielle Anwendung des Ultraschalls (Ultraschallreinigung -entgasung u.a.m.) macht die Schallintensitäten von einigen Watt pro cm^2 erforderlich, was um das 10^5 fache die lautesten Schalle übertrifft, die von unserem Ohr ohne Schmerz aufgefasst werden können. Es ist z.B. möglich, die Schallintensität von ca. 10^5 W/cm^2 zu erhalten. (1) Wenn die Intensität von 1 W/cm^2 dem Schalldruck von 1,7 at entspricht, so lassen sich unter Anwendung der modernen Technik Schalldrücke von hunderten at und Beschleunigungen, die 10 Millionen g. übertreffen, erhalten.

Bei so hohen Intensitäten lassen sich in Flüssigkeiten und teilweise in Gasen qualitativ neue Erscheinungen beobachten. Ich meine dabei nicht diejenigen Erscheinungen, die sich von sogenannten Approximationen zweiter Ordnung der nichtlinearen Akustik umschreiben lassen und die bereits verhältnismässig gut untersucht sind, wie z.B. die Änderung der Kurvenform der sich fortpflanzenden Schallwelle und die damit zusammenhängende nichtlineare Schallabsorption, – Schallstrahlungsdruck, Bildung der von Schall verursachten Ströme. In der vorliegenden Arbeit handelt es sich um viel größere Erscheinungen, die aber im geringeren Maße untersucht wurden. Die vorliegende Übersicht bezweckt den Kreis der zu untersuchenden Erscheinungen zu umreißen und zu versuchen, diejenigen Probleme aufzuwerfen, die vor allem zu lösen sind.

Es ist interessant zu vermerken, daß abgesehen davon, daß diese Erscheinungen im ungenügenden Maße untersucht wurden, findet ein Teil dieser Erscheinungen eine verhältnismässig breite Anwendung in der Praxis (Verstäubung von Aerosolen, Koagulation, verschiedene Anwendungsgebiete der Ultraschallkavitation). Es ist aus diesen Beispielen zu ersehen, daß die Physik in ihrer Entwicklung den praktischen Forderungen nachsteht.

I. Schallströme und der akustische Sprudel

1. *Akustischer Sprudel.* Das akustische Sprudeln wurde anscheinend zum ersten Mal im Bronzegefäß beobachtet, das von geschickten Meistern des alten Chinas gefertigt wurde. Reibt man die Gefäßhenkel mit den im Wasser angefeuchten Händen, so beginnt das in

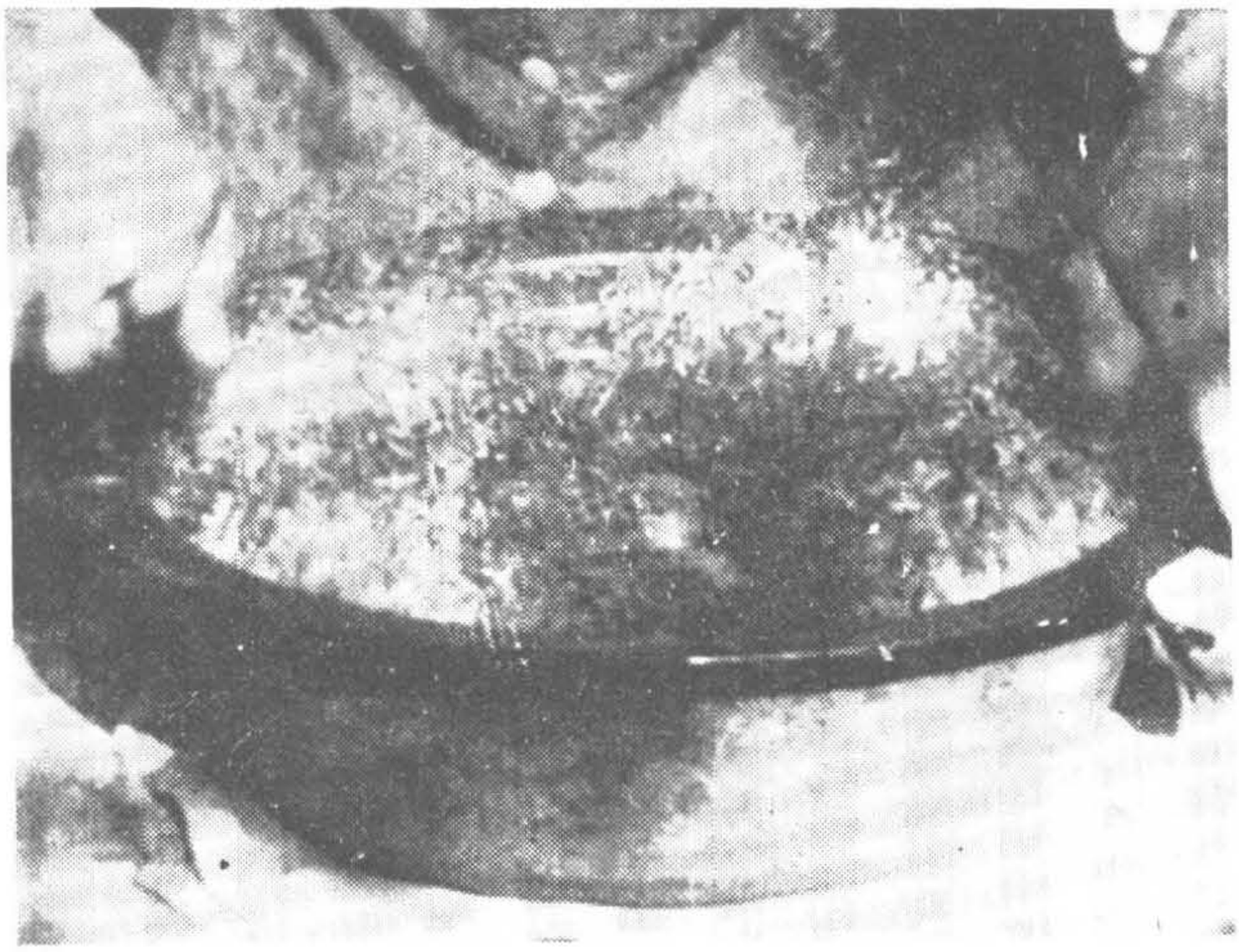


Fig. 1. Sprudeln des Wassers im chinesischen Gefäß.

diesem Gefäß enthaltene Wasser heftig zu sprudeln (Fig. 1). Die beim Reiben der Gefäßhenkel entstehenden Relaxationsschwingungen werden durch den Gefäßboden an das Wasser übertragen. Es ist anzunehmen, daß der Sprudel als Resultat der Bildung von intensiven Strömen entsteht, da bei der Dicke der Wasserschicht, die viel geringer ist, als die Wellenlänge, (während des Sprudels strahlt das Gefäß einen tiefen helenden Ton mit einer Frequenz von 200–400 H_z aus) das Entstehen von hohen Schalldrücken kaum zu erwarten ist. Die senkrecht zur Wasseroberflöche gerichteten Stromkomponenten bilden vielzahlige Sprudelchen.

In diesem Zusammenhang ist es angebracht die Ergebnisse, des Werkes [2] anzuführen, bei welchem eine Hochfrequenzaufnahme der Entstehung eines Sprudels an der Grenzfläche zwischen Wasser und Öl über die Stirnseite des Magnetostriktionsschallsenders, der mit einer Frequenz von 24 kHz arbeitete, durchgeführt wurde. Die Aufnahmefrequenz betrug 2000 Filmbilder pro Sekunde. Im Moment der Schallschaltung bildete sich auf der Oberfläche des Senders eine Wolke von Bläschen, die sich unter der Wirkung des Stromes zur Grenzfläche hob. Die Deformation der Grenzfläche begann etwas früher, als die Bläschenwolke die Oberfläche erreicht hatte. Der Zusammenhang dieser Erscheinungen in der Zeit ist aus der Fig. 2 zu ersehen; auf der Abszissenachse ist die Zeit in msek, auf der Ordinatenachse – der, von der Oberfläche des Senders abgezählte Abstand aufgetragen. Die waagerechte Punktlinie 1 entspricht der Grenzfläche zwischen zwei Flüssigkeiten; die Punkte auf der Linie 3 sind die Lagen der Bläschenwolke in verschiedenen Zeitmomenten, die mit Hilfe der Aufnahme bestimmt wurden. Die Wolke, die sich mit einer Geschwindigkeit von 0,65 m/sek fortpflanzt, erreicht in 14,5 m/sek. die Grenzfläche. Etwas früher, in 12 m/sek., beginnt die Deformation der Grenzfläche (Linie 2), die mit einer Geschwindigkeit von 0,42 m/sek. vor sich geht. Nimmt man an, daß die Wolke in ihrer Fortpflanzung, dem Strom etwas nachbleibt, so beträgt die Geschwindigkeit des Letztgenannten (seine Fortpflanzung ist mit der Punktlinie 4 gezeigt) 0,85 m/sek. Entstand die Deformation der Grenz-

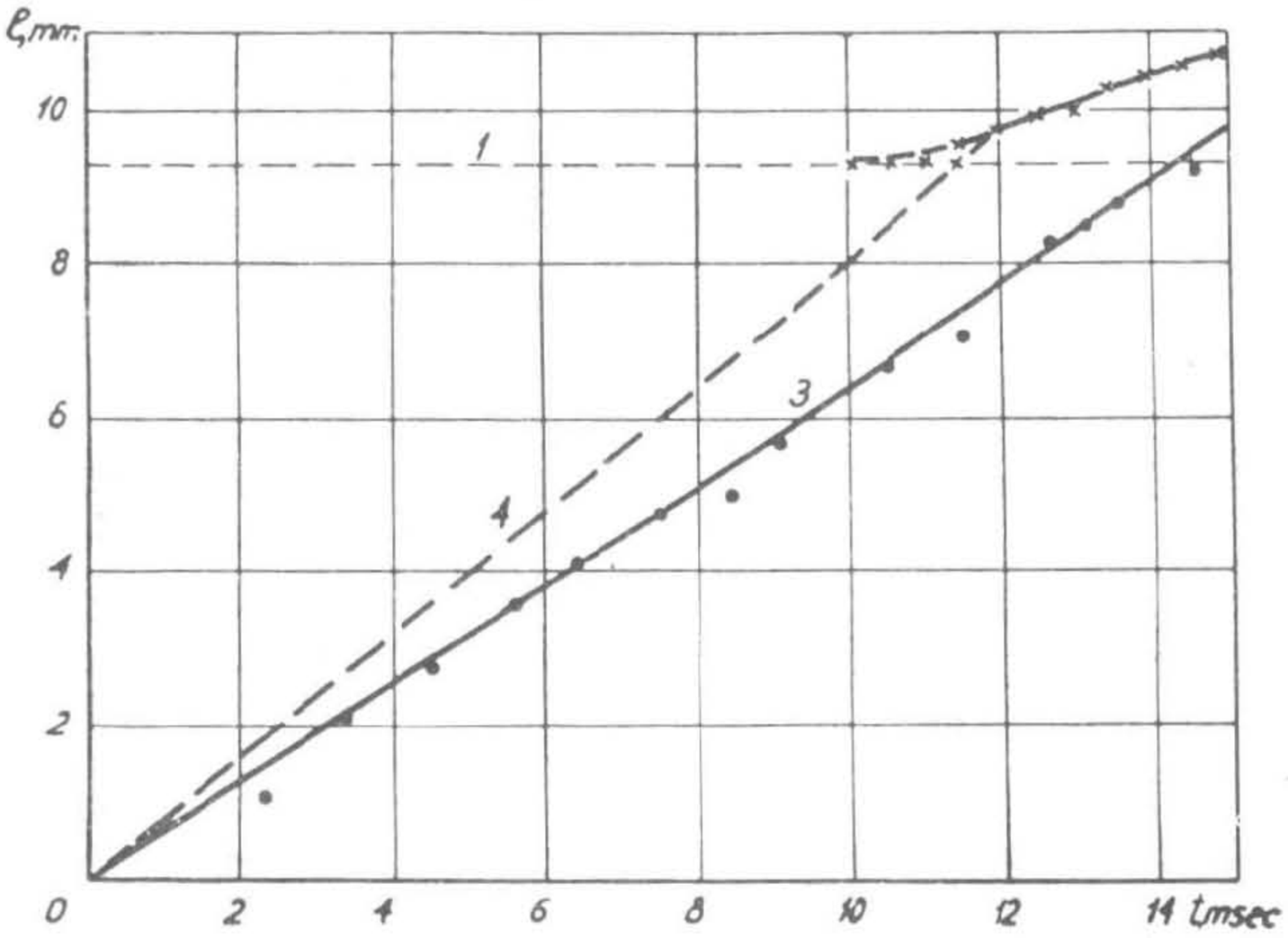


Fig. 2. Zusammenhang zwischen der Deformation der Grenzfläche zweier Flüssigkeiten mit dem Schallstrom.

fläche unter der Wirkung des Schallstrahlungsdrucks, der sich, wie bekannt, mit einer Schallgeschwindigkeit fortpflanzt, so wäre der Effekt schon in 10^{-5} sek. zu beobachten. Wir haben hier ein anschauliches Beispiel der Deformation (des Sprudelns), das als Ergebnis des Stromes und nicht des Schallstrahlungsdruckes entsteht. Der Sprudel ist hier eine natürliche Fortsetzung des Flüssigkeitsstromes – des Schallwindes.

2. *Zerstäubung der Flüssigkeit.* Der Ultraschallsprudel ist anscheinend einer der bekanntesten qualitativen Effekte, welche die Ultraschalle von hoher Intensität begleiten, die bereits



Fig. 3. Bildung von Aerosol im Ultraschallsprudel; die Verschiebung des Strahles ist zu sehen.

in klassischen Experimenten von Wood und Loomis beobachtet wurden. Ungeachtet dessen ist dieser Effekt sehr schwach untersucht. Sehr interessant ist eine der Nebenerscheinungen und zwar der Effekt der Flüssigkeitszerstäubung in dem Ultraschallsprudel [3]. Wie die Hochfrequenzaufnahme mit einer Frequenz von 2000 Filmbildern pro Sekunde zeigte wird das Aerosol am Anfang (am Stamm) des Sprudels ausgeschieden (Sieh Figur 3). In einigen Fällen wurden Abweichungen und sogar Verschiebungen des Wasserstrahls beobachtet, die von der Rückwirkung des Auswurfs von Aerosol zeugen. Eingehender wurde diese Erscheinung mit Hilfe der Aufnahme mit einer Frequenz von 4000–6000 Filmbildern pro Sekunde untersucht. Diese Untersuchung zeigte daß der Auswurf von Aerosol einen unterbrochenen explosiven Charakter aufweist. Die Analyse einzelner Filmbilder zeigt (Figur 4), daß die Tröpfchen, die früher durchsichtig waren, vor dem Auswurf des Aerosols das Licht diffusiv reflektieren.

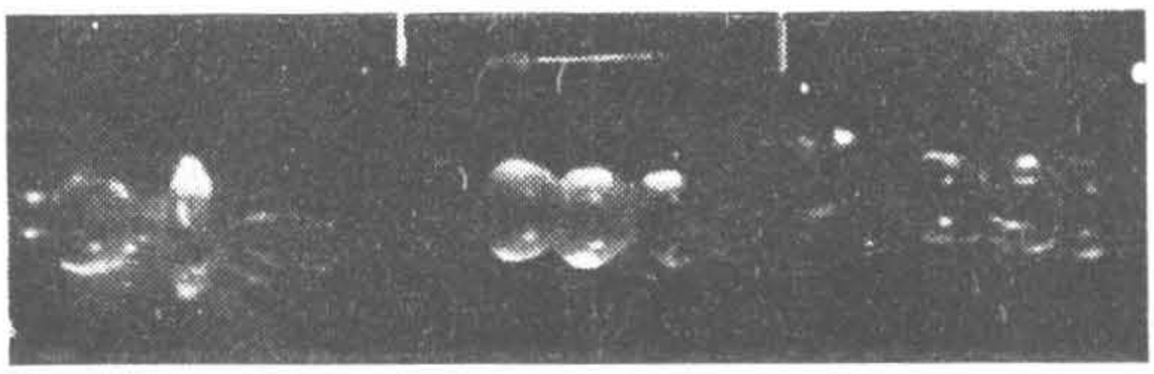


Fig. 4. Die Aufhellung der Strahlkugelchen und der Auswurf vom Aerosol.

Es gelang im weiteren festzustellen, daß wenn der Sprudel von oben nach unten gerichtet ist, zum Beispiel das Ultraschall in den Flüssigkeitsstrahl eingeführt wird, der aus einer Öffnung im Gefäßboden herausfließt, geht die Ausscheidung von Aerosol praktisch kontinuierlich vor sich.

Die Physik des Prozesses der Bildung von Aerosol ist noch nicht geklärt; wie bekannt existieren zwei Hypothesen. Die erste erklärt die Zerstäubung der Flüssigkeit durch die Bildung der Oberflächenwellen von einer großen Amplitude, was zum Abreißen der Tröpfchen von Wellenkämmen führt.

Die zweite gründet auf den Stoßwellen, die beim Zuklappen der Kavitationsbläschen entstehen. Nimmt man die erste Hypothese an, so wird die Monodispersität des Aerosols und die Frequenzabhängigkeit der Bläschenausmaße verständlich. Die zweite Hypothese erklärt den explosiven Charakter des Prozesses, kann aber die Monodispersität nicht erklären. Es wäre am Platze die analoge Lage mit der Erklärung des Mechanismus der Ultraschallemulgierung anzuführen, in der dieselben zwei Hypotzesen figurieren. [4] Es ist höchst wahrscheinlich, daß die zugrunde dieser zwei Erscheinungen liegenden Mechanismen sich als analog erweisen.

Die Erscheinung »der Explosion«, des flüssigen Tröpfchens in einem hochintensiven Schallfeld läßt sich auch unter folgenden Bedingungen beobachten. Es ist bekannt, daß die kleinen Flüssigkeitströpfchen in der Luft, in den Schallwellenbäuchen hängen können. In unserem Experiment hingen die Tröpfchen in der Luft in dem, mittels eines fokussierenden Konkavsenders und eines Flachreflektors gebildeten Feldes. Die Schallfrequenz betrug 20 kHz, die Feldintensität einige W/cm^2 und die Aufnahmefrequenz – 1000 Filmbilder pro Sekunde. Es hat sich erwiesen, daß die hängenden Tröpfchen ohne sichtbaren Grund ab und zu explodieren. Es gibt zur Zeit keine Annahmen um den Mechanismus dieser Erscheinung zu klären.

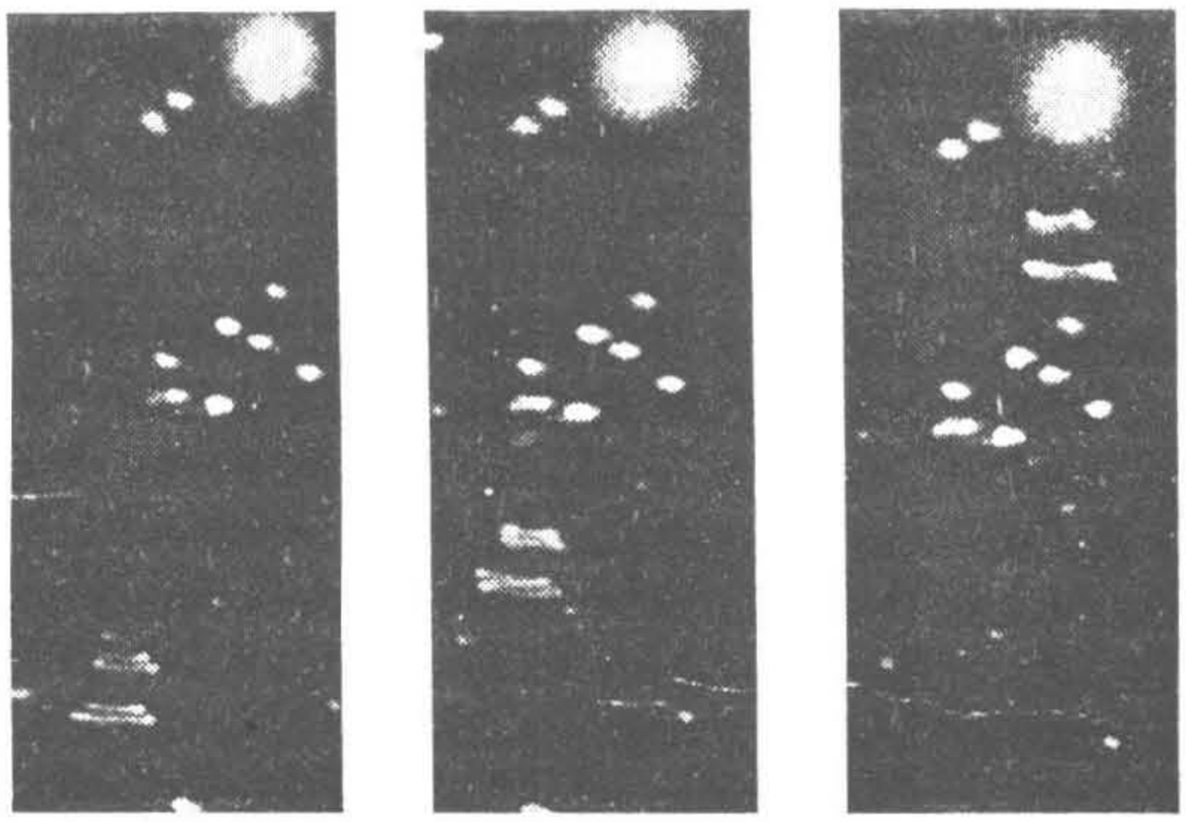


Fig. 5. Gepaarte Teilchenaggregate im Schallfeld.

3. *Akustische Koagulation.* Die bekannte Erscheinung der Koagulation kleiner Teilchen in Flüssigkeiten und Gasen ist, aller Wahrscheinlichkeit nach, auf die hydrodynamische Effekte zurückzuführen, die mit intensiven Schallfeldern zusammenhängen. Es sind dabei zwei Fälle zu unterscheiden. Einmal, wenn die Rede von der Koagulation der Gasbläschen in der Flüssigkeit ist, dann ist der Hauptfaktor – die, zwischen den schwingenden Bläschen entstehenden, Bjerknes-Kräfte [5]. Die Koagulation der festen Teilchen, die sich in der Flüssigkeit befinden, sowie der festen Teilchen und der flüssigen Bläschen, die sich in Gasen im suspendierten Zustand befinden, läßt sich durch diese Kräfte nicht erklären, da die Schwingungen ihres Volumens zu gering sind. Dieser, besonders verbreitete Fall der Koagulation, hat z.B. noch keine befriedigende physikalische Erklärung gefunden.

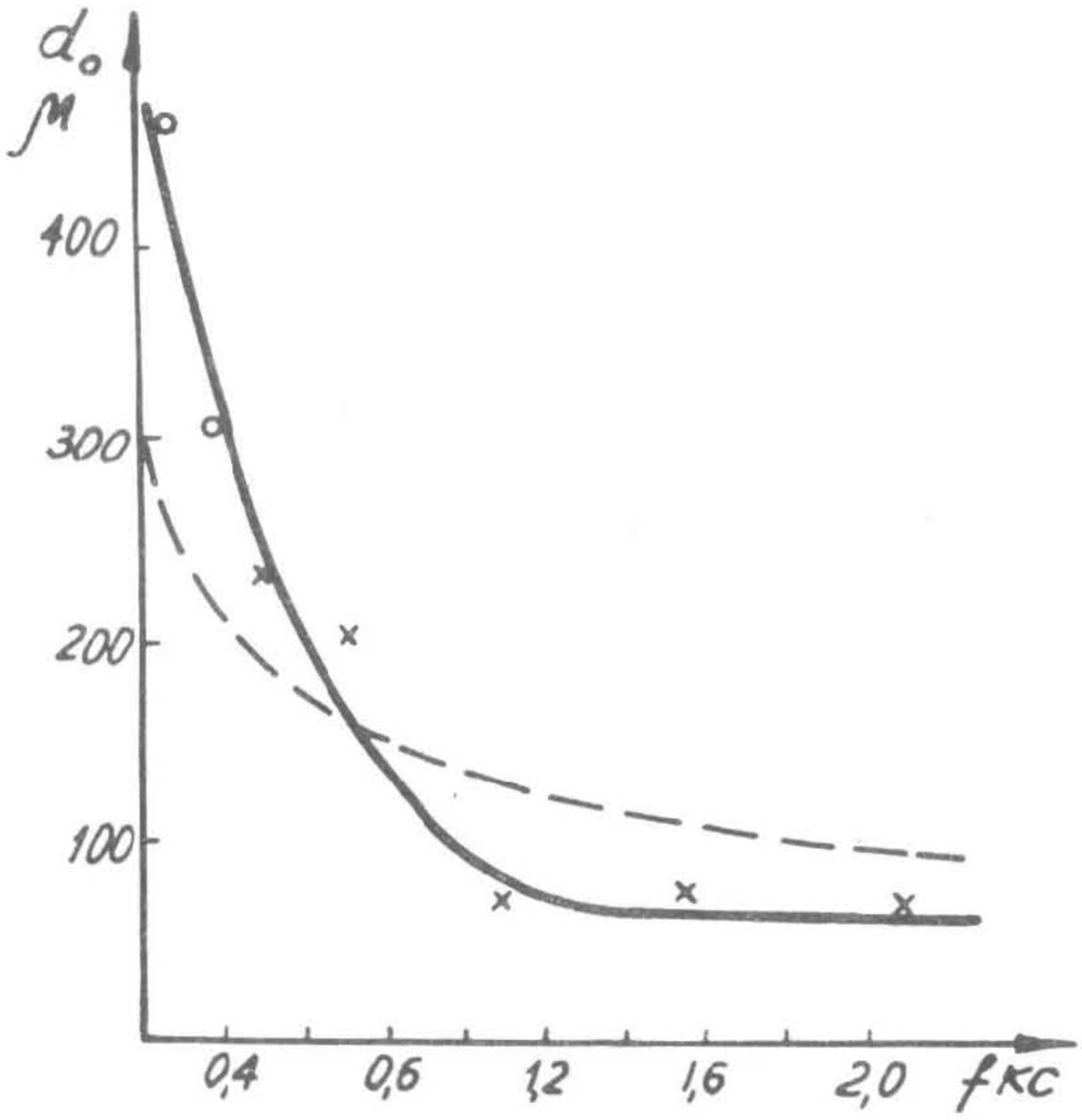


Fig. 6. Frequenzabhängigkeit des Abstandes zwischen den Teilchen des Aggregats; mit der Punktlinie ist die verdoppelte Größe der Grenzschicht gezeigt.

Vor kurzem wurde in unserem Laboratorium eine neue Erscheinung nachgewiesen, die Bildung im Ultraschallfeld von stabilen Paaren, Dreien und Aggregaten, die eine Anzahl von Teilchen enthielten. [6] Auf der Fig. 6 sind einige solche Aggregate gut zu sehen, die sich im Schallfeld gebildet haben. Die Abstände zwischen einzelnen Teilchen eines solchen Aggregates sind annäherend der verdoppelten Dicke der Grenzschicht gleich und nehmen auch mit der Frequenzzunahme ab (sieh Fig. 6).

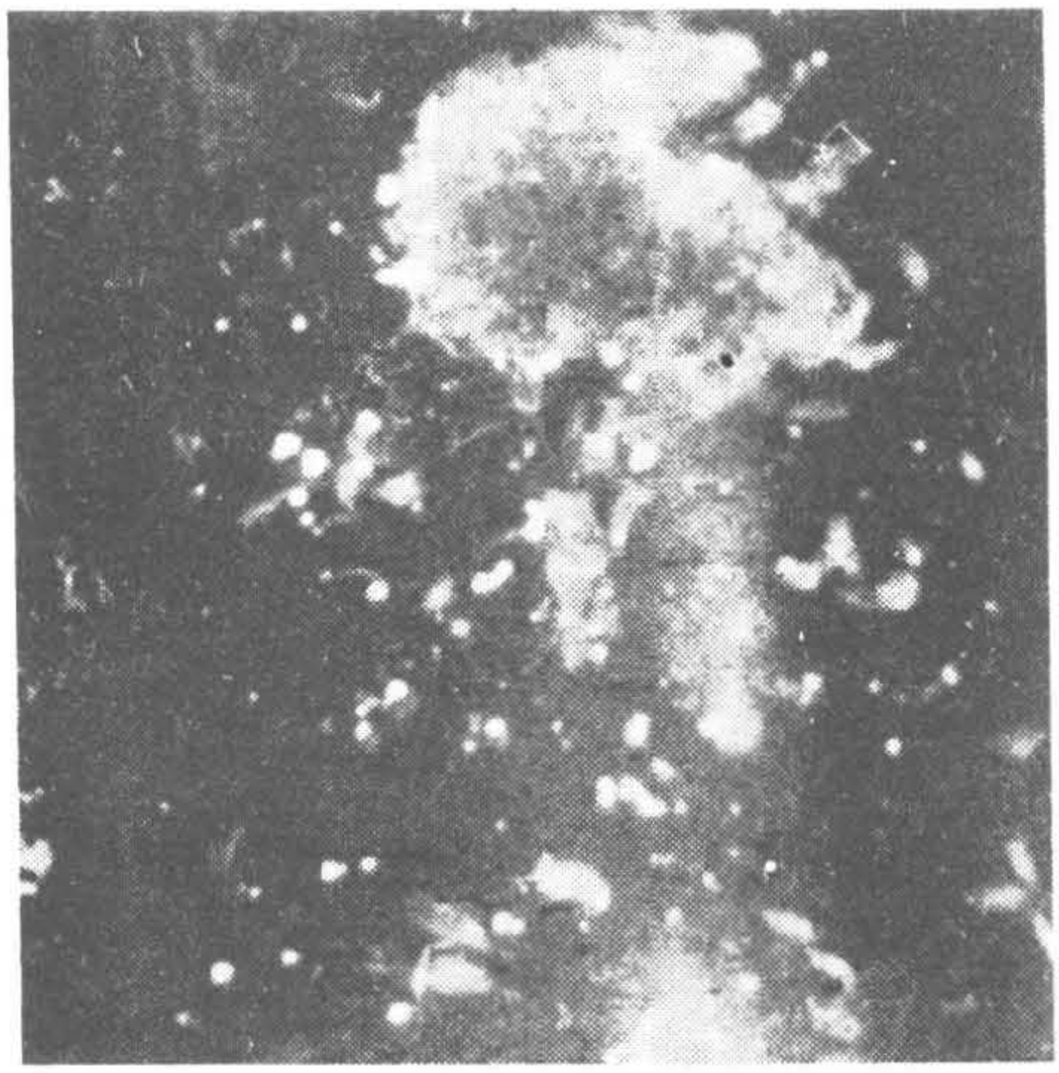


Fig. 7. Klümpchen fester Teilchen, das sich in Flüssigkeit unter der Wirkung des Ultraschalls bildete.

Die Bildung eines großen Klümpchens fester Teilchen in der Flüssigkeit ist aus der Fig. 7 gut ersichtlich. Das sind Leichen der kleinen Cyclopen die im starken Ultraschallfeld eingegangen sind. Die akustischen Bedingungen waren so, daß eine gut zu beobachtende starke Koagulation unerwartet entstanden war. Nach einiger Zeit zerfällt das Klümpchen anscheinend unter der Wirkung derselben Ströme.

Am Abschluß läßt es sich sagen, daß außer den in diesem Abschnitt erwähnten hydrodynamischen Strömen von großem Ausmaß auch die sogenannten Mikroströme, die in der letzten Zeit von Prof. Nyborg und seinen Mitarbeitern erfolgreich untersucht wurden, von großer Bedeutung sind. [7]

II. Kavitation

Dem Problem der Ultraschallkavitation sind viele Untersuchungen gewidmet. Die überwiegende Mehrheit dieser Untersuchungen ist dem Stadium des Lebens des einzelnen Kavitationsbläschens gewidmet. Es wurde die Entstehung der Kavitation [8], die Schwingungen des Bläschens [9], die von diesem hervorgerufenen Stoßwellen [10], die Ultraschalllumineszens [11] u.a.m. untersucht. Und obwohl hier noch nicht alles klar ist, ist doch Einiges schon bekannt. Interessant ist auch die andere Seite dieser Erscheinung, deren Entwicklung erst im Beginn steht.

Die ersten, die ihre Aufmerksamkeit diesem Aspekt der Erscheinung gewidmet haben, waren die japanischen Kollegen [12]. Die Rede ist von den Prozessen, die in der kavitierenden Flüssigkeit im ganzen vor sich gehen. Das Kavitationsbläschen entsteht nie allein (wenn wir das nicht bestreben). Für die Kavitation ist die Entstehung einer Kavitationswolke kennzeichnend, die aus einer Anzahl von Bläschen besteht. Das Vorhandensein einer solchen Wolke kann die akustischen Eigenschaften der Flüssigkeit, an sich beeinfließen. Das Wasser, in dem die Kavitation entstand, besitzt schon andere Parameter. Das, im Ergebnisse der Schwingungen und des Zuklappens der Bläschen entstehende Kavitationsrauschen ermöglicht es über einige Eigenschaften der Kavitation zu schließen. Jedes einzelne Kavitationsbläschen ist ein eigenartiger Wandler der Schallenergie, indem er einen Teil der auf seine Bildung aufgewandten Feldenergie als Stoßwelle, die von Hochfrequenzrauschen begleitet wird, zurückgibt. Die Gesamtheit der Mikrostoßwellen kann als Quelle der, für die Lösung praktischer Aufgaben nützlichen Energie dienen.

1. *Änderungen der Flüssigkeitseigenschaften bei der Entstehung der Kavitation.* Die Bildung der Kavitationsbläschen verletzt die Kontinuität der Flüssigkeit. Sind die Ausmasse der Bläschen im Vergleich zu der Wellenlänge gering, (was am häufigsten vorkommt), so kann die Flüssigkeit mit Bläschen als ein neues Medium mit anderen physikalischen Mitteleigenschaften untersucht werden. Da das Vorhandensein oder Fehlen der Bläschen und ihre Ausmasse von dem momentanen Wert des Schalldruckes abhängen, muß das äquivalente Medium nicht lineare Eigenschaften besitzen.

Der Prozeß der Schallstrahlung in ein derartiges Medium läßt sich dadurch kennzeichnen, das in der Verdünnungsphase, wenn die Bläschen erscheinen, die Flüssigkeit auf Kosten des Bläschendehnens nachgiebiger wird und der Strahler, indem er sich bewegt, eine kleinere Arbeit verrichtet, als in der Verdichtungsphase, wenn sich die Bläschen zugeklappt haben und die Oberfläche des Strahlers auf den Wellenwiderstand der reinen Flüssigkeit arbeitet. Die ausgestrahlte Gesamtenergie, verringert sich, also in Vergleich zu der Ausstrahlung in die kontinuierliche Flüssigkeit bei derselben Schwingungsgeschwindigkeit der Strahleroberfläche. Das ist aus der von uns ermittelten [13] Abhängigkeit der ausgestrahlten Leistung Wa (gemessen von elektrischer Seite) von dem Quadrat der Schwingungsgeschwindigkeit v_m bei einer Frequenz von 21,4 kHz im destillierten und gekochten Wasser gut ersichtlich (Kurve 1 Fig. 8). Es ist zu ersehen, daß die beim Fehlen der Kavitation linear wachsende ausgestrahlte Leistung ihr Anwachsen beim Erscheinen der Kavitation verlangsamt; auf einem verhältnismässig großen Abschnitt bleibt die ausgestrahlte Leistung, trotz dem Anwachsen der Schwingungsgeschwindigkeit, konstant. Erst bei sehr hohen Geschwindigkeiten nimmt die Leistung zu, aber viel langsamer, als auf dem Anfangsabschnitt.

Obwohl in einem nicht linearen Medium, wie eine Kavitationsflüssigkeit, der Begriff der akustischen Impedanz, die als Verhältnis momentaner Werte des Schalldrucks zu der Schwingungsgeschwindigkeit untersucht wird, keinen physikalischen Sinn hat, läßt sich der Begriff einer äquivalenten in der Zeit durchschnittlichen Strahlungsimpedanz einführen, die als:

$$\bar{R} = 2\frac{Wa}{v_m^2}$$

untersucht wird.

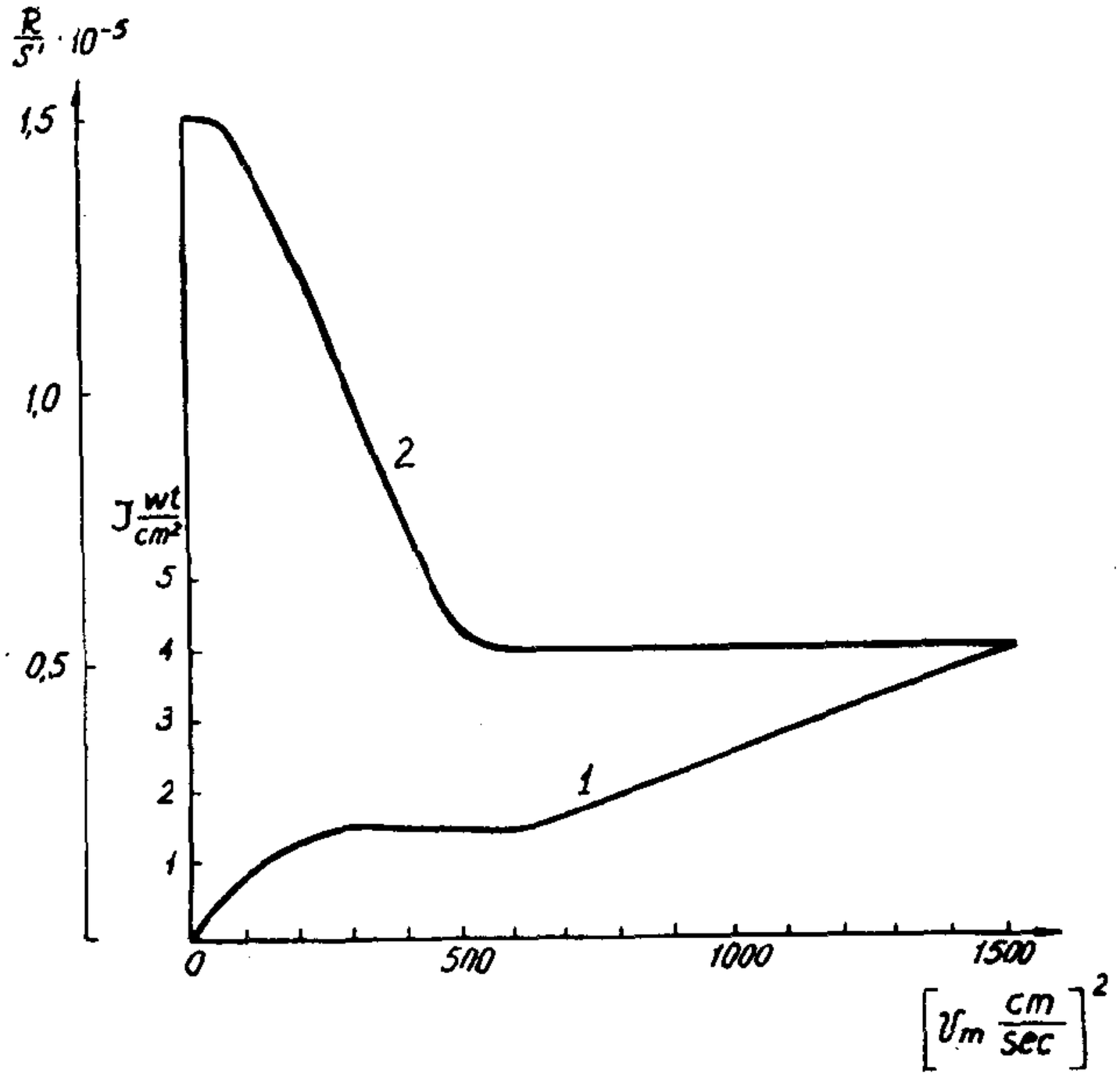


Fig. 8. Schallstrahlung in die Flüssigkeit beim Vorhandensein der Kavitation; Kurve 1 - ausgesandte Leistung; Kurve 2 - Strahlungswiderstand.

Der Berechnungswert dieser aus der Kurve 1 erhaltenen Größe ist auf Fig. 8, Kurve 2 dargestellt. Es ist zu ersehen, daß dieser äquivalente Widerstand sich beim Erscheinen der Kavitation zu verringern beginnt; seine Verringerung kann bedeutend sein (um das 3–5 fache). Daraus folgt vor allem, daß es nicht möglich ist, wie es viele machen, die von dem Strahler abgegebene Leistung im Vorkavitationsregime zu messen und dann die uns interessierenden Werten um das Kavitationsregime zu extrapolieren. Diese Methode ist leider sehr verbreitet.

Wird aber die Flüssigkeit, und das kommt am häufigsten bei verschiedenen praktischen Verwendungen des Ultraschalls vor, keiner speziellen Entgasung unterworfen, so bilden sich neben den Kavitationsbläschen auch Gasbläschen, die sowohl auf Kosten der Gasdiffusion, als auch auf Kosten der Koaleszenz und des Zusammenfließens der kleinen Bläschen in die großen wachsen. Solche Gasbläschen, die sich nicht zuklappen, können zu einer wesentlichen Verringerung der, in der Zeit konstanten Kompressibilität der Flüssigkeit führen und den beschriebenen Effekt verstärken. Das bezieht sich auf die Bläschen, deren Ausmasse größer sind, als die der Resonanzbläschen, da nur solche Bläschen eine zusätzliche Kompressibilität hervorrufen. Die Resonanzbläschen, ohne die Reaktionseigenschaften des Mediums zu ändern, ergeben nur eine zusätzliche Absorption. Was die Bläschen anbetrifft, die kleiner sind, als die Resonanzbläschen, so steigern sie etwas die durchschnittliche Flüssigkeitsdichte auf Kosten der angeschlossenen Masse; die Größenordnung dieser Zunahme ist noch nicht untersucht worden.

Für eine eingehende Untersuchung dieser interessanten Erscheinung müßte man die tatsächliche Veränderung der Flüssigkeitskompressibilität messen, die durch die Erscheinung

der Kavitationshöhlen in verschiedenen Wellenphasen bedingt ist. Das könnte man, z.B. mittels der Messung mit der Impulsmethode auf hohen Frequenzen der momentanen Werte der Geschwindigkeit der Schallfortpflanzung in der Flüssigkeit erreichen, in der die durch eine niederfrequente Anstrahlung hervorgerufene Kavitation stattfindet. Solche Experimente wurden aber meines Wissens noch nicht verwirklicht.

Was die Mittelcharakteristiken des Strahlungsprozesses anbetrifft, so ist es zu vermerken, daß die Kurvenform des Schalldruckes in der Zeit zu errechnen, wäre nur in dem Falle möglich, wenn die Gesetze der Bildung, der Schwingung und des Zuklappens der Kavitationsbläschen bekannt wären. Man kann aber den Begriff des effektiven Schalldrucks auf der Strahleroberflächen einführen, indem dieser Druck als eine Reaktion des nicht linearen Mediums auf den sinusförmig schwingenden Sender definiert wird.

$$P_{eff} = V_m \frac{\bar{R}}{S}$$

Der auf diese Weise ermittelte Wert $P_{eff} = F(V_m)$ ist auf Fig. 9 angeführt (ununterbrochene Kurve). Sie hat ein charakteristisches Aussehen, indem sie zwei Extremumpunkte und einen fallenden Abschnitt zwischen diesen aufweist. Auf derselben Figur sind mit der Punktlinie die Experimentalkurven der Änderung des Schalldrucks gezeigt, die mit Hilfe von einem Breitbandmeßempfänger aufgenommen wurden. Der Lauf dieser Kurven wiederholt qualitativ die Berechnungskurve; es weist dieselben Extremen und fallende Abschnitte auf. Analoge Experimentalergebnisse für einzelne Harmonischen, beim Vorhandensein der Kavitation wurden im Werk [12] ermittelt.

Wie bekannt, entspricht der fallende Abschnitt auf der Kurve der Abhängigkeit der momentanen Werte des Schalldrucks von der Schwingungsgeschwindigkeit dem negativen Widerstand, d.h. das im System Selbstschwingungen entstehen können. Unsere Kurve auf Fig. 9 zeigt das Verhältnis nicht momentaner, sondern effektiver Werte. Dabei läßt sich

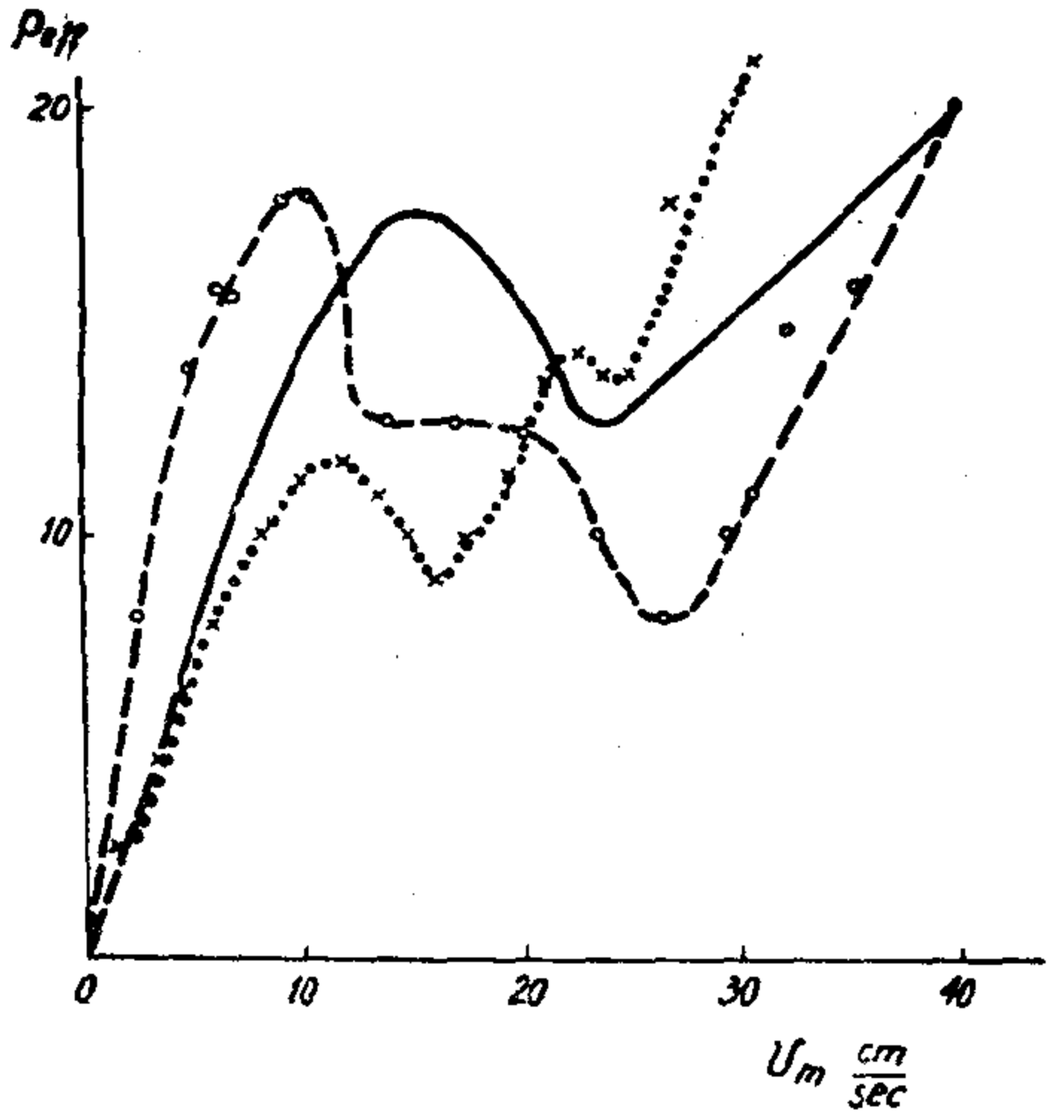


Fig. 9. Abhängigkeit des effektiven Schalldrucks von der Schwingungsgeschwindigkeit der Strahlersoberfläche. Ununterbrochene Kurve - Berechnungswert; Punktlinie - Experimentalwert.

aber das Vorhandensein des fallenden Abschnitts in einigen Fällen eben so auslegen. Es wäre in diesem Zusammenhang, interessant das Werk [15] zu erwähnen; die Verfasser haben das Vorhandensein der Selbstrelaxationsschwingungen in der kavitierenden Flüssigkeit festgestellt. Es ist anzunehmen, daß die Entstehung dieser Schwingungen auf den oben erwähnten fallenden Abschnitt der Abhängigkeit des effektiven Schalldrucks von der Schwingungsgeschwindigkeit zurückzuführen ist.

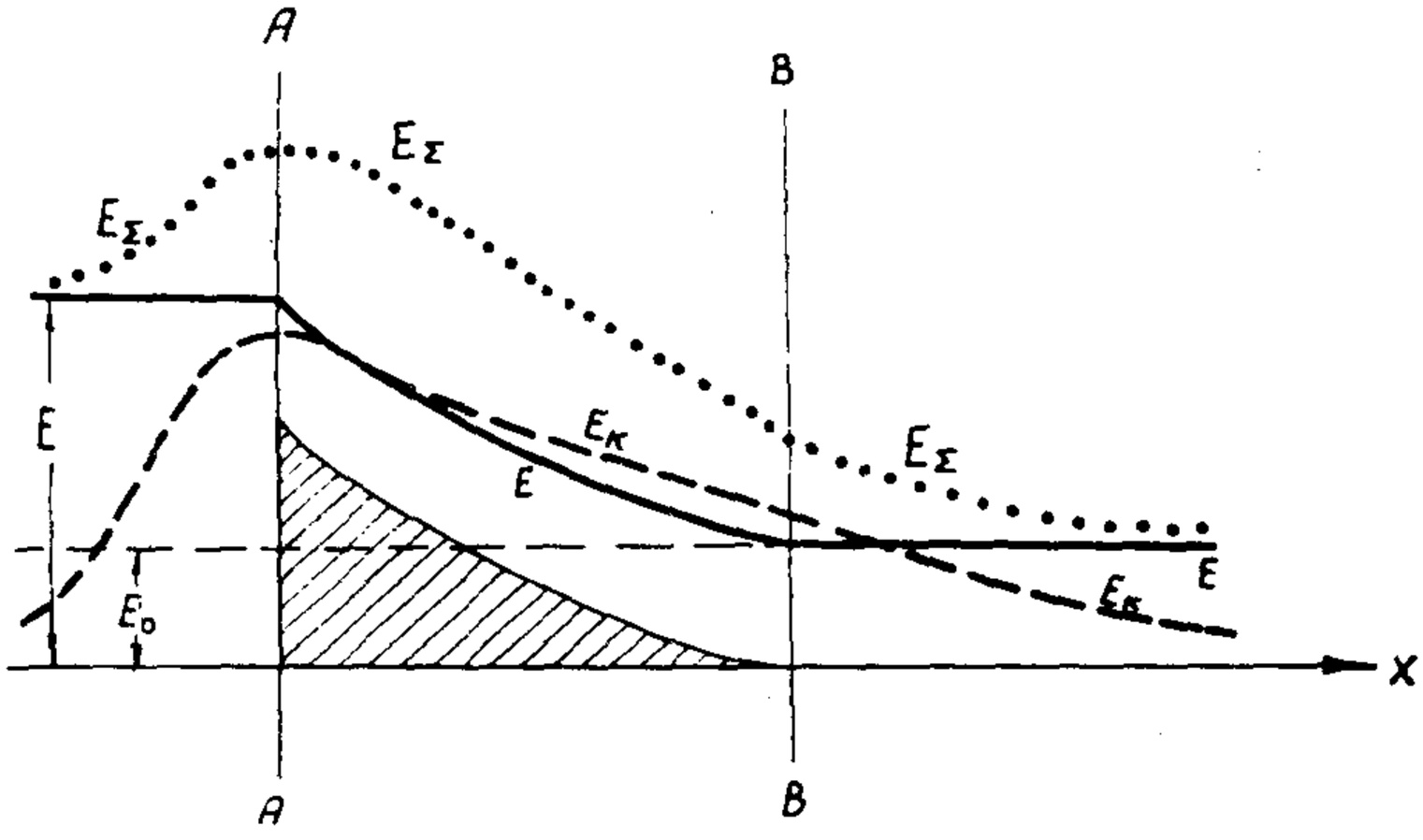


Fig. 10. Schematische Energiebilanz bei der Kavitation.

2. *Die Energetik des Kavitationsprozesses und das Kavitationsrauschen.* Das Entstehen der Kavitation ändert wesentlich die energetische Bilanz des Schallfeldes. Unsere Kenntnisse auf diesem Gebiete ermöglichen es leider nicht, ein quantitatives Bild aufzubauen; wir sind gezwungen, uns mit der qualitativen Untersuchung zu begrenzen. Zum Beispiel nehmen wir den einfachsten Fall des Schallfeldes in der fortschreitenden Flachwelle beim Vorhandensein von der Kavitation. Nehmen wir an (Siehe Fig. 10.), daß auf der Achse X sich die sinusförmige Flachwelle mit einer Energiedichte E von links nach rechts fortpflanzt; mit der Punktlinie ist die Schwelle der Entstehung der Kavitation E_0 bezeichnet. Links von der Linie A–A ist ein konstanter Druck aufgelegt, der die Entstehung der Kavitation hindert). Sobald die Schallwelle die Grenze A–A überschreitet, ruft sie das Erscheinen von Kavitationsbläschen hervor. Da darauf eine bestimmte Energie des Schallfeldes aufgewandt wird, so nimmt die Energiedichte der Schallwelle verhältnismäßig schnell ab und dadurch nimmt auch die Intensität der Bildung von Kavitationsbläschen ab. Nachdem die Energie in der Welle die Kavitationsschwelle E_0 (Linie B–B) erreicht, hört die Bildung der Kavitation auf.

Jedes Kavitationsbläschen ist ein eigenartiger Energiewandler. Einen Teil der Energie des primären Schallfeldes, die auf seine Bildung und Anwachsen aufgewandt wurde, gibt er als eine Stoßwelle zurück, die bei seinem Zuklappen entsteht; ein Teil der von ihm akkumulierten Energie wird außerdem in das Licht als Sonolumineszenz umwandelt.

In dem Falle, wenn in der direkten Nähe von dem sich zuklappenden Bläschen sich die Grenze eines Festkörpers zieht oder Einsprengungen einer anderen Flüssigkeit sich befinden, kann die Stoßwelle eine mechanische Arbeit vollziehen, z.B. kann eine Kavitationserosion der festen Grenze, Bildung der Emulsion, Zerreißung der großen Moleküle a u.a.m.

zustandekommen. Ist das Medium homogen, so geht die mechanische Energie der Stoßwelle teilweise in die Wärme über. Das Zuklappen des Kavitationsbläschens ruft jedoch in beiden Fällen das Kavitationsrauschen hervor, dessen Spektrum, wie bekannt, aus einer Anzahl der Ober- und Untertonen und weißem Rauschen besteht.

Da die Bläschen fast immer für Punktquellen zu halten sind, strahlen sie in alle Seiten aus. Die Summierung dieser Strahlung verursacht auch das Mittelniveau des Kavitationsrauschens, dessen Feld sich über die Grenzen des Gebietes der Kavitationsentstehung fortpflanzt, das auf Fig. 10 als schraffierter Abschnitt gezeigt ist. Die Energiedichte dieses Feldes ist durch die Linie E_k gekennzeichnet. Indem sich diese mit dem Feld des Ausgangsschalles E summiert, bildet es das summarische Feld E_Σ das sich mit Hilfe des Breitbandempfängers nachweisen läßt.

Obwohl die Felder E und E_k ihrer Spektralzusammensetzung nach, unterschiedlich sind, lassen sich ihre Energien nicht trennen, da sie beide die Spektralkomponente mit einer Frequenz des Ausgangsschalles enthalten. Eine solche Trennung wäre aber sehr wünschenswert um über die Eigenschaften des entstandenen Feldes zu schließen, da die Felder E und E_k sich noch dadurch unterscheiden, daß *das Feld des Kavitationsrauschens E_k die sekundäre Kavitation, wie groß auch seine Intensität ist nicht hervorrufen kann.* Dies ist darauf zurückzuführen, daß für die Bildung eines Kavitationsbläschens ein bestimmter Wert des negativen Druckes zu schaffen ist und die Energie des sich zuklappenden Bläschens als Stoßwelle-Druckwelle zurückabgegeben wird. Also steigert das Feld des Kavitationsrauschens E_k, indem es sich mit dem Grundfeld E energetisch summiert, die Fähigkeit des Feldes neue Kavitationsbläschen zu bilden nicht.

Für eine approximative Trennung der Felder, kann man aus E_Σ mit Hilfe eines Filters die ganze Frequenzkomponente ausschneiden, die gleich Frequenz vom Grundschall ist. Die Energie des Restens E_k^1 wird sich von E_k durch die Komponente mit der Grundfrequenz, die sich subtrahieren läßt, und durch die Energie der Harmonischen des Grundschalls, die dabei zu summieren sind, unterscheiden. Es gibt z.Z. keine Möglichkeit, diese Fehler zu ermitteln.

Der Wert E_k^1 kann angewandt werden, um über die Stärke der Kavitationserosion zu schließen, da der Schalldruck des Kavitationsrauschens von der Zahl der Kavitationsbläschen und der Geschwindigkeit deren Zuklappens abhängig ist, d.h. von denselben Werten, die die Kavitationserosion bestimmen [15]. Auf diesem Wege läßt sich aber die Grenze der Kavitationszone nicht feststellen, da die Stärke des Kavitationsrauschens, sowie auch die des Feldes E_E sich bei der Überschreitung der Grenze wenig ändern; diese Überschreitung läßt sich nur nach der Änderung der hochfrequenten Komponenten des Kavitationsrauschens nachweisen.

Anscheinend, darf man nicht alle Kavitationsbläschen für gleich halten; einige dieser, die in der Zone einer hohen Energiedichte des Ausgangsfeldes entstehen, können bis auf große Halbmesser ausgedehnt werden und g folglich eine Stoßwelle von größerer Leistung hervorrufen. Der Übergang von der Grenze A–A zu der Grenze B–B läßt sich also nicht nur durch die Verringerung der Dichte der gebildeten Bläschen kennzeichnen, sondern auch durch die Verringerung der Effektivität jedes einzelnen Bläschens.

3. *Verzerrung der Feinstruktur des durch die Kavitation hervorgerufenen Schallfeldes.* Die Kavitation, die ihrer Natur nach, ein statischer Effekt ist, ruft eine Verzerrung der Fein-

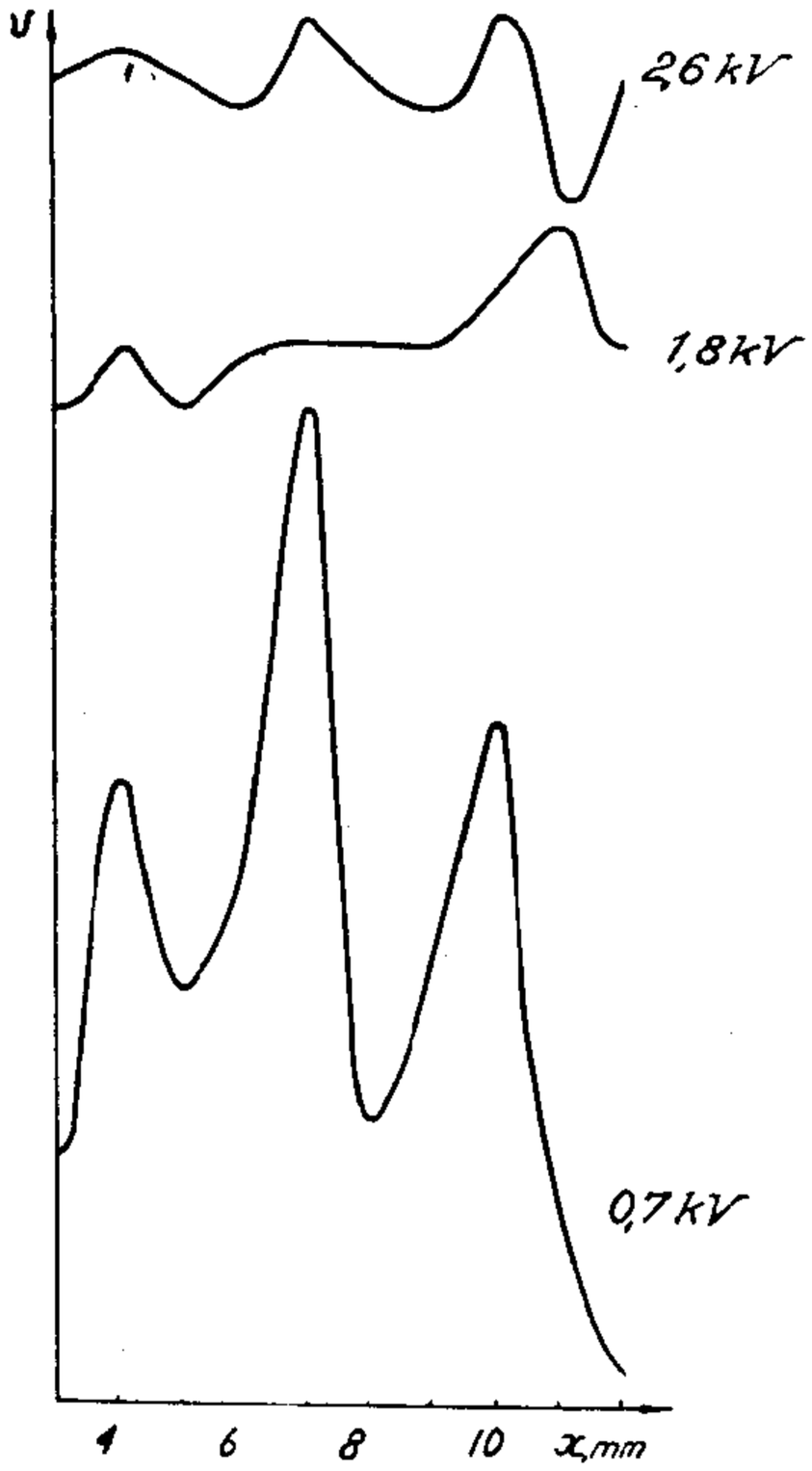


Fig. 11. Verteilung des summarischen Schallfeldes auf der Fokalfläche des Konzentrators bei verschiedenen Spannungen.

struktur des Schallfeldes hervor, darunter die Verzerrung der Struktur der Brennzone in fokussierenden Systemen. Auf Fig. 11 sind die Verteilungen des Schalldruckes auf der Brennfläche des Höchstleistungsultraschallkonzentrators dargestellt, [16] die mittels einer Breitbandschallsonde auf einer Frequenz von 0,5 MHz, mit verschiedenen, an die Quarzplatten zugeführten Spannungen gemessen wurde. Es ist zu sehen, daß mit der Entstehung und Entwicklung der Kavitation, die Fokalzone, die auf der ersten Kurve, die dem Fehlen der Kavitation entsprach so deutlich umrissen war, zerstört wird. Bei hohen Spannungen, erhält man ein stark verschwommenes Feld, das sich praktisch auf einige Zentimeter von der Fokalachse verbreitet.

Doch fehlt wie es oben bereits erwähnt wurde, die Verteilung des Rauschens mit der Verteilung der Kavitationserosion nicht zusammen. Die Letztgenannte läßt sich experimentell als eine Abnahme von dem Gewicht einer im Kavitationsfeld befindlichen Probe bestimmen. Auf Fig. 12 ist ein Vergleich auf der Brennebene der auf diese Weise bestimmten Kavitationserosion 1 und des summarischen Schallfeldes 2 gezeigt. Als Probe wurde ein Aluminiumzylinder von einer Länge von 2,0 mm und einem Durchmesser von 1,5 mm genommen. Auf derselben Figur ist das Kavitationsbereich 3 bezeichnet. Es ist zu sehen, daß der Rauschpegel sich nicht nur innerhalb des Kavitationsbereiches, sondern auch weit über seine Grenzen hinaus wenig ändert, während die Erosion an der Grenze des Bereiches schnell abnimmt.

Die Bildung der Kavitationszone in fokussierenden Systemen unterscheidet sich etwas von

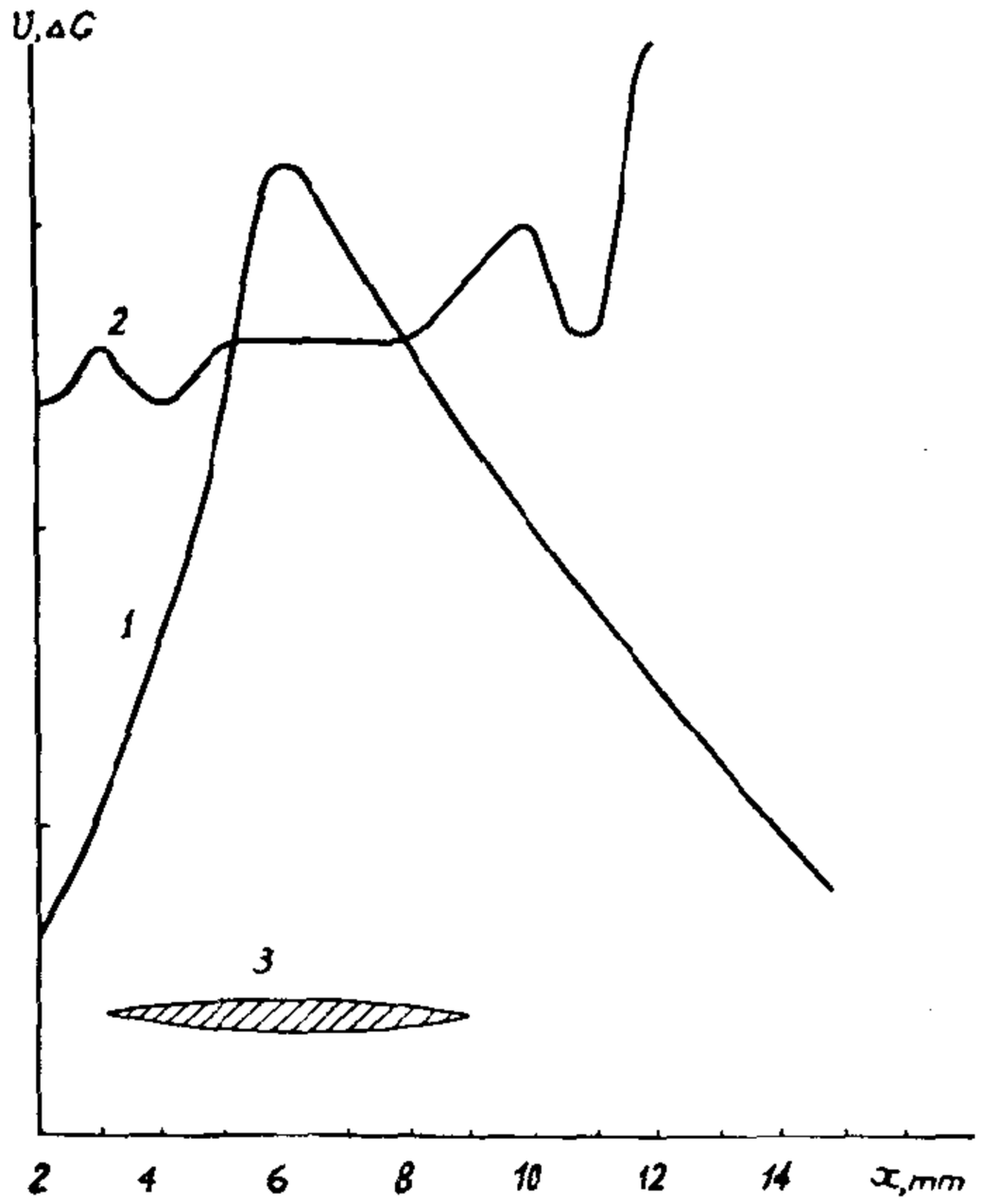


Fig. 12. Verteilung des summarischen Schallfeldes und der Kavitationserosion auf der Fokalfläche des Konzentrators.

dem oben erwähnten Fall der Flachwelle, da die Energiedichte mit der Annäherung an das Zentrum der Fokalzone zunimmt. Diese Zunahme kompensiert einigermaßen die Abnahme der Energie des Primärfeldes, die auf die Bildung der Kavitation aufgewandt wird. Im Ergebnisse, entsteht in der Fokalzone eine Kavitationswolke mit deutlich umrissenen Grenzen und einer gleichmäßigeren Verteilung der Kavitationsbläschen, als es für eine Flachwelle der Fall ist.

4. *Einige Eigentümlichkeiten der Kavitation beim Ultraschall von Höchstintensität.* Zum Erhalten der Ultraschallschwingungen von Höchstintensität wurde im Akustischen Institut der Akademie der Wissenschaften der UdSSR eine spezielle Anlage entwickelt – ein Konzentrator der Ultraschallschwingungen von hoher Leistung. Da diese Anlage schon mehrmals beschrieben wurde [1, 16, 17], führen wir hier nur einige Zahlenangaben an, die das Schallfeld in der Fokalzone dieser Anlage kennzeichnen. Bei einer Frequenz von 0,5 MHz, bei einer Spannung auf den erregenden Quarzplatten von 4,0 kV (diese Spannung läßt sich bis auf 7,0 kV steigern) kann die durchschnittliche Dichte des Schallenergiestromes (beim Fehlen der Kavitation) 30.000 W/cm^2 betragen. Im Zentrum des Brennpunkts erreicht dieser Wert 120 000 W/cm^2, was einer Amplitude des Schalldrucks von 600 at, einer Amplitude der Schwingungsgeschwindigkeit von 40 m/sek, einer Amplitude der Verschiebung von 13 μ und einer Amplitude der Beschleunigung von $1{,}2 \cdot 10^7$ g entspricht. Bereits im Jahre 1951 wurde von Neppiras die Annahme geäußert, daß die Zunahme der Energiedichte eines Ausgangsfeldes zur Zunahme der Energie der sich beim Zuklappen eines Kavitationsbläschens bildenden Stoßwelle nur so lange führt, bis die Zeit des Zuklappens mit der Halbperiode des Grundfeldes nicht vergleichbar wird. Bei einer weiteren

Zunahme werden die Bläschen mit geringeren Geschwindigkeiten oder überhaupt nicht zuklappen.

Diese Annahme wurde von uns mit Hilfe des oben beschriebenen Konzentrators experimentell geprüft. Es wurden dabei die Werte der Kavitationserosion und die Intensität der Sonolumineszenz in der Fokalzone unter Abhängigkeit von der Spannung auf dem Konzentrator gemessen. Unter denselben Bedingungen wurden auch mittels der Mikroaufnahme die Halbmesser der schwingenden Kavitationsbläschen R_m gemessen. Es hat sich dabei erwiesen, daß die größten Halbmesser der Bläschen sich mit der Spannung auf dem Konzentrator in einem linearen Zusammenhang befinden.

$$R_m = 0{,}6 \cdot 10^{-3}\ V\ /\mathrm{kV}/$$

Die Zeiten des Zuklappens der Bläschen wurden nach der bekannten Approximationsformel von Rayleigh ermittelt, die für das Wasser das Aussehen

$$\tau = 0{,}9\,\frac{R_m}{\sqrt{Po}}\ \text{hat,}$$

wobei Po – der hydrostatische Druck ist. Auf Fig. 13 sind die Experimentalergebnisse angeführt; die Kreischen entsprechen der Kavitationserosion, die Dreiecke – der, mit Hilfe von einem Elektronenvervielfacher gemessenen Lumineszenz. Beide Werte haben ein deutlich ausgesprochenes Maximum, das dem Werte $\tau/T/_2$ entspricht der nah zu Einz ist. Bei einer weiteren Zunahme der Spannung nehmen beide Werte schnell bis auf Null ab. Eine derartige Abhängigkeit konnte bei einem konstanten Schalldruck (d.h. der Zeit des Zuklappens) und einer veränderlichen Frequenz erhalten werden. Aber bei niedrigen Frequenzen mußten darauf Schalle von sehr großen Intensitäten angewandt werden. Eine Überschlagsbewertung ergibt in unserem Fall, daß das Verhältnis $\tau/T/_2 = 1$ (unter Bedingungen unseres Experiments) der Amplitude eines Schalldrucks von ca. 40 at entspricht, was mit einer Intensität der Flachwelle von ca. 500 W/cm^2 übereinstimmt.

Bei einer Verringerung des hydrostatischen (atmosphaerischen) Drucks, nimmt der Halb-

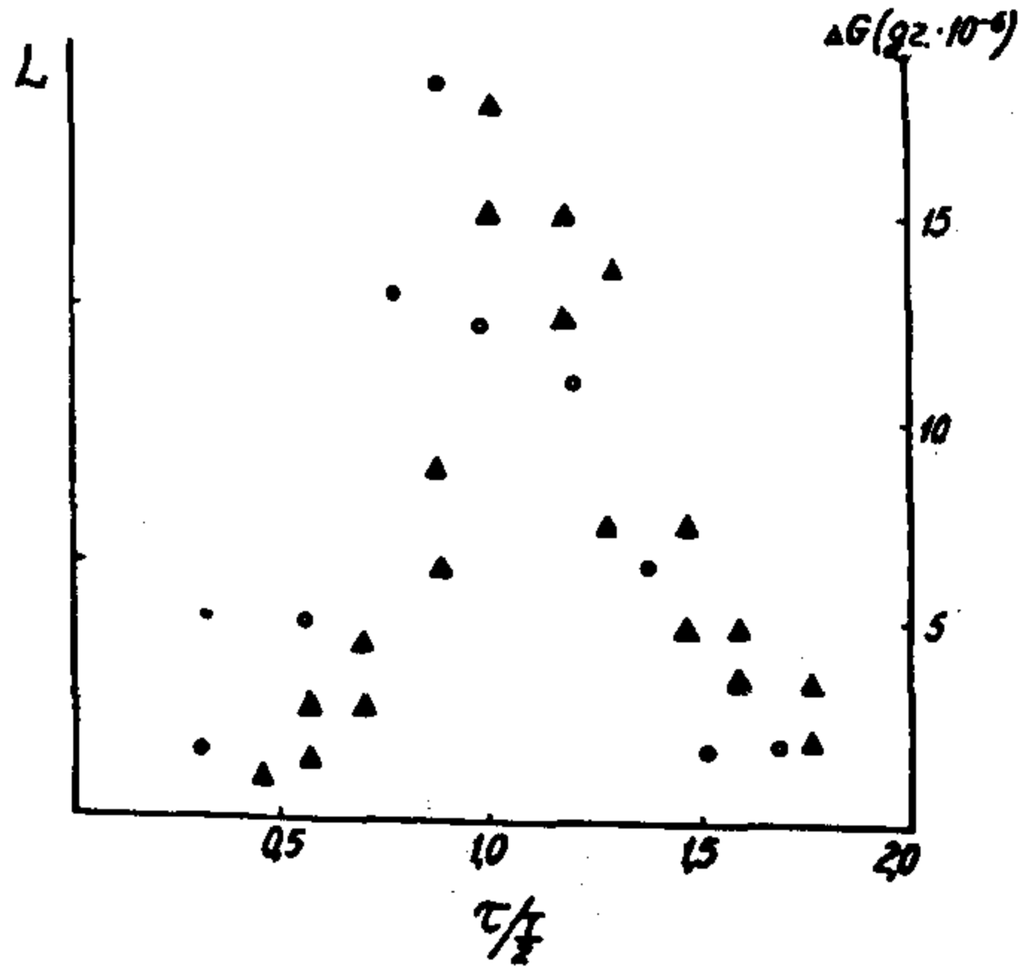


Fig. 13. Abhängigkeit der Kavitationserosion (Kreischen und der Intensität der Ultraschallumineszenz (Dreiecke) von dem Verhältnis der Zeit des Zuklappens zu der Halbperiode der Schwingung.

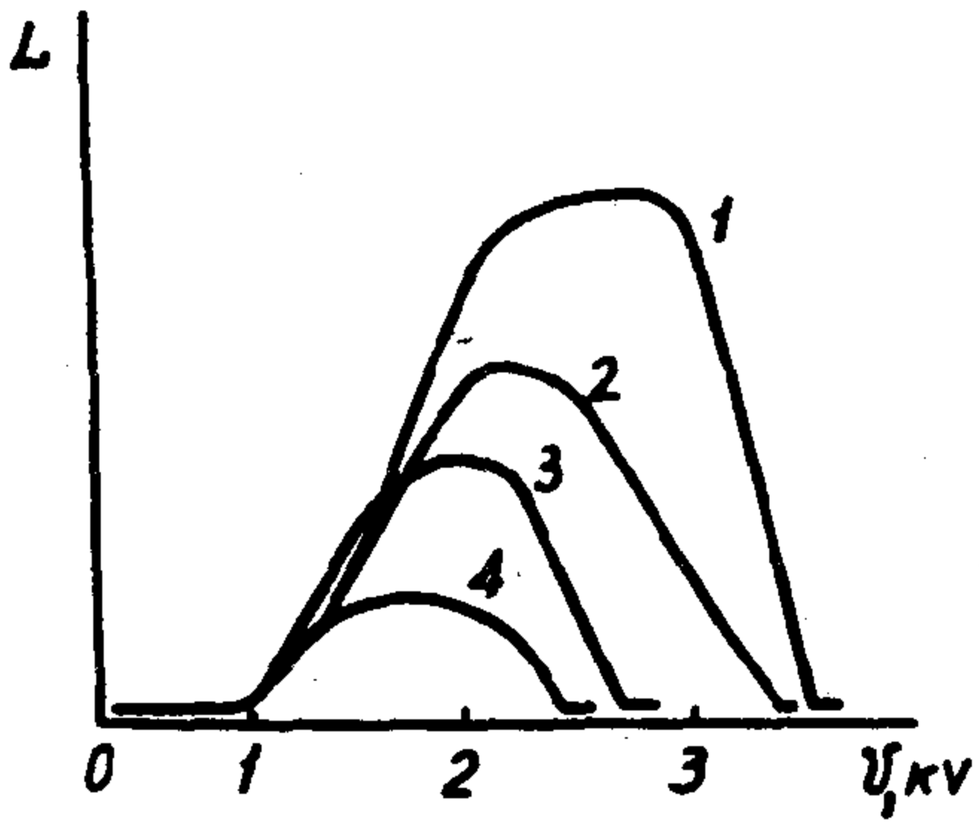


Fig. 14. Abhängigkeit der Lumineszenzintensität von der Spannung auf dem Konzentrator bei verschiedenen hydrostatischen Drücken.

messer des Kavitationsbläschens, wie das aus der Rayleigh – Formel folgt, zu und die Bedingung kann bei geringeren Schalldrücken eintreten. Es ist zu erwarten, daß dabei auch die beim Zuklappen von Kavitationsbläschen abgegebene Energie abnehmen wird.

Die Experimentaluntersuchung der Ultraschallumineszenz unter Abhängigkeit von der Spannung auf dem Konzentrator bei verschiedenen atmosphärischen Drücken haben diese Annahme bekräftigt. Die entsprechende Kurvenschar ist auf Fig. 14 angeführt, aus der es gut ersichtlich ist, wie das Lumineszenzmaximum mit der Abnahme des konstanten Druckes in der Größe abnimmt und sich in die Seite geringerer Spannungen verschiebt. Bei einem konstanten Druck, der unter 0,3 atü war, ließ sich die Lumineszenz überhaupt nicht nachweisen.

Die maximale Lumineszenzintensität, wie es aus der Kurve 1, Fig. 15 zu ersehen ist, nimmt mit der Steigerung des konstanten Druckes linear zu. Die Spannung, bei der dieses Maximum eintritt, hängt ebenfalls von dem konstanten Druck linear ab (Kurve 2, Fig. 15). Es ist zur Zeit noch keine Theorie entwickelt, die es ermöglichen könnte, diese Erscheinung von der quantitativen Seite zu bewerten.

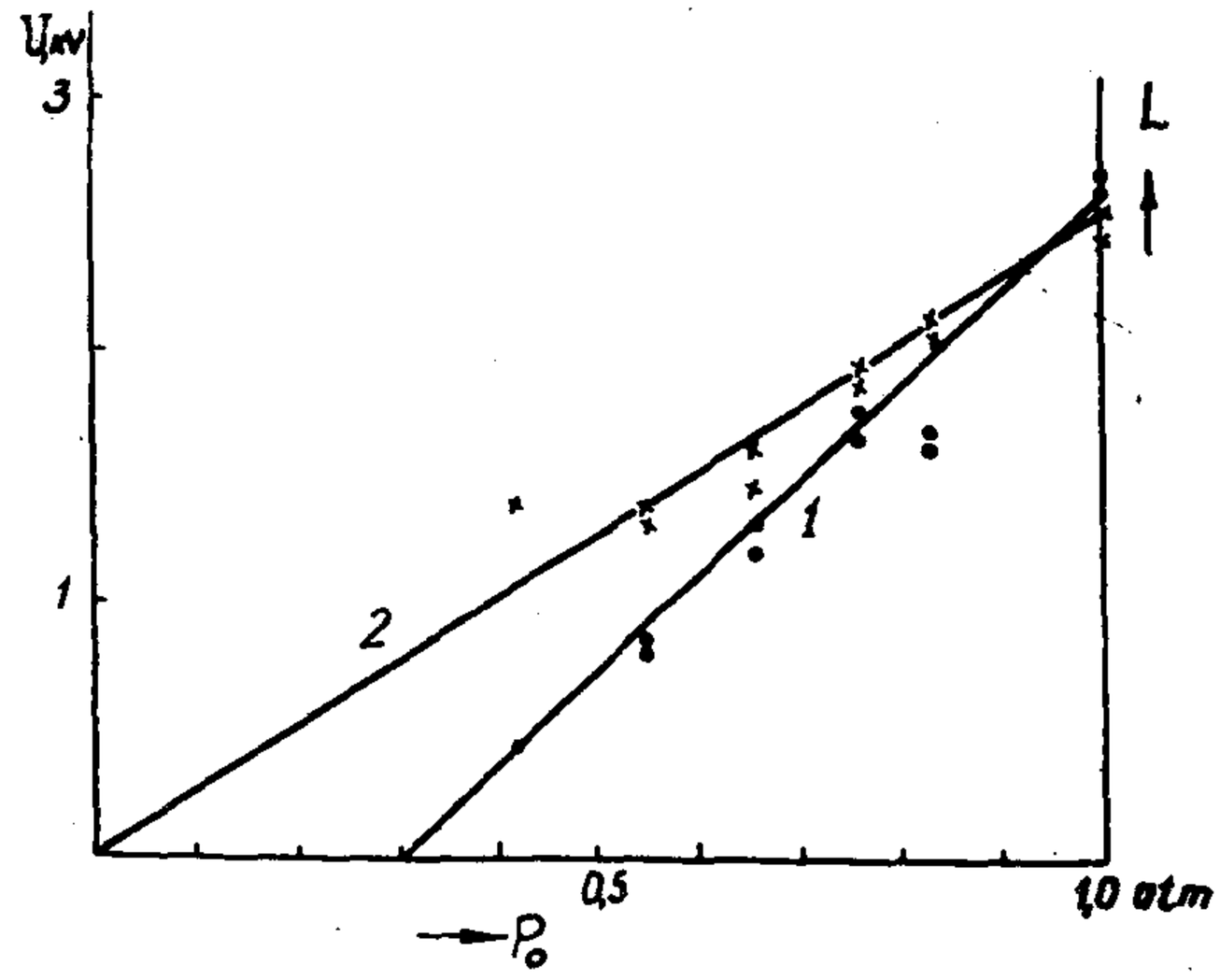


Fig. 15. Abhängigkeit der größten Lumineszenzintensität (1) und der Spannung, bei der die Maximale Lumineszenz beginnt (2) von dem hydrostatischen Druck.

LITERATURVERZEICHNIS

1. L. D. Rosenberg, La generation et l'étude des vibrations ultra-sonores de très grande intensité. Acustica *12*, 1, 40–49, 1962.
2. L. D. Rosenberg, L. O. Makarow, Über die Ursachen der Deformation der Flüssigkeitsoberfläche unter der Wirkung des Ultraschalls. Ber. d. Akad. d. Wissensch. d. UdSSR, *114*, 2, 275–276, 1957.
3. L. D. Rosenberg, O. K. Eknodiosjanz, Über die Kinetik der Ultraschallnebelbildung. Akust. Ztschr. UdSSR, *6*, 3, 370–373, 1960.
4. S. A. Nedushij, Untersuchung des Prozesses der Bildung von Emulsionen, der durch die Wirkung der Schall- und Ultraschallschwingungen hervorgerufen wurde (Übersicht). Akust. Ztschr. UdSSR, *7*, 3, 275–284, 1961.
5. W. F. Kasanzew, Bewegung der Gasbläschen in der Flüssigkeit unter der Wirkung der Bjerknes-Kräfte im Schallfeld. Ber. d. Akad. d. Wissensch. d. UdSSR, 129, 1, 64–67, 1959.
6. O. K. Eknodiosjanz, L. N. Burawov, Über das Verhalten der Aerosolteilchen im Schallfeld. Akust. Ztschr. UdSSR, *7*, 4, 492–493, 1961.
7. W. L. Nyborg, Small-scale acoustic streaming in liquids. J.A.S.A. *28*, 6, 1237–1242, 1956. Acoustic streaming near a boundary. J.A.S.A. *30*, 4, 329–339, 1958. Sonically induced microstreaming applied to a surface reaction. J.A.S.A. *31*, 6, 706–711, 1959.
8. D. Sette, Sonic cavitation and radiation. Proc. III intern. Congr. on Acoustics Stuttgart 1959, v.l. 330–333 (1961).
   D. Lieberman Radiation-induced cavitation. Phys. Fluids. *2*, 4, 466–468, 1959.
9. E. A. Neppiras, B. E. Noltingh, Cavitation produced by ultrasonics, Proc. Phys. Soc. *63B*, 674–685, 1950, *64B*, 1032–39, 1951.
   E. Meyer, Some new measurements on sonically induced cavitation, J.A.S.A. *29*, 1, 4–8, 1957.
10. W. Güth, The formation of pressure by cavitation. Cavitation in Hydrodynamics, London H.M.S.O., 1956, VII–X. Zur Entstehung der Stoßwellen bei der Cavitation. Acustica, *6*, 6, 526–531, 1956.
11. P. Günter, E. Heim, G. Eichorn, Phasenkorrelation von Schallwechseldruck und Sonolumineszenz. Z. Angew. Phys. *11*, 7, 274–277, 1959.
    E. Meyer, H. Kuttruff, Zur Phasenbeziehung zwischen Sonolumineszenz und Kavitationsvorgang bei periodischer Anregung. Z. Angew. Phys. *11*, 9, 325–333, 1959.
    I. E. Elpiner, Ultraschallumineszenz (Übersicht). Akust. Ztschr. *6*, 3–15, 1960.
    P. Jarman, Sonoluminescence: A discussion J.A.S.A. *32*, 11, 1459–1462, 1960.
12. Y. Kikuchi, Recent results of research and development in the field of ultrasonic in Japan. Proc. III. Intern. Congr. on Acoustics. Stuttgart, 1959, v. II, 1193–1206, 1961.
13. L. D. Rosenberg, M. G. Sirotjuk, Über die Schallausstrahlung in die Flüssigkeit beim Vorhandensein von Kavitation. Akust. Ztschr. UdSSR, *6*, 4, 478–481, 1960.
14. M. Degrois, B. Badilian, CR Acad. Sci. *254*, 231–233 8.1.1962. *254*, 837–839 29.1.1962. *254*, 1213–1215 12.II–1962.
15. A. S. Bebtschuk, J. N. Borissow, L. D. Rosenberg, Zur Frage über die Kavitationserosion. Akust. Ztschr. UdSSR, *4*, 4, 361–362, 1958.
16. L. D. Rosenberg, M. G. Sirotjuk, Hochstleistungsultraschallkonzentrator. Akust. Ztschr. UdSSR, *5*, 206–210, 1959.
17. L. D. Rosenberg, M. G. Sirotjuk, Fokussierender Ultraschallkonzentrator. Autorenschein der UdSSR. Nr. 124217, 1959.

On the Perception of Sound and Speech

by J. F. SCHOUTEN*

Introduction

The many subjects dealt with at this conference may be grouped under various headings. Many speakers dealt with the physical phenomena of sounds occurring in the outside world: vibrations, periodic oscillations, noise, transmission of sound in the air and in solid materials etc., all physical phenomena, rigidly to be described by mathematical formulae in terms of the properties of physical matter.

Others dealt with the properties of the human ear, that marvellous product of biological sub-miniaturization, which is able to pick up the faintest of vibrations, to sort them out into part vibrations of different kinds and to set up a pattern of nervous impulses transmitted to our brain.

Still others dealt with sound as a phenomenon within our mind: the impressions of sound which we can recognize and memorize and which enable us to relate these, our sound impressions, to the objects, the animals or the human beings physically producing these sounds.

This last morning of invited papers is particularly devoted to subjects falling under this third heading.

In this last lecture I intend to dwell upon our subjective world of sound in its relation to the physical world we have to become aware of through our sense of hearing.

We are not usually aware of any discrepancy between a physical world outside and a subjective or perceptual world inside. On the contrary: we see the image of a person and take it for granted that this very same person must be present in the flesh. So far so good. But please look at the head of your nearest neighbour and compare its huge size with the minute dimension of the head of somebody sitting some 10 or 20 feet beyond. Has the latter person dwindled to dwarflike dimensions by an obscure illness? "Oh no" you say indignantly, "he just *looks* smaller because he is further away".

Granted, but take the lesson to be learned from this. By a queer trick of our cameralike eye we obtain a completely distorted impression of the outside world, such e.g. that far away persons and objects look smaller. And by an additional, equally queer trick of our perceiving mind we are not usually aware of this distortion and subconsciously interpret the dwarflike image of the person as a life-size person at a larger distance.

* Instituut voor Perceptie Onderzoek, Insulindelaan 2, Eindhoven, the Netherlands.

Can you now imagine why a blind person, even whilst trying to understand what vision is, cannot possible understand that objects look smaller merely by being further away and that the visual image of a table presents us with four legs of different size?
If, when walking, we hurt our leg, we conclude that a physical object exists obstructing our stride. This is justifiable in 99 out of 100 cases but it can also happen that a sudden pain in our leg cannot be attributed to a collision with a material object and is caused by something happening within our own leg.
And, finally, if we hear a voice uttering a sentence we take it for granted that the person with that particular voice has indeed spoken those very words. We are right again, 99 out of 100 times, but hallucinations occur even with normal persons. We may hear a voice speaking words and find out to our bewilderment that the voice came from inside rather than from outside.
So let us be extremely careful and strictly distinguish between the physical world of sound outside and the perceptual world within us. Our perceptual world presents us essentially with a make-believe world, a very believable one at almost all times, but not always a very true one.
I hope to give you some examples this morning of physical sounds and their perceptual counterparts. I further hope to give you an idea how the relations between the two provide us with important clues regarding the properties of the human ear as well as those of our perceptual system.
I shall first give you some examples of our faculties of frequency analysis and then of our faculties of time analysis.

Subjective Frequency Analysis

Musical chord

I shall first let you hear three tones produced sequentially. Physically they are described as 3 sinusoidal oscillations of ascending frequency as shown on the left of Fig. 1. Subjectively we hear them as 3 pure tones of different pitch as indicated on the right.
I shall then strike the three frequencies together as a chord. You will observe the familiar effect that even in the chord we can still detect the presence of the three tones, just as indicated in the musical notation, which is essentially a graph on a vertical frequency scale. This represents, in fact, Ohm's acoustical law, stating that the human ear performs some sort of Fourier-analysis, viz. that the human ear dissects a sound into part sounds of different pitch corresponding to its physical analysis into oscillations of different frequency. As far as physical analysis is concerned we are at liberty to describe the sound either as 3 different time patterns or as 3 different frequencies. Subjectively, however, it turns out that the three tones of different pitch, when sounded together as a chord, keep their individuality and do not merge into a completely new sound.
The separable items in our physical world are called: objects. Simularly the separable items in our perceptual world are called: percepts. The chord still contains its 3 constituent tones as separate percepts just as 3 eggs together are seen as 3 different eggs.
You will notice, incidentally, that in the wave pattern of the physical chord it is very difficult to see at first sight that it is the summation of the three constituent wave patterns. In

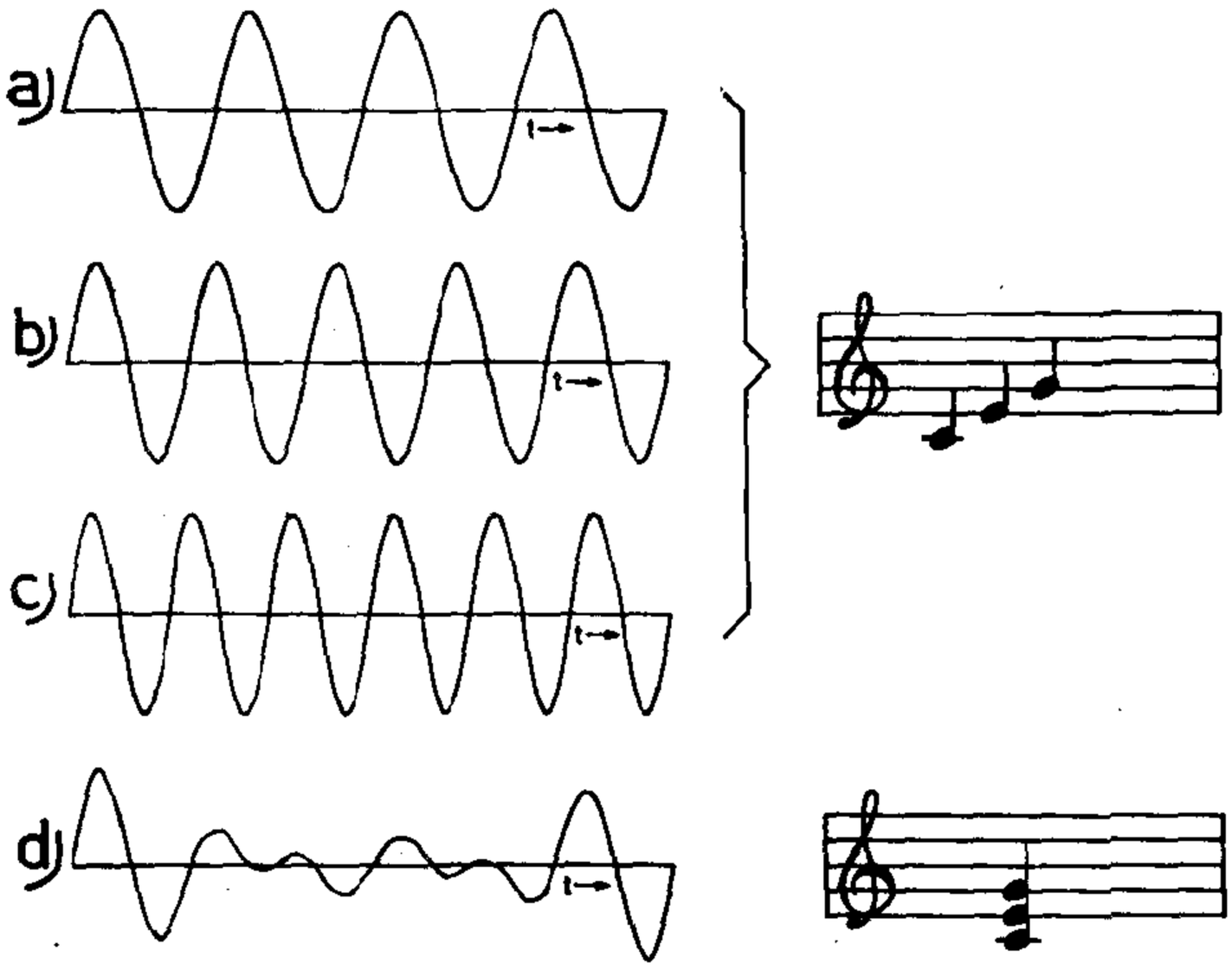
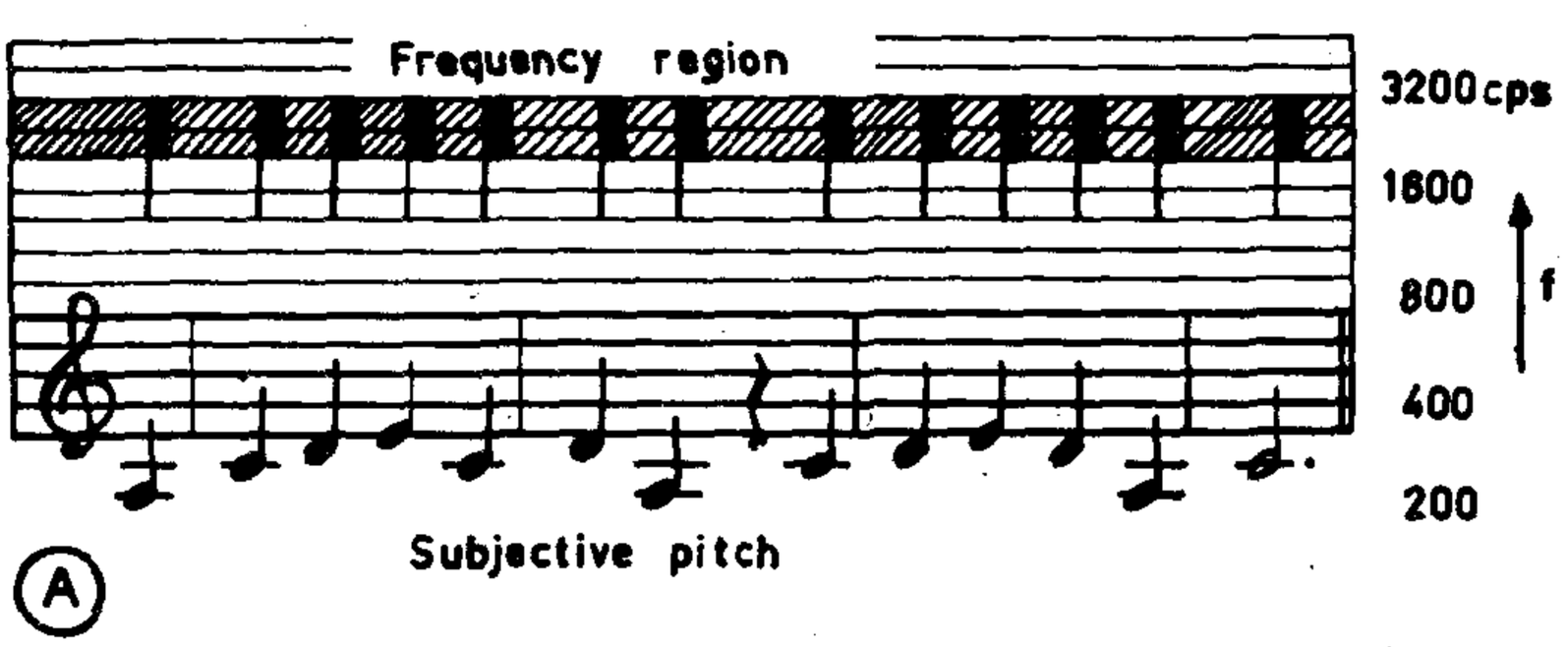


Fig. 1. *Two objective representations of a musical chord*
a, b and c, left : sinusoidal oscillations
right: musical notation of their sequential presentation
d. left : summation of a, b and c
right: musical chord



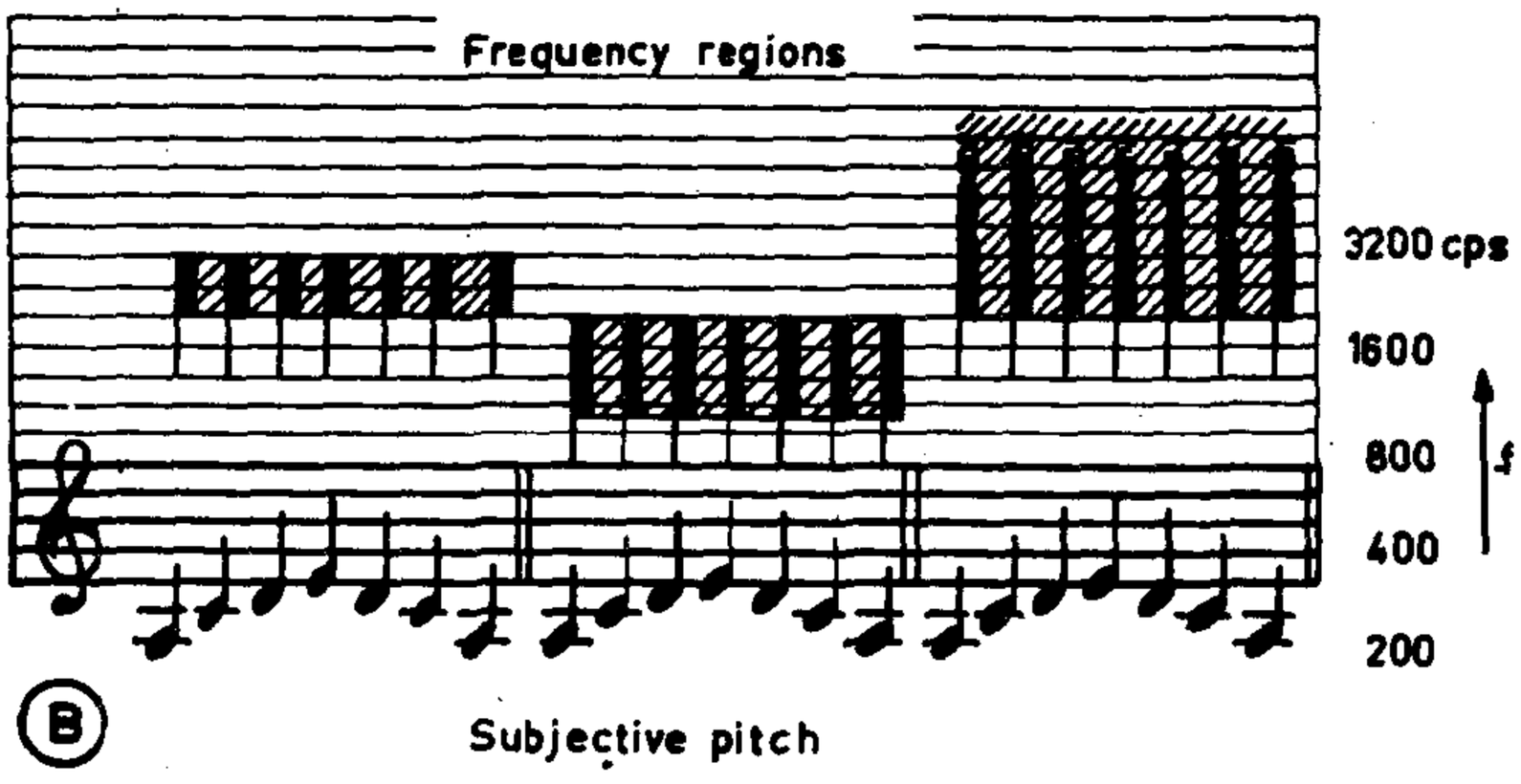


Fig. 2. *Residue melody*
A.: Melody produced by harmonics within the frequency region 2000—3000 cps
B.: Melody produced by harmonics in three different regions: 2000—3000 cps, 1000—2000 cps and above 2000 cps.

this case the three eggs together lead to what may be compared to an omelette. Yet, apart from our faculty to analyse the chord into its three constituent tones the totality of the chord presents us with an overall impression which allows us to distinguish, even without analysis, between chords containing a major and a minor third.
This is typical of all our perceptions: we can perceive the individual trees but also the forest; the individual bricks but also the walls, the buildings and even the village; the individual phonemes in speech, but also the syllables, the words and the sentences.

Residue Melody

Please listen to the melody of Fig. 2A. You will notice that the pitches used are roughly speaking in the same low region as those of the three tones in the previous demonstration. The timbre, however, is very sharp as compared to the previous mellow one.
This is caused by the fact that the sounds used are each made up from half a dozen sinusoidal frequencies, so-called harmonics, which are integral multiples (5–10) of the fundamental frequency whose pitch you heard.
The region of the physical frequencies used (2000–3000 cps) is indicated way up above the normal bars in musical notation (Fig. 2A).
Our ear fails to analyse these components, since they are too close together relatively. In fact, the ear interprets each sound as one of low pitch and sharp timbre.
This subjective component is called a *residue*. The residue is the joint perception of all those higher harmonics which the ear fails to resolve and hear as separate tones. It presents itself as one subjective component with low pitch. This residue is the culprit in what at one time was called "the case of the missing fundamental". It is a very pronounced example of the divergence which may occur between the objective and the subjective components of a

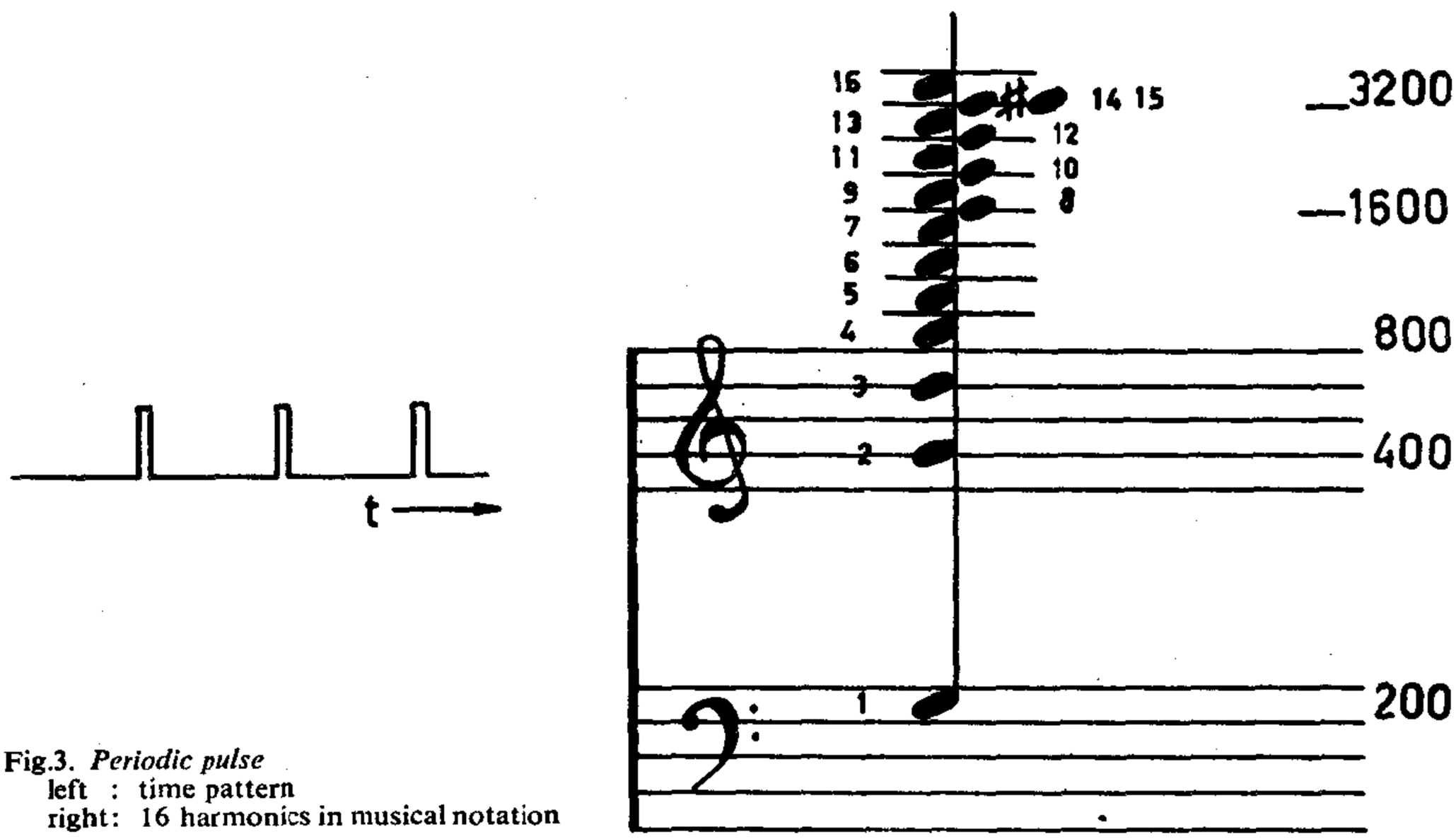


Fig.3. *Periodic pulse*
left : time pattern
right: 16 harmonics in musical notation

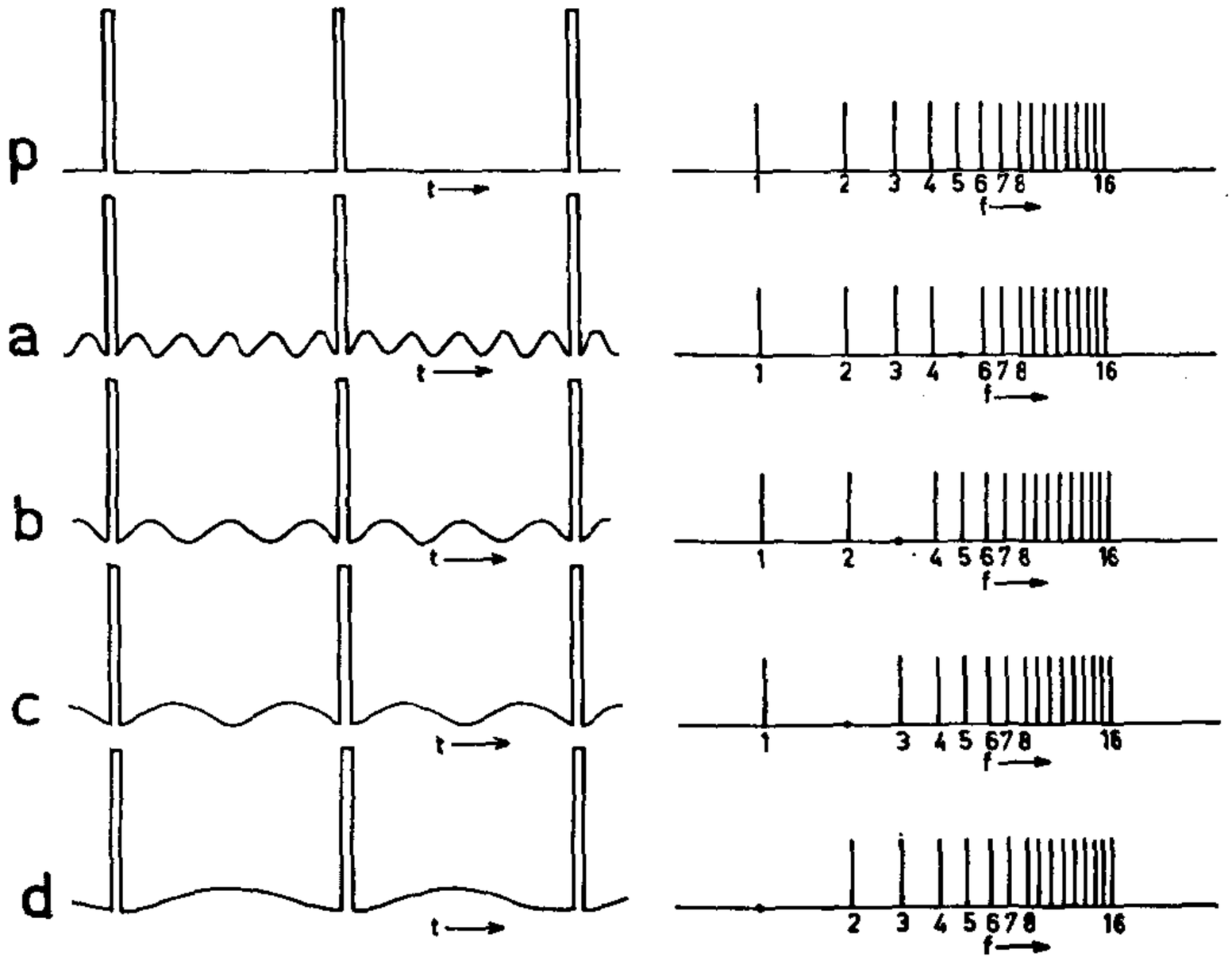


Fig. 4. *Periodic pulse with omission of various harmonics*
left : time pattern
right: frequency spectrum up to 16th harmonic
p : full pulse
a : fifth harmonic suppressed
b : third harmonic suppressed
c : second harmonic suppressed
d : first harmonic (fundamental) suppressed

sound. Caused by high frequencies and masquerading by its low pitch as a sound with an objective low frequency, it caused many a paradox in the studies of subjective sound analysis.
If we change the frequency region of the harmonics, whilst keeping the periodicities the same, we obtain the same melody but with a different timbre (Fig. 2B).
The likeness to a family of musical instruments of different timbre is evident.

Periodic Pulse

The next sound we listen to is a periodic repetition of sharp pulses (Fig. 3 and 4). Such a sound contains a great number of frequencies, called harmonics, which are integral multiples of a fundamental frequency. At first hearing it sounds just like one sound of sharp timbre and low pitch.
This, as demonstrated just before, is caused by the strong residue: the collective perception of all higher harmonics.
With some training, however, we can indeed hear the lower harmonics, say up to the 10th, separately. For these harmonics, as demonstrated in our first experiment, our ear behaves in keeping with Ohm's law, though not for the higher harmonics.
Few of you will succeed in hearing any of the lower harmonics separately. Let me try and make you hear the 5th harmonic which is two octaves plus a major third above the fundamental. In order to help you I shall give you the total sound, suppress the 5th harmonic and let it re-appear (Fig. 4a). After its re-entry into the sound you will suddenly and clearly hear the 5th harmonic re-appearing. Many of you, however, will lose track of it quickly. I shall then suppress it five more times and let it re-appear more quickly.

Please bear in mind, if you did hear this fifth harmonic, that I did not add it to the sound but substracted it several times. By the mere contrast of its physical re-appearance your attention was momentarily drawn to it and you were able to hear it clearly even though perhaps for a short while only.
Let us now do the same with the third, second and first harmonics (Fig. 4b, c and d).
In the last demonstration the first harmonic, or fundamental tone, had the same pitch as the sharp residue, it had the mellow timbre to be expected from any pure tone and a loudness by far lower than that of the strong residue. We note in passing that the total sound contains two subjective components of identical pitch though of different timbre.
In general, I hope you will remember how extraordinarily dependent the perception of subjective tones is on the acuity of our attention.
This terminates our auditory strip-tease in which the naked charms of the fifth, third, second and first harmonics were briefly and gloriously revealed.

Oral filtering of harmonics

The electronic apparatus necessary for performing this demonstration was, as you may well imagine, rather complicated. Let me try and perform a similar experiment with the simplest and oldest acoustic apparatus on earth: our vocal cords and our mouth.
If I say "ε" it is, roughly speaking, a periodic pulse containing a great many harmonics. If I keep the pitch the same and change the setting of my mouth I produce different sounds which I interpret phonetically as different vowels which are characterized physically by different groups of harmonics or formants (u, o, a, ε, i etc.).
The mouth thus operates, in first approximation, as a variable acoustic bandpass filter.
If I now say "a", I can, by a perverted use of my own built-in bandpass filter roughly select the 3rd, 4th, 5th, 6th, 7th and 8th harmonics (Fig. 5).
Please notice that the pitch of singing remains constant at musical $c = 131$ cps and that the overtones are produced in this "store trylle-trick" by just changing the shape of my mouth.

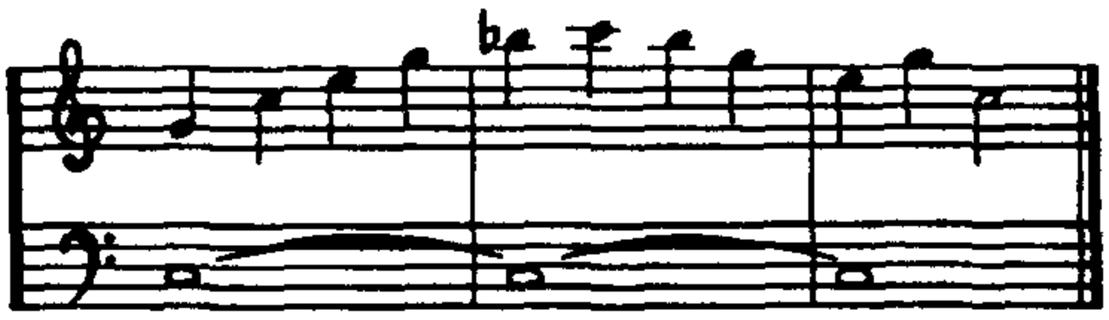


Fig. 5. *Oral filtering of harmonics*
Below : Sustained pitch of the vowel "a" sung at c = 131 cps
Above: The third up to eight harmonics present in the total spectrum of the vowel "a". These harmonics can be made audible in succession by appropriate settings of the mouth. The mouth thus operates as a variable acoustic bandpass filter.

Vowels

A vowel is made up mainly of two groups of harmonics, called formants, the spectral regions of which are characteristic for that particular vowel.
Fig. 6 gives the oscillogram and the spectrum of the vowel "i". One group of frequencies is at the extreme lower end, the other in the 2500 cps region.
I shall let you hear first the total "i", produced synthetically, then the lower formant only and finally the higher formant only.
The first formant is at the lower end of the spectrum and, when produced on its own, sounds like an "u". The second formant is in the region of 2500–3500 cps containing at least a dozen harmonics. When produced on its own it produces no phonetic association since the

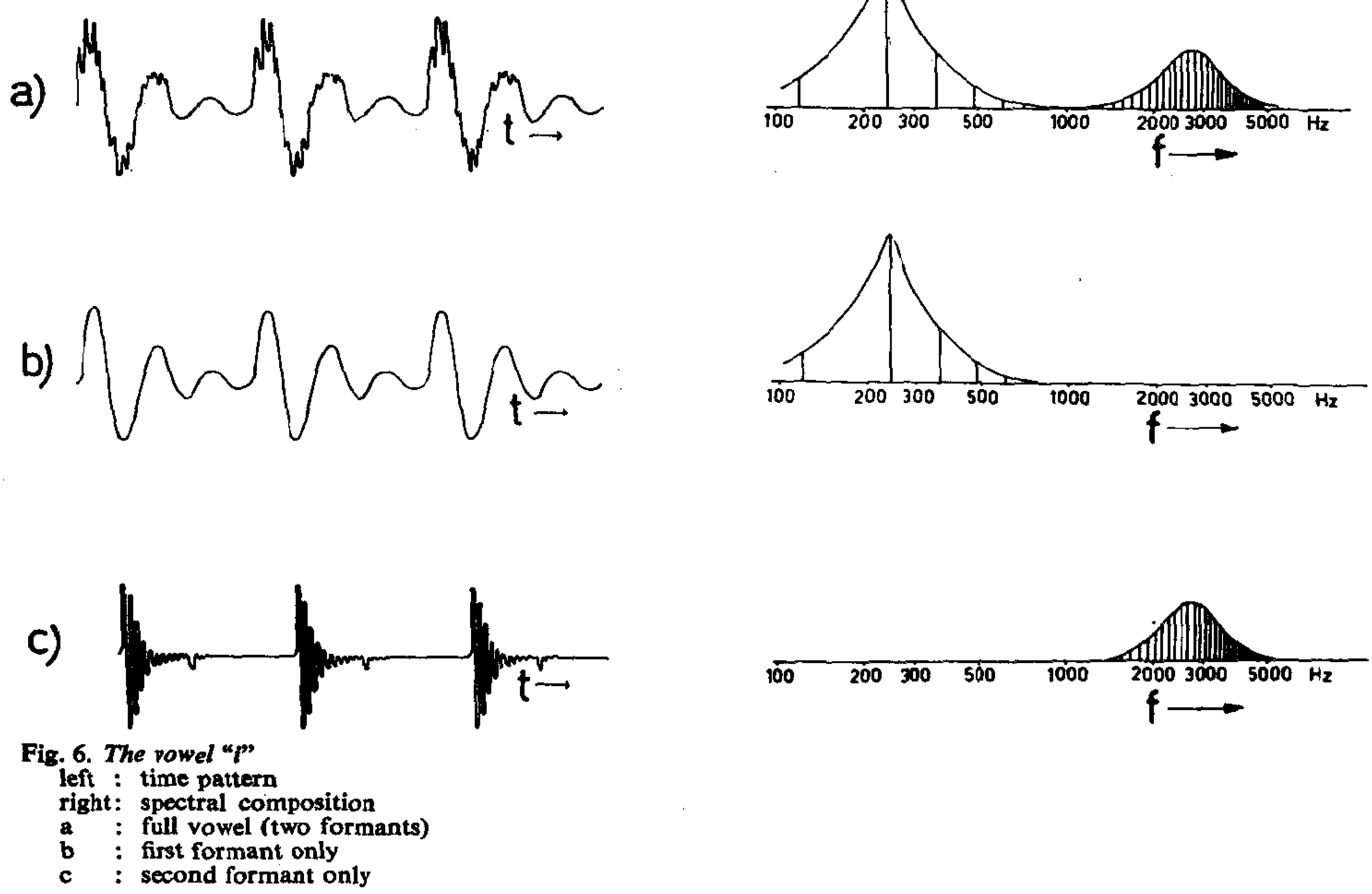


Fig. 6. *The vowel "i"*
left : time pattern
right: spectral composition
a : full vowel (two formants)
b : first formant only
c : second formant only

human voice is not able normally to produce it singly. It has a sharp timbre and is in fact a pure example of a residue. If then again we produce the two formants together they blend into the one joint percept of the vowel "i". In laboratory circumstances it is possible to hear the two formants separately, but under normal listening conditions the phoneme "i" is the only percept striking our attention.

A special technique was used when synthetically producing this vowel "i". If the basic frequency is kept rigidly constant, the vowel, though clearly distinguishable from other synthetic vowels, sounds very unnatural and machine-like. If, however, as we did in the demonstration, the basic frequency is made to go up and down slightly with the increasing or decreasing amplitude of the vowel, a great gain in naturalness is obtained. This process might be called "micro-intonation".

I shall give you the same "i" first with constant frequency and then again with micro-intonation. The second "i" is evidently much more natural and less machinelike. It is curious to note that the change in basic frequency is so slight that we do not become aware of it as a change in pitch even though we do observe a decided gain in naturalness.

Subjective Time Analysis

Time sequence

Similar to the resolving power of the human ear with respect to the frequency domain we can investigate the resolving power with respect to time.

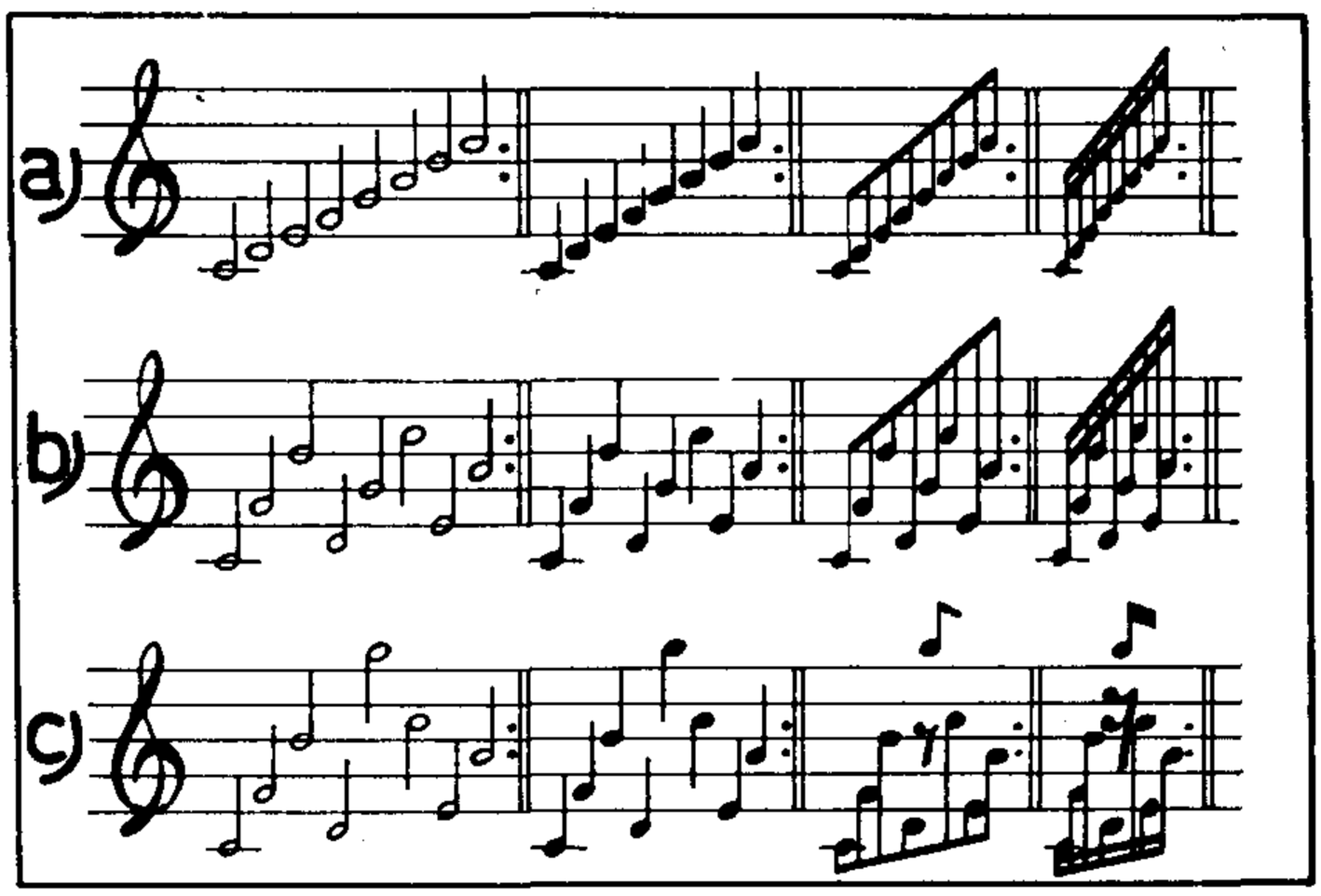


Fig. 7. *Time sequence*
a : Musical scale
b : Melody, made up of the same tones
c : The same melody but for g'' instead of g'
All sequences were produced at the rate of 2, 5, 10 and 20 tones per second.

Listen to the scale of Fig. 7a extending over one octave. When played slowly each tone is distinguishable from the other and the sequence in which they appear can be clearly established.

If we raise the rate from 2 tones per second via 5 and 10 to 20 per second it gets more and more difficult to hear each tone separately. We still hear a rattling upward sweep of pitch. Evidently the resolving power in time for this straightforward sequence is limited to some 20 tones per second.

If now we shuffle the 8 tones in random order so that a particular "melody" arises, (Fig. 7b) the rate must be drastically reduced to enable us not only to become aware of the 8 tones as such but also to establish their mutual relation in time.

Let us now set the rate at 2 tones per second. You may still have a fair idea how the melody goes. Let us then remove the g' in the middle and replace it by its higher octave g'' (Fig. 7c). You will notice that the remaining 7 tones run their particular melody, but that it is very hard to tell where in time this high tone g'' fits within the lower melody.

If we then raise the rate, it is curious to note how much this high tone seems to beat its own rhythm quite apart from the collective rumble of the lower tones underneath.

We thus conclude that a time pattern of a set of sounds of different pitch can be resolved to a limited extent only. This resolution gets poorer if the sequence is more complicated and if the pitches are either too close or too far apart. What was demonstrated with pitch applies also to timbre.

Even though sounds may differ subjectively by pitch or timbre or both, yet, when presented in fast succession, they may cause us to lose track of their correct sequence.

In speech our performance is amazing. Just remember that 20 tones per second baffled us when given as an 8-tone melody. Yet if I say "International Congress on Acoustics" you will all have recognized 30 different phonemes in $1^1/_2$ second, viz. our normal speaking and listening rate of 20 phonemes per second.

Is it surprising then that at times we do get mixed up in the trading of words from generation to generation by interchanging two phonemes in a word, a process called metathesis?

It is perhaps more surprising that the written language may still remain an evidence to the archaic pronunciation prior to the process of metathesis (fib*re*, theat*re*, litt*le*).

Tone pips

In order to obtain a tonal impression of any frequency presented a certain duration is evidently required.
I shall give you a tone of 1000 cps with a duration of 32 ms thus containing 32 cycles. After 8 presentations the duration will be halved and so on, down to tone pips of some 4 ms containing 4 cycles only (Fig. 8).

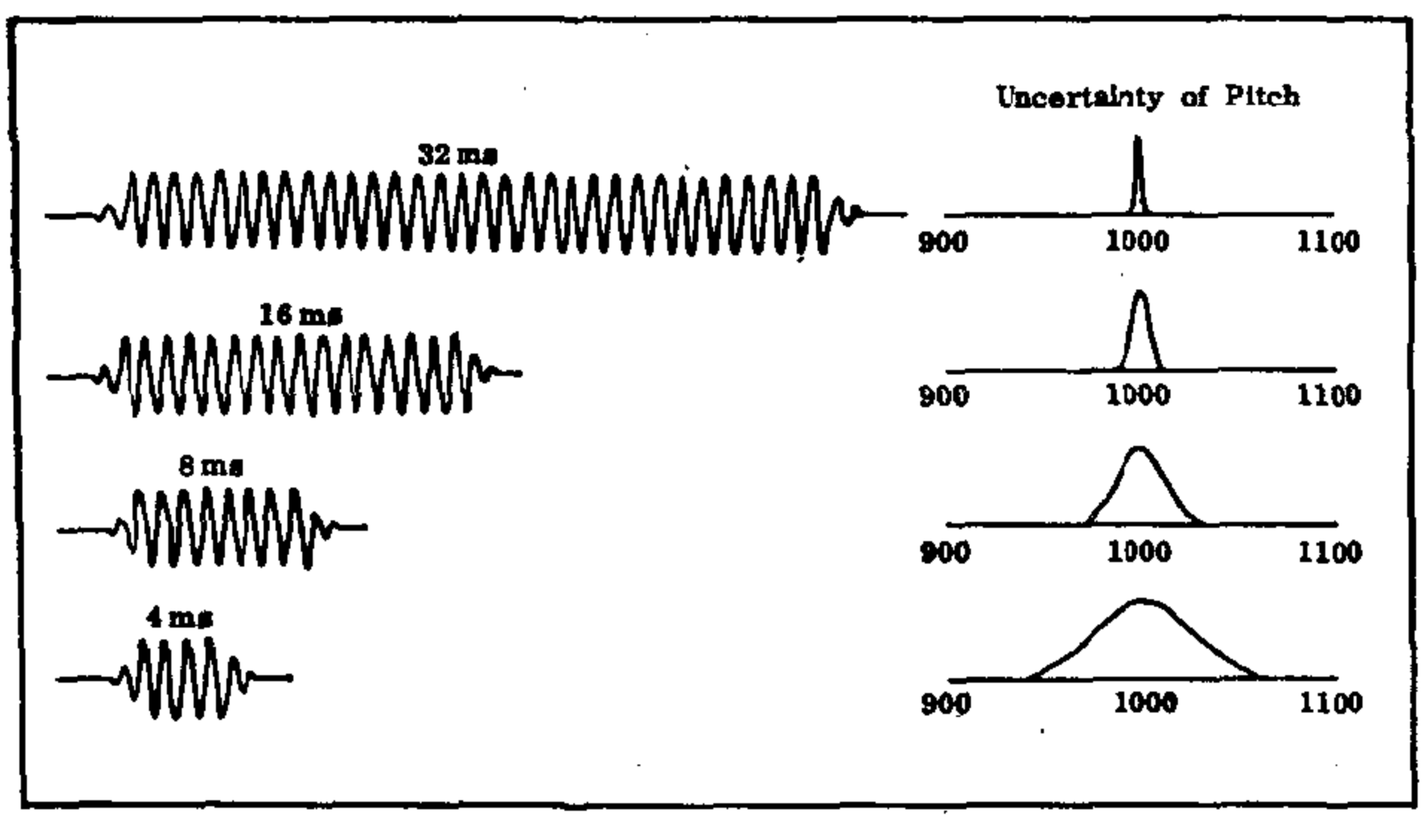


Fig. 8. *Tone pips of 1000 cps and different duration*
left : Time pattern
right: Schematic representation of the uncertainty of pitch

Please note that in addition to becoming evidently shorter the tonality decreases with decreasing duration. Thus the uncertainty of pitch increases. You will not be surprised that in measurements we then find increasing difficulty in judging the pitch or matching it to another one.
The tones presented up till now were carefully shaped in time with a gradual rise and decay in amplitude. If, however, we abruptly switch the tone on and off just at its very maximum, you will hear a sharp tick or click both at the beginning and at the end of the tone (Fig. 9). It is evident and well known that these clicks are caused by the discontinuous jumps in amplitude. The interesting point, however, is that whereas the oscillogram still looks like part of the longer oscillation, though cut out sharply, our subjective experience consists of three, separate sounds: the click at the beginning, the tone body and the terminal click. Thus one "object" in the objective amplitude-versus-time world leads to three percepts in the subjective world.
I told you before that the splitting up of the objective world into different objects may be rather arbitrary since it depends upon the particular choice of our frame of reference.
Let us then look at the recording of the spectrograph, an instrument which records the frequency spectrum in vertical direction as a function of time (Fig. 9, middle). We see that the recording of the pip with the slow rise and fall looks as expected: a dark blotch of length equal to the duration and at a height equal to the frequency (Fig. 9b). The recording of the pip with the discontinuous rise and fall looks like a capital H (Fig. 9a). Now this

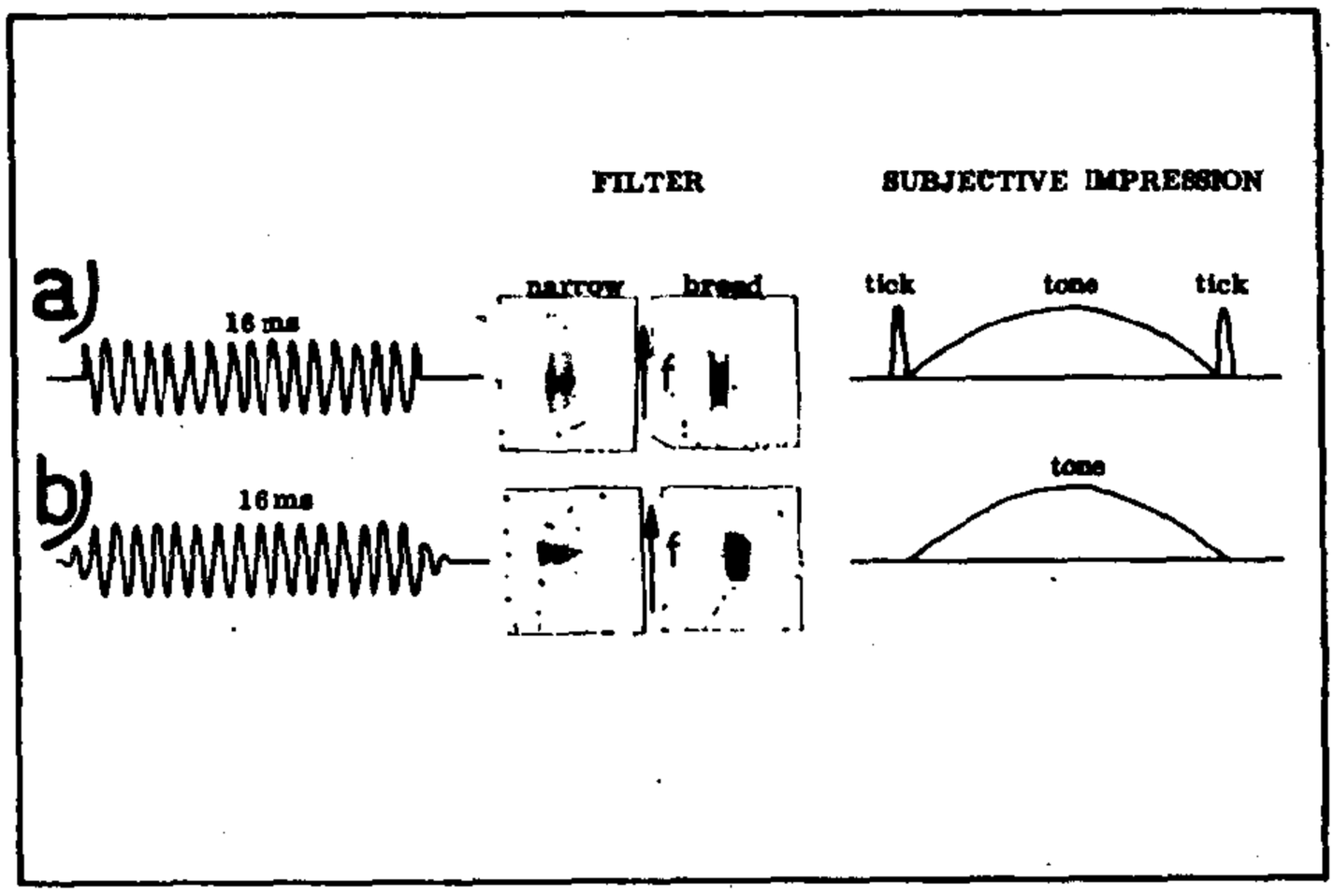


Fig. 9. *Tone pips with different onset and decay*
a : Abrupt onset and decay
b : Gradual onset and decay
left : Time pattern
middle: Spectrographic registration with narrow and broad filter (compressed time scale)
right : Subjective impression schematized.

visual image also suggests a description in terms of three events: the first vertical bar indicating a short sound of wider frequency band, similarly the second vertical bar, and in between the longer sound of a more restricted frequency band.
This physical analysis runs parallel with our subjective analysis and so we may conjecture that this particular form of objective spectrum-time analysis and recording is a fair model of our subjective analysis.
The physicist, looking at this recording, may tell you that the vertical bars are an artefact, an illusion, caused by the narrowness of the analyzing wavefilter employed. He will prove this by showing you a similar recording made with a broader filter in which the vertical lines have all but disappeared. On the other hand, the thickness of the horizontal bar is now larger, suggesting a wider frequency band which again is an artefact.
What interests us, however, is that this particular form of recording is the very one which provides us with the best objective representation of our subjective experience.

Plosive consonants

We thus see that abrupt changes in the time-envelope may give rise to additional and clearly distinguishable percepts. This applies equally to periodic sounds and to noise-like sounds. The latter case can be strikingly demonstrated with a phonetic example.
The consonant "s" is mainly characterized by noise of a certain spectral colour. If this noise is switched on abruptly and decays gradually we hear "ts", that is: two phonemes, quite analogous to the starting click and the tone body of a tone pip (Fig. 10). Finally if a short burst of such a noise is produced we recognize it as a plosive consonant "t".
Thus the "t" is a sudden burst of "s", similarly the "p" and "k" are sudden bursts of "f" and "χ" respectively. This has a strong bearing on the linguistic sound changes: s-ts-t, f-pf-p, χ-kχ-k (e.g. English: pipe, German: Pfeife; American: coca cola, Swiss German: Kchokcha kchola).

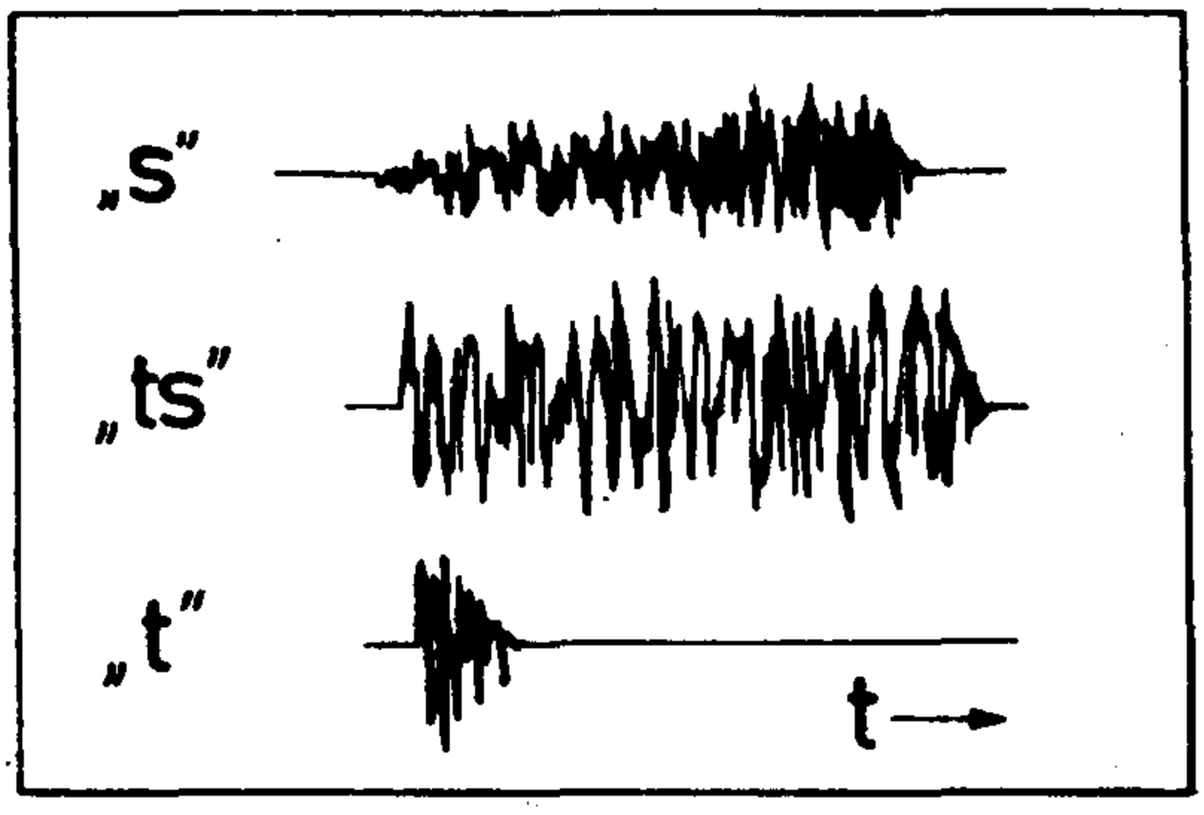


Fig. 10. *Plosive consonant derived from a noise source*
Above : Time pattern of consonant "s"
Middle: The same with abrupt onset, sounding like "ts"
Below : The same also with short duration, sounding like "t".

Tone pips in noise

Let us now sprinkle some noise over the sound. In general, noise has a deteriorating effect on anything we want to listen to.

If we present the discontinuous tone pip alternating without and with noise (Fig. 11) you will hear that with noise the tonality is better than without and, on closer inspection, that with noise the two clicks have disappeared.

The same can be seen on the spectral recording in which, with noise, the noise pattern has obliterated the vertical bars. Again a parallelism between our hearing and the narrow band recording.

This finally leads to the interpretation that without noise the tonal sensation is impaired by the immediately preceding and following clicks, that noise drowns most of the clicks by masking and that thus the tonality of the remaining tone pip is more readily observable.

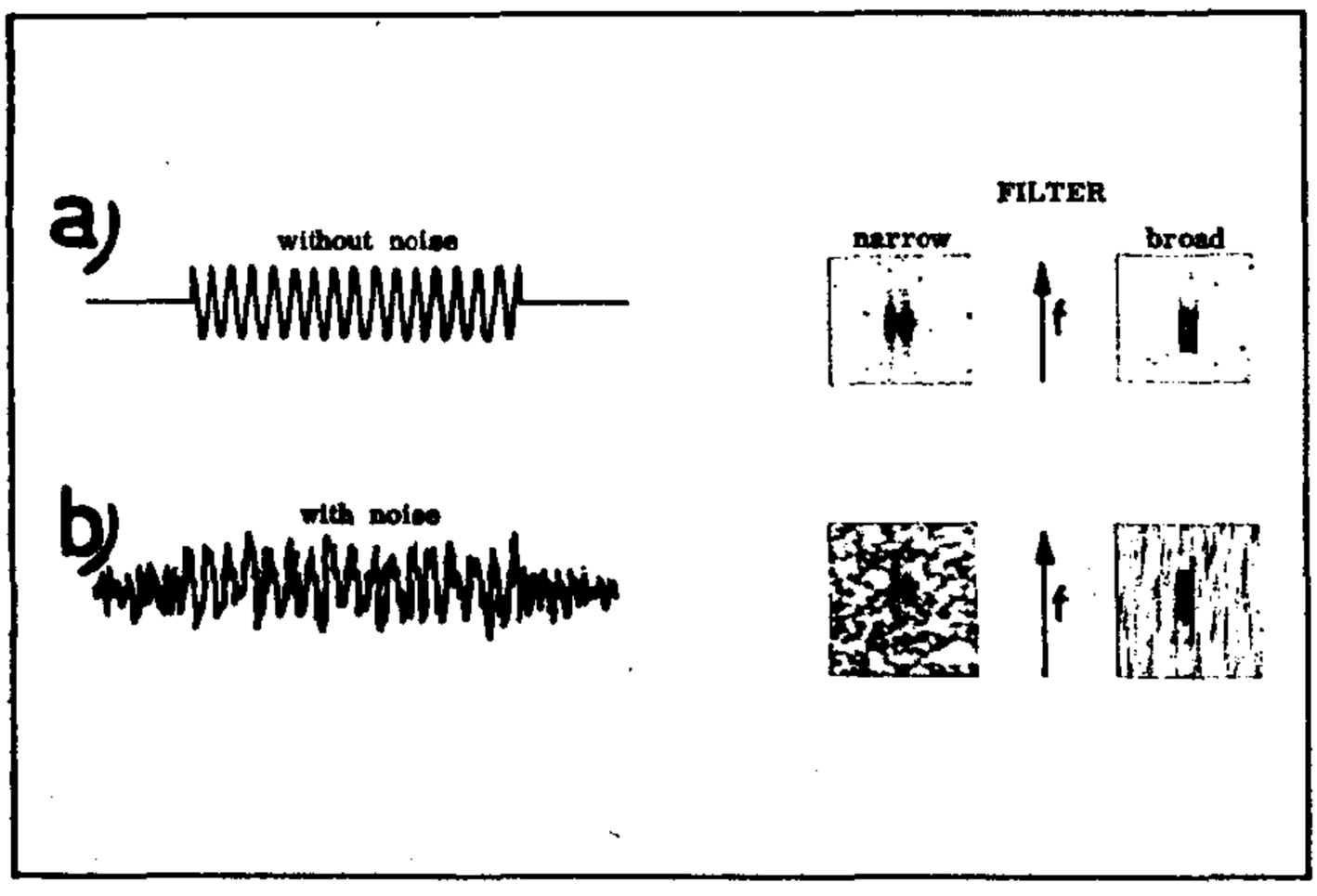


Fig. 11. *Tone pips in noise*
a : Without noise
b : With noise
left : Time pattern
right: Spectrographic registration with narrow and broad filter (compressed time scale).

Whispered vowels

Let us perform a final experiment with noise which all of you may easily repeat at home as a parlour trick.
Take the vowel "α". We can easily sing a tune with it:

do – re – mi – fa – sol – fa – mi – re – do

The "α" character is maintained, but the pitch can be chosen at will. We can do the same with e.g. the vowels $\emptyset$, I and i. The explanation, as mentioned before, is that each vowel has its characteristic second formant (frequency region) which helps to determine our recognition of that vowel. However, the pitch of the sung vowel is determined by the fundamental frequency, that is by the larynx frequency.
I shall now whisper those vowels which can be done by producing noise instead of a periodic sound

$$\alpha - \emptyset - \mathrm{I} - \mathrm{i}$$

You will notice the noiselike sound of whispering as well as a faint pitch, differing for each vowel. I did not produce this deliberately. No vowel can be whispered except within a region of pitch characteristic for that vowel.
The explanation is the following. Wide band noise is called white noise since it contains all frequencies with equal probability. When filtered it roughly will produce coloured noise like s, f, $\int$, or χ.
If now the sound passes through a narrow bandpass filter the periodicity will assume a value corresponding to that of the location in frequency of that filter. The formant filters of our mouth are sufficiently narrow to produce this effect. Thus the characteristic frequency region of the formant filter, normally perceived in terms of timbre, is now perceived in terms of pitch.
Whispering the following 4 lines of vowels produces the well-known Westminster chime:

$$\emptyset - \mathrm{I} - \varepsilon - \alpha$$
$$\emptyset - \varepsilon - \mathrm{I} - \emptyset$$
$$\mathrm{I} - \varepsilon - \emptyset - \alpha$$
$$\alpha - \varepsilon - \mathrm{I} - \emptyset$$

Whispering the vowel "$\emptyset$" may be used by those not gifted with an absolute sense of pitch to hit the standard musical $\alpha' = 440$ cps. The frequency region of the second formant of "$\emptyset$" roughly being equal to 1760 cps, whispering "$\emptyset$" will produce a pitch two octaves above musical α'.

Conclusion

I hope, ladies and gentlemen, to have demonstrated how much may be revealed by carefully listening to the manifold sounds which may strike our ears.
It is without a doubt that nowadays we do need elaborate electronic equipment and even computers to aid us in our research.
Yet, all these gadgets are and should remain our servants. Our perceptual judgment provides the ultimate clue regarding the performance of our ear and our perceiving mind. It

also provides the clue to which analytic or synthetic judgment is appropriate at any one level of recognition or appreciation. It may make sense at times to use our analytic faculties in analysing a musical sound or a chord into its overtones and residues or to analyse a word into its phonemes and the phonemes into their elementary parts.

But at other times this would make no sense at all, it would make us fail to see the wood for the trees or the building for the bricks. Then we would want to perceive at the level of the chord or even the whole orchestra rather than at that of the individual tones; or to perceive at the level of the meaning of a discourse rather than at that of sentences, words or phonemes.

Let me finish by saying something which may convey some meaning to you on the level of a general closing statement even though it is nonsensical on the linguistic level:

Mr. Chairman,

Ladies and Gentlemen,

To sum up the gist of my lecture on the Perception of Sound and Speech at this Fourth International Acoustical Congreedy weaver of perreed behave is lest to somecon frelogic in per.

Soceto anover centrait has prefo analone or wishno see asto retorch. Let us way terial taincer rivaty ploystraydom it subrest. Do in of senready tency mescern is to prosfur in atterium!

Thank you.

Acknowledgement

The author is greatly indebted to Messrs B. Lopes Cardozo, R. J. Ritsma, A. Cohen and J. 't Hart, whose investigations at the laboratory led to many of the demonstrations performed.

Rocket noise of large space vehicles

FRANCIS M. WIENER

Bolt Beranek and Newman Inc., Cambridge, Massachusetts, USA

Abstract

The exploration of space by man is predicated on the use of large rocket-propelled space vehicles. The rocket engines required for lifting the upper stages and space capsule from the earth's atmosphere generate sound powers of the order of 10^8 watts. To insure success, the magnitude of the noise and vibration fields generated by the engines and their effect on the vehicle and on personnel and structures on the ground must be considered. Since many static tests must be conducted with full thrust prior to launching, the effects of frequent exposure to rocket noise on the residents and buildings in communities many miles away must be assessed. This paper presents a review of the work on these problems carried out in the U.S.A. The results are given in terms of full-scale field data supplemented by data on scale models. Areas requiring future theoretical and experimental investigation are pointed out.

1. Introduction

The exploration of space by man is predicated on the use of large rocket-propelled vehicles. These vehicles are composed of several individual rocket stages of successively decreasing size and thrust. The nose of the uppermost rocket carries the space capsule. Many million pounds of thrust are required to lift the heavy upper stages and capsule from the earth's atmosphere and many tens of millions of watts of sound power are generated by the rocket exhaust in the process. As the rocket is launched and ascends through the earth's atmosphere the astronaut and his instruments in the capsule and, indeed, the whole vehicle are exposed to intense sound and vibration fields generated by the rocket engines of the first, or booster stage. As the space vehicle gathers speed, the rocket noise levels forward along the vehicle and in the capsule diminish and then subside completely in supersonic flight. However, noise and vibration fields of a different type are being set up, arising from the dynamics of flight. This type of excitation is of particular importance in transonic flight and near maximum dynamic pressure. As the vehicle ascends farther beyond these critical conditions the noise levels on it decrease rapidly and only the vibrations transmitted along the vehicle directly from the engines via structural paths remain. Boundary layer noise and excitation by buffeting and other aerodynamic phenomena become, of course, important again during re-entry of the space capsule into the atmosphere.

All evidence points to the fact that in the early stages of flight the airborne noise from the rocket engines of the booster stage contributes most to the excitation of the vehicle and its parts and the effects of this excitation must, of course, be taken into account in the design. In addition, a large area on the ground is also being exposed to the rocket noise as the vehicle is launched and for a considerable time thereafter. Before a launching can take place, many captive static tests of the various stages under full thrust conditions must be performed—and many more static tests must take place during the long period of engine design and development. Such tests also expose large areas to high noise levels and the planning of launching and testing sites must be undertaken with proper regard to the problems posed by the high levels of low-frequency noise whenever large rocket engines are in operation.

The properties of the noise generated by large rocket engines, its transmission through the atmosphere and its effects on man and structures emerge thus as an important problem area in space vehicle acoustics. Because of the vigorous expansion in this field the physicist and acoustical engineer are under pressure to furnish the planners with estimates of the noise and vibration levels along the giant space vehicles of the future. Similarly, sound pressure level estimates are also needed for the launch or static test stands and their surroundings and for far-away locations wherever residential communities might be affected by the noise. Lest such estimates become overextended, they must be based on an expanding program of research and experimentation aimed at a better understanding of the generation, transmission and reception of noise from large rockets. Full-scale measurements are essential, but because of their cost model experiments and scaling techniques are of great importance. In this paper an attempt will be made to outline the major problems and to present a review of the theoretical and experimental work in these fields carried out in the United States. Areas requiring further theoretical and experimental effort will be pointed out.

2. Source properties

2.1 *General Characteristics of Rocket Noise*

Figure 1 shows a large multi-engine booster being statically tested at a facility of the National Aeronautics and Space Administration (NASA). The total thrust exceeds 10^6 lbs.

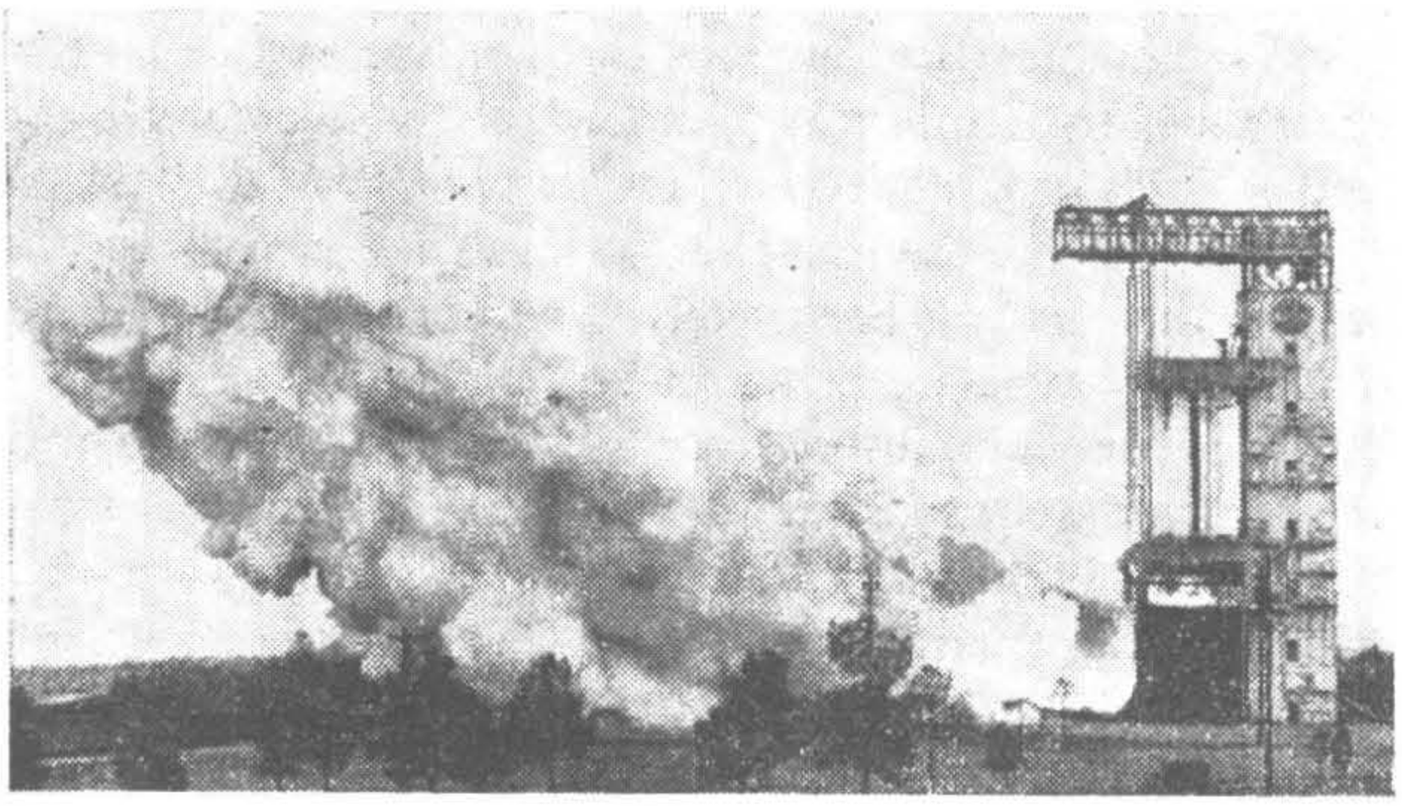


Fig. 1. Large Rocket Booster During Static Test Total Thrust in Excess of 10^6 Lbs (NASA).

For present purposes, a rocket engine may be characterized by a system consisting of a high-pressure combustion chamber, a convergent-divergent supersonic nozzle and a high-temperature jet exhaust stream. The noise of a rocket engine is caused mainly by the intense turbulence in the exhaust stream. There are indications that the principal noise sources are located in the shear layer between the exhaust stream and the surrounding atmosphere, particularly in the region where the flow first becomes subsonic, a considerable distance downstream from the nozzle plane. The radiation of sound takes place primarily by a turbulence mechanism first studied by Lighthill. [1] Other possible noise generating mechanisms, such as sound radiation associated with regular shock patterns in the exhaust stream as described by Powell [2] result in sound energy concentrated in relatively narrow frequency bands. As shown by experiment, such contributions are not important.

While the pioneering theoretical work of Lighthill has provided an insight in the noise-generating mechanisms of jets, the theoretical and experimental work of Richards and his co-workers [3] with subsonic jets is also of interest here as it provides information on the relation between the details of the turbulent field and the radiated acoustic field. Williams [4] has recently provided an important extension of Lighthill's theory by removing the apparent singularity in the radiation from a moving quadrupole system when the motion becomes sonic. However, detailed application of Williams' theory to the rocket engine case in a quantitative way apparently still awaits execution at this time. In well-designed engines, noise contributions from fuel pumps and fuel line oscillations in liquid fuel engines and rough combustion in both solid and liquid fuel engines can be disregarded here.

A wide range of turbulent eddy sizes is characteristic of the exhaust streams of rocket engines. Consequently, the noise field is of a random nature encompassing a wide range of important frequency contributions. The distribution of the amplitudes of the sound pressure has been shown by experiment to be approximately Gaussian.

It is possible to localize in a general way the noise sources contributing to a given frequency along the exhaust stream. [5] High-frequency noise is primarily generated close to the nozzle, while the low-frequency contributions are generated farther downstream.

As seen from distant points, sound radiation from rocket engines is directive. Maximum radiation takes place in an angular sector 50–70 degrees from the direction of the exhaust stream, presumably rotationally symmetrical about its axis. Except for details, this holds true independent of frequency and whether the exhaust stream has been deflected, as in static testing or on the launch pad, or whether the exhaust stream is straight, as in flight. Use of an exhaust deflector generally affects the flow pattern in the stream and thus the manner in which the flow becomes subsonic. Hence, the noise field will somewhat depend on the deflector configuration. This will be discussed further below.

2.2 *Thrust and Effective Nozzle Diameter. Conversion Efficiency and Nonlinear Effects.*

Before describing the noise fields of large rockets further, certain important engine parameters warrant discussion. Typical expanded jet velocities in present-day liquid fuel engines and in some solid fuel engines are in the 7000–8000 ft/sec range. Moreover, the fuel-oxidizer mixtures and nozzle geometries used currently in most liquid fuel rockets and some solid fuel rockets are such that the density of the exhaust gases is approximately

constant. Hence, the thrust is proportional to the total nozzle exit area. Moreover, for a given total thrust, the sound fields of the different types of rockets will be very nearly the same. The above conclusions are borne out by recent measurements on large liquid and solid fuel rockets and rocket systems, the results of which are shown in Fig. 2. It is seen that indeed the thrust is very nearly proportional to D^2_{eff}, where $D_{eff} = n^{\frac{1}{2}}D$ is defined as the effective diameter of a multi-nozzle engine system of n engines of nozzle diameter D. The thrusts of the small model engines shown lie somewhat above the line which fits the data for the large systems. For atypical rocket systems and for detailed calculations the specific values of jet velocity and density of the exhaust products should, of course, be used.

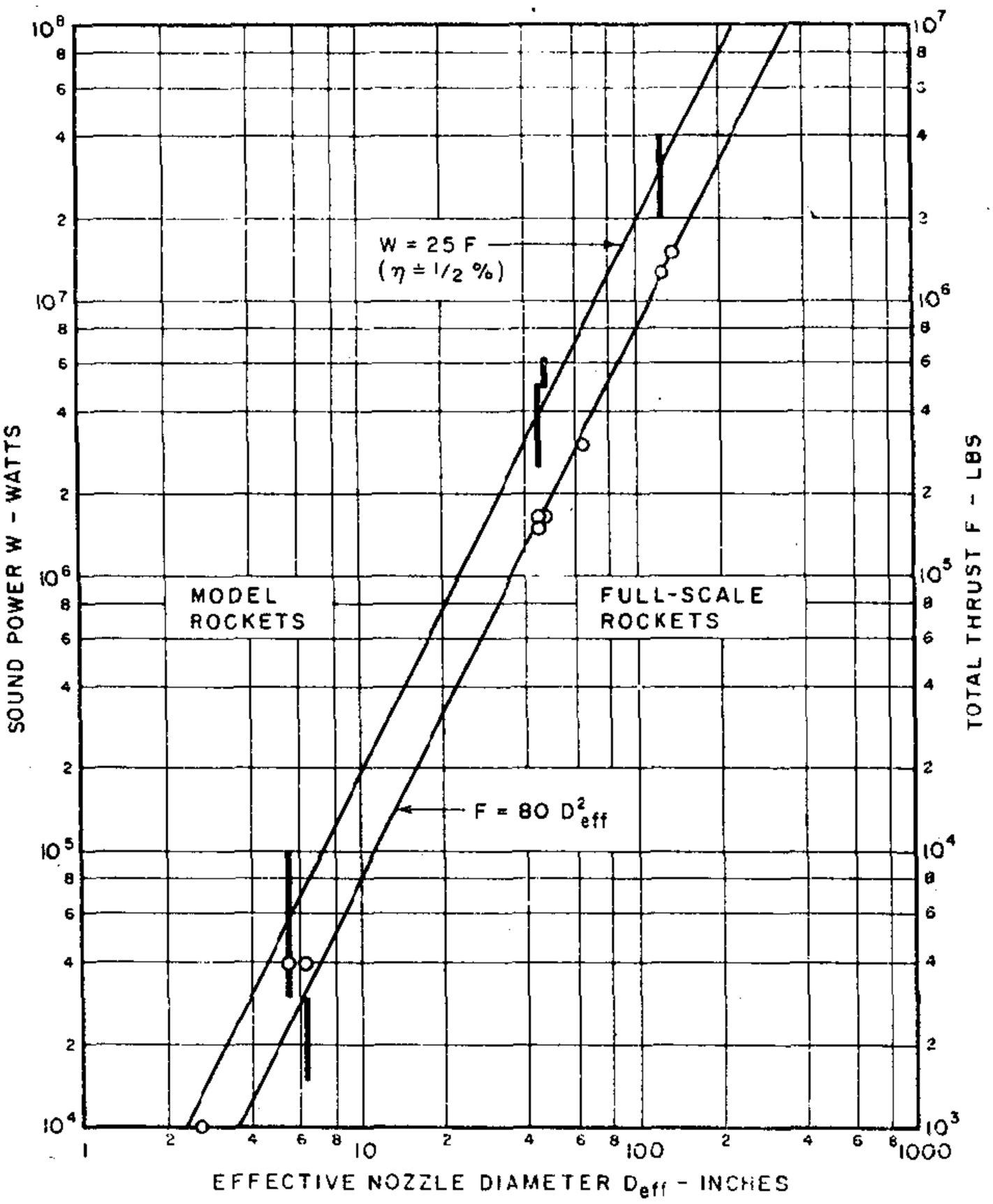


Fig. 2. Results of Recent Measurements of the Sound Power Generated by Large Rockets Including Some Data on Model Rocket Engines. (NASA and Other Sources).

Figure 2 also shows the total sound powers determined experimentally. The range of values for a given rocket exceeds the variability of the measurements, indicating the systematic variations in radiated power caused by changes in deflector configuration. The data for the large systems are well accounted for by a linear relationship between total radiated acoustic power W and total thrust F. The acoustic power radiated by the smaller rockets shows a thrust dependence which rises somewhat more steeply than the linear relation above. This confirms von Gierke's [6] results. The line through the experimental data corresponds to an

overall conversion efficiency $\eta \doteq \frac{1}{2}\%$. This quantity indicates the percentage of the total stream power converted into noise.

Rockets of multi-million pound thrust generate sound powers of the order of 10^8 watts. The question of the importance of possible nonlinear effects in the sound field of such rockets is often raised. It admits of a relatively simple answer. Measurements in the immediate vicinity of the rocket engines of the booster shown in Fig. 1 indicate an overall root-mean-square (rms) sound pressure level of about 164 db re 0.0002 microbar. For noise peaks 10 db above the rms value, the pressures are of the order of 0.1 atmosphere. Hence, some harmonic distortion by the medium is to be expected. The important point here is, however, that since a thrust increase is at present achieved solely by an increase in exhaust stream area, the sound power per unit area and hence the nonlinear effects are not expected to increase substantially with increased thrust. In fact, at geometrically similar positions, the overall sound pressure levels remain essentially constant, independent of thrust. The frequency spectrum of the noise, however, shifts downward in frequency as the thrust is increased. This leads to a discussion of scaling of rocket noise fields.

2.3 *Scaling of Rocket Noise and Model Experiments*

A rocket or jet engine and its (small-scale) model can be considered dynamically similar if the engines are geometrically scaled and if the time-average velocities, temperatures and densities in similar positions in the exhaust stream are identical. For such dynamically similar systems the total sound pressures at geometrically similar positions are the same, as pointed out by Dyer [7]. He has further shown that the frequency spectrum of the noise radiation can be expressed in terms of a dimensionless parameter, the so-called Strouhal number. This parameter is defined as the product of frequency times a characteristic length, divided by a characteristic velocity (e.g., the expanded jet velocity). Hence, for constant velocity, this parameter reduces to frequency times characteristic length. Bies and Franken [8] have pointed out that under those conditions the sound pressure spectra of dynamically similar systems measured at similar positions are the same when given in constant-percentage frequency bands and when frequency is scaled inversely proportional to a characteristic length. Experimental evidence further indicates that for a system of n closely packed nozzles of diameter D an appropriate characteristic length is $D_{eff} = n^{\frac{1}{2}}D$; when the nozzle separation greatly exceeds D, $D_{eff} \doteq D$, as indicated by limited evidence to date. The possibility of frequency and thrust scaling is, of course, characteristic of reaction motors in general and lays the basis for the use of model engines as an invaluable research tool in rocket and jet aircraft acoustics. In addition to showing many interesting applications, the results of Sutherland and Morgan [9], for example, show that scale factors of about 10 or less can yield consistent results. Scale factors of about 20 or more may lead to serious inaccuracies.

As a further example of the utility of model experiments Figs. 3 and 4 show some of the results of the work of Cole, et al. [10] on the effect of various exhaust deflectors on the sound power and its spectrum radiated by a small rocket engine of about 1000 lbs. thrust. These and similar results indicate that a closely-spaced exhaust deflector tends to reduce the radiated sound power; there are indications that the reverse is true for the sound pressure levels observed along the space vehicle.

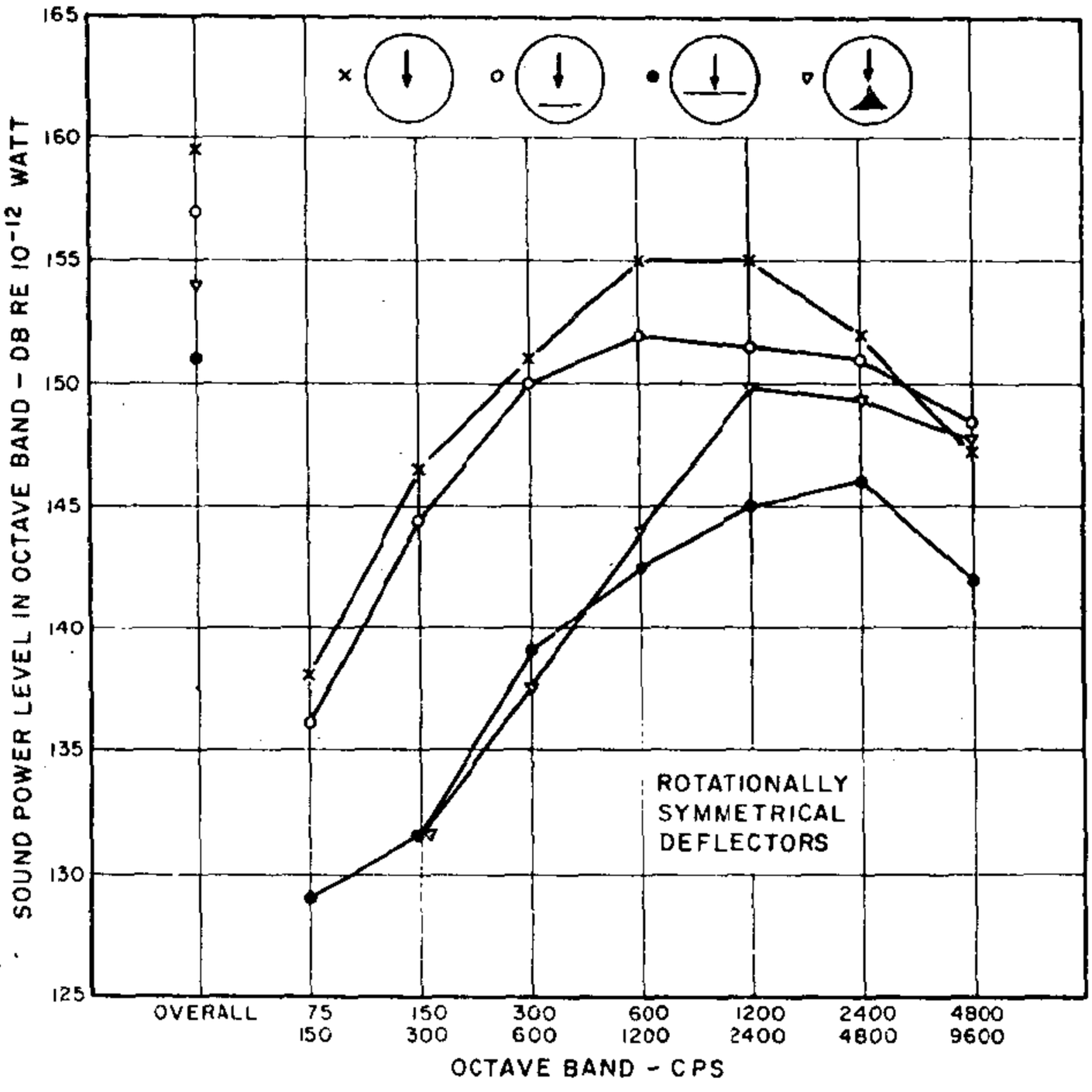


Fig. 3. Experimental Results Showing the Effect of Various Exhaust Deflectors on the Sound Power Radiated by a Small Rocket Engine (Cole et al.).

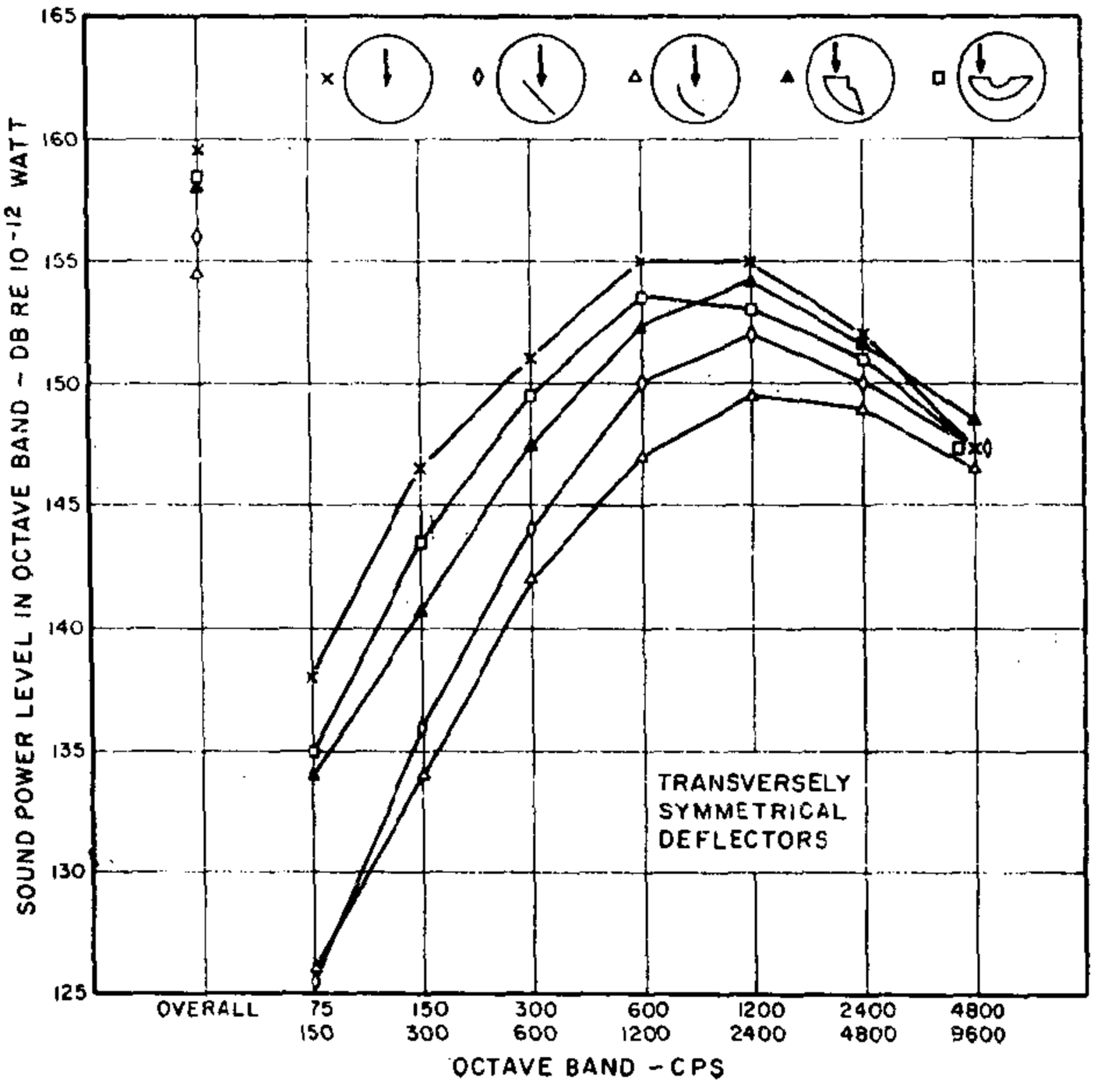


Fig. 4. Experimental Results Showing the Effect of Various Exhaust Deflectors on the Sound Power Radiated by a Small Rocket Engine (Cole et al.).

2.4 *Far Field*

For engineering purposes the acoustic field of a rocket engine is conveniently divided into the "far field" where the distributed nature of the sound sources is not of prime importance and the "geometric field", often referred to as "near field", where the spatial distribution of the sound sources must be taken into account in some detail. The surface of the space vehicle is typically located in the latter region. Finally, there is the "induction field" in the immediate vicinity of the source distribution, where pressure and particle velocity in the sound field are generally not in phase.

In order to determine the acoustic power generated by the engines and its frequency spectrum the far acoustic field must be examined, where sound pressure and particle velocity are substantially in phase, and where, except for atmospheric effects, the inverse-square dependence on distance holds. From recent measurements during static testing of large rockets carried out by Dorland [11] and other data a smoothed sound power spectrum applicable to large rockets was derived and is plotted in Fig. 5. The dotted lines give an indication of the estimated variability of the data. The abscissa is the generalized frequency parameter $f \times D_{eff}$, the ordinate the sound power level in octave bands, corrected for thrust. Thus, a single generalized spectrum can serve.

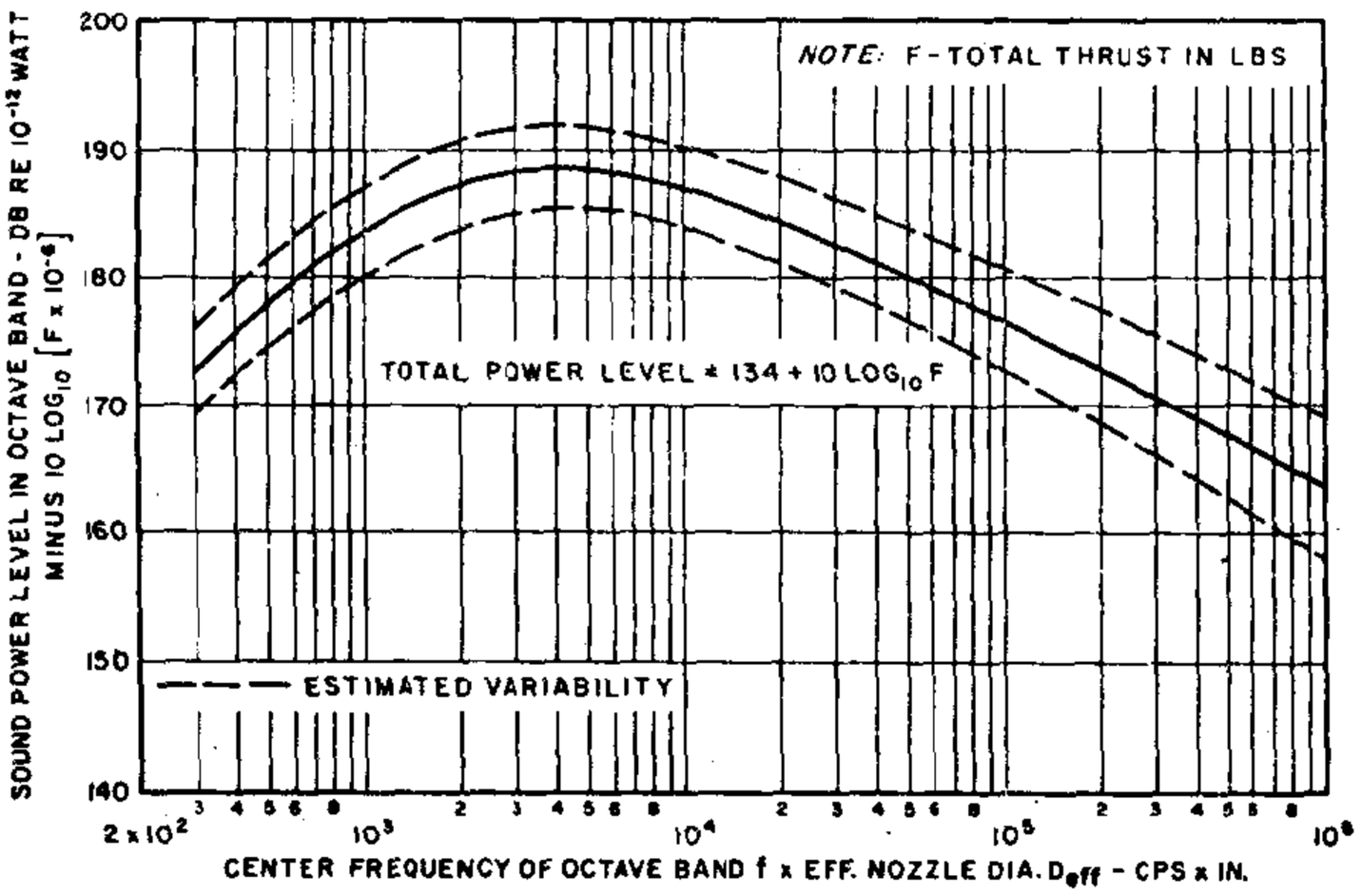


Fig. 5. Smoothed Generalized Sound Power Spectrum Function For Large Rockets.

Experience to date indicates that the directional characteristic of large rocket engines is a function that varies comparatively slowly with angle and is approximately independent of frequency in a broad sector containing the direction of maximum radiation. Sound pressure levels in the sector of maximum radiation are 3–5 db higher than the space average. Figure 5 can therefore be used to estimate the sound pressures and their spectra in the far field. Clearly, estimates of the sound pressure levels for distances beyond about one mile must take into account sound attenuation in the atmosphere and other propagation effects.

Knowledge of the acoustic power and spectrum of a rocket in flight is clearly of importance in predicting the magnitude and spectra of the sound pressures and their time histories on the ground after launching. It is not known in detail how the sound power and spectrum of

the rocket engine source change as the rocket gathers speed after launch. In addition to the possible effects of the change of atmospheric parameters on the radiated sound power there is the effect of the considerable forward speed of the vehicle. On the premise that the reduced relative velocity in the shear layer of the exhaust results in reduced sound radiation, Powell [12] has suggested that the total sound power level radiated by a typical rocket in flight with speed v ft/sec be reduced by the term $30 \log_{10} [1 - v/7500]$. This correction predicts only a small effect on the sound pressures. More work is needed on this problem to test and verify the above relation. Finally, there is also an expected Doppler shift in the noise spectrum radiated by the vehicle engines in flight. Lack of accuracy in current prediction and measurement techniques has so far prevented its isolation from other variables. Based on these concepts the expected overall sound pressure levels due to the launching of a space vehicle of over the 10^6 lbs thrust on a typical trajectory were calculated. The excess attenuation in the atmosphere was estimated by assuming an attenuation coefficient of 1 db/mile independent of distance. Although the atmospheric propagation effects have thus been oversimplified, the graphs in Figs. 6 and 7 show reasonable agreement between prediction and the experimental data obtained by NASA. The fluctuations of the measured sound pressure level, especially those measured several miles away are noteworthy; they

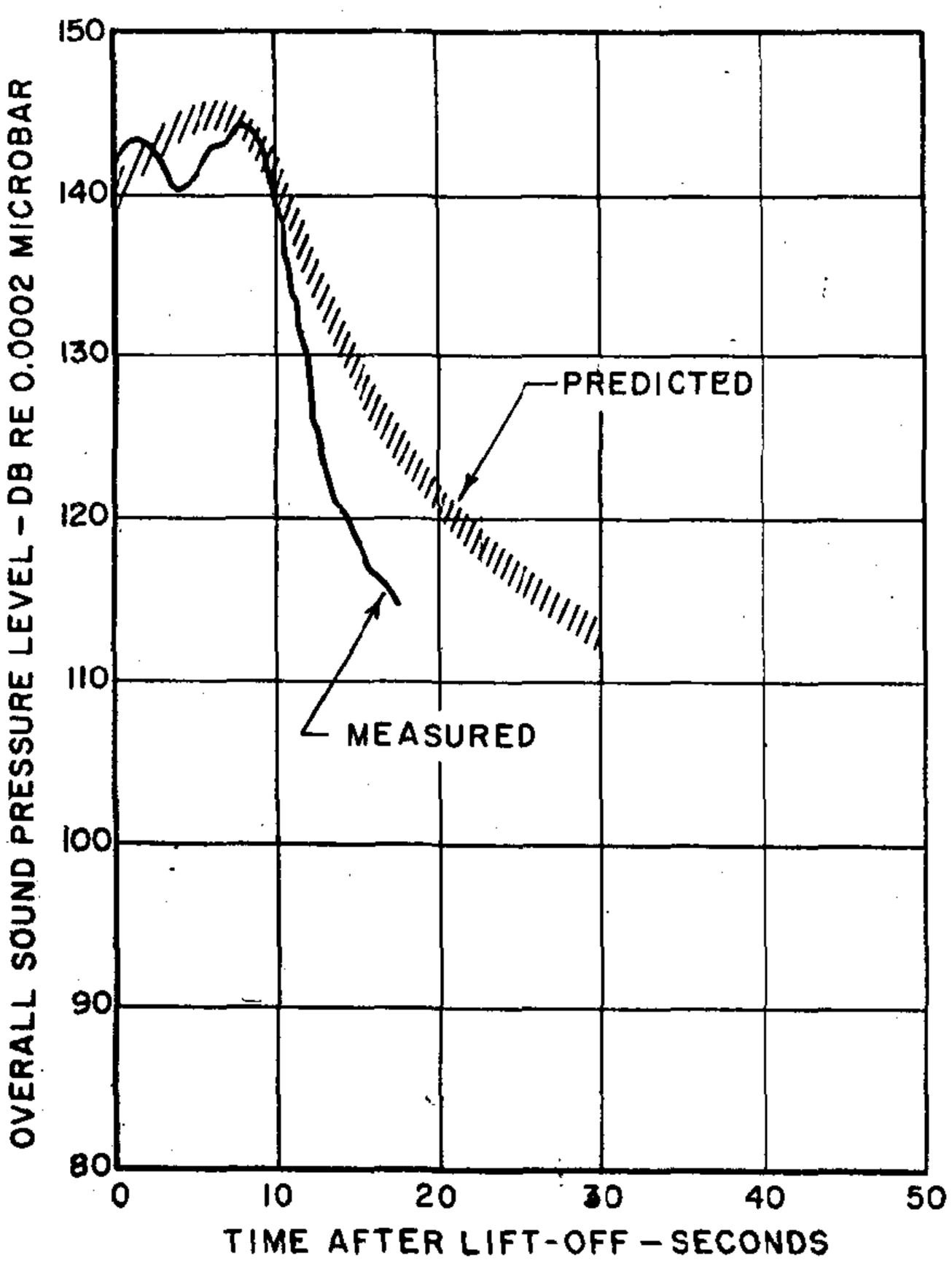


Fig. 6. Predicted and Measured Time History of the Sound Pressure Levels Close-In on the Ground from the Launching of a Large Space Vehicle.

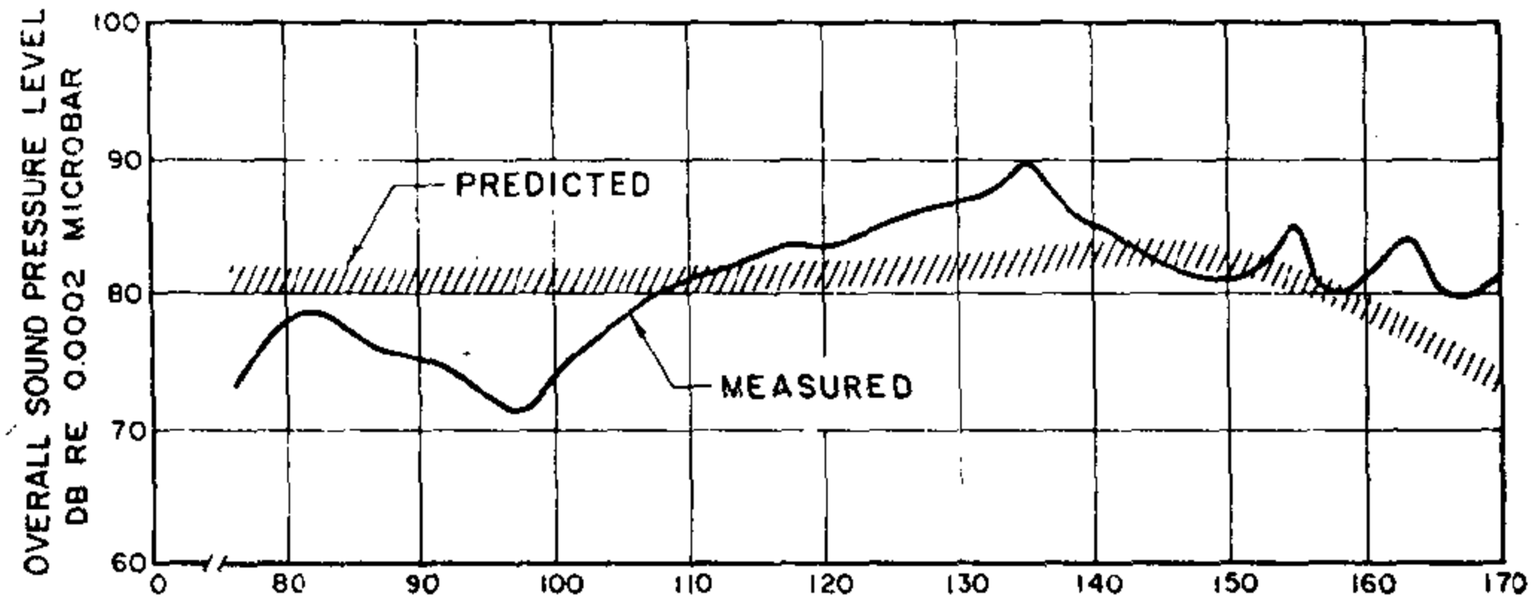


Fig. 7. Predicted and Measured Time History of the Sound Pressure Levels Far Away on the Ground from the Launching of a Large Space Vehicle.

are characteristic of sound having traveled far through the atmosphere. This important subject of propagation of low-frequency rocket noise through the atmosphere will be discussed in some detail in another section.

2.5 *Geometric Field*

The subject of sound-induced vehicle vibrations is complex and an almost autonomous discipline in itself and many workers are currently active in this field. Its treatment in detail is beyond the scope of this presentation. Surveys of this field have been provided by Dyer [13] and Powell [14]. Here, we wish to discuss briefly the acoustic excitation. Clearly, the levels and spectra of the rms sound pressures at various points along the vehicle and at the space capsule are of central importance. In addition, the spatial correlation of the sound pressures over the vehicle surface or the surface of a sub-structure of interest are important in that the response of the structure depends on it. Both the longitudinal and angular correlation of the sound pressures near large space vehicles have been investigated by Dyer et al. [15] Magnitudes, spectra and correlations of the sound pressures in the vicinity of the vehicle are known to be affected by the deflector configuration, as is the distribution of the sound sources along the exhaust stream. Experimental determinations of the apparent source location have been made by observing the amplitude and phase relations of the sound pressures with an array of microphones spaced along the vehicle axis. From such measurements during static tests of a vehicle at a NASA installation the apparent source locations were calculated. These results and other data are summarized in the smoothed functions shown in Fig. 8 in generalized coordinates. It is seen that the high-frequency energy comes mostly from locations near the nozzle plane, whereas the low-frequency energy comes from farther down-stream. This general result is true whether the rocket is in free flight or whether the exhaust stream is deflected as during static testing or before lift-off.

Attempts at calculating the pressure field near the vehicle from the source distribution and other parameters have not been very successful to date. Estimates of the sound pressure levels at various positions along the vehicle are at present best derived from direct measurement. Figure 9 shows a smoothed spectrum function of sound pressure near the nose of the vehicle before lift-off. Since the shapes of large rocket vehicles do not vary greatly, the results can be expressed in terms of a scaled distance $L = 12\text{–}15\ D_{eff}$ and in terms of a single function whose ordinate is independent of thrust. The abscissa is the generalized frequency parameter $f \times D_{eff}$.

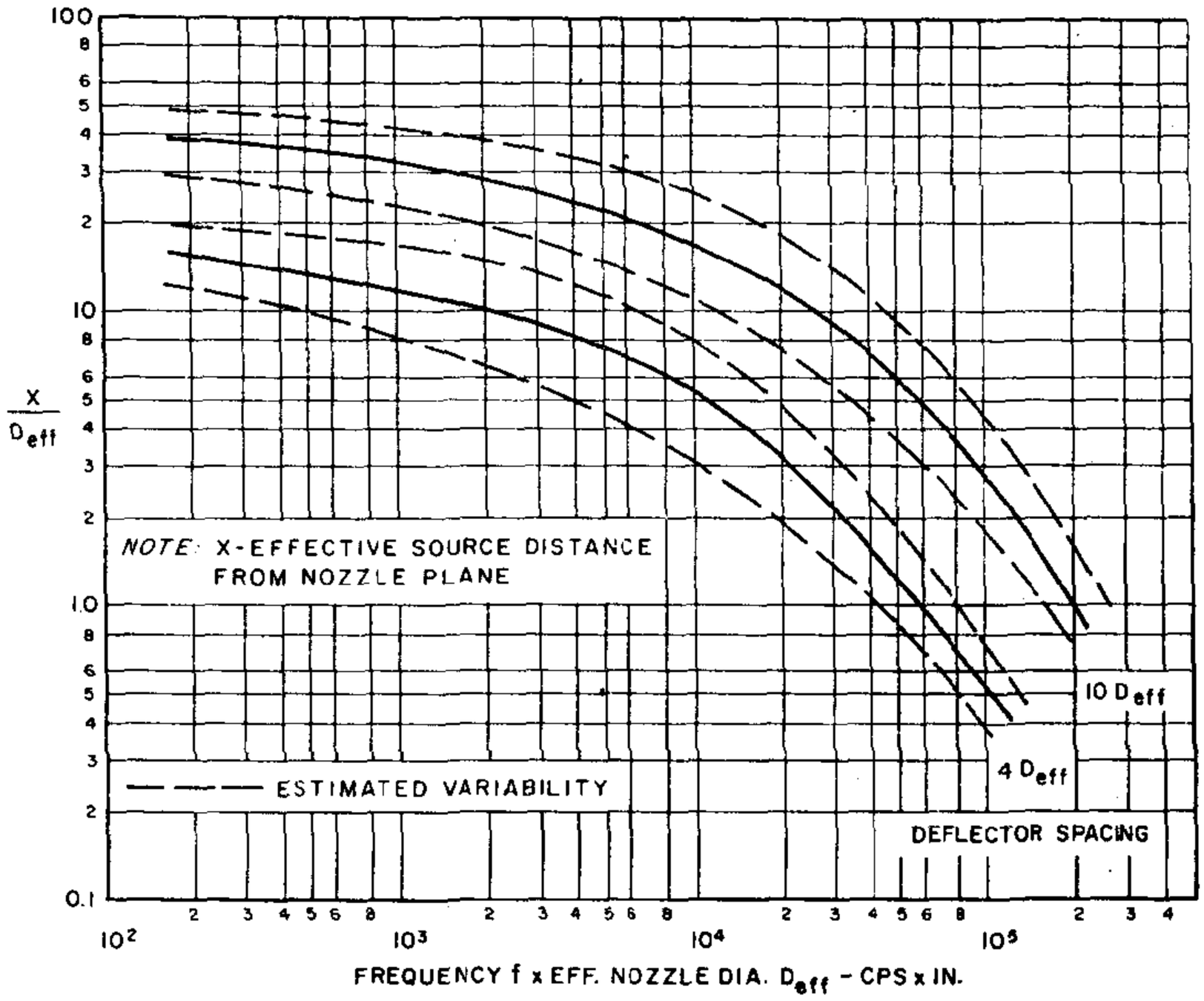


Fig. 8. Smoothed Relation Between Source Location and Generalized Frequency Parameter During Static Testing of Large Rockets.

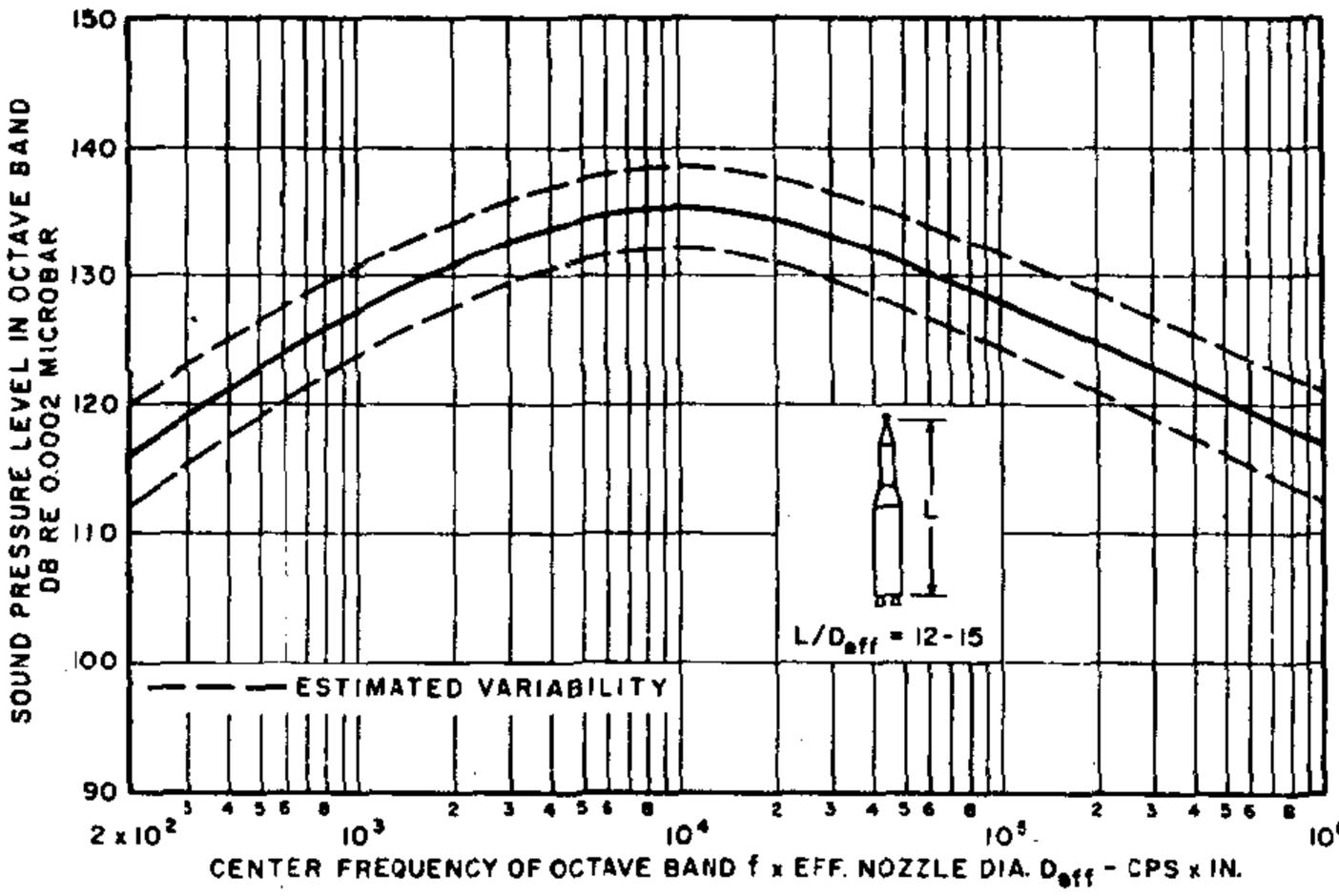


Fig. 9. Smoothed Sound Pressure Spectrum Function Near the Nose of Large Space Vehicles Before Lift-Off.

Experience has shown that the sound pressure levels along the vehicle are largest during static testing or before lift-off. After lift-off the absence of the deflector structures reduces the sound pressures in the geometric field. In addition, Powell [16] has estimated that the sound pressure levels along the vehicle are further reduced by a term $20 \log_{10} (1-M)$ where M is the Mach number of flight. To this should be added the decrease in radiated sound power estimated earlier. The correlation in the pressure field is generally reduced by the forward motion of the vehicle because of the expected decrease in apparent wavelength.

3. Noise reduction at the source

3.1 *General*

Because of the enormous sound powers generated by present-day and future rocket boosters serious consideration is being given to methods of noise reduction at the source. This is particularly important for static testing. First, because of the much larger frequency of static firings as compared with the frequency of launchings and second, because atmospheric effects are more likely to lead to sound reinforcement when the source is near the ground. In-flight noise suppressors for large space vehicles do not appear feasible at this time.

Decreasing the velocity of the exhaust gases is the basic objective of direct noise reduction because the velocity shear is the source of the turbulence and hence of the noise. To be effective, an appreciable velocity reduction must be achieved. This can be accomplished by induction of secondary air into the exhaust stream or by injecting water spray into it, or by a combination of the two. A very effective way of introducing large quantities of water into the exhaust is by the use of a large body of water into which the exhaust stream issues. Finally, a dissipative muffler can be employed.

With diligence and proper design any or all of these methods can be translated into practice to silence small rocket engines in model tests. To apply them to full-scale rockets in the multi-million pounds thrust class presents formidable practical difficulties, because of the large sizes and costs involved. No full-scale working muffler for rocket engines has been built as yet to the best of our knowledge.

3.2 *Noise Reduction by Water Spray and Secondary Air.*

In static testing of large rockets the flame deflector is often watercooled, with the surplus water issuing as a spray into the jet stream. Typical flow rates equal those of the flow rates of the exhaust gases, and no appreciable noise reduction is observed. This is in accordance with the results of model tests [17, 18]. From these and other tests Galloway [19] has compiled a summary curve which shows the reduction of the sound power generated in terms of the conversion efficiency η, the fraction of the stream power converted into noise, as a function of quantity of water injected. Figure 10 shows that a reduction in η of about 20 calls for a weight flow of injected water spray one order of magnitude higher than the weight flow rate of propellant. For one of the large contemporary boosters a water rate flow of 3 to 4 $\times$ 10^5 gallons per minute would be required. This implies almost prohibitively expensive piping and pumping facilities, in addition to a large supply.

When the exhaust stream of the rocket is directed into a basin filled with water appreciable noise reduction can be obtained, provided the basin is large and deep enough. Recent experiments [20] with a model rocket engine the exhaust stream of which was directed into a 90-degree deflector and thence, through a water-filled duct into a large water-filled tank, show sound power level reductions of about 20 decibels or more. However, the total water volume required at present corresponds to over 30–40 times the weight of the propellant. In addition, free hydrogen tends to accumulate in the tank from the exhaust gases of liquid fuel engines with consequent danger of explosion. Moreover, the turbulent water must be kept from splashing back into the engine area. While basically a promising method for

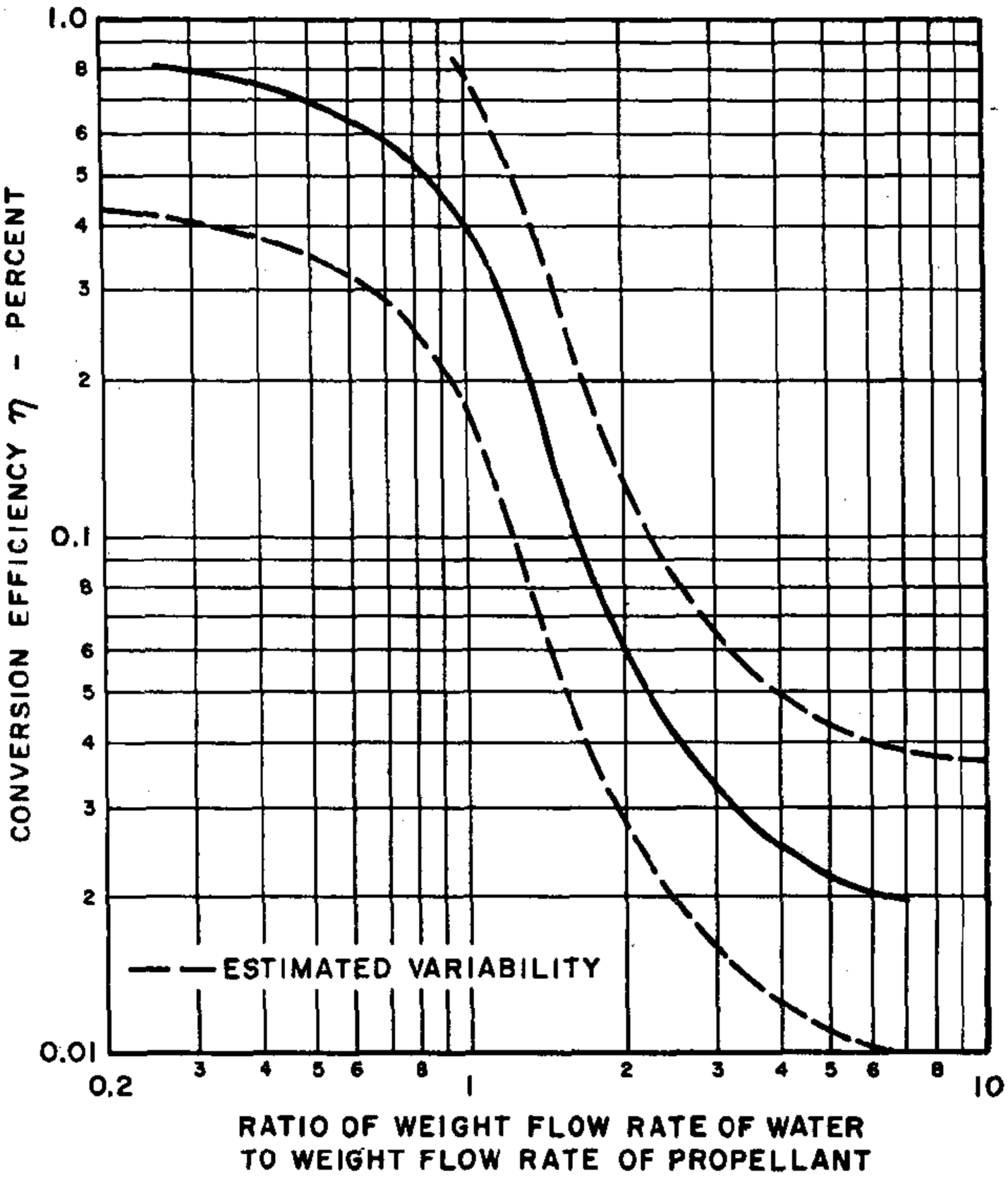


Fig. 10. Function for Estimating the Effectiveness of Water Spray Injected Into Rocket Exhaust (Galloway).

silencing very large rockets, much work remains to be done before a full-scale installation can be attempted.

If secondary air were used very large quantities would also be required to reduce the temperature of the exhaust gases significantly. Model tests have shown that augmenter and induction tubes must be carefully designed lest they also augment the noise.

3.3 *Noise Reduction by Dissipative Mufflers*

Although the design and performance of ducts containing sound absorbing liners is well known, ducting of high-speed, high-temperature rocket exhaust gases of large volume presents great problems. Ordinary sound absorbing fibrous materials are not capable of withstanding these severe conditions. Special ceramic or other materials which can withstand the effects of high-velocity turbulent flow of high temperature may be developed. Another approach is to utilize the nonlinear behavior of orifices in the presence of flow. Such orifices may be provided by a muffler equipped with perforated plate linings. Westervelt [21] has shown that the acoustic reactance of an orifice decreases in the presence of steady flow, while the resistance increases with the speed of the steady flow. Work is underway to explore this avenue of approach further and find possible limitations in its appli-

cability to the problem at hand. Its main advantage, of course, is that a perforated plate can withstand the turbulent forces and high temperature of the exhaust stream far better than any known fibrous material.

4. Propagation of rocket noise through the atmosphere

4.1 *General*

Current estimates of the mean sound pressure levels generated by a rocket of 10^7 lbs thrust at a distance of 10 miles indicate sound pressure levels in the 100–110 db range (re 0.0002 microbar). The spectrum peak is estimated to be in the vicinity of 10 cycles per second. The predicted total sound power level is about 204 db re 10^{-12} Watt. Thus the problem of propagation of rocket noise through the atmosphere acquires new dimensions in terms of greatly increased source power, distance and the low frequencies involved. Data [22—25] on the interaction between audible sound, radiated over conventional distances by conventional sound sources, and the meteorological parameters of the atmosphere are of limited value here. One must look for guidance to studies of sound propagation from sources of large spatial extent and great power. The work on the propagation of sound from explosions [26—28] and perhaps from lightning [29] appears more relevant for furnishing guidance here. On examining this literature one finds that sound refraction caused by the variations of air temperature and wind speed and direction with height significantly affects the magnitude of the pressures at large distances. Indeed, experiments show that sound refraction effects are important also in the propagation of rocket noise along the ground. [30] Ray acoustics [31] applied to a layered atmosphere with constant gradients of sound velocity indicates that the effective sound velocity variation with height in the first several thousand feet above ground should be considered. To assess sound propagation from rockets in flight the profile slopes at higher altitudes appear relevant. Since the spectrum of rocket noise consists of contributions extending to the medium and high audible frequencies, the dissipative processes in the atmosphere also play a role. These are primarily due to molecular absorption. There is also scatter attenuation caused by inhomogenieties in the atmosphere along the transmission path. Moreover, considerable fluctuations of the received sound pressure level about the mean value are characteristic of long transmission paths. The effect of the sound transmission anomalies may be expressed in terms of an excess attenuation, i.e., the attenuation of the mean sound pressure level in excess of the level reduction in an ideal, homogeneous atmosphere at rest. There is no comprehensive theory of sound propagation in a real atmosphere available at present and in order to estimate the sound pressure levels at distant points recourse must be had at present to empirical procedures. To obtain the estimates of excess attenuation as a function of frequency necessary for planning purposes and to explain sound propagation data on rocket noise which are only recently becoming available, dissipative attenuation effects and excess attenuation (positive or negative) due to refraction effects are superimposed to obtain their combined effect. This is, of course, an oversimplification. Available data on the magnitude of these dissipative effects pertain generally to measurements on conventional sound sources over comparatively short distances. These data must be extrapolated to low frequencies and large distances. Although this procedure has not been without some success in explaining

recent results of sound propagation measurements of rocket noise, there is a great need for a fresh and comprehensive theoretical approach, supplemented by measurements of rocket noise and their spectra at large distances, accompanied by the relevant meteorological observations.

4.2 *Molecular Absorption and Scatter Attenuation*

This type of excess attenuation is commonly quantified in terms of an attenuation coefficient which is independent of distance. Kneser [32] has done pioneering theoretical work in the investigation of sound absorption in gases and elaborate laboratory investigations of molecular absorption in air have been made recently. [33, 34] Horiuchi's data are especially appropriate here in that they extend to low audio frequencies (300 cps). Results of older experimental investigations have been summarized by Nyborg and Mintzer. [35] There is disagreement among the results, especially at low and high humidities. However, the linear relation between frequency and the maximum of the absorption coefficient (as humidity is changed at constant temperature) seems well established. Of course, the results of field investigations conducted in a real atmosphere are always contaminated by the presence of sound scattering, sound refraction and other effects. When sound propagation takes place from air to ground or from hilltop to hilltop, refraction effects caused by wind and temperature gradients are smaller than when sound is propagated along the ground. Even so, the measured effective attenuation coefficients are larger than those obtained in the laboratory.

Ingard and Wiener [36] have estimated the scatter attenuation in the atmosphere assuming back-scattering of sound by eddies of size comparable with the wavelength. Based on measured turbulence spectra near the surface they found only a slight dependence on

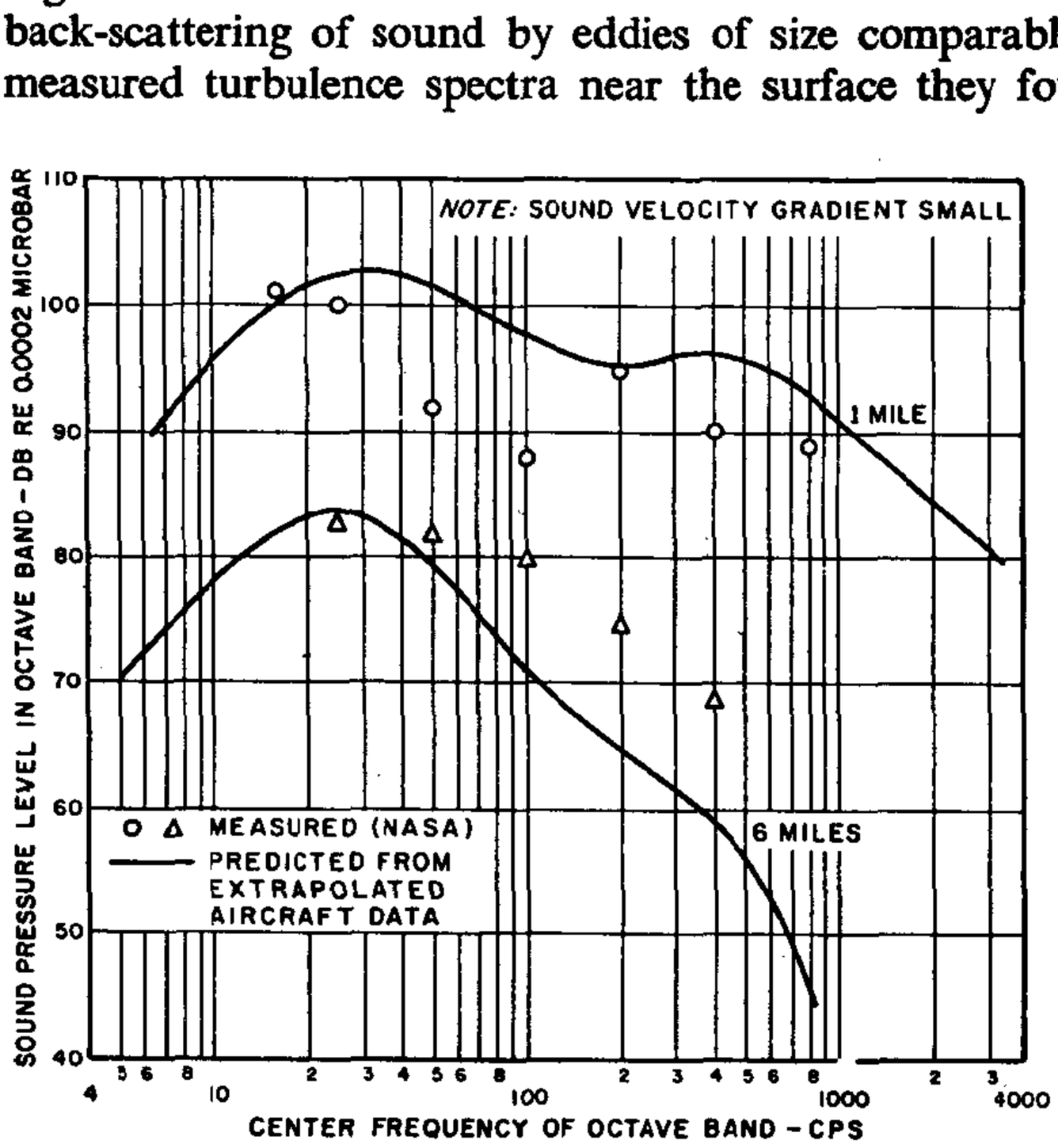


Fig. 11. Comparison of Measured and Predicted Rocket Noise Spectra. Sound Propagation Along the Ground.

frequency and wind speed. This is in general agreement with the meager experimental evidence available but in disagreement with Lighthill's [37] calculations which predict much larger values and stronger dependences.

The extrapolation of experimental values of sound attenuation data from flyovers of aircraft for the purpose of estimating the attenuation of rocket noise in the absence of pronounced refraction effects cannot be but somewhat arbitrary at this stage, awaiting confirmation by experiment and additional theoretical work. Figure 11 shows a comparison between the noise spectra on the ground obtained at two distances by direct measurement and those estimated. The noise was generated by a rocket of more than 10^6 lbs thrust held stationary on the pad. Since in this case the average sound velocity gradient in the first 3000 ft was relatively small ($\sim 3 \times 10^{-3}$ sec^{-1}) estimates based on extrapolation of data from air-to-ground propagation of aircraft noise were used to predict the mean sound pressure spectra at 1 and 6 miles in Fig. 11.

Application of the above procedure for predicting the excess attenuation of sound from a rocket in flight in the absence of large sound velocity gradients is also appropriate, since present evidence indicates that at the very low frequencies molecular absorption is not controlling, even though the atmosphere above 30,000 ft is often very dry which may theoretically lead to molecular absorption peaks. Figure 12 shows the results of measurements of the sound pressure levels on the ground at a distance of 10 miles at the time of maximum overall levels when a large space vehicle was launched. Meteorological measurements showed that the slopes of the lower portions of the sound velocity profile were relatively small. Another spectrum estimate, due to v. Gierke and based principally on an extrapolation of the results of the measurements by Sabine [38] is also shown in Fig. 12.

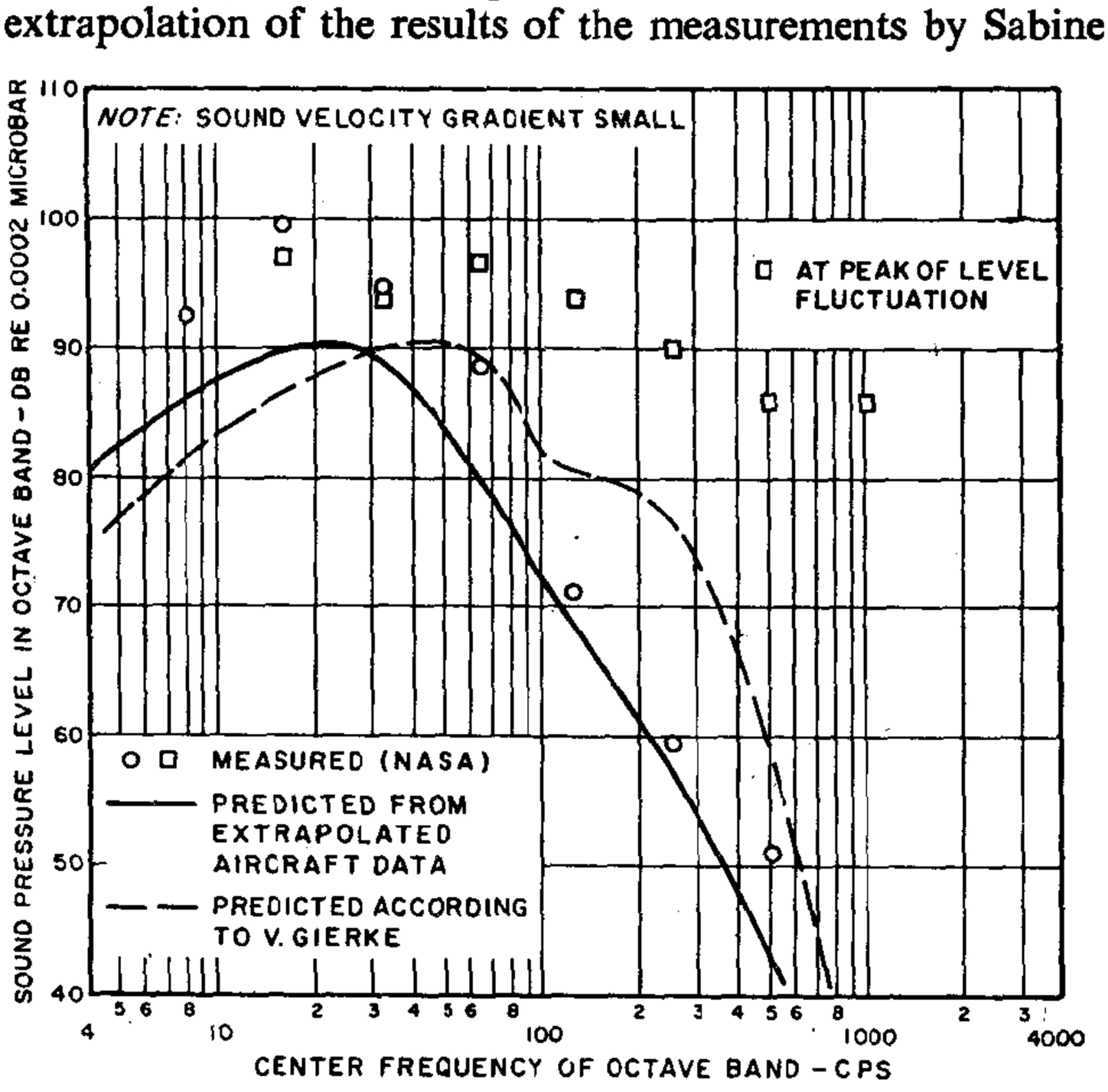


Fig. 12. Comparison of Measured and Predicted Maximum Rocket Noise Spectra on the Ground After Launching, 10 miles from the Pad.

Note the considerable divergence of the predicted values. In addition, there is also shown a spectrum measured during a pronounced fluctuation peak which occurred during another launching under nominally similar conditions. The divergence of the data gives an indication of the magnitude of the effect of fluctuation phenomena in the atmosphere. Clearly, much additional work is called for here to account for the observed differences and improve the estimation procedures.

4.3 *Sound Refraction*

Recent measurements of the noise levels from large rocket engines during static testing propagated along the ground over large distances have shown a large dependence on the shape of the variation of the effective velocity of propagation of sound with height (sound velocity profile). This quantity, at a given height, in a given direction, is equal to the speed of sound at the temperature at the point in question in still air, and added to it the vector component of the wind at that height in the direction considered. In general, and as predicted by simple ray acoustics, positive slopes of this profile tend to decrease excess attenuation, often to the point of resulting in sound pressure levels exceeding those predicted by the inverse-square relation. Conversely, negative profile slopes increase excess attenuation. Figure 13 shows a scatter diagram of the excess attenuation measured at 1 mile plotted against the average slope of the sound velocity profile in the first 4000 ft. The profile data were obtained from balloon ascents conducted at a relatively short distance from the site and within a short period of time of the sound pressure measurements performed at a NASA installation during static testing of a large space vehicle. [39] Note the (negative) correlation shown by the data. Work is underway to investigate the effect of sound refraction on excess attenuation in a more detailed and quantitative way.

It appears now that sound refraction is not as important for a rocket in flight as it is when the rocket is on the pad. This is plausible because, generally speaking, the slopes of the sound velocity profile under given conditions tend to decrease on the average with height.

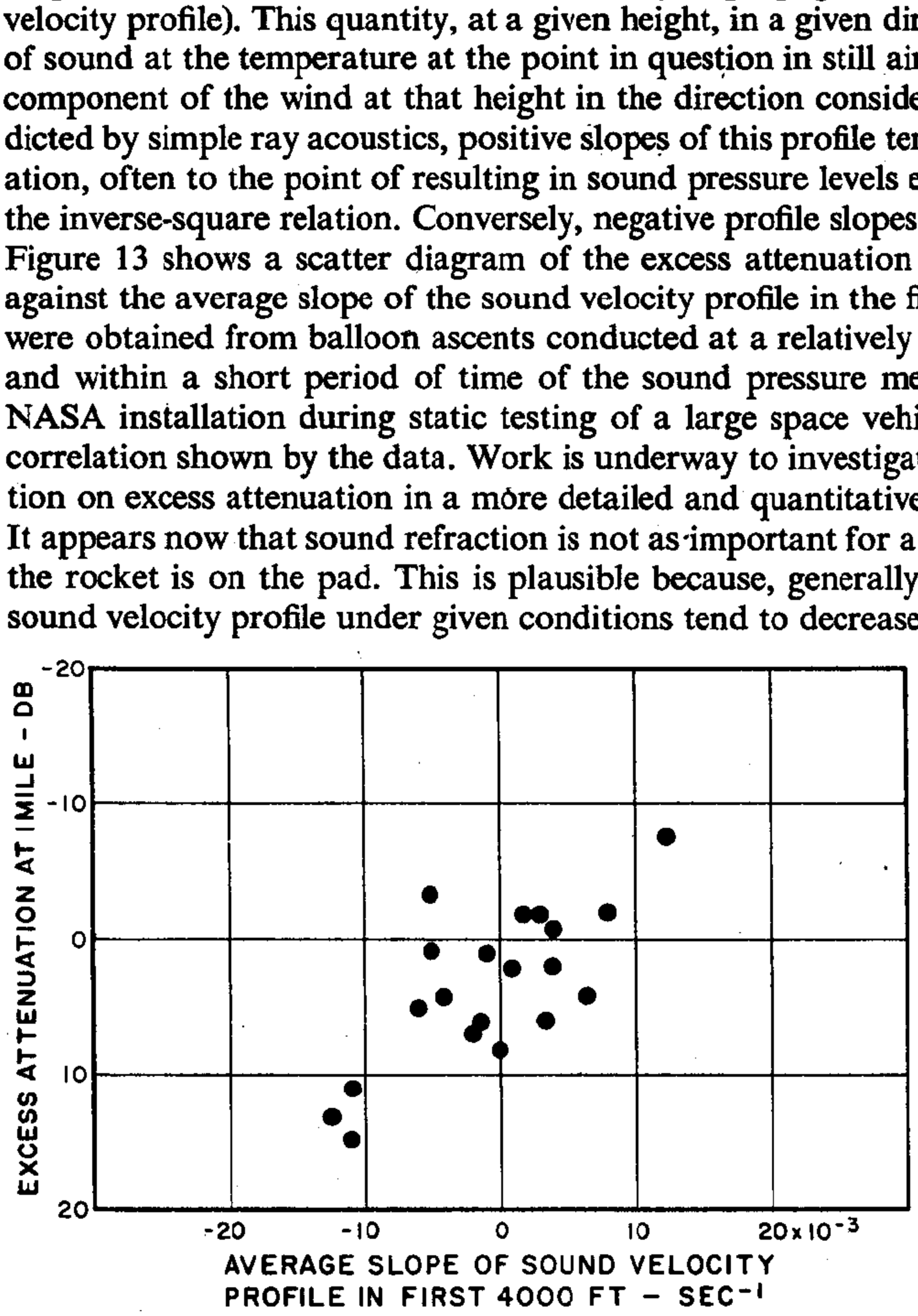


Fig. 13. Scatter Plot Relating Excess Attenuation and Average Initial Slope of the Sound Velocity Profile During Static Testing of a Large Space Vehicle (NASA).

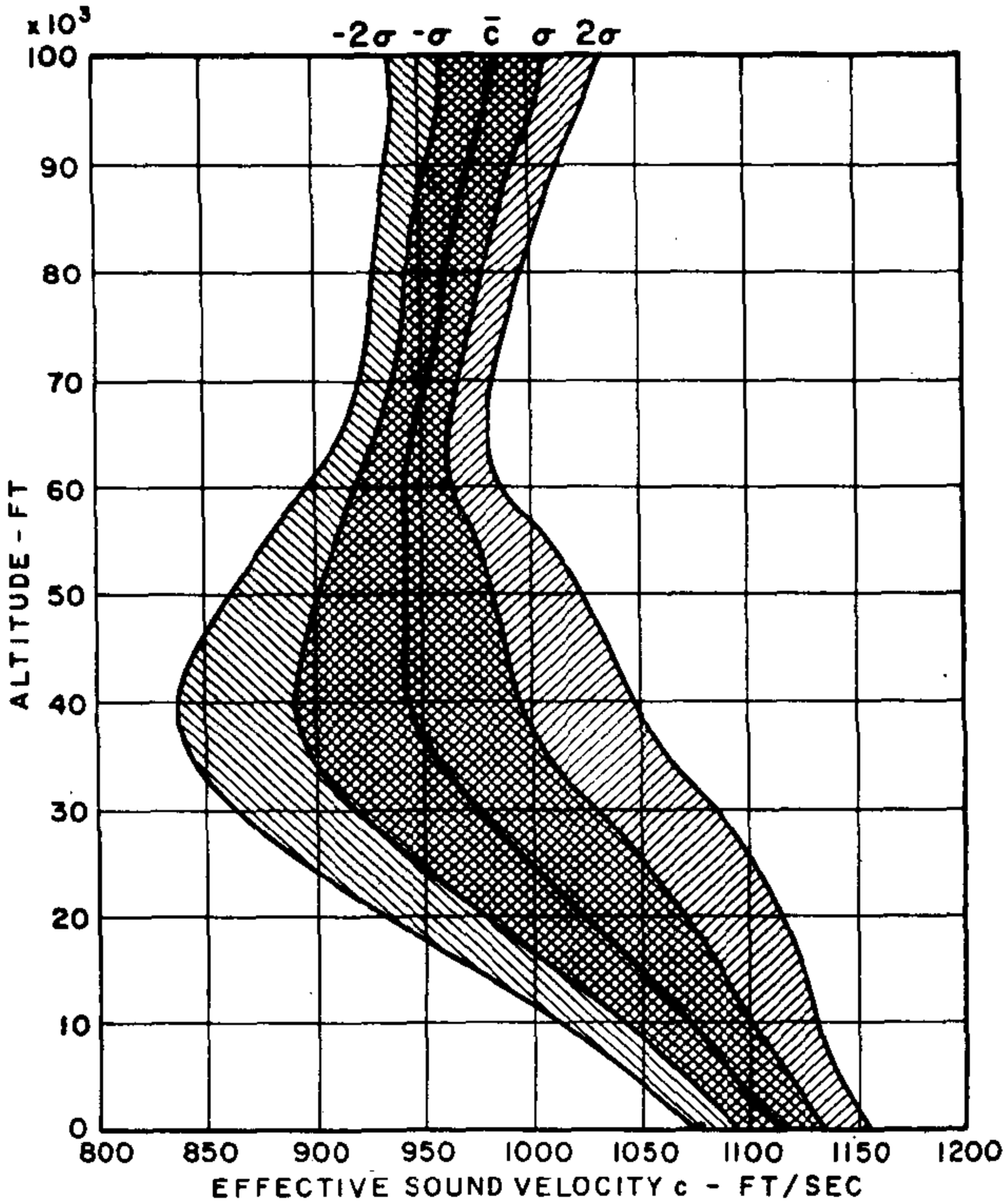


Fig. 14. Seasonal Probability Distribution of the Sound Velocity for a Given Site and Direction.

Occasionally, measurements show indications of locally positive profile slopes in a certain layer. This may give rise to locally increased sound pressure levels on the ground at a distance related to the height of the profile irregularity. [40] Experimental verification is difficult because of the often transient nature of the phenomenon and the time lag and lack of detail of the meteorological information obtained with present techniques from balloon ascents.

Seasonal distributions of the sound velocity profile for a given site are of great value to the planner in assessing the likelihood of occurrence of sound refraction effects at that site. Such a profile distribution is given in Fig. 14 which shows the mean effective sound velocity $\bar{c}$ and the contours for $\bar{c} \pm \sigma$ and $\bar{c} \pm 2\sigma$ where σ is the standard deviation of the data. For typical sites in the U.S., σ is controlled by the variability of the wind data and not by the variability of the temperature profiles.

5. Avoidance of damage or complaint

5.1 *General*

The problems of how to maintain the life function of the astronaut in flight are many and complex. They are the subject matter of space medicine and allied branches of the life

sciences and hence outside the scope of this paper. A comprehensive review of the effects of vibration on man was made by Goldman [41]. There is no comprehensive study of the various effects of intense low-frequency and infra-sonic noise on man available at this time. This presentation will deal primarily with the effects of rocket noise on people and structures on the ground. Effective exposure times are then of the order of minutes for a single launching or static test. Excitation by direct structure-borne vibrations is negligible in most instances, as vibrations are rapidly attenuated as they travel through the soil away from the pad. In conclusion, a brief account is given of the interaction between rocket noise and space vehicle structures and equipment.

5.2 *Effects of Rocket Noise on People*

In addition to producing temporary or permanent deafness and pain, excessive levels of rocket noise may result in a variety of other direct effects, such as disorientation, dizziness and nausea, circulatory and respiratory changes, impairment of vision and psychomotor efficiency through excessive noise-induced vibrations of the organs concerned. There is another set of symptoms such as irritability, tiredness, impairment of intellectual performance and of the ability to communicate by voice. A comprehensive review of these and other effects due to excitation by noise from jet aircraft engines was undertaken some time ago under sponsorship of the Office of Naval Research of the U.S. Navy [42]. The results of the above-mentioned study show that, by and large, the tolerable noise levels in humans exposed to jet aircraft engine noise with their whole body and without ear protection are set by the sensitivity and fragility of man's ear. However, many of the results of this study do not pertain directly to the very low-frequency components important in rocket noise; hence, at present one must extrapolate to those low frequencies in the absence of extensive data pertaining directly to rocket noise.

It is known that both the threshold of pain in the auditory system and the levels which can be tolerated without appreciable temporary hearing loss increase as the frequency of the stimulus is lowered. Ward et al., [43] have established the essentially inverse relationship between exposure time and maximum tolerable sound pressure level in terms of a temporary hearing threshold shift criterion measured two minutes after exposure. Present indications are that at the very low frequencies, not examined by Ward et al., even higher sound pressure levels are tolerable for a given exposure time than those estimated from Ward's data. Hence, it is not at all certain that the tolerable limits of very low-frequency noise in terms of their overall effects on humans are still provided by the hearing mechanism. Future research must provide the answer to this question.

Current estimates of maximum tolerable octave band levels of rocket noise based on hearing damage risk criteria recently proposed by Kryter et al., [44] suggest that maximum octave band levels of noise from large rocket engines should not exceed 135–140 db (re 0.0002 microbar) for daily exposures of a few minutes. Frequently, safety regulations considering other hazards than noise (e.g. accidental explosion of the rocket propellant mixture) limit the rocket noise levels to which humans on the ground are directly exposed below those prescribed by the hearing damage risk criterion.

Residential communities, albeit at large distances from static test and launching sites, are often exposed to levels of rocket noise of considerable magnitude. Although short of caus-

ing damage to residential structures the noise intrusion may elicit reactions of annoyance, anger or fear in the residents. Moreover, the noise may interfere with normal activities. In formulating a community noise criterion for rocket noise, results of studies of the effects of noise intrusion from jet aircraft are of limited value because of the differences in the character of the noise, its duration and other factors. The meager evidence available at present indicates that sound pressure levels of 100–105 db (re 0.0002 microbar) in the frequency octaves below about 100 cps in the spectrum of the noise from large rockets will result in some adverse community reaction. Clearly, the magnitude and character of this reaction will be conditioned by the economic and emotional connection of the residents with the noise-producing activity.

5.3 *Effects of Rocket Noise on Building Structures*

Exposure of building structures to rocket noise may cause structural damage. Structures of practical interest in the damage problem are complicated and their response to acoustic excitation is correspondingly so, as many resonant modes are excited simultaneously and coincidence effects [45] may occur. However, at low frequencies it may be reasonably assumed that the bandwidth of each mode is less than the frequency spacing between modes so that a mode-by-mode analysis is appropriate. At least two mechanisms of damage in panels are of importance: 1) bending stresses in a structure such as a window which exceed the endurance stress of the glass and 2) support forces exceeding the holding power of the nailing around the perimeter of a panel. The damage criteria, i.e. the critical octave-band sound pressure levels which contain the fundamental resonance frequency of the panel and which are not to be exceeded, are found to increase 6 db per frequency octave for the stress mechanism, and 3 db per octave for the force mechanism [46]. With typical material constants and sizes, the analysis shows the damage potential of the same order of magnitude in both cases. A different interpretation of the stress mechanism by Regier and Hubbard [47], who considered structures designed for a constant static pressure, leads to critical octave-band sound pressure levels independent of frequency. When failure is appro-

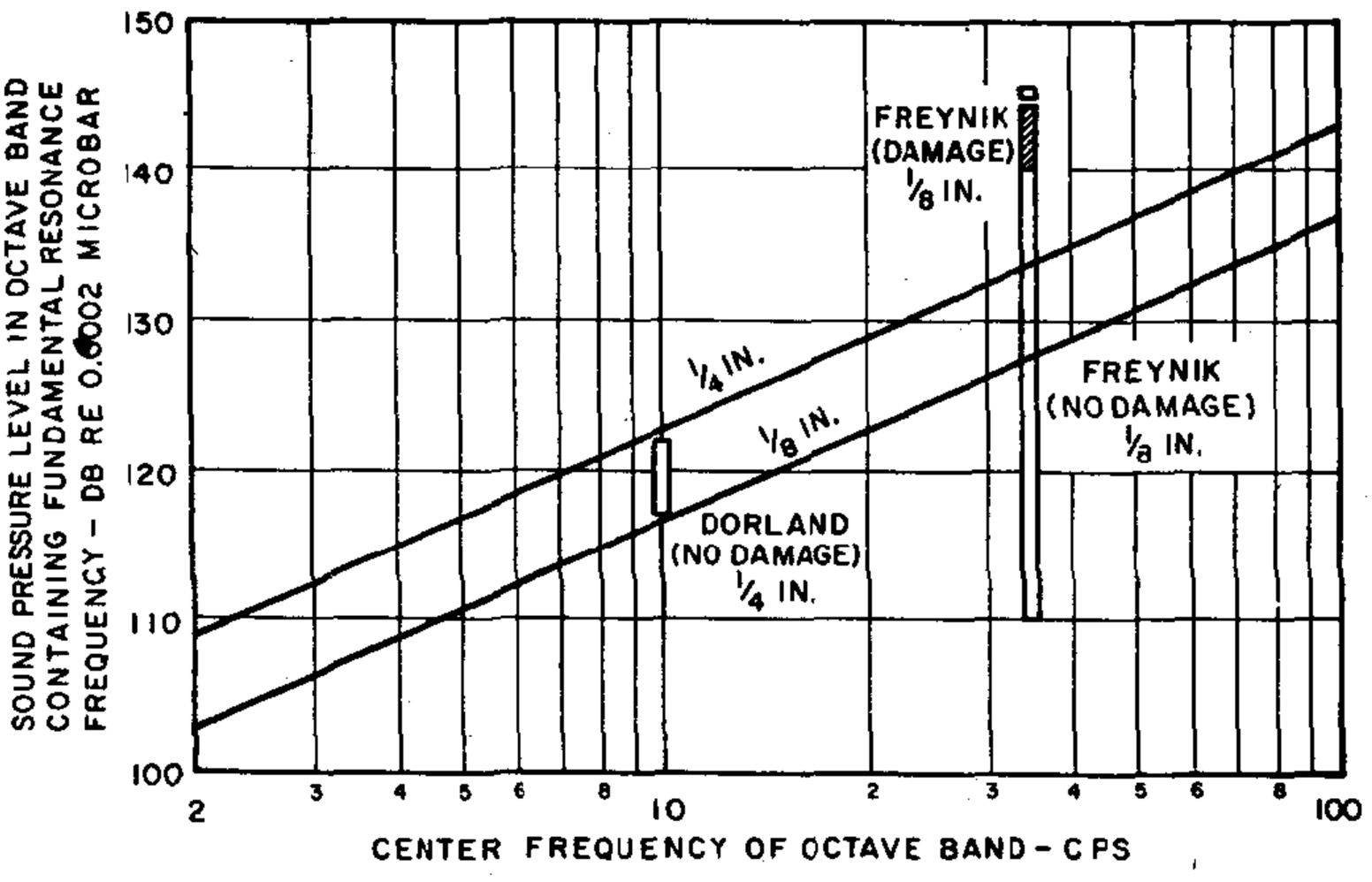


Fig. 15. Tentative Damage Criterion for Windows (Dyer and Wiener) and Experimental Results (Dorland, Freynik).

ached by increasing the excitation, the strains increase less rapidly than predicted by the linear analysis. Hence, the above estimation procedures imply a certain conservatism in the engineering sense. [48]

Figure 15 shows a plot of the critical sound pressure levels calculated for $^1/_4$-inch and $^1/_8$-inch glass panels according to Ref. 46. Experimental results for two different windows exposed to random noise are also shown [48, 49]. These and other data are seen to be consistent with the proposed damage criterion if a conservative point of view is taken, i.e., sound pressure levels above the critical ones may or may not cause failure. Below them they do not. Further work is indicated to obtain a better insight into the sound-induced damage processes and develop better criteria.

5.4 *Effects of Rocket Noise on Space Vehicle Structures and Equipment*

Measurements of the acceleration of typical panels of the outer skin of a space vehicle during static tests have demonstrated that their response is determined primarily by the incident rocket noise field and not by structurally transmitted engine vibrations. A detailed discussion of the response of such structures and the allied fields of structural fatigue and vibration control is far beyond the scope of this presentation. Many investigators are active in these areas and important contributions have recently been made, leading to a better understanding of the physical processes involved and to improved estimation procedures of the sound-induced response of panel and panel-like structures [50-52].

Sound fields of two general classes are encountered in space vehicles. Panels and structures on the vehicle surface are exposed directly to the sound field of the rocket engines at nearly grazing incidence, except in the vicinity of the engines. Devices situated inside the vehicle such as electronic and other instruments located in an instrument canister or an astronaut inside his space capsule are exposed to a reverberant sound field. The magnitude and spectrum of the interior sound field are determined by the response of the skin panels or canister walls to the exterior sound field and their ability to re-radiate sound into the interior spaces. These two phenomena are, of course, closely inter-related. Typical space vehicle substructures of interest here are panels with stringers, stiffeners, brackets and the like. They are sufficiently complicated in their mechanical behavior that a simple mode-by-mode approach, as was outlined in the preceding section, is no longer appropriate. A statistical treatment using energy concepts has been advanced by Lyon and Maidanik [53] and applied with success.

Structures or instruments located on the inside of the vehicle skin are exposed to the reverberant sound field transmitted from the engines to the inside. In addition, the sound-induced vibrations of the skin are transmitted to the instruments or equipment package by structural paths. Estimating the response of an assemblage of electronic and other components to such a combined sound and vibration excitation is a complex problem incapable at this time of a general solution. Several recent studies, however, [54, 55] are providing an increasingly better understanding of the problem. One important finding, for example, concerns small electronic components, such as capacitors, resistors, vacuum tubes and transistors. Damage or malfunction of such components is almost always due to the (sound-induced) vibrations, transmitted from the supporting chassis or panel, rather than due to the direct effect of the sound field on the component proper.

6. *Conclusion*

The author is aware, of course, that this presentation represents a compromise in time and space, rather than a complete discussion of the rocket noise problem. Nonetheless, it is hoped that this presentation has left the reader both with an appreciation for the progress made in this field and with a glimpse of the variety of challenging problems still awaiting solution.

7. *Acknowledgments*

The author is indebted to the great many workers in the field on whose results this paper is based. Special thanks are due Drs. Ira Dyer, Peter A. Franken and Richard H. Lyon of Bolt Beranek and Newman, whose help in preparing the material in certain sections and organizing others was invaluable. For several years the author and his colleagues at Bolt Beranek and Newman have worked with personnel of the National Aeronautics and Space Administration and its predecessor agency on rocket noise problems. Many of the data presented here have arisen from this joint effort. In particular, the author wishes to thank Messrs. C. C. Thornton, W. D. Dorland and R. N. Tedrick of the George C. Marshall Space Flight Center of the National Aeronautics and Space Administration for their help in making available the results of their tests. The assistance rendered by Dr. Harrison E. Cramer of the Round Hill Field Station of the Massachusetts Institute of Technology in collecting and interpreting meteorological data is gratefully acknowledged.

REFERENCES

1. M. J. Lighthill, Proc. Roy. Soc. (London), Ser. A., *211*, 564–587 (1952) and *222*, 1–32 (1954).
2. A. Powell, J. Acoust. Soc. Am., *25*, 385–389 (1953).
3. E. J. Richards in *Proceedings of the Third International Congress on Acoustics, Stuttgart 1959*, L. Cremer, Ed., Vol. II, Elsevier Publishing Co., Amsterdam, London, New York, Princeton (1961).
4. J. E. F. Williams, J. Acoust. Soc. Am., *33*, 1675 (1961) (A).
5. I. Dyer, Chapter 9 in *Random Vibration*, S. H. Crandall, Ed., John Wiley and Sons, Inc., New York and London, (1959).
6. H. v. Gierke, Chapter 33 in *Handbook of Noise Control*, C. M. Harris, Ed., McGraw-Hill Book Co. Inc., New York, Toronto, London (1957).
7. I. Dyer, loc. cit.
8. D. A. Bies and P. A. Franken, J. Acoust. Soc. Am., *33*, 1171–1173 (1961).
9. L. C. Sutherland and W. V. Morgan, NOISE Control, *6*, 6–12, May-June (1960).
10. J. N. Cole, R. T. England and R. G. Powell, Wright Air Development Division, Technical Report 60–6, Wright-Patterson Air Force Base, Ohio (1960).
11. W. D. Dorland, Characteristics of the Saturn Space Booster as a Noise Source, Paper L25, Fourth International Congress On Acoustics, Copenhagen (1962).
12. A. Powell, J. Acoust. Soc. Am., *30*, 1048 (1958).
13. I. Dyer, loc. cit.
14. A. Powell, Chapter 8 in *Random Vibration*, S. H. Crandall, Ed., John Wiley and Sons, Inc., New York and London (1959).
15. I. Dyer, P. A. Franken and E. E. Ungar, NOISE Control, *6*, 31–40, 51, January-February (1960).
16. A. Powell, J. Acoust. Soc. Am., *30*, 1048 (1958).
17. J. N. Cole, et al., loc. cit.
18. NASA, unpublished results.

19. W. J. Galloway, private communication.
20. NASA, unpublished results.
21. P. J. Westervelt, J. Acoust. Soc. Am., *23*, 347–348 (1951).
22. U. Ingard, J. Acoust. Soc. Am., *25*, 405–411 (1953).
23. W. L. Nyborg and D. Mintzer, Wright Air Development Center, Technical Report 54–602, Wright-Patterson Air Force Base, Ohio (1955).
24. F. M. Wiener and D. N. Keast, J. Acoust. Soc. Am., *31*, 724–733 (1959).
25. H. J. Sabine et al., Wright Air Development Center, Technical Report 57–353, Wright-Patterson Air Force Base, Ohio (1958–1961).
26. E. F. Cox, H. J. Plagge and J. W. Reed, Bull. Am. Meteor. Soc., *35*, 95–103 (1954).
27. P. Rothwell, J. Acoust. Soc. Am. *28*, 656–665 (1956).
28. E. F. Cox, Phys. of Fluids, *1*, 95–101 (1958).
29. W. J. Remillard, Office of Naval Research, Technical Memorandum 44, Acoustics Research Laboratory, Harvard University, Cambridge, Mass. (1960).
30. R. N. Tedrick, Determination of Zones Subject to Meteorological Focusing, Paper O12, Fourth International Congress on Acoustics, Copenhagen (1962).
31. D. I. Blokhintzev, *Acoustics of a Nonhomogeneous Moving Medium*, Moskva, Leningrad (1946) English Translation in Technical Memorandum 1399, National Advisory Committee on Aeronautics, Washington, D. C., (1956).
32. H. O. Kneser, in *Handbuch der Physik*, S. Flügge, Ed., Springer-Verlag, Berlin, Göttingen, Heidelberg (1961).
33. E. J. Evans and E. N. Bazley, Acustica, *6*, 238–244 (1956).
34. I. Horiuchi, J. Acoust. Soc. Am., *29*, 772 (1957) (A).
35. W. L. Nyborg and D. Mintzer, loc. cit.
36. U. Ingard and F. M. Wiener, J. Acoust. Soc. Am., *30*, 670, (1958) (A).
37. M. J. Lighthill, Proc. Camb. Phil. Soc., *49*, 531–551 (1953).
38. H. J. Sabine, et al., loc. cit.
39. NASA, unpublished results.
40. R. N. Tedrick, loc. cit.
41. D. E. Goldman, Chapter 11 in *Handbook of Noise Control*, C. M. Harris, Ed., McGraw-Hill Book Co. Inc., New York, Toronto, London (1957).
42. H. W. Ades et al., Benox Report, University of Chicago, Chicago Ill., (1953).
43. W. D. Ward, A. Glorig and D. L. Sklar, J. Acoust. Soc. Am., *31*, 791–794 (1959).
44. K. D. Kryter, A. Z. Weisz and F. M. Wiener, J. Acoust. Soc. Am., *34*, 383–391, (1962).
45. L. L. Beranek, Chapter 13 in *Noise Reduction*, L. L. Beranek, Ed., McGraw-Hill Book Co., Inc., New York, Toronto, London (1960).
46. I. Dyer and F. M. Wiener, J. Acoust. Soc. Am., *33*, 1682 (1961) (A).
47. A. A. Regier and H. H. Hubbard, NOISE Control, *5*, 13–19, 58, September (1959).
48. H. S. Freynik, Jr., J. Acoust. Soc. Am., *34*, 720 (1962) (A).
49. W. D. Dorland, NASA, unpublished results.
50. M. A. Heckl et al., Aeronautical Systems Division, Technical Note 61–122, Wright-Patterson Air Force Base, Ohio (1961).
51. Y. K. Lin, J. Aerospace Sci., *29*, 67–75 (1962).
52. G. Maidanik, J. Acoust. Soc. Am., *34*, 809–826 (1962).
53. R. H. Lyon and G. Maidanik, J. Acoust. Soc. Am., *34*, 623–639 (1962).
54. D. U. Noiseux et al., Wright Air Development Division, Technical Report 60–754, Wright-Patterson Air Force Base, Ohio (1960).
55. D. U. Noiseux et al., Aeronautical Systems Division, Technical Report 62–614, Wright-Patterson Air Force Base, Ohio (1962).

A Decade of Musical Acoustics

by ROBERT W. YOUNG

U. S. Navy Electronics Laboratory, San Diego 52, California

Musical acoustics is that part of acoustics relevant to music. It includes the science of vibrating strings, air columns, plates and rods. The acoustics of a room in which music is performed certainly influences the effect—and architectural acoustics is in a sense a part of musical acoustics. The loudspeaker that generates sound as the output of an electrical musical instrument comes within the broad scope of musical acoustics. What the ear hears is closely related to one's appreciation of music—and psychoacoustics joins with musical acoustics.

The topics mentioned are also, of course, parts of other branches of acoustics. For example, a spectrometer specifically developed for the analysis of musical sounds may turn out to be an extremely useful instrument with which to measure the sound spectrum of a noisy airplane.

The present review of musical acoustics, however, must be restricted, and it is therefore limited to musical instruments, studies of intonation and scales, and the relation of publications on musical acoustics to acoustics as a whole. Even so, this review of a decade of musical acoustics is only illustrative. Much relevant work is not cited specifically, but some hints are given where other recent publications on musical acoustics can be found.

Electrical Music

A wide gamut of musical sounds [4, 19, 37, 53, 61] can be generated or further processed [65, 80] by the aid of electricity. Instruments in which the original signals are generated electrically and under the control of an organ-like keyboard, are now quite widely known; during the past decade new features incorporated in commercial instruments include electronic timbre change during onset of a tone, electronic vibrato, percussion sounds that can be sustained, and electronic choral effects. Electrical solo instruments afford the performer an opportunity to concentrate on the nature of a single tone, its onset and decay, even its *subharmonic* [72] content.

Ideas of randomness, under the guidance of certain selection rules, are being exploited in the composition [85] of music; information theory [62] is being applied. High-speed digital computers are being employed to investigate the processes of composition; to illustrate such capabilities, the *Illiac Suite* for string quartet was composed [24] by a computer oper-

ating on random numbers plus some rules of counterpoint, rhythmic patterns, dynamic markings, and probabilities for melodic intervals. Moreover, a music synthesizer has been developed which does not require a composer to have personal skill in playing all instruments and which can be used in conjunction with a computer that operates on probabilities (63) of rhythms and note sequences to give the composer the opportunity to listen directly to music composed automatically within a broad framework he prescribes.

Still another way of producing music by the digital computer is to use it to generate *any* desired waveform. The prospect of such complete control is exciting indeed, because one can create extremely arbitrary waveforms to test hypotheses about the perception of sounds (musical and otherwise) that were hitherto unavailable. As to this means of composing music, "Present limitations (53) lie in lack of understanding of what waveform will produce a given subjective effect rather than in lack of ability to create a specified waveform. There is no limitation whatever in speed of execution. A computer is very easy to use. The electronic equipment (the computer and output equipment) have been constructed once and for all. There are no soldering irons, tape splicing, or even knob-twisting involved. No manual dexterity or velocity is required. Instead one writes down and gives to the computer a sequence of numbers."

These numbers are used to generate successive electrical impulses of the proper amplitudes and time intervals, to reconstruct (53) the desired waveform. It turns out that if one wants to include components in this waveform with frequencies up to 10 000 c/s, then 20 000 pulses will have to be generated in each second. After the square-topped electrical pulses are fed to a suitable 10 000-c/s low-pass filter, an amplifier, and a loudspeaker, they ultimately appear as the desired sound.

The total quantity of numbers required is prohibitive, however, if they must be individually specified; for example, some 36 million numbers would be needed for Beethoven's Fifth

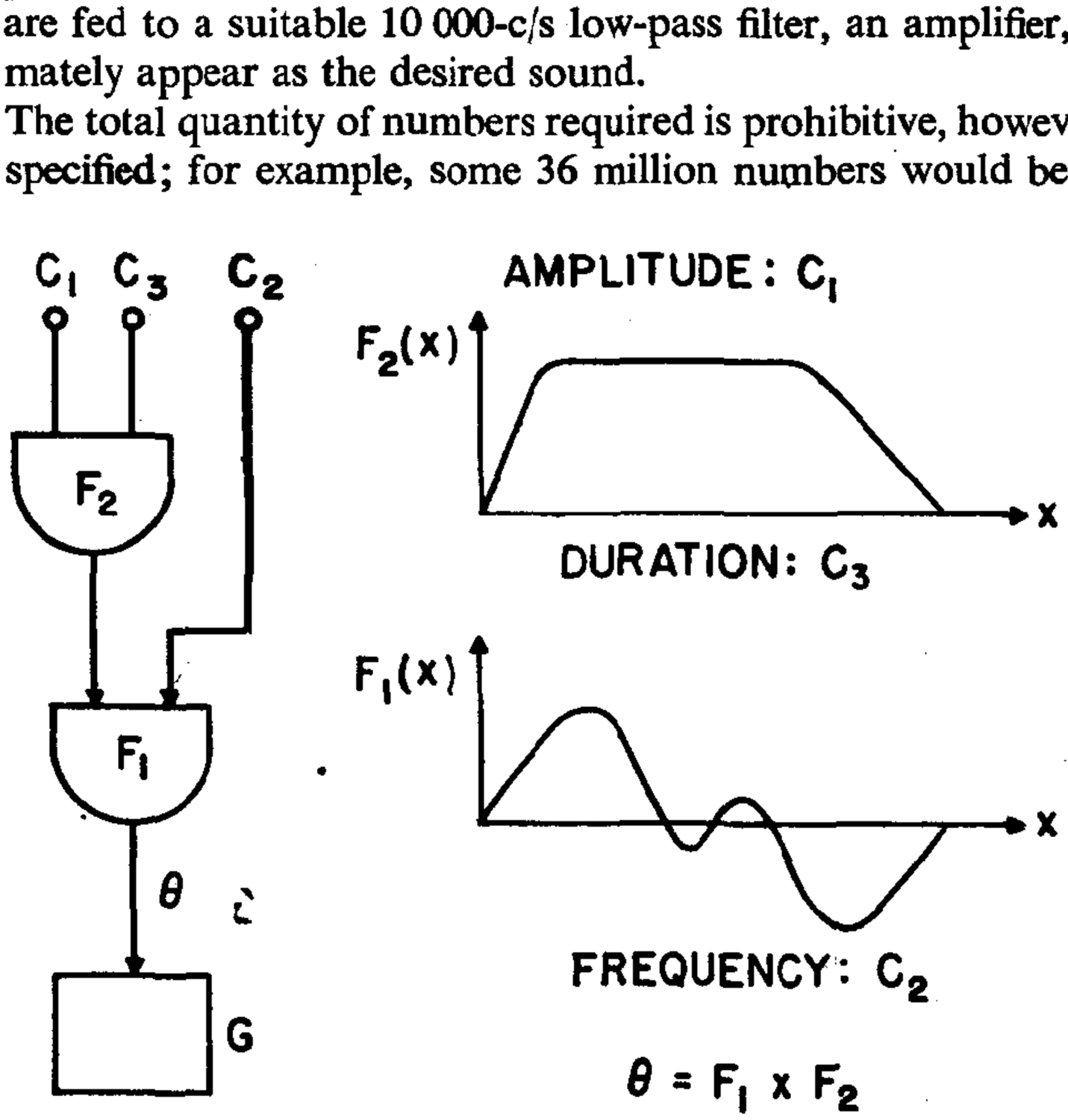


Fig. 1. Elements of a compiler equivalent to a simple musical instrument, to produce a sequence of θ digits.

Symphony. How, then, does one manage the digital computer? Part of the answer is very like the classical procedure of writing a musical score for an orchestra, wherein the positions of the notes on the staves constitute instructions to musicians when and how to play, except that cards are punched to turn on, at proper times, the computer equivalents of musical instruments. Such an equivalent is called a compiler [52] which produces the sequence of numbers equivalent to the sound of a particular musical instrument.

Figure 1 indicates the functioning of such a compiler. G is a generator which leads to an acoustic output, in the manner mentioned above, from an input of Θ digits. Now the magnitudes represented by these digits are determined by the products of two functions F_1 and F_2. As suggested at the lower right of Fig. 1, $F_1(x)$ is periodic but the waveform is quite arbitrary; the frequency at which this function is repeated is determined by control signal C_2, which itself need not be constant. $F_2(x)$ is an envelope control for some given attack and decay characteristic such as sketched at the upper right of Fig. 1; the maximum amplitude is established by the number C_1 and the total duration by C_3.

In the simplest situation the composer calls for a given musical note by use of five numbers. Three of these are C_1, C_2, and C_3; the other two numbers respectively select which instrument is to play, and when. The composer thus has readily available to him many new kinds of sounds; he is now challenged to learn how to call for them at the proper times.

Singing

For our oldest musical instrument—the singing voice—further evidence has been reported during the past decade on formants—those frequency regions where the partials have particularly large amplitudes. The formants near 400 or 500 c/s and near 2800 c/s have been found [70] to be about the same regardless of what tone or what vowel is being sung. Formants in the singing voice differ, moreover, in both width and frequency location from the formants of the speaking voice; it has been suggested [70], therefore, that the initial part of a vowel is sung with the resonant cavities of the head adjusted as for a speech sound but then readjusted quickly to the appropriate singing position.

Quantitative evidence has been published [10] recently showing that, along with the greater dynamic range employed by the experienced singer, high-numbered partials of significant amplitude are more likely to be found in the voice of the professional singer than in the voice of an inexperienced one. The distribution of partials changes with mood; measurable components as high as 8000 c/s have been observed [22] in an "aggressive" baritone voice. Also, the average frequency of a steady tone is likely to be more nearly constant [10] when sung in a normal room than if sung in an anechoic room.

There have been two theories about voice production: one that the resonating cavities function simply to modify the amplitudes of harmonic components of sound generated by the vocal cords, the other that the cavities themselves produce independent sounds as a consequence of shock excitation by the puffs of air from the vocal cords. The second theory leads to partials that tend to be inharmonic. Recent research [10] reveals that *both* harmonic and inharmonic partials are present, particularly during the initiation of the sound.

The vibrato is recognized as an important element of beauty in the singing of European music, although it appears to be lacking [18] in Indian music. The vibrato may consist of

modulations in frequency, amplitude, and/or waveform that may hardly be noticed as such. A typical rate of modulation is about 6 per second and the total extent of the frequency change as much as a semitone. Much of the earlier evidence showed the frequency and amplitude modulations to be in phase; current experiments, [10] however, indicate that for some professional singers a phase difference of 180° sometimes exists.

In one recent experiment [71] various sounds were presented to the singer's ears via headphones. When the external stimulus was a frequency-modulated tone, the singer synchronized her own vibrato with it. When the stimulus was a band of filtered white noise, the singer's vibrato became less regular.

Other experimenters [17, 84] have returned the singer's voice to her ears with more or less delay in the sound introduced electrically. Except for those accustomed to singing on the stage to the accompaniment of an orchestra, an increased delay was found to result in a decreased vibrato rate.

Such evidence plus other facts about the ear lead to the conclusion [17, 71, 84] that the common vibrato rate of 6 or 7 per second is a consequence of basic physiological time-constants. If the vibrato were much slower it would be sensed as a disagreeable tremolo, and if it were much faster it would give instead the impression of a group of tones.

Woodwind Instruments

According to very elementary theory the tone of the clarinet is supposed to consist of only odd numbered partials. Experiment, however, has long demonstrated that the even numbered partials are often of significant amplitude. During the past ten years further theory and experiment have been reported [14] in support of an explanation for the appearance of the even numbered partials based on an assumption that the aperture between the tip of the clarinet reed and mouthpice is never completely closed. A recent experimenter, however, who measured by a phototube the light transmitted through the reed aperture, found [3] that with vigorous blowing the reed closes the aperture completely during half the cycle but that with soft blowing the reed executes an almost sinusoidal motion that leaves the aperture partially open all the time.

In a recent contribution [7] to the acoustics of instruments with lateral side holes, it has been shown that, as a consequence of covered finger holes, the only musically useful bores for woodwinds are those of the Bessel horn family; furthermore, the bores are essentially limited to cones and cylinders. Also, it was demonstrated [8] that a "proper choice of hole dimensions permits the closed-hole part of the bore to behave in all essentials like a smooth walled cylindrical bore of altered diameter in which waves propagate with slightly lowered but constant velocity."

A characteristic feature of a woodwind instrument is the small hole toward the upper end of the instrument—the "register" hole—that is opened when higher notes are played. Certain irregularities of intonation result from the use of this hole.

Figure 2 depicts some average measurements [90] of the intonation [59] of a clarinet. The notes were played in ascending chromatic order, with a pause of about 2 seconds between each note. The measurements were made with a Stroboconn, an instrument with 12 rotating

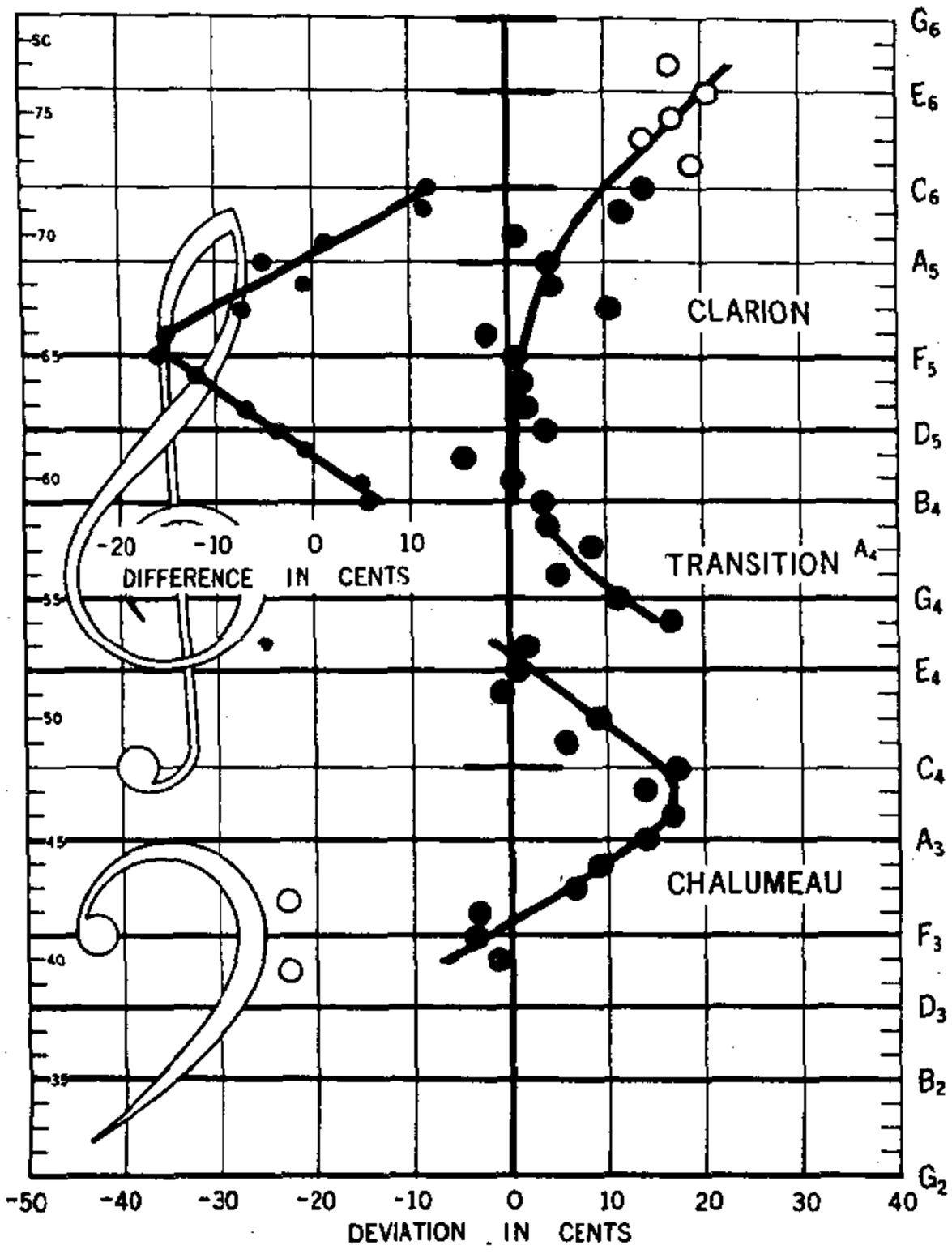


Fig. 2. Average intonation of a soprano clarinet, and differences between intonation of corresponding notes in the clarion and chalumeau registers.

disks originally called a chromatic stroboscope, by which one can measure directly how sharp or how flat a given note is, as a deviation in cents from the corresponding note in the equally tempered scale based on the *A* of 440 c/s. Remember that there are 1200 cents to an octave, and the interval between 440 and 441 c/s is 4 cents.

In this kind of graph, horizontal lines represent the musical staff and ledger lines, but they are spaced unequally to make the scale of chromatic intervals uniform. Note names appear at the right; they are further identified by a subscript for each octave. The usual tuning note is here called A_4. A purely numerical scale is also available at the left of of the figure, representing the number of semitones counted upward from $C_0 = 16.35$ c/s.

Different parts of the playing range of the clarinet are given distinctive names: chalumeau, transition, clarion. These are marked in the figure. Usually the highest of the chalumeau register is said to be E_4, but for the present purposes it seems proper to include F_4 in the chalumeau register, and to add C_6 to the clarion register. Above the clarion register is a series of high notes marked by open circles. Points are plotted for the notes as they are written for the clarinet; since this clarinet is in $B^\flat$ the actual sounds are a whole tone lower than indicated by the figure.

Notes of the clarion register are obtained with the same holes open as for corresponding notes a twelfth lower in the chalumeau register; the nominal frequency ratio is 3:1 between

corresponding notes. The shift from the chalumeau to clarion register is accomplished by opening the register hole.

The function of the register hole is to insure that a velocity antinode will occur in its vicinity. If the register hole is not situated exactly at an antinode the resulting tone is sharpened an amount increasing with the displacement from the antinode. The consequences of this effect can be seen by inspection of the *differences* in intonation between corresponding notes of the chalumeau and clarion registers. These differences have been plotted at the upper left of Fig. 2. For example, the interval between F_4 and C_6 is 12 cents greater than an equally tempered twelfth: a point is therefore plotted at + 12 on the "difference" scale and on the line for C_6.

The difference is a minimum near F_5; on either side of this note the differences increase. One concludes, therefore, that for the note F_5 the register hole is located at the velocity antinode, and that the antinodes for higher notes occur above the register hole and antinodes for lower notes occur below the register hole.

As a consequence of this sharpening due to the register hole, if the notes of the clarion register are to be nearly in tune with equal temperament, the notes of the chalumeau register must be relatively sharp in the general vicinity of B_3; Fig. 2 shows this to be the case.

Brass Instruments

The actual shapes of the air columns in brass wind instruments are not readily described by well known mathematical functions, so that a theory to give the natural frequencies of such instruments must be adaptable to horns of arbitrary shapes. Recently, by an approximation to the actual shape by a manifold of short exponential horns, the acoustical impedances of a trombone and fluegelhorn were calculated [86] with a high-speed digital computer; theoretical evidence was obtained in support of a long-standing suspicion that the frequency of the fundamental mode of vibration of a fluegel horn in $B^\flat$ is really near $A^\flat$. When impedance calculations and frequency measurements were made [38] on a system consisting of a small chamber, a cylindrical section, and a Salmon horn, the resonances were found in general to be inharmonic; the effect of the chamber, however, can be balanced against a suitable widening of the horn to produce harmonically related resonances. On the experimental side, the resonance frequencies of good and bad trumpets were measured [26] by the help of a high-impedance electroacoustic driver; the inharmonic resonance frequencies so obtained were about 5% greater than those produced by a player in the normal manner. Quantitative evidence was recently reported [1] on the formants in the sound of a cornet played with four different mutes. In general, a mute tends to reduce the amplitudes of the lower-frequency partials; it further causes a "notch" in the sound spectrum in the vicinity of 300 c/s that is related to resonance in the mute.

Compromises in the intonation of brass wind instruments are necessary if a chromatic scale is to be played with only three valves. The difficulties of intonation of the brass wind instruments have been placed [81] next to those of the clarinet. It can be demonstrated readily, on the simple assumption that the frequency varies inversely as length of instrument, that if the three valve tubes are adjusted individually so as to lower the intonation by 2, 1, and

3 semitones respectively, then the notes obtained with valves in combination should be quite sharp—specifically, 55 cents sharp. Measurements [13.n] of intonation of trumpets and cornets do indeed reveal notes that are relatively sharp when played with the three valves in combination, but the amount is less than 40 cents, instead of 55 cents. From such evidence it can be shown [89] that the frequency varies roughly as the inverse 3/2-power of the actual length of the cornet.

As a consequence of this more marked length dependence, the compromises in intonation of cornets and trumpets need not be as severe as long supposed. Further, by application of the principal of least squares, it has been recently demonstrated [89] that the discrepancies in intonation for equal temperament can be minimized by adjusting the valve slides so that the tones produced individually are about 5, 1, and 16 cents flat respectively; then the tones that must be obtained by the three valves in combination are only about 16 cents sharp. Such deviations are almost within the normal variability of solo playing. When the notes are tested in ascending order, for each combination of valves in order, and the tests are then repeated, the standard deviation of the various measurements on any one note is likely to be 7 or 8 cents; that is, two-thirds of the measurements are likely to fall within 8 cents of a given average value. These are differences in intonation that occur without the knowledge of the player. Whether this kind of variability exists when brass instruments are played in an orchestra remains to be demonstrated.

Violin

Investigators of the past decade have continued to use response curves of violins in their search for physical characteristics that are correlated with musical excellence. Such response curves delineate as functions of frequency, perhaps up to 10 000 c/s, the airborne sound pressure that results from mechanical excitation of one kind or another. The excitation may be provided, for example, by an electrodynamic driver [44] attached to the bridge of a violin. In addition to response at the driving frequency, evidence [35] of nonlinear response can be observed; subharmonic series of resonances have been identified. [25] A wide-band mechanical excitation [13.q] can be had by an impulse produced by the striking of a 7-gram steel ball against a bracket attached to the bridge of a violin.

A mechanical bow may be run steadily while a narrow-band sound analyzer is swept slowly over the frequency range of interest. Some experimenters prefer to produce the sound by normal bowing; in this case magnetic tapes loops are likely to be used to prolong the sound and numerous analyses [55] are made for the original sound displaced by small frequency intervals.

From such evidence it appears that in violins of good quality, the response is relatively greater in the range from about 250 to 1000 c/s and from 2000 to 3000 c/s than is the case for mediocre violins; a broad minimum in response near 1500 c/s prevents a good violin from having an undesirable nasal quality. A violin of good carrying quality [55] has its maximum in response at low or medium frequency.

The large response near 300 c/s has been traced to the resonance of the air filled cavity open to the air via the *f*-holes. For various reasons, long *f*-holes seem to be better [73] than short ones, but changes in shape and area are difficult to make in view of the traditional appear-

ance of a violin. A succession of closely spaced resonances found in old Italian violins leads to improved efficiency even though the resonances are less damped [43] than is the case for modern violins.

Many of the resonances at higher frequencies can be traced to the complex modes of vibration of the body of the violin. With mechanical excitation these vibration shapes, frequencies, and phase relations have been plotted out [20, 32, 64] in detail. The resonances shift [2] somewhat with humidity. Some correlation has been observed [28] between the nodal lines on belly and back; a nodal line of every mode passes near the sound post. Frequencies of the normal modes for violas seem to be about 0.8 to 0.9 those of violins as one might expect from the relative sizes, but there does not appear [29] to be similar correspondence for the relatively thicker violoncello.

The vibration patterns depend, of course, on the thickness of the plates; thinning near the edges of the top plate of a viola had the beneficial effect of increasing the general response [73] some 2 dB. Resonances of the plates can be tested [46, 79] with an electrodynamic driver or by tapping them before the instrument is assembled. When the resonance frequencies of the top plate alternate between those of the back plate, and when the intervals are not greater than a whole tone an instrument is judged good. [25]

Varnish and kind of wood [45] have been, of course, subject to question for a long time. Spruce traditionally used for the top plate of a violin, is found [55] to possess internal damping that increases more rapidly with frequency than is the case for other kinds of wood; likewise soft varnish [55] increases the high frequency damping more than does hard varnish.

In addition to the well-recognized transverse motion, a string can vibrate longitudinally and also in torsion. The frequency of the torsional motion is lower [39] than the fundamental frequency of transverse motion, but the longitudinal motion introduces a generally inharmonic component mingled with the upper partials. Now the frequency of longitudinal motion is essentially independent of tension, but it varies with material [38] and with humidity. Consequently this longitudinal component may be a harmonic on some days yet not on others—with a resultant variation in the quality of the sound.

The early appearance of high frequency components in the transient sound of a violin can be related [55] to the fact that higher partial vibrations are produced with only slight force by the bow; theory on this subject was recently modified [30] by inclusion of a term to account for yielding of the bridge. Other recent theory [33] indicates that, except for bowing at the center of the string, steady periodic motions of a bowed string must be preceded by an aperiodic motion. The theory also leads to the observation that a definite change in tone quality is to be had by increasing the speed of bowing along with decreasing the bowing force, because a simultaneous increase of both would simply increase amplitude of all partials in the same ratios. Extensive experiments [11] on strings, however, lead to the observation that an increase in bowing speed, other factors being constant, results in a relative reduction of the amplitudes of the higher partials whereas an increase in bowing force results in a marked increase in the relative amplitudes of the higher partials.

Finally, it should be remembered that the impression of the tone of a violin depends upon the ear, and nonlinear distortion evident in both good and poor violins [35] affects the envelope of the violin sound that is ultimately judged [76] by the ear.

Organ and Piano

The revival of interest in baroque organs, started some 40 years ago and accentuated by the rebuilding of organs [41, 42] after World War II, has inspired a practical curiosity as to what, physically, characterizes the sounds of all kinds of organs. Acoustical features of the church containing an organ have also been considered. [31, 69]

Significant technical peculiarities [41] of baroque organs are the direct mechanical connection between key and air valve (the tracker action), windchests on which all the pipes for a given key are mounted together, and relatively low wind pressures. All of these features affect the transient sounds. "In spite of their evanescent nature, the view is now held [68] that it is these transients which enable the listener to distinguish the sounds of different musical instruments or between two of the same class . . . The transient . . . ought to be exhibited as a characteristic alongside the steady-state spectrum."

Experiments [12, 48] demonstrate that when the air is admitted to a pipe rapidly, the initial transient contains noise-like and inharmonic components which are almost absent when the air is admitted slowly. The closing transient depends primarily on the cavity under the valve.

Pipes mounted on a common windchest are subject to coupling [40] principally via the surrounding air and only secondarily via the windchest. The coupling is greatest between pipes having the same fundamental frequencies and the coupling increases with decreasing separation. Consequently, if one wants to attain a fusion of the sounds from different pipes associated with the same key on the organ, then these pipes should be mounted as near to each other as possible.

Recent measurements of the resonances of organ pipes curiously reveal [58] a stronger damping at the fundamental resonance than for the second or third resonance. Quantitative measurements [58] of the phase angle for reflection at both the mouth and open end show a curved dependence upon frequency, which means that the resonances are inharmonic; furthermore, at the higher frequencies there is almost complete transmission so that the partials in the sound of the organ pipe extend only up to some limiting frequency. [57] Since the phase angle depends upon the "height of mouth"—a dimension that is changed [66] in the voicing of organ pipes—it follows that the voicing modifies the inharmonic relations [16] among the resonances. Moreover, the different materials, of which organ pipes are made, contribute differently [49] to the inharmonic components.

Inharmonic partials [87] likewise characterize the tones of a piano. Extensive subjective tests with tones produced by a synthesizer [21] have demonstrated that the "warmth" of piano tones is due to the inharmonicity of the partials, and the partials for strings below middle *C* must be inharmonic for the sounds to be piano-like. The tests also showed that, for piano-like sounds an octave or so below middle *C*, the attack time should be less than 0.1 sec (meaning the time for the sound pressure to attain 0.9 of its maximum value).

General tests proposed [60] for judging the quality of piano tones include measurements of the decay of the partials measured in several octave bands (three keys may be struck simultaneously for some tests) and, most importantly, measurements of the initial transients. Finally, in regard to the piano it may be mentioned that the efficiency of transfer of energy from hammer to string has been found [23] to be maximum when the striking point is at

about $^1/_8$ the length of the string and the hammer is twice as heavy as the string; for a bass string, however, the hammer is relatively light and the transfer of energy is relatively inefficient.

Scales and Intonation

Perhaps unique in the history of musical acoustics was a 3-day international colloquium at Marseille in 1958, devoted solely to the musical scale and to the physics of musical instruments based on vibrating strings or air columns. A complete record of the conference has been published [13], but unfortunately it has not been abstracted and papers on specific topics cannot be located readily.

The number of theoretical musical scales is myriad, and proposed new scales [74] and methods of tuning [75] are to be compared with those of earlier years. Such comparison is greatly aided by a concise historical survey of tuning and temperament published [5] in the past decade.

Some experiments [6] with violin tones suggest that there are really two scales, one the melodic scale of Pythagoras and the other the harmonic one of Aristoxenes often called a "just" scale. In another recent experiment, however, diatonic scales played as melodies on an electronic organ were preferred [82] in equal temperament rather than in just tuning. Moreover, with a variety of chords and tonal sequences played [51] on a piano, the "stretched" tuning usually followed by piano tuners—which in the midfrequency range suggests the Pythagorean scale—was preferred over strict equal temperament. When the intervals used by famous choirs were measured, [47] the major third was typically sung a little larger than even a Pythagorean third; moreover, the variability in intonation was such that the average deviation of an interval from its median size was often more than 20 cents.

A significant step toward good intonation was the international agreement [27], in 1955, on the frequency of 440 c/s for the tuning *A*. Music contains notes other than *A*, however, and questions still remain as to how the intonation of these notes should be judged.

Different physical vibrators, such as strings or air columns, are employed in musical instruments, so that each kind of instrument exhibits its own characteristic irregularities [13.n, 56, 78] in intonation, such as those mentioned earlier for the clarinet and cornet. Also, the player of a wind instrument can intentionally change the intonation over a range of 10

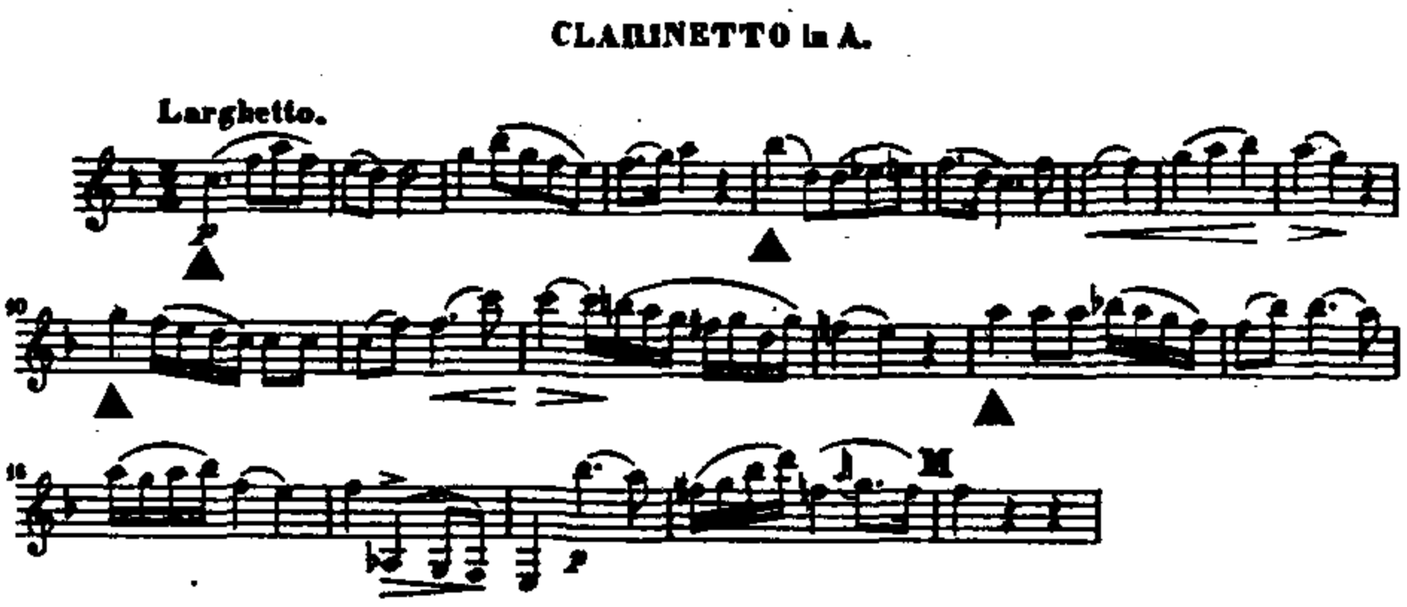


Fig. 3. Music played by clarinet from Quintet No. 6 by Mozart.

or 20 cents or more. There is indeed the possibility that different intonation is desirable [54] in different musical contexts. Even under controlled conditions [50] different sections of an orchestra may be tuned differently—discrepancies of 10 or 15 cents are not uncommon. The unison strings on the piano are not tuned exactly in unison; tests [34] indicate that a preferred musical effect is attained when there is a difference of 1 or 2 cents among the three strings that are nominally tuned to the same frequency. Moreover, discrepancies greater than 5 cents are common even when musicians try [15] to tune their instruments in unison. All in all, exact unisons are unlikely.

Most tests of intonation of instruments have been made without accompaniment and, in the case of the clarinet mentioned above, in simple chromatic order. Does a player perform differently when he plays a melody? What one really wants to know is the intonation that exists during an actual performance.

Intonation measurements were made [13,n] during the playing of the first twenty measures of the second movement (the Larghetto) of the Quintet No. 6 by Mozart for clarinet, two violins, viola and violoncello. The clarinet music is shown in Fig. 3. Notice that the four notes marked with triangles are played after rests; otherwise the notes are played in a very connected fashion. This melody was played in the Key of *F* as shown in Fig. 3, and also in the Key of *E* and in the Key of *G*; all this was done with a $B^{\flat}$ clarinet without and with piano accompaniment appropriately transposed.

Measurements were made with an arrangement that provided an automatic record. The pointer of the chromatic stroboscope was connected by a wire to the pen of a Brüel and Kjaer recorder and a special mechanism was constructed to lift the pen from the paper until the experimenter depressed a bar. While the music was being played the experimenter adjusted the chromatic stroboscope to the intonation of the quarter and half notes in the music and then recorded the measurements simply by depressing the bar; the resulting

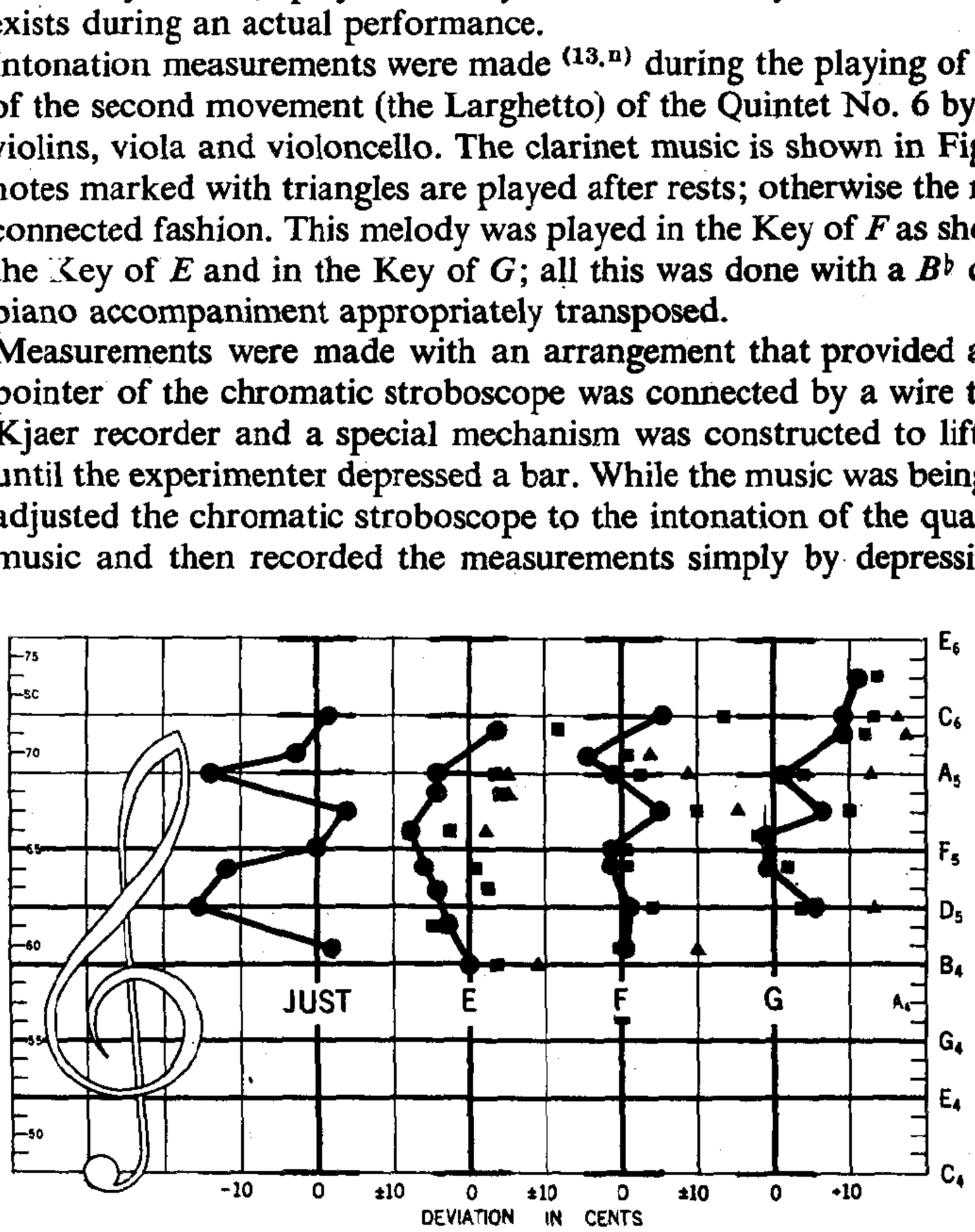


Fig. 4. Average intonation of quarter and half notes in the Larghetto as played by a clarinet, with the music in the Keys of E, F and G: see large circles connected by heavy lines. Smaller squares represent the intonation of the same notes but as measured in chromatic tests. For each key, triangles signify the average intonation of the four notes played just after rests. The pattern marked "Just" displays the deviations from equal temperament for a fixed just scale in the Key of F.

marked tape could be read leisurely at a later time. Eight sets of measurements were made without accompaniment and five sets with accompaniment.

For the experiments with the accompaniment, the piano was tuned intentionally 10 cents below the standard equal temperament based on the *A* of 440 c/s. There was a slight indication that when the clarinetist played with the piano he lowered the intonation of all notes by an average amount of about 2 cents, but he certainly did not come down to unison with the piano. Therefore measurements for the music in a given key were averaged together, whether with or without accompaniment. The average results are displayed as circles in Fig. 4. The notes names are given at the right and deviations from equal temperament are plotted horizontally.

The small squares in Fig. 4 represent the averages for the chromatic tests described before. In general, these notes were sharper than those played in the melody.

It is evident that the same pattern of intonation was not followed in each key. Rather it appears that the intonation followed the trend as observed in the chromatic tests: the circles tend to parallel the squares. The note G_5, for example, was consistently sharp whether it occurred as the tonic in the Key of *G* or the supertonic in the Key of *F*.

The pattern at the left in Fig. 4 marked "Just" displays the deviations from equal temperament for a fixed just scale in the Key of *F*. The third in such a scale (the note *A*) is at—14 cents in comparison with a third in equal temperament, and the sixth (the note *D*) is at —16 cents in comparison with the equally tempered sixth. On a graph like this the just scale pattern of intonation has the same shape regardless of the key. There is little in the experimentally observed patterns of intonation that suggest conformance with the pattern for fixed just intonation.

Early in these tests it was observed that notes played after rests were consistently sharp. Data for the notes after the rests were therefore averaged separately, and are represented by triangles in Fig. 4. The average amount of sharpening was 10 cents, and it can be seen in the figure that the triangles lie consistently to the right of the respective circles by nearly this amount.

In brief, this preliminary experiment with clarinet intonation indicates that notes following rests were played relatively sharp by 10 cents, and that adjustments were not made for the key in which a simple melody was played, either with or without a slightly flat piano accompaniment. Irregularities in intonation appear to result from the clarinet itself and not from the musical context.

Publications in Musical Acoustics

As scientific publications of all kinds become more voluminous—and this has certainly happened during the past decade—one seems to notice fewer papers on musical acoustics. Is the number actually decreasing or are the papers merely concealed by the abundance of papers on other subjects?

Fortunately, it is perhaps easier to locate papers on acoustics than those in some other branches of science. Many of the papers on acoustics are published in specialized acoustical journals and furthermore, acoustical papers published elsewhere are also listed in these

journals. Although there are indeed omissions from these lists one can find in them a great many of the papers on musical acoustics as well as on acoustics as a whole.

Table I summarizes the number of papers on acoustics published during the past ten years in the most widely distributed acoustical journals. The inclusive dates of publication are shown. The number of papers in musical acoustics is also shown, and the ratio of these to the total (expressed as a percentage). *Musical acoustics* is here defined in the limited sense as including only musical instruments, the voice, and scales and related matters of intonation. Two items require further explanation. The "Contemporary Papers" are acoustical papers listed in the Journal of the Acoustical Society of America but published elsewhere. Secondly, Gravesaner Blätter is not a general acoustical journal but is dedicated to music, electroacoustics, and related problems of science; nevertheless, only part of its papers are counted under musical acoustics with the present restricted definition.

In round numbers, the table indicates that 20 000 papers on acoustics were published during the past ten years, of which only two per cent deal with musical acoustics.

In order to obtain some idea of whether this relation is changing, similar statistics were compiled for the time before World War II. Information about "Contemporary Papers" is not available but, as shown near the bottom of the Table I, statistics for this period are available from three journals. Notice that, in these three journals, 75 papers on musical acoustics were published in six years—an average of 12 per year; 13% of the papers in these journals dealt with musical acoustics. During the past ten years, in Acustica and the Journal of the Acoustical Society of America the average was only 10 papers per year on musical acoustics.

TABLE I.

Papers published in acoustical journals, including letters to the editor and original reviews.

Inclusive Years	Journal Name	Papers		% Music
		All	Music	
1952–1961	*Acustica*	746	48	6
1955–1961	*Akusticheskiĭ Zhurnal*	478	3	1
1952–1961	*"Contemporary Papers"*	17400	320	2
1955–1961	*Gravesaner Blätter*	163	75	46
1952–1961	*Journal of Acoustical Society of America*	2100	53	3
1952–1961	*Journal of Acoustical Society of Japan*	318	3	1
	Total (estimated, less duplicates)	20000	400	2
1936–1941	*Akustische Zeitschrift*	170	20	12
1936–1941	*Journal of Acoustical Society of America*	310	51	16
1932–1939	*Revue d'Acoustique*	117	4	3

Figures 5 gives on a logarithmic plot some of the details of the publications records of these journals, and affords an insight into the steadily increasing rate of publication during the past decade. The papers presented at the First International Congress on Acoustics and published in the 1954 volume of Acustica have not been included in the graph for that journal. Instead, the proceedings for each of the three Congresses have been prorated over three years respectively, and have been added in the totals represented by the top curve.
As an aside, it may be mentioned that the review [67] of musical acoustics given at the first Congress can be located readily by author and subject, in lists of acoustical publications, because it also appeared in Acustica and was treated as a regular journal paper. In contrast, the status of musical acoustics as described [77] at the Second Congress and the

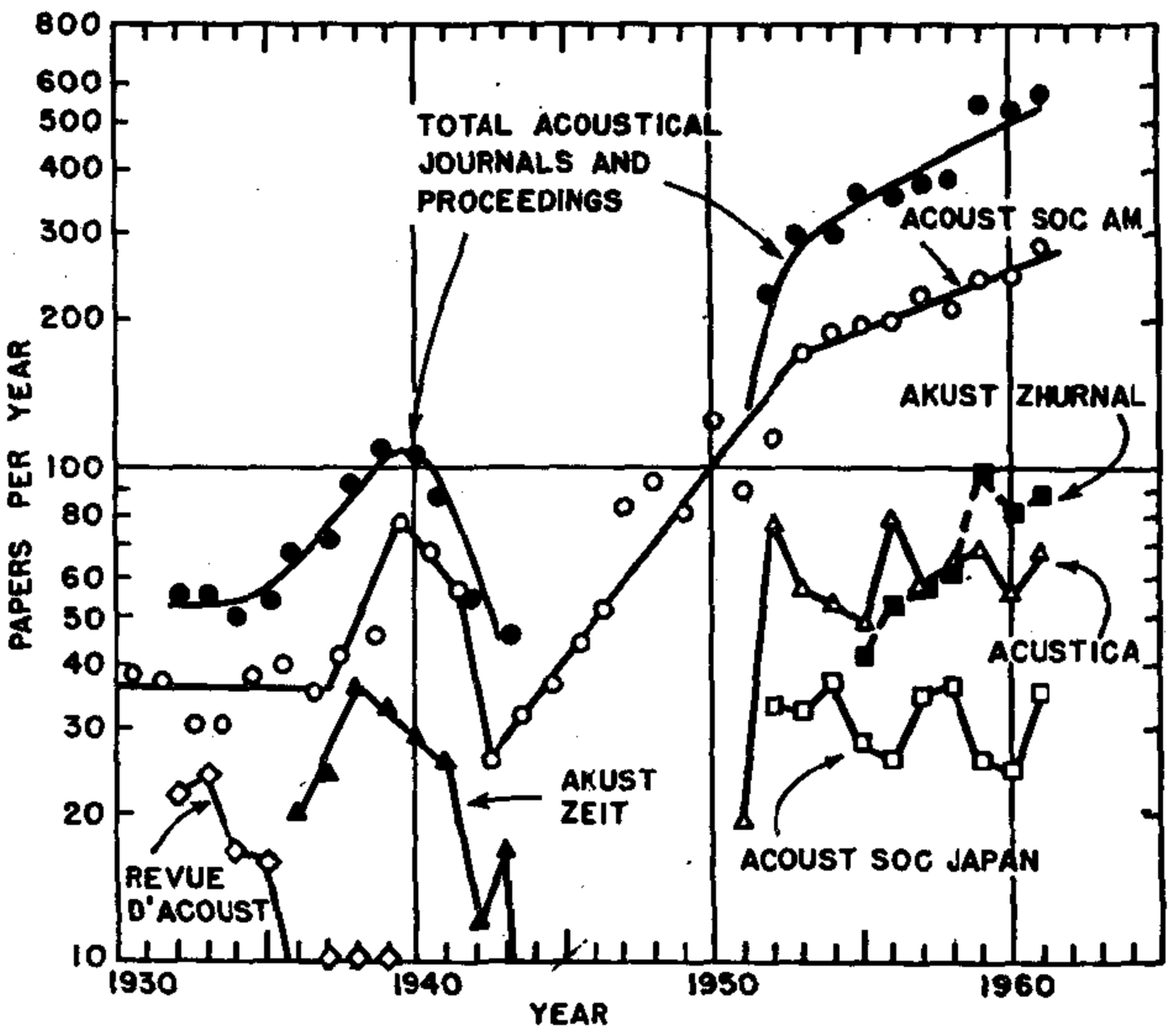


Fig. 5. Numbers of papers per year including letters to editors published in acoustical journals since 1930. The totals include also papers published in the Proceedings of the first three International Congresses on Acoustics, prorated per year.

appraisal [9] of the acoustical environment for an organ, as well as all the other papers printed in the Proceedings, are not listed among the "Contemporary Papers" on acoustics nor in Physics Abstracts nor in Physikalische Berichte. Unless some different procedure is followed for the Proceedings of the Third ICA, the several hundred papers on acoustics from that Congress are likewise destined for bibliographic oblivion.
It is evident from Fig. 5 that since 1953 there has been a rather steady increase in the rate of publication of acoustical papers as a whole, the rate increasing about 7% per year which means a doubling in 10 years. Thus, if one compares the total over 100 papers in these journals in 1940 with the 1960 total of about 500 papers, it becomes evident that acoustical papers now appear at more than four times the 1940 rate and that these two rates twenty years apart are consistent with the recently observed yearly increase. Such a trend suggets that the growth of acoustics is a consequence of some world-wide influence only interrupted by World War II.

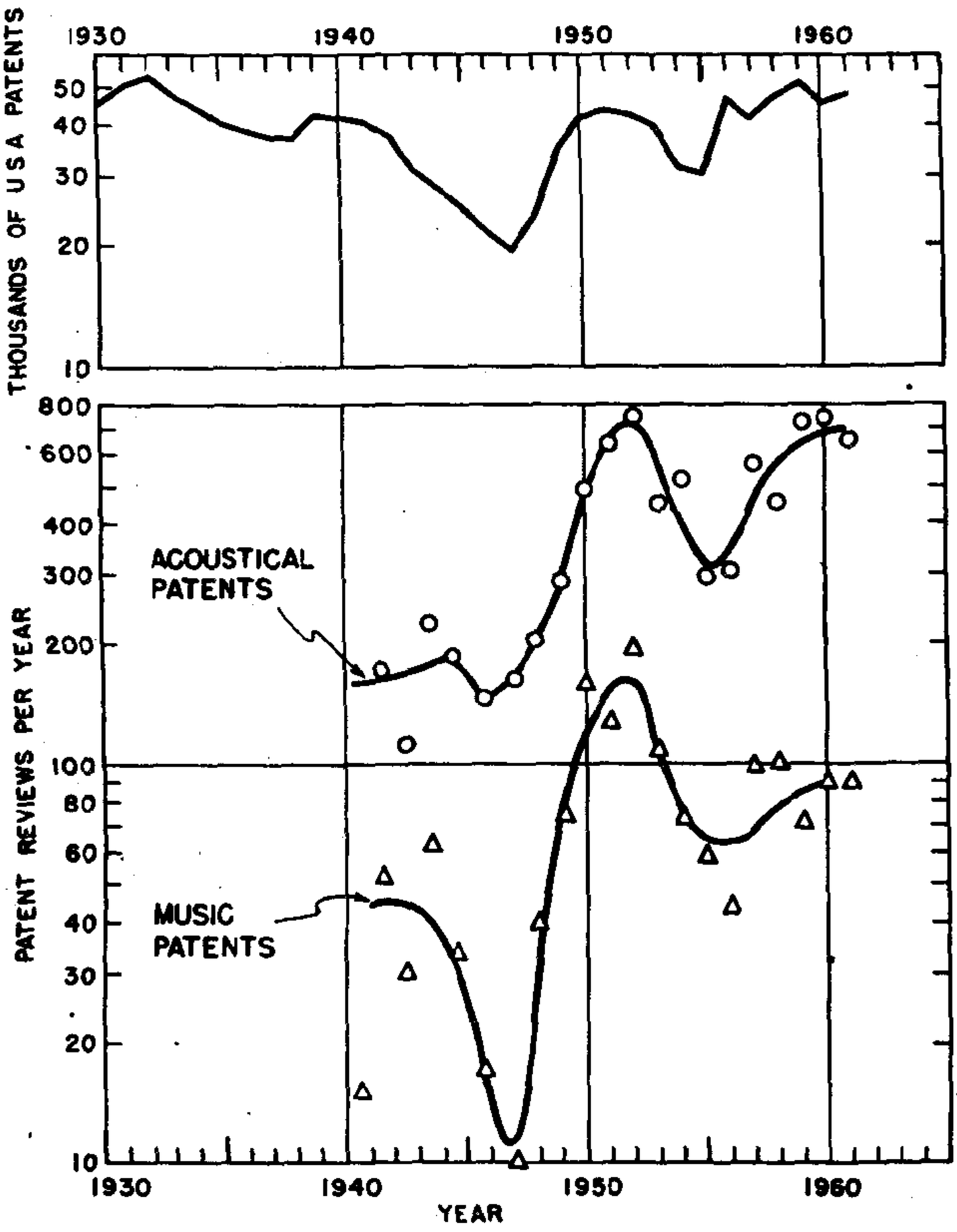


Fig. 6. Number of patents issued per year by the United States Patent Office since 1930, the number of these patents reviewed each year in the Journal of the Acoustical Society of America, and the number of the patents reviewed pertaining to music.

What happened to acoustical technology during this period? World-wide statistics for acoustical patents are not readily available but probably the situation may be judged adequately from United States patents reviewed in the Journal of the Acoustical Society of America. Figure 6 shows, at the top, the numbers of all kinds of patents issued each year by the United States Patent Office; note the steady decline that accompanied World War II about 1940. At the center, the graph shows numbers of acoustical patents reviewed each year, and the graph at the bottom shows the patents reviewed on musical instruments. Observe that the output of acoustical patents about 1950 increased at a much faster rate than did the output of patents of all kinds, and musical instrument patents increased at an even faster rate. In view of the delay of several years between an application for a patent and the appearance of a review in the Journal, it may be concluded that the marked increased in reviews of musical patents, about 1950, simply reflects the pent-up ideas on musical instruments that could not be developed until after World War II.

It follows from Fig. 6 that now about 1.5% of all patents issued by the United States

Patent Office are acoustical patents, and 12% of these concern music. About 1940, for comparison, 0.4% of all patents issued were acoustical but 25% of these acoustical patents pertained to music.

Conclusions

The patent statistics show that from 1940 to 1960 the rate of appearance of acoustical patents as a whole increased by a factor of four while the rate for musical patents doubled. By contrast, publications on musical acoustics (as indicators of scientific activity), are continuing to appear still at about the same rate as they did 25 years ago, and they have not shared the 4-fold increase of acoustical papers as a whole. The outlook for musical acoustics is no better now than it was nine years ago when a similar analysis [88] led to nearly the same relationships among patents and publications; it continues to be true that, to provide meaningful guidance for future technology, "research on musical acoustics must be revived, for the benefit of those who will listen to music in years ahead."

Research of the past decade has produced evidence of inharmonic components in the sound of every traditional musical instrument, including the voice. Even though such inharmonic components may disappear from the steady sound they are most significant in the initial transient that distinguishes one musical instrument from another.

Variability in intonation is typical of all musical instruments, whether unconscious or intentional by the use of vibrato or strings or pipes not tuned exactly in unison. Intervals identified by simple frequency ratios are to be found in a statistical sense only, if at all. From all sides comes the evidence that uncertainty in frequency is both a practical and artistic [83] requirement in music.

In brief, an essential acoustical element of music is noise—more specifically, organized noise in the sense of randomness within limits. Although sine waves continue to serve as simple conceptual elements of musical sounds, the sounds of actual musical performance appear to approach stochastic composites *within limits*. A challenge of musical acoustics of the next decade is to measure and interpret those limits.

REFERENCES

1. James E. Ancell, "Sound Pressure Spectra of a Muted Cornet", J. Acoust. Soc. Am. 32 (1960) 1101–1104.
2. Ulrich Arns, "Eine neue Art objektiver Qualitätsfeststellung von Geigen", Gravesaner Blätter Heft 7/8 (1957) 92–116.
3. John Backus, "Vibrations of the Reed and Air Column in the Clarinet", J. Acoust. Soc. Am. 33 (1961) 806–809.
4. H. Badings and J. W. de Bruyn, "Electronic Music", Philips Tech. Rev. 19 (1958) 191–201.
5. J. Murray Barbour, *Tuning and Temperament*, Michigan State College Press, East Lansing, 1953, Second Edition.

6. Mehdi Barkechli, "Mesure des Intervalles Harmonique de la Gamme à Partir de la Sensation Subjective de Consonance", Acustica 2 (1952) 243–250.
7. A. H. Benade, "On Woodwind Instrument Bores", J. Acoust. Soc. Am. 31 (1959) 137–146.
8. A. H. Benade, "On the Mathematical Theory of Woodwind Finger Holes", J. Acoust. Soc. Am. 32 (1960) 1591–1608.
9. E. Power Biggs, "A Musician and His Acoustical Environment", Proc. Second Int. Cong. Acoustics, American Institute of Physics, 1957, p. 50–52.
10. Alf Bjorkland, "Analyses of Soprano Voices", J. Acoust. Soc. Am. 33 (1961) 575–582.
11. B. Bladier, "Contribution à l'étude des cordes du violoncelle", Acustica 11 (1961) 373–384.
12. R. S. Caddy and H. F. Pollard, "Transient Sounds in Organ Pipes", Acustica 7 (1957) 277–280.
13. F. Canac, Editor, *Acoustique Musicale*, Editions du Centre National de la Recherche Scientifique, Paris VII, 1959.
    a. M. Barkechli, "Quelques idées nouvelles sur la consonance", p. 21.
    b. A. D. Fokker, "Les cinquièmes de tons: les subtilités des lignes mélodiques et l'enrichissement des harmonies", p. 29.
    c. M. Barkechli, "Évolution de la gamme dans la musique orientale", p. 39.
    d. J. Chailley, "Le problème de l'harmonique 7 devant l'historie musicale", p. 47.
    e. R. Tanner, "Le rôle de la fusion dans l'appréciation de justesse des accords", p. 49.
    f. F. Winckel, "Influence des facteurs psycho-physiologiques sur la sensation de consonance-dissonance", p. 61.
    g. J. Brillouin, "Réflexion sur les problèmes d'acoustique musicale", p. 75.
    h. R. Tanner, "Le problème des rapports simples. Notion de psycharithme", p. 83.
    i. R. Dussaut, "Proposition de quelques réformes en théorie acoustico-musicale", p. 103.
    j. P. Riety, "Le vocabulaire d'acoustique d'après la formation du langage musical", p. 111.
    k. L. Pimonow, "La détermination de la qualité des instruments de musique au moyen de spectre sonore", p. 119.
    l. T. Tarnóczy, "Recherches sur le spectre de l'orgue", p. 135.
    m. Dom J. Kreps, "Les synthèses sonores de l'orgue à tuyaux", p. 145.
    n. R. W. Young, "Sur l'intonation de divers instruments de musique", p. 169.
    o. W. Lottermoser, "L'examen acoustique des violons dans le Physikalish-Technische Bundesanstalt", p. 185.
    p. B. Bladier, "Contribution à l'étude du violoncelle", p. 193.
    q. A. V. Rimski-Korsakov, "Recherches sur le timbre des violons et guitares et sur l'excitation des vibrations d'une anche d'harmonium", p. 203.
    r. Z. Carrière, "Entretien d'un tuyau à anche libre", p. 215.
    s. J. Guittard, "Mesures d'impédance acoustique et de fréquences de résonances par différents systèmes de tuyaux sonores et de résonateurs", p. 221.
    t. J. Mercier, "Étude de la stabilité des oscillations entretenues dans un tuyau sonore couplé à un tuyau mort", p. 231.
    u. M. Grützmacher, "Le spectre d'un son de cloche", p. 239.
    v. E. G. Richardson, "L'acoustique des cors et des trompes", p. 245.
    x. R. Husson, "La physique du pavillon pharyngo-buccal dans la phonation", p. 253.
14. R. G. Chatterji, "Theory of the Clarinet", Proc. Nat. Acad. Sci. India A21 (1952) 261–267.
15. John F. Corso, "Unison Tuning of Musical Instruments", J. Acoust. Soc. Am. 26 (1954) 746–750.
16. H. Danzer, "Stationäre Schwingungen der Orgelpfeifen", Zeit. Physik 162 (1961) 516–541.
17. J. A. Deutsch and J. K. Clarkson, "Nature of the Vibrato and the Control Loop in Singing", Nature 183 (1959) 167–168.

18. B. C. Deva, "Vibrato in Indian Music", Acustica 9 (1959) 175–180.
19. Alan Douglas, *The Electrical Production of Music*, Philosophical Library, New York, 1957.
20. F. Eggers, "Untersuchung von Corpus–Schwingungen am Violoncello", Acustica 9 (1959) 453–465.
21. H. Fletcher, E. Donnell Blackham, and Richard Stratton, "Quality of Piano Tones", J. Acoust. Soc. Am. 34 (1962) 749–761.
22. D. B. Fry and Lucie Manén, "Basis for the Acoustical Study of Singing", J. Acoust. Soc. Am. 29 (1957) 690–692.
23. William H. George, "The Acoustics of the Piano", a lecture given 9 October 1957 for, and later printed by, The Institute of Musical Instrument Technology, London.
24. L. A. Hiller, Jr. and L. M. Isaacson, *Experimental Music*, McGraw Hill, New York, 1959, p. 178.
25. Carleen M. Hutchins, Alvin S. Hopping, and Frederick A. Saunders, "Subharmonics and Plate Tap Tones in Violin Acoustics", J. Acoust. Soc. Am. 32 (1960) 1443–1449.
26. J. Igarashi and Masaru Koyasu, "Acoustical Properties of Trumpets", J. Acoust. Soc. Am. 25 (1953) 122–128.
27. International Organization for Standardization Recommendation R 16, "Standard Tuning Frequency (Standard Musical Pitch)", 1955.
28. Y. Kahasi and Y. Tokita, "On the Acoustical and Vibrational Studies of Violin (III)", J. Acoust. Soc. Japan 8 (1952) 15–17.
29. Y. Kahasi and Y. Tokita, "On Acoustical and Vibrational Studies on Violas and Violoncellos", J. Acoust. Soc. Japan 11 (1955) 19–23.
30. K. C. Kar, "Note on the Minimum Bowing Pressure", Ind. J. Theoretical Phys. 2 (1954) 47.
31. L. Keibs und W. Kuhl, "Zur Akustik der Thomas-kirche in Leipzig", Acustica 9 (1959) 365–370.
32. A. Keller, "Untersuchung eines Resonanzkastens auf Grund seiner Schwingungsformen", Acustica 10 (1960) 372–380.
33. Joseph B. Keller, "Bowing of Violin Strings", Communications on Pure and Applied Mathematics 6 (1953) 483–495.
34. Roger E. Kirk, "Tuning Preferences for Piano Unison Groups", J. Acoust. Soc. Am. 31 (1959) 1644–1648.
35. Hermann Kurz, "Die Nichtlinearen Verzerrungen von Geigenresonanzböden", Acustica 3 (1953) 148–152.
36. Th. Lange, "Die Eigenfrequenzen von Trichtern", Acustica 5 (1955) 323–330.
37. Hugh Le Caine, "Electronic Music", Proc. IRE 44 (1956) 457–478.
38. Emile Leipp, "Le degré hygrométrique de l'air ambiant: paramètre sensible du spectre d'une corde harmonique en boyau", Compt. rend. 249 (1959) 2014–2016.
39. Emile Leipp, "Les paramètres sensibles des instruments à cordes", doctoral thesis, Faculty of Sciences, University of Paris, 20 May 1960.
40. W. Linhardt, "Über die akustische Kopplung der Pfeifen bei der Orgel", Acustica 11 (1961) 297–301.
41. W. Lottermoser, "Acoustical Design of Modern German Organs", J. Acoust. Soc. Am. 29 (1957) 682–689.
42. W. Lottermoser, "Organ Building on an Acoustical Basis", Gravesaner Blätter Heft 11/12 (1958) 147–157.
43. W. Lottermoser, "Das Ausgleichsverhalten von Geigen und seine Beziehungen zu der Resonanzkurve", Acustica 8 (1958) 91–98.
44. W. Lottermoser und W. Linhardt, "Beitrag zur akustischen Prüfung von Geigen und Bratschen", Acustica 7 (1957) 281–288.
45. W. Lottermoser und J. Meyer, "Über die Möglichkeiten einer Dendrochronologie von altitalienischen Geigen", Instrumentenbau Zeit. 12 (1958) 295–297.

46. W. Lottermoser und J. Meyer, "Resonanzen von Geigendecken und -Böden", Instrumentenbau Zeit. 13 (1959) 185–189.
47. W. Lottermoser und J. Meyer, "Frequenzmessungen an gesungenen Akkorden", Acustica 10 (1960) 181–184.
48. W. Lottermoser und J. Meyer, "Steuerbare elektrische Tastentraktur für Orgeln", Acustica 11 (1961) 117–118.
49. W. Lottermoser und J. Meyer, "Über den Einfluss des Materials auf die klanglichen Eigenschaften von Orgelpfeifen", Metall 2 (1962) 108–111.
50. W. Lottermoser und H. J. von Braunmühl, "Beitrag zur Stimmtonfrage", Acustica 5 (1955) 92–97.
51. D. W. Martin and W. D. Ward, "Subjective Evaluation of Musical Scale Temperament in Pianos", J. Acoust. Soc. Am. 33 (1961) 582–585.
52. M. V. Mathews, "An Acoustic Compiler for Music and Psychological Stimuli", Bell Sys. Tech. J. 40 (1961) 677–694.
53. M. V. Mathews, J. R. Pierce and N. Guttman, "Musical Sounds from Digital Computers" Gravesaner Blätter Heft 23/24 (1962) 119–125.
54. Hermann Meinel, "Über den Normstimmton", Wissenschaftliche Annalen 5 (1956) 2–15.
55. H. Meinel, "Regarding the Sound Quality of Violins and a Scientific Basis for Violin Construction", J. Acoust. Soc. Am. 29 (1957) 817–822.
56. H. Meinel, "Musikinstrumentenstimmungen und Tonsysteme", Acustica 7 (1957) 185–190.
57. D. M. A. Mercer, "The Effect of Voicing Adjustments on the Tone Quality of Organ Flue Pipes", Acustica 4 (1954) 237–239.
58. J. Meyer, "Resonanzeigenschaften offener Labialpfeifen", Acustica 11 (1961) 385–396.
59. Jürgen Meyer, "Über die Messung der Frequenzskalen von Holzblasinstrumenten", Das Musikinstrument Heft 10 (1961) 614.
60. J. Meyer and W. Lottermoser, "Über die Möglichkeiten einer klanglichen Beurteilung von Flügeln", Acustica 11 (1961) 291–297.
61. W. Meyer-Eppler "Elecktronische Music", *Klangstruktur der Musik* 1955, pp. 133–158.
62. A. Moles, "Informationstheorie der Musik", Nachr. Technik Fachberichte 3 (1956) 47–55.
63. Harry F. Olson and Herbert Belar, "Aid to Music Composition Employing a Random Probability System", J. Acoust. Soc. Am. 33 (1961) 1163–1170.
64. G. Pasqualini, "Récents résultats obtenus dans l'étude électroacoustique de la caisse harmonique des instruments à archet", Acustica 4 (1954) 244–249.
65. Jacques Poullin, "Musique Concrète", *Klangstruktur der Musik*, 1955, pp. 109–132.
66. A. Rakowski and E. G. Richardson, "A Spectral Analysis of the Voicing Process", Gravesaner Blätter Heft 15/16 (1960) 55–58.
67. E. G. Richardson, "Electro-acoustics Applied to Musical Instruments", Acustica 4 (1954) 212–217.
68. E. G. Richardson, "The Transient Tones of Wind Instruments", J. Acoust. Soc. Am. 26 (1954) 960–962.
69. Albert R. Rienstra, "Acoustical and Organ Design for Church Auditoriums", J. Acoust. Soc. Am. 29 (1957) 783–788.
70. S. N. Rzhevkin, "Certain Results of the Analysis of a Singer's Voice", Soviet Physics—Acoustics (Ak. Zh.) 2 (1956) 215–220.
71. G. G. Sacerdote, "Researches on the Singing Voice", Acustica 7 (1957) 61–68.
72. Oskar Sala, "Subharmonische elektrische Klangsynthesen", *Klangstruktur der Musik*, 1955, pp. 89–108.
73. F. A. Saunders, "Recent Work on Violins", J. Acoust. Soc. Am. 25 (1953) 491–498.

74. Theodore E. Simonton, "New Integral Ratio Chromatic Scale", J. Acoust. Soc. Am. 25 (1953) 1167–1175.
75. A. L. Leigh Silver, "Equal Beating Chromatic Scale", J. Acoust. Soc. Am. 29 (1957) 476–481.
76. E. Skudrzyk, "Betrachtungen zum Musikalischen Zusammenklang", Acustica 4 (1954) 249–253.
77. A. M. Small and D. W. Martin, "Musical Acoustics: Aims, Problems, Progress and Forecast", Proc. Second Int. Cong. Acoustics, American Institute of Physics, 1957, pp. 68–75.
78. Donald W. Stauffer, "Intonation Deficiencies of Wind Instruments in Ensemble", master's thesis, The Catholic University of American Press, Washington, D. C., 1954.
79. L. Suominen, "Untersuchungen über die Schwingungseigenschaften des Geigenkörpers", Acoustica 8 (1958) 363–370.
80. V. A. Ussachevsky, "Processes of Experimental Music", J. Audio Engr. Soc. 6 (1958) 202–207.
81. Martin Vogel, *Die Intonation der Blechbläser*, Gesellschaft zur Förderung der systematischen Musikwissenschaft, Düsseldorf, 1961.
82. W. D. Ward and D. W. Martin, "Psychophysical Comparison of Just Tuning and Equal Temperament in Sequences of Individual Tones", J. Acoust. Soc. Am. 33 (1961) 586–588.
83. Fritz Winckel, "Die Trübung als Wesenselement der Kunst", *Klangstruktur der Musik*, Verlag für Radio-Foto-Kinotechnik, 1955, p. 36.
84. Fritz Winckel, "Die Bedeutung des Vibratos in der Musik", Gravesaner Blätter Heft VI (1956) 40–47.
85. Iannis Xenakis, "Stochastic Music", Gravesaner Blätter Heft 23/24 (1962) 169–184.
86. F. J. Young, "The Natural Frequencies of Musical Horns", Acustica 10 (1960) 91–97.
87. R. W. Young, "Inharmonicity of Piano Strings", Acustica 4 (1954) 259–262.
88. Robert W. Young, "Twenty-Five Years of Musical Acoustics", J. Acoust. Soc. Am. 26 (1954) 955–959.
89. Robert W. Young, "Optimum Lengths of Valve Tubes for Brass Wind Instruments", J. Acoust. Soc. Am. 34 (1962) 716.
90. R. W. Young and J. C. Webster, "The Tuning of Musical Instruments (III) The Clarinet", Gravesaner Blätter Heft 11/12 (1958) 182–186.