REPORTS
of
THE 6TH INTERNATIONAL CONGRESS ON ACOUSTICS

Editor in Chief
Dr. Y. KOHASI
ACOUSTICAL SOCIETY OF JAPAN

I

GP
OPENING LECTURE and GENERAL PROGRESS PAPERS

A

PHYSIOLOGICAL and PSYCHOLOGICAL ACOUSTICS

TOKYO, AUGUST 21-28, 1968
Published for the International Council of Scientific Unions with the Financial Assistance of UNESCO.

UNESCO Subvention — 1966 — DC / 2.1 / 414 / 26

The texts of the reports contained in this book remain the property of their authors and may not be reproduced without their permission.

Printed in Japan
FOREWORD

These Reports were submitted and were read at the 6th International Congress on Acoustics held in Tokyo, Japan, 21-28 August 1968.

The International Congress on Acoustics are sponsored by the International Union of Pure and Applied Physics of UNESCO, and organized by the Acoustical Society of Japan under the auspices of the International Commission on Acoustics of IUPAP.

The technical program of the Congress Meeting comprised about 450 invited and contributed papers, delivered in 13 sessions throughout the week. Although every contributed paper was limited to four printed pages, the large number of lectures made it necessary to split the works into following six volumes.

Volume I  GP  General Progress Papers
           A  Physiological and Psychological Acoustics

Volume II  B  and  C  Speech

Volume III D  Electroacoustics
            E  Architectural Acoustics

Volume IV  F  Noise including Aeroacoustics
            G  Mechanical Vibration

Volume V   H  Physical Acoustics
            J  Molecular Acoustics

Volume VI  K  Ultrasonics
            L  Underwater Sound
            M  Bioacoustics
            N  Musical Acoustics

August, 1968

J. SANEYOSHI

Chairman of Organizing Committee
INTERNATIONAL COMMISSION ON ACOUSTICS

President
I. Malecki               Poland

Secretary
F. Ingerslev              Denmark

Members
M. L. Brekhoskikh        U. S. S. R.       D. Sette                Italy
J. Frenkel               Belgium          J. Saneyoshi            Japan
B. Lindsay               U. S. A.         E. Zwicker              Germany

Corresponding Members
T. Tarnoczy              Hungary        F. Kolmer                Czechoslovakia
D. W. Robinson            England         R. Lehmann              France

THE 6TH INTERNATIONAL CONGRESS ON ACOUSTICS

Chairman of Organizing Committee
J. Saneyoshi           President of the Acoustical Society of Japan

Chairman of Executive Committee
K. Awaya               Ex-Vice President of the Acoustical Society of Japan

General Secretary
J. Igarashi           Vice President of the Acoustical Society of Japan

Editorial Committee
Y. Kohashi             Chairman
Y. Wada                Vice Chairman
Y. Arai     M. Nagata
R. Araki           K. Negishi
I. Ebihara       O. Nomoto
H. Fujisaki       M. Okujima
H. Fujita            M. Onoe
R. Hattori        H. Sakai
J. Igarashi      T. Suzuki
Y. Kato              R. Teranishi
M. Koyasu           Y. Tokita
E. Mori              T. Yasuda
## OPENING LECTURE

Acoustics and acoustical industries in Japan  
*Jun'ichi Saneyoshi*  

## GENERAL PROGRESS PAPERS

| GP-2-1 | Fluid induced vibration  
*R. E. D. Bishop* | 3 |
| GP-2-2 | Der Impuls-Schaltest und seine raumakustische Beurteilung  
*W. Reichardt* | 11 |
| GP-2-3 | Current topics of speech research  
*G. Fant* | 21 |
| GP-3-2 | Ultrasonic methods of investigating dislocations and point defects of crystalline solid bodies  
*L. G. Merkulov* | 23 |
| GP-3-3 | Acoustical measurement of chemical relaxation in electrolyte solutions  
*Konrad Tamm* | 25 |
| GP-4-1 | Objective measurements in room acoustics  
*Heinrich Kuttruff* | 39 |
GP-4-2 Radiation induced nucleation and ultrasonic cavitation
_D. Sette_ .................. GP-51

GP-4-3 Psychophysics and the measurement of loudness
_S. S. Stevens_ ............... 63

GP-6-1 Computers in acoustics: Symbiosis of an old science and a new tool
_M. R. Schroeder_ ............ 71

GP-6-2 The perception of timbre
_J. F. Schouten_ .............. 89
Opening Lecture

ACOUSTICS AND ACOUSTICAL INDUSTRIES IN JAPAN

Jun'ichi Saneyoshi.
President of the Acoustical Society of Japan

It is my honour and pleasure to have the opening lecture. Most of you, the participants in this Congress from abroad, must be visiting Japan for the first time, and must be interested not only in the Congress itself but also in many things about this country. It may be supposed that foreign participants have a little knowledge about acoustics in Japan, because a few literature has been published in foreign languages. From these viewpoints the contents of my opening lecture are selected.

I. Outline of the Acoustical Society of Japan

The Acoustical Society of Japan that has organized the sixth I. C. A. is introduced at first. The Society was established 32 years ago in 1936.

1. Member  At present the number of regular members is approximately 1300.
Regular members consist of specialists in many fields, i.e. electrical engineers, architects, physicists, chemists, physiologists, physicians, psychologists, and others. The Society has 88 sustaining members, too.

2. Journal  The journal of the Society is issued bimonthly, and each issue has three original papers in average. Each paper is written in Japanese, however, it contains a fairly long English abstract of 600 words, and title and explanation of
figures and illustrations in the paper are written in English, in order to make the papers understandable for foreign readers to some extent. Some of letters to editor is entirely written in English. To each author of the best three papers Satō Prize is awarded every year.

3. Spring and Autumn Meeting. The Society holds meetings twice a year, in spring and autumn. In each meeting about 150 papers are reported and discussed. Preprints of the papers are prepared with a photographic printing process. The papers are written in only Japanese, but English translation of the title is written in foot note of each paper.

4. Research Committees. The Society sustains seven research committees, i.e. noise control, architectural acoustics, speech, evaluation of sound quality, electro-acoustics, ultrasonics and chemical acoustics. Each committee holds symposium several times a year. All members of the Society have the right of participating in the symposium.

5. Committee for Investigation of Acoustical Standards. This committee, having many subcommittees, cooperates with International Electrotechnical Commission (IEC) and International Standard Organization (ISO). The committee also prepares draft of some Japanese Industrial Standards (JIS) answering request of the Ministry of Trade and Commerce.

6. Local Chapters. The Society has two local chapters, Kansai and Tohoku chapter. The former is in Kyoto-Osaka district and the latter is in Sendai city. The Sendai Symposium on Acoustoelectronics and the Speech Symposium Kyoto have been taken care of by members of each local chapter, respectively.

7. Short History of the Society. On 15th April of 1936 fifteen members assembled and established the Acoustical Society of Japan, and elected Dr. Tashimoto, a seismologist, to the President. These 15 members were specialists of many fields, physics, physiology, psychology, architecture and mechanical engineering, however, they commonly loved good sound and hated bad sound or noise. Accordingly the purpose them was to realize good sound by mutual communication and cooperative research among them breaking barriers between different kinds of special field.
This small society began to hold meeting once a month. In autumn of the same year the Society planned a large scale meeting to which members and non-members were invited to give a short lecture for report of their research. This meeting was successful, and number of members began to increase rapidly.

Meanwhile a committee for standardization of acoustical terminology started, and it finished its work in three years. The first number of the journal was issued in 1938, but the publication was interrupted from 1944 to 1950 because of poverty after the War.

In the confusion of defeat the Society could only hold small meeting often. The biannual meeting for report of research was revived in 1946.

Dr. K. Sato, since he was elected to the president, made much effort to reconstruct the Society. He reopened publication of the journal, and inaugurated the system of research committees. Dr. Sato had special interest in noise problem, so he let noise research committee to standardize noise meter and to measure noise at many places in city. Besides, he recommended and helped Tokyo Municipal Government to establish Noise Prevention Act.

During Dr. Sato's presidency of six years the foundation and system of the Society were settled down. Since that time the Society has been developing favourably.

II. Modernization and Industrialization of Japan

It would be better to review briefly the modernization and industrialization of this country as the background of acoustics and acoustical industry in Japan.

1. Meiji Restoration This year, 1968, is a commemorative year for us, because our country abolished isolation policy and began to modernize herself just one hundred years ago, in 1868. Before that time Emperor was living in Kyoto, and he had no connection with politics. On the other hand, a government belonging to Tokugawa family existed in Tokyo, and this Tokugawa Government was controlling many feudal lords distributed in whole country.
The Tokugawa Government had been continuing isolation policy for two and a half centuries, it had had no diplomatic relation with foreign countries, and had forbidden people foreign trading, travelling and communication. The Government had been becoming weak in the middle of 19th century, restored voluntarily its sovereign power to the Meiji Emperor in 1868. However some civil war occurred, this restoration finished almost in peace. We call it Meiji Restoration, not Meiji Revolution.


Young and excellent Meiji Emperor organized a new government in Tokyo assisted by many young and ambitious statesmen. The new government adopted radically new policy of modernization. Because they had been astonished at the progress of European civilization and they were afraid of colonization by foreign countries.

Strong military power was naturally a goal of the policy. But it was not so urgent, because Japan was very much far distant from powerful countries. Accordingly, the policy was, instead of importing a large quantity of weapons, to strengthen industrial and economical power as the base of military power. Along this policy the Meiji Government decided to learn frankly European civilization from European countries and United States, and to reform all systems of politics, education, military, social service and so on. Some of people was confused, but most part of nation followed the policies of the new government.

Public services such as telegraph, post, railroad were opened and operated by the government several years after the restoration. Telephone service was started also by the government in 1890. The network of these public services extended gradually all over the country.

Heavy industries such as mining, iron manufacture, ship building, military arsenal were also established by the government. Some of them were transferred to private enterprises later. Light industries were brought up generally as private enterprises. Among them silk industry earned much foreign money exporting special product of this country. Cotton industry played similar role following silk industry.

The government was also very eager to popularize education. It encouraged
cities, towns and villages to establish many primary schools. Higher grade schools of many kinds were established and managed by the government. The first university was established in Tokyo by the government in 1883.

At the beginning of above mentioned each public services, industries and schools the government invited from foreign countries eminent experts and professors paying very high salary. Japanese people learned eagerly from those teachers miscellaneous techniques and mastered them in relatively short period.

3. Constitution and People

The constitution of Japan was proclaimed in 1890, 23 years after the Meiji restoration, and Diet (House of Representatives and House of Peers) was opened in the same year. Thereby this country became constitutional monarchy, but, a tendency of strong government were continuing until 1944, the end of the second World War. Meanwhile, democratic atmosphere grew up among some of educated people, but it was suppressed by the pressure of military authority and government until completely smashed in nineteen thirties.

Most of the people had been working very diligently under the slogan of "Rich Country and Strong Military Force", since Meiji restoration till 1944. It may be said that the rapid growth of industries were supported by diligence of people, diffusion of education, encouragement and leadership of the government and import of technology and science.

On the other hand, Japanese traditional arts and handicraft were disregarded by the government, and they were preserved among the people. It may roughly be said that the policy of modernization was nearly equal to Europeanization.

After the 2nd World War, a radically new and democratic constitution was prepared on agreement between occupying country and Japan, and proclaimed in 1946. During the period of poverty and hunger after defeat occupying forces recommended or ordered many reformation of political, financial, social and educational systems along the direction to democracy. The reason why our country accepted those reformation with a few resistance might be that most of Japanese people had been hating oppressive military authority and bureaucracy of their own. Democracy in this country, however it was rather enforced by foreign forces in this way, has been settling down in the mind of people gradually.
4. Reconstruction and Development after the War. It is very difficult to describe and analyze the rapid expansion and development of industries after the War. But my personal view is as follows. Nearly all cities were burnt and all public service facilities were heavily damaged. Nearly 90 billion people had to live in narrow territory of four islands. The reconstruction of burnt or destroyed properties, themselves required industries a huge amount of production of many kinds. To respond the extraordinary huge demand survived industries had to extend or establish factories with new machines and equipments. This generated indirect demands successively through many industries of near all kinds. Supported by these amplified demands and abundant labour of relatively well educated and diligent engineers and labourers industries continued rapid expansion.

In the early years of reconstruction we were troubled by shortage of foods and capital. Fortunately we could have aid of foods from occupying country, and people saved money well even in low standard of living. This money saving contributed to new investment of industries. At the investment managers of some industries imported the newest technology and constructed efficient up-to-date factories. Export of products from those factories made import of raw materials possible. Another reason of rapid industrial development might be that we spent a very little man power and money for armament.

5. Problems. Some of foreign people may consider our industrial and economical development as wonderful or mysterious. However, we have many problems and difficulties, such as small amount of natural resources and narrow area of lend. Japan imports high percentage of raw materials and fuel, and considerable amount of foods. In order to make these import possible Japan must export enough much quantity of industrial products.

Rising standard of living is requiring much more consumption of industrial products and energy. This fact needs much more import of raw materials and fuel, then more and more increase of export become neccessary. In order to make such a extension of export possible, it is neccessary to raise level of technology and science and to improve constitution of enterprises.

Since Meiji Restoration Japanese people learned many things of all sort from
European civilization and culture. This frankly learning attitude contributed to rapid development during 100 years. But on the other side we have something like a dislocation in traditional culture and in the way of thinking. The way of thinking has changed remarkably, however, it will not become the same as that of European people who have originated European civilization. I think these are the origin of confusions in this country at present.

But I feel that some new culture of ours is already growing among the confusions, and I wish Japan will contribute with her own new culture to whole world much more in future.

III. Acoustics and Acoustical Industries

In the following part of this lecture acoustical products in Japan and research works related to them will be described. Because I was originally an electrical engineer, and it is difficult to pick up eminent purely scientific works in acoustics fairly.

1. Musical Instruments. The oldest acoustical industry in this country is production of musical instruments. Typical and common instruments of Japanese traditional music are "shamisen", "koto" and "shakuhachi". Shamisen is a kind of string instrument with three strings which are not rubbed with bow but snapped with a large plectrum. Koto is a kind of harp with 13 strings stretched on a wooden hollow body. Shakuhachi is a wind instrument made of bamboo with seven holes.

These instruments have been produced by handicraft until now. Only a few scientific research has been made on the traditional instruments.

In school education the Meiji Government forsook traditional music and adopted European music. In all of many elementary schools singing was taught with the aid of reed organ or pianoforte. Responding this demand reed organ has been produced since late 19th century and the production of piano was begun in nineteen twenties. At present a large number of pianos is produced in large scale factories in Shizu-
oka Prefecture, and a large number of pianos is exported.

Other instruments of all kinds used in orchestra, brass band and jazz band are produced except pipe organ. Recently electric guitars, electric organs and pianos are produced and sold.

2. **Telephone Transmitter and Receiver** The oldest electroacoustical products were telephone transmitter and receiver. Production of them had been begun soon after the establishment of telephone service. But they were manufactured faithfully to drawings, and tested by only hearing test method until 1933. Quantitative expression concerning the performance of receiver in the specification was only the weight of it. For too light receiver could not operate hook switch of a telephone set.

In order to improve telephone set, a research group led by Dr. T. Hayasaka in Electric Communication Laboratory (belongs to Telegraph and Telephone Public Corporation) began exhaustive investigations and fundamental researches in 1944, just after the War. As the results Type 4 telephone set that has nearly flat response from 300 to 2,200 Hz was designed. Engineers belonging to telephone set makers participated in the trial manufacturing, and this type has been regularly produced by five makers and practically used since 1953. Telephone articulation in this country was remarkably improved by this success.

This research group developed more improved Type 600 telephone set in 1960, in which flat response frequency range was extended up to 3,400 Hz. The number of telephone subscribers in Japan has been increasing from nearly 2 million at 1953 to nearly 10 million at present. This means that nearly all of telephone sets used in this country are thus improved Type 4 or Type 600.

In parallel to the design of Type 4 telephone set Dr. Hayasaka and his colleagues finished the development of a new standard reference system for the purpose of calibration, or test of sensitivity, of telephone transmitter, receiver and microphones. This new system, developed in Japan and U.S.A. independently, is based on reciprocity principle, and it does not need any primary standard transducer. So Dr. Hayasaka named it as "mutual calibration system".

At present, the primary standard mutual calibration apparatus authorized by
the Government is maintained at the Electrotechnical Laboratory that belongs to the Ministry of Trade and Commerce. Similar apparatuses are possessed and maintained by telephone set makers and other laboratories as the secondary standard mutual calibration apparatuses. These secondary standard apparatuses must be compared with the primary standard apparatus periodically. Noisemeters are calibrated with the secondary standard apparatus, too.

3. **Loudspeaker.** Radio broadcasting was opened in 1925. Most part of radio receivers manufactured since that time until 1946 were equipped with so called "magnetic cone speaker". This type loudspeaker had a paper cone driven at the center by a balanced armature type unit through a short metal rod. It had poor response in lower frequency range, but was cheap and simple. Moving coil or dynamic speakers were produced for high quality radio receivers, phonographs, public address systems and cinema theaters.

Since 1947 dynamic type speakers have been manufactured mainly. As shown in Table 1, annual production of loudspeakers reaches 88 million units, and these are attached to TV receivers, radio receivers, tape recorders, phonographs and public address systems, or exported as parts.

### Table 1. Annual Production of Audio Apparatuses and Parts in 1967.

<table>
<thead>
<tr>
<th>Item</th>
<th>Amount of money</th>
<th>Number of articles</th>
<th>Percentage of export (money)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$10^6$ Yen</td>
<td>$10^6$ $</td>
<td>$10^6$</td>
</tr>
<tr>
<td>T. V. receiving set</td>
<td>311,766</td>
<td>876</td>
<td>7.03</td>
</tr>
<tr>
<td>Radio receiving set</td>
<td>117,295</td>
<td>325</td>
<td>?</td>
</tr>
<tr>
<td>Tape recorder</td>
<td>82,113</td>
<td>228</td>
<td>7.19</td>
</tr>
<tr>
<td>Stereo-phonograph</td>
<td>61,547</td>
<td>171</td>
<td>2.05</td>
</tr>
<tr>
<td>Loudspeaker</td>
<td>15,406</td>
<td>42.7</td>
<td>88.58</td>
</tr>
<tr>
<td>Microphone</td>
<td>3,394</td>
<td>9.4</td>
<td>8.07</td>
</tr>
<tr>
<td>Record player unit</td>
<td>2,794</td>
<td>7.8</td>
<td>.38</td>
</tr>
<tr>
<td>Phonograph pick up</td>
<td>1,682</td>
<td>4.7</td>
<td>7.01</td>
</tr>
<tr>
<td>Earphone</td>
<td>584</td>
<td>1.6</td>
<td>13.51</td>
</tr>
</tbody>
</table>
4. Microphone  In the early years of radio broadcasting imported double button type and Reisaz microphones were used. Mr. Marumo, who belonged to Japan Broadcasting Corporation (NHK, Nippon Hosou Kyokai), succeeded in heat treatment of noiseless carbon powder, and the Corporation made Reisaz type microphones and used them for more than ten years. Meanwhile velocity type microphones were produced and they expelled unstable carbon microphones.

After 1947, microphones of many types, moving coil, condenser and piezo type microphones were imported and manufactured. Together with the development of other usages of microphone, such as tape recorder, hearing aid and handy public address system, production of small and light weight microphones has been increased.

5. Tape Recorder. At Electrical Communication Laboratory of Tohoku University Prof. Nagai and Mr. Igarashi had been studying magnetic recording with alloyed metal of several kinds since 1935. They were aiming at noise suppression by perfect demagnetization at erasing using alternating magnetic field of 50Hz. This method was successful in noise suppression. But they tried incidentally to superpose H.F. current on voice coil of a recording head, and they were surprised by unexpected fine results; erasing without special head or coil, good linearity and low noise.

Mr. Shuka applied this method of H.F. biasing on paper tape coated with iron oxide in 1946, and he got great success in commercialization of useful and handy tape recorders. This was the first step of the development of Sony Corporation.

6. Hydrophone  Japanese Navy and Prof. Nukiyama (Tohoku University) began investigation and fundamental research of hydrophone and passive sonar system in 1926. But the navy authority impatiently decided to adopt German system, and the first equipment was manufactured in this country in 1933. Since then all submarines of the navy were equipped with this hydrophone system until 1943.

7. Active Sonar  Active sonars, that detect a submarine by catching ultrasonic echo, were manufactured since 1933, and destroyers and submarine chasers of the navy were equipped with them gradually. The transducer of the sonar was the Langevintype that consisted of thick steel discs and many quartz plates in early
period, but it was substituted later by magnetostriction transducers invented and developed by Prof. Nukiyama, Aoyagi, Kikuchi and Fukushima at Tohoku University.

8. **Fathometer** Ultrasonic fathometers with magnetostriction transducer and recorder were manufactured since 1937, and they were installed on merchant ships and warships. For the purpose of deep sea sounding Dr. Hashimoto invented a new multiple pulse emission system, in which pulses are emitted with the time interval corresponding to 1000 meters. Using a fathometer of this system Hydrographic Bureau measured a depth of 10,400 meters in the Pacific Ocean in 1940. Complying request of Hydrographic Bureau a shallow water precision fathometer was developed by me in 1951, in which frequency was raised to 70 kHz and the recorder was driven by a synchronous motor and tuning fork oscillator. The error was so small, within ±0.1 m, that this type of fathometer has been used for sounding of harbour and sea shore.

9. **Fish Detector** After the war some of sonar engineers who lost job were very earnest to develop a fish detector. They made experiments on board together with progressive fishermen using small type fathometer. At last fishermen recognized the effectiveness of it, and ultrasonic fish detector began to diffuse whose through the country since 1949. Many makers designed and sold fathometer of many types each fitted miscellaneous requirements of fisheries of many kinds.

In early period main object was dense shoal of sardine, but later, a single body of large fish, such as tuna fish or salmon, could be recorded by new types with higher frequency (over 200 kHz) and with shortened interval of pulses Japanese people eat fishes very much, and many Japanese fishery boats equipped with fish detector go out far away even into the Atlantic Ocean.

10. **Flow Detector** Flow detector has been actually used since 1950, and the usage becomes very broad. For users of many kinds a handbook has been published.

11. **Ultrasonic Cleaning Apparatus.** Watch makers used it at first for the cleaning of parts of wrist watches. At present, very many manufacturers of several decades are producing cleaning apparatuses of many kinds that meet requirements of
# ACOUSTICS AND ACOUSTICAL INDUSTRIES IN JAPAN

## Table 2. Annual Production of Ultrasonical Apparatuses in 1967.

<table>
<thead>
<tr>
<th>Item</th>
<th>$10^6$ Yen</th>
<th>$10^5$ $\text{$}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fish detector</td>
<td>2,426</td>
<td>6.75</td>
</tr>
<tr>
<td>Cleaning apparatus</td>
<td>678</td>
<td>1.88</td>
</tr>
<tr>
<td>Flaw detector</td>
<td>407</td>
<td>1.07</td>
</tr>
<tr>
<td>Fathometer</td>
<td>83</td>
<td>0.23</td>
</tr>
<tr>
<td>Miscellaneous</td>
<td>2,151</td>
<td>5.97</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>5,747</strong></td>
<td><strong>15.92</strong></td>
</tr>
</tbody>
</table>

miscellaneous usages.

As the transducers of them $\text{BaTiO}_3$, nickel and ferrite are used, but the percentage of ferrite transducer, developed by Prof. Kikuchi, is increasing. The Association of Electronics Industries established a general standard specification of Ultrasonic cleaners in 1967. In which the high frequency electric power flowing into the transducer at a normal working condition is adopted as the rating of a cleaning apparatus. As a special high frequency powermeter utilizing Hall's effect has been developed already.

12. **Medical Ultrasonics** Ultrasonic therapy in which megacycle ultrasonic waves are irradiated through skin surface into tissue does not commonly used in this country. Diagnosis by ultrasonic echo is investigated and actually used by many physicians in many fields of medical science. Operation by focussed ultrasonic waves is also studied. Japanese Society of Medical Ultrasonics was established in 1965. The Society holds biannual meeting, and the papers read at each meeting are published as proceedings in English.

13. **Sound Absorbing Materials for Architecture.** Dr. K. Sato constructed a large reverberation chamber in Kobayashi Research Institute of Physics in 1956. Since then absorption coefficient of newly developed materials by makers have been measured in this reverberation chamber and the date is indicated in the catalogue of each material.
14. **Design and Construction of Auditorium** The first acoustically designed auditorium was Ohkuma Memorial Hall of Waseda University. The designer Dr. T. Sato is one of the initial members of the Acoustical Society. After the War many auditoriums have been constructed, for example, public halls in many cities and large theaters. The acoustic characteristics of those auditoriums are fairly good in general, because it became a custom that acoustical design and testing must be entrusted to specialists.

At the end of my lecture I would like to make some excuse. I have been obliged to leave from acoustical works since just two years ago when I was appointed to the President of Tokyo Institute of Technology. After that I was very busy and exhausted until I became ill in this May and resigned the President of T. I. T. Fortunately I could restore my health before the Congress, but I had to write the draft of this lecture in a hospital without referring many literatures and materials. These are the reason why this lecture is not forward-looking but rather historical one that missing latest information. I would be very happy if my lecture could help foreign participants to understand our country.
There has been an enormous advancement of scientific research and technical development for the last 50 years. Over that period the number of research workers engaged in R and D works is considered to have increased as much as 40 times and it amounts now to about 6 million. The world R and D annual expenditures totalled to approximately 40 billion of US Dollars per year/apart from those allocated on military research/.

Against that general background acoustics has had a rather modest though ever increasing position. Whereas prior to the first world war the number of persons dealing with acoustics had not exceeded 300, today it may be estimated that there are about 23,000 active acousticians throughout the world. That increase has been somewhat higher than the average rate of manpower growth as related to the scientific research and technological development as a whole. Taking into account the technical and administrative staff, the total number of people employed in the acoustic laboratories approaches 50 thousand.

Parallel to the quantitative growth there have occurred some profound changes in the structure and organization of acoustics.

The term "acoustics" denoting a group of scientific specializations had been used for the first time in the 17th century, but it was not earlier than in the middle of the 19th century when there followed an integration of acoustics as scientific discipline, and in particular, a close co-operation between the representatives of
physical and physiological acoustics.

Discoveries and inventions in the field of ultrasonics had brought new branches of acoustics into being in the inter war period and above all that of underwater and molecular acoustics.


At the same time the first acoustical societies, or acoustic divisions of the physical associations were established in France, Germany, England and US. In the USSR it was the Acoustical Commission attached to the Academy of Sciences that was responsible for pursuing researches.

The Second World War had brought about a set back of editorial and organisational activities, as well as a narrow advancement of some branches of acoustics and electroacoustics, closely related to the military purposes.

But already in the years 1947-1948 on the background of the post-war reconstruction there was a development of applied acoustics, and increased importance of the fundamental researches.

In 1950's two new acoustical / periodicals of international publicity were founded: "Acoustica"/1950/ and "Akusticheskij Zhurnal"/1955/. In the last decade several new periodicals of local importance have appeared, as for example, the Archives on Acoustics in Poland. An increased number of publications in the field of acoustics as shown at the Table 1 is an evidence of advancement of researches and rapid technical development.

It is quite natural that the importance of national and international acoustical organisations has been growing.
There are, at present, in Europe 17 acoustical societies with about 2,100 members. Two of them Groupement des Acousticiens de la Langue Francaise/GALF/ and the Scandinavian Acoustical Society have covered several countries through their activities.

The Japanese Acoustical Society includes more than one thousand members, and the corresponding American Society embracing US and Canada has 4,150 members. It is also worth mentioning here the Latin American Acoustical Group associating 63 members. An acoustical society is being organized in India.

In the Socialist Countries, the representative bodies of acoustics are either Scientific Councils or Committees affiliated to Academies of Sciences. They have been set up in Czechoslovakia, GDR, Hungary, Poland and Rumania, whereas in the USSR, on account of the large size of the Country, there have been in existence two scientific councils, one for acoustics and another one for ultrasonics, both attached to the Academy of Sciences.

It was just soon after the war when it appeared to be necessary to establish closer contacts among acousticians throughout the world. In 1951 the International Commission on Acoustics /ICA/, was founded as a special commission of the International Union of Pure and Applied Physics /IUPAP/. The first chairman of the ICA was elected Professor Richard Bolt, subsequently in the years:1957-66 that post was held by Professor W. Furrer, since 1951 the duties of the Secretary of the Commission have been successfully performed by Professor F. Ingeralev.

In the past period the activities of the ICA had been confined mainly to organizing world acoustical congresses which were held every three years. Each of them was an important step in the development of acoustics.

The number of participants and that of countries represented, as well as the number of submitted reports have been growing each congress as it is indicated on the Table 2. It must be emphasized that an increasing number of small conferences and symposia devot-

<table>
<thead>
<tr>
<th>Congress</th>
<th>I</th>
<th>II</th>
<th>III</th>
<th>IV</th>
<th>V</th>
</tr>
</thead>
<tbody>
<tr>
<td>Participants</td>
<td>319</td>
<td>760</td>
<td>1213</td>
<td>1328</td>
<td>1594</td>
</tr>
<tr>
<td>Countries</td>
<td>22</td>
<td>19</td>
<td>22</td>
<td>33</td>
<td>39</td>
</tr>
<tr>
<td>Reports</td>
<td>83</td>
<td>291</td>
<td>329</td>
<td>331</td>
<td>449</td>
</tr>
</tbody>
</table>
ed to some special acoustical issues have been arranged in recent years. In 1967 15 acoustical conferences were held in different parts of the world.

The acoustical problems have been also taking an increasing share at various congresses on non-acoustical topics, either purely scientific or technical ones. For instance, at the first Congress of Non-Destructive Testing in 1955 the reports on the ultrasonic control accounted for 15%, whereas at the Sixth Congress held in 1966 the corresponding share grew up to 32%.

The Standardization work is considered to be an important platform in the international scientific and technological co-operation. It has been pursued by the International Organization on Standardization /ISO/ as related to acoustics and mechanical vibrations; and by the International Electrical Commission /IEC/ in the field of electroacoustics.

The development of world acoustics is not only of quantitative or organizational nature. Important evolution has been going on in the very structure of scientific research. There have been some rapid advances in some fields of acoustics, whereas in some other branches rather slow progress may be observed. In view of a great number of works a statistical estimate may be applied here and the final classification of topics leads to some interesting conclusions.

In the Table 3 you can see the division of topics in the world bibliography in the period of the last forty years /1929-1963/.

<table>
<thead>
<tr>
<th>Table 3. Topics in the world bibliography share in percentages</th>
</tr>
</thead>
<tbody>
<tr>
<td>--------</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>3</td>
</tr>
<tr>
<td>4</td>
</tr>
<tr>
<td>5</td>
</tr>
<tr>
<td>6</td>
</tr>
<tr>
<td>7</td>
</tr>
<tr>
<td>8</td>
</tr>
<tr>
<td>9</td>
</tr>
<tr>
<td>10</td>
</tr>
<tr>
<td>11</td>
</tr>
</tbody>
</table>
The following classification has been introduced:

1. Architectural and building acoustics;
2. Physiological and psychological acoustics;
3. Electroacoustical equipment;
4. Musical acoustics and musical instruments;
5. Noise control;
6. Acoustics of speech;
7. Ultrasonics;
8. Mechanical vibrations and shocks;
9. Underwater acoustics;
10. Physical acoustics;
11. Standardization;
12. Miscellaneous and general topics.

In the last years there has been an increased interest in physical acoustics, and in the problems of mechanical vibrations; until 1960 the number of articles in physiological acoustics had been constantly growing. On the other hand the share of the musical and architectural acoustics has been decreasing.

Table 4 presents the distribution of topics in reports delivered at the Congresses of ICA. At the two first Congresses the distribution had been rather incidental, but at the last three Congresses the proportions between the major subjects were much like in the world literature. A more detailed analysis shows that starting from 1962 some new items have appeared within the topics of the ICA Congresses: the coupled acoustical and electrical fields, acoustics of aircraft noises and the cybernetic approach to the bioacoustics. More and more attention has been paid to the use of computers for the acoustical purposes.

Table 4. Topics at the ICA and other Congresses Share in percentages.

<table>
<thead>
<tr>
<th>Congress</th>
<th>ICA I</th>
<th>II</th>
<th>III</th>
<th>IV</th>
<th>V</th>
<th>ASA1967</th>
<th>USSK1968</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>10,8</td>
<td>8,5</td>
<td>19,2</td>
<td>11,8</td>
<td>14,0</td>
<td>2,9</td>
<td>8,7</td>
</tr>
<tr>
<td>2</td>
<td>10,8</td>
<td>13,5</td>
<td>12,7</td>
<td>14,5</td>
<td>10,7</td>
<td>19,2</td>
<td>5,3</td>
</tr>
<tr>
<td>3</td>
<td>36,1</td>
<td>11,0</td>
<td>16,8</td>
<td>12,1</td>
<td>13,1</td>
<td>12,9</td>
<td>9,1</td>
</tr>
<tr>
<td>4</td>
<td>18,1</td>
<td>7,8</td>
<td>-</td>
<td>1,8</td>
<td>2,9</td>
<td>-</td>
<td>3,6</td>
</tr>
<tr>
<td>5</td>
<td>-</td>
<td>13,8</td>
<td>8,3</td>
<td>13,0</td>
<td>9,6</td>
<td>8,3</td>
<td>16,0</td>
</tr>
<tr>
<td>6</td>
<td>-</td>
<td>7,4</td>
<td>5,6</td>
<td>10,3</td>
<td>8,0</td>
<td>11,7</td>
<td>7,4</td>
</tr>
<tr>
<td>7</td>
<td>14,1</td>
<td>6,7</td>
<td>8,6</td>
<td>7,8</td>
<td>10,5</td>
<td>4,1</td>
<td>25,4</td>
</tr>
<tr>
<td>8</td>
<td>-</td>
<td>5,3</td>
<td>6,7</td>
<td>10,5</td>
<td>16,8</td>
<td>12,1</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>-</td>
<td>3,6</td>
<td>1,2</td>
<td>1,8</td>
<td>2,0</td>
<td>14,1</td>
<td>1,6</td>
</tr>
<tr>
<td>10</td>
<td>9,7</td>
<td>27,7</td>
<td>22,3</td>
<td>20,2</td>
<td>18,7</td>
<td>6,7</td>
<td>8,0</td>
</tr>
</tbody>
</table>
A comparison of scientific activities of the two largest communities of acousticians in the US and USSR may be of interest.

The Table 4 gives a comparison of reports presented at the Meeting of Acoustical Society of America in November, 1967, and the Second All-Union Acoustical Conference held in Moscow in February 1968. The Table 5 presents a comparison between the topics of the Journal of Acoustical Society of America and those of Acoustitchesky Zhurnal for the period 1964-1966. Without embarking upon more detailed analysis it may be said that the American acousticians have shown more interest in physiological acoustics /3/, and mechanical vibrations /8/, as well as in underwater acoustics /9/, whereas the USSR acousticians have published a larger number of papers on physical acoustics /10/, ultrasonics /7/, and noise control /5/.

It is still worth while to notice the trends of technological progress. With that aim in mind the patents issued in the US in the years 1949-1958 and 1959-1963 have been compared in the Table 6. It is obvious that most patents concern the electro-acoustical equipment, but the share of innovations in noise control and mechanical vibrations has been constantly increasing whereas the number of patents in the field of architectural and musical acoustics has been decreasing. That fact does coincide with the above discussed

<table>
<thead>
<tr>
<th>Table 5. Topics of Jour. Acous.</th>
<th>JASA</th>
<th>A. Zh.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>6,4</td>
<td>5,2</td>
</tr>
<tr>
<td>2</td>
<td>18,8</td>
<td>2,3</td>
</tr>
<tr>
<td>3</td>
<td>4,8</td>
<td>4,1</td>
</tr>
<tr>
<td>4</td>
<td>2,3</td>
<td>1,2</td>
</tr>
<tr>
<td>5</td>
<td>0,8</td>
<td>4,7</td>
</tr>
<tr>
<td>6</td>
<td>6,0</td>
<td>1,2</td>
</tr>
<tr>
<td>7</td>
<td>10,1</td>
<td>23,5</td>
</tr>
<tr>
<td>8</td>
<td>12,3</td>
<td>7,8</td>
</tr>
<tr>
<td>9</td>
<td>19,6</td>
<td>5,2</td>
</tr>
<tr>
<td>10</td>
<td>16,9</td>
<td>40,3</td>
</tr>
<tr>
<td>11</td>
<td>2,0</td>
<td>4,7</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 6. Patents issued in USA</th>
<th>1949-1958</th>
<th>1959-1963</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>5,1</td>
<td>2,7</td>
</tr>
<tr>
<td>2</td>
<td>1,3</td>
<td>2,9</td>
</tr>
<tr>
<td>3</td>
<td>55,9</td>
<td>49,6</td>
</tr>
<tr>
<td>4</td>
<td>21,9</td>
<td>14,9</td>
</tr>
<tr>
<td>5</td>
<td>2,4</td>
<td>6,8</td>
</tr>
<tr>
<td>6</td>
<td>0,9</td>
<td>0,8</td>
</tr>
<tr>
<td>7</td>
<td>4,0</td>
<td>4,2</td>
</tr>
<tr>
<td>8</td>
<td>0,9</td>
<td>11,7</td>
</tr>
<tr>
<td>9</td>
<td>7,6</td>
<td>6,4</td>
</tr>
</tbody>
</table>
trends of development of acoustics.

On the basis of historical background and taking into account the present state of acoustics, we may anticipate its future development as a scientific discipline and branch of technology. When assuming an average rate of research and development we may presume that in fifteen years about 30,000 acousticians will have been engaged in the field of acoustics in the world as compared with the total figure of 120,000 of all those employed in the laboratories dealing with acoustics.

In the next years to come a number of new subjects of acoustics such as sonochemistry, bioacoustics and hyperson sound technique will certainly advance and find a broader practical application.

The parts of technical acoustics will become more and more specialized while being bound to some rather remoter fields as e.g. medicine, metallurgy of building engineering. Thus the question arises whether the integration process which had taken place a hundred years ago will not be opposed by the present process of specialization, and whether acoustics will preserve its unity. The matter in question has been several times discussed at the ICA meetings. The general opinion was, that within the framework of the national societies and on the international level it is rather useful to aim at integration of the three main sections: audioacoustics, ultrasonics and mechanical vibrations. The specialized branches of acoustical technology have got their common basis in the fundamental researches, and the delimitation of borders inside acoustics has proved to be artificial and not precise. It should also be remembered that the division of acoustics would mean practically that the important achievements of acoustic societies may be wasted and the ties between the acoustics as a scientific discipline and its numerous technical applications may be loosened.

Acoustics has been and will remain an inter-disciplinary science, and branch of technics. And that is where its weakness lines but the forces and prospects of development are inherent too. The Sixth Acoustical Congress at which the activities developed by the sections have been combined in one organic whole may indicate how important and how useful is the problem of integration of acoustics.

A very spectacular example of that sort of integration of acoustics and its growing significance is the activity developed by the Acoustical Society of Japan. Out of the lot of the Congress reports the papers delivered by the Japanese acousticians have accounted for 36%, and as to the most modern sections of acoustics that share has been even higher. It is an obvious evidence of the rank held by the Japanese acoustics in the world science and technology.

Therefore, at present, on behalf of the International Commission on Acoustics I would like to express my great acknowledgment and admiration for your activities, our dear friends, and at the same time I have the pleasure to express my best and
sincere thanks to the Society, and above all to the Organizing Committee headed by Professor Saneyoshi and to its secretary general Professor Igarashi for their taking up the heavy duties of organising the Congress and for their warm hospitality and friendship offered to all of us already in the first hours of our stay.
Introduction

It is not easy to write a general paper with any valid claim to being a critical survey of the literature on "Fluid Induced Vibration" - the title I have been given. The subject is simply too vast. What I propose to do, therefore, is limit my comments to "the deterministic vibration of solid (though not rigid) systems caused by the permanently attached flow of otherwise undisturbed fluids". In other words this paper relates to "classical flutter" and to transient oscillations associated with gust penetration. Needless to say these two problems have received most attention in aeronautics as special forms of aerelastic phenomena. But it is very probable that, in the next few years, we shall see a considerable volume of work in the closely related field of hydroelasticity.

The decision to limit the discussion in this way has not been reached without very considerable reflection. While, almost inevitably, a readable paper on this topic must be an introduction as well as a survey, it can be argued that just such an introduction is needed. Again, it can be argued that the subject I have chosen is hardly relevant to acoustics, being mainly concerned with oscillations of rather low frequency in aircraft and other light structures. While this is certainly so, it may be that acousticians would like an introduction to the field because there is no reason to suppose that it will never have much relevance to their work. Traditionally acousticians have taken a lively interest in mechanical vibration.

By its very nature, aerelasticity is a specialised field of study that demands considerable mathematical skills (as well as physical insight). It is not, perhaps, a subject that can be described as readily accessible to (say) professional mechanical or civil engineers or to acousticians or naval architects; yet it must inevitably be of increasing importance outside aeronautics. Those who wish to acquaint themselves with this field without having to make continual and specific reference to aeronautical practice will find useful introductory articles in the Handbook of Engineering Mechanics [1], [2]. It must be admitted, however, that aerelasticity is seldom if ever treated in any very general manner. Perhaps it is the urgency and vital importance of the subject that have forced writers to address themselves always to specific problems. This state of affairs means that I must find a method of introducing this field before attempting to survey it.
FLUID INDUCED VIBRATION

R.E.D. Bishop

General Considerations

It is natural, in this field, to regard a mechanical structure as a "system" upon which forces act, and by which motions are executed. That is the approach that has almost universally been adopted and it is the standpoint from which we can best proceed*. The configuration of the system may be represented by generalised coordinates which may be arranged in the form of a column matrix q. Each element of q represents a small departure from a steady state of motion in an equilibrium configuration. There may be at most six coordinates devoted to a specification of the velocity and orientation of the system in space and the remainder will normally represent distortions, which are either imposed (as at the rudder of a ship) or are caused involuntarily. We shall not be concerned here with distortion coordinates of the former type. The selection of coordinates for a particular analysis offers considerable scope for skill and experience.

It is necessary to describe the mass, internal damping and stiffness distributions of the system. These may be specified in terms of the chosen coordinates through matrices \( A, B \) and \( C \) respectively. Of course the selection of coordinates and the subsequent determination of the structural matrices \( A, B \) and \( C \) can represent considerable tasks in their own right (e.g. see ref.[5]).

Generalised forces are applied to this structure. Typically, we are concerned with small departures from the steady forces that are associated with steady flow in the absence of distortion. It will be convenient to divide these force perturbations into two general classes. In the first place there are forces \( \mathbf{f} \) which are applied to the system as a consequence of its departure from steady motion (i.e. of its oscillation) in the surrounding fluid. Secondly, there are forces \( F(t) \) that can be thought of as being imposed by some external agency, as for instance by some non-uniformity in the flow of the surrounding fluid. By means of Lagrange's equation we may now relate the behaviour of the structure to the forces and moments applied to it and arrive at the familiar matrix equation

\[
\mathbf{Aq} + \mathbf{Bq} + \mathbf{Cq} = \mathbf{F} + \mathbf{F}(t)
\]

This is to be thought of as a set of linear equations with constant coefficients.

Consider for a moment a conventional aircraft and suppose that some step disturbance in the flight condition (such as a sudden change of incidence) were to occur. This disturbance would cause a downwash velocity over some or all of a lifting surface and this downwash would not be compatible with the pressure distribution prevailing at the surface when the step took place. The surface would therefore shed vorticity thereby modifying the pressure distribution over it until, after some time has elapsed, conditions tend to the appropriate new steady state. In other words the matrix \( q \) is dependent on the time histories of the various generalised coordinates. Furthermore, any step disturbance of one generalised coordinate will in general generate changes in the generalised forces at all the other generalised coordinates so that each element in the column matrix \( q \) is the sum of a "direct" term and also \( n-1 \) "cross" terms where \( n \) is the number of degrees of freedom. The evaluation of these time dependent direct and cross functions of generalised force arising from unit step disturbances at the generalised coordinates \( \mathbf{q} \) may be referred to loosely as the "Wagner problem".

Provided that the time variation of the displacements \( q \) is "slow" and "small", it may legitimately be assumed that the flow of the fluid around the system is always that which relates to the instantaneously prevailing motion. In this event it can validly be assumed that the elements of \( q \) can be regarded as linear combinations of

*An alternative approach might conceivably be of value [4]. It would involve the ideas of "impedance matching" between the system and the flowing fluid.

—GP-4—
This means that equation (1) may be written in the alternative form

\[(\lambda^2 + \Delta)\ddot{y} + (\beta + \delta)\dot{y} + (\sigma + \eta)y = \Phi(t)\]  

(2)

where \(\lambda, \beta, \delta, \sigma, \eta, \Phi\) contain constant elements for a given regime. These can conveniently be called the inertia, damping and stiffness matrices associated with the given fluid flow.

This equation has the familiar appearance of that governing forced oscillation of a simple linear system. But whereas the matrices \(A, B, C\) have certain well known properties\(^*\), the combined (bracketed) matrices do not in general possess them.

This is for small and slow motions. There is a second special case in which the Wagner functions do not have to be considered. As one would expect, a small steady harmonic variation of \(g\) produces steady harmonic fluid forces of the same frequency, though there will in general be frequency dependent phase differences between the motions and the forces. The result of this is that the form (2) of our general equation of motion is again valid though the elements of \(\lambda, \beta, \delta, \sigma, \eta\) are now frequency dependent.

Equations (1) and (2) form a convenient basis for our discussion. It must not be imagined however that these equations will necessarily be used in the forms given in any specified problem. Nor, of course, can it be assumed that linear theory will always be adequate.

**Classical Flutter**

If an aircraft flies fast enough it will either commence to oscillate or become violently distorted, and so destroy itself. The phenomenon is known as classical flutter and it is essentially spontaneous. That is to say there is no question of an independent function \(\Phi(t)\). Nor, within the meaning of the word "classical", is there any question of periodic detachment of flow. In terms of linear theory, equation (2) now has a null matrix on the right hand side. The matrices \(\lambda, \beta, \delta, \sigma, \eta\) are essentially speed dependent and such is their form they introduce the possibility of unstable solutions of the homogeneous equations of motion.

The linear theoretical approach to flutter analysis has existed now for a number of years. It is both widely known and generally accepted in the aeronautical world. Moreover - with reservations which further research can confidently be expected to remove - it can be said that the theory is quite satisfactory. Broadly speaking, the approach can be thought of as being centred on the equations that have been quoted in this paper. Equations of motion are written for a selected system, which equations include terms representing aerodynamic forces. The technique is then to examine the stability of solutions of the equations by seeking stability boundaries.

But instability cannot yet be satisfactorily "explained" in physical terms. That is to say it is not possible to foresee instability (as it is to foresee resonance, for example); nor is it possible adequately to explain what happens as a system becomes unstable as a result of increasing speed. Considerable progress has been made on the mathematical side, however, and we now have a much improved grasp of the mathematical implications of the onset of flutter [6].

\(^*\)They are all symmetric and non-negative definite for a linear passive system in which \(A = 0 = C = \delta\).
A number of well recognised techniques now exist for the calculation of the aerodynamic terms, so long as they relate to systems that bear some resemblance to conventional aeronautical configurations. But none are so highly developed as to be cut and dried. So far as sub-sonic flight is concerned, there appear to remain few serious problems provided it is legitimate to neglect viscosity, thickness and local shock wave effects. The transonic range is still particularly difficult. Understanding of the aerodynamics in the supersonic region is gradually increasing but it remains a fact that rather little is known about the unsteady aerodynamics of the hypersonic range.

The aerodynamic terms used in flutter calculations are to some extent checked by measurements made with wind tunnel models oscillating in rigid body modes (i.e. by derivative measurements). But in flutter calculations, one is concerned with distortion modes. There remains some question therefore as to whether or not methods of calculation of flutter derivatives are sufficiently reliable when applied to modes of distortion.

In the recent past, much work on the measurement of derivatives at supersonic, transonic and high subsonic speeds was believed to be open to serious criticism, on the grounds of "wall interference". This is particularly true of the transonic range and without any doubt wall interference must still be regarded as potentially one of the major problems of flutter analysis. This defect in wind tunnel testing techniques has to some extent been removed by recent work in which the phenomenon of wall interference has been analysed, and to some extent explained [7].

One of the main drawbacks with flutter analysis is the continuing lack of a physical "feel". It is still necessary to proceed blindly without a real understanding of what quantities matter and what are relatively unimportant. This is now giving rise to many papers of a basic sort whose purpose is to elucidate the mechanism of instability. It is perhaps a commentary on the urgency and difficulty of the practical problems of flutter suppression in the past that this modern fundamental work goes right back to first principles (e.g. see ref. [8]).

Another drawback of the standard approach to flutter is that it fails to show what corrective action should be taken in order to increase flutter speeds out of the range of flying capability of a given aeroplane. (It is to be noted that although flutter is of aerodynamic origin, it is almost invariably countered by structural and not aerodynamic modification.) Here again the shortcomings are well known and a good deal of work has been done in an effort to remove them (e.g. see ref. [9]).

Yet another drawback in current flutter analysis techniques is the division that has grown up between those who study the handling characteristics of aircraft, treating the vehicle as a rigid body (which performance may have some allowance made for flexibility), and those whose concern is with flutter. It is surely true that a much more unified approach to the dynamics of aircraft will emerge in future.

In this connection, the newcomer to this field may well be perplexed by what is in effect a historical accident. A rough and ready explanation may not be out of place. If equation (2) is written in the form

\[ A \ddot{\mathbf{q}} + B \dot{\mathbf{q}} + C \mathbf{q} = \mathbf{Q} \]  

(3)

and solutions of the form

\[ \mathbf{q} = \mathbf{\psi} e^{\lambda t} \]  

(4)

are sought, there emerge the homogeneous algebraic equations

---GP-6---
\[(A\lambda^2 + B\lambda + C)\gamma = 0 \quad (5)\]
\[\gamma = 0 \quad (6)\]

The vector \(\gamma\) may be partitioned in the form

\[\{\gamma_{\text{rigid}} \cdot \gamma_{\text{distortion}}\} \text{ or } \{\gamma_x \cdot \gamma_d\}.\]

With this partitioning, equation (6) becomes

\[
\begin{bmatrix}
\gamma_{rr} & \gamma_{rd} \\
\gamma_{dr} & \gamma_{dd}
\end{bmatrix}
\begin{bmatrix}
\gamma_x \\
\gamma_d
\end{bmatrix} =
\begin{bmatrix}
0 \\
0
\end{bmatrix}
\quad (7)
\]

Those who study the stability of flight are concerned with this equation in some form or other, but mainly with regard to \(\gamma_x\). In flutter analysis, on the other hand, one is concerned with \(\gamma_d\). Unfortunately the closeness of the two types of investigation is not reflected in the literature.

It is of interest to enquire what comparable problems arise with ships. It transpires that oscillatory instability is not a common phenomenon although uni-directional instabilities are becoming more so with large ships having high block coefficients. Further, it is usually quite unnecessary to introduce distortion coordinates in the matrix \(\gamma\) when it relates to a ship or submarine.

In recent years there has been renewed interest in the stability of flow of fluids in pipes. Paidoussis has examined [10] the problem of a flexible slender cylinder surrounded by a fluid flowing in the longitudinal direction. The instability caused by flow through horizontal [11] and vertical [12] tubes has also been investigated.

In the writer's view this work - together with the early work of Brooke Benjamin cited in refs. [10] to [12] - is of a more profound significance than one might at first suspect. It may well turn out that the shortcomings of the aeroelastician's usual approach can be avoided by the adaptation of some of the ideas suggested by this work on pipes. In this connection, the impedance matching technique of ref. [4] might prove useful. Certainly the standard matrix approach is not particularly satisfying from a physical point of view.

Aside from aircraft, ships, pipes and such structures as the blades of turbomachinery and of helicopter rotors, few mechanical structures oscillate in fluid flows without periodic detachment of flow. The retention of an attached flow (or of a flow whose point of detachment remains always at one point of the surface) can be difficult to achieve. Attempts have been made (e.g. see ref [11]) to explain certain forms of panel flutter in terms of attached flow, but it seems to be still open to question whether or not such analysis is adequate.

Transcendent Loading

An aircraft flying in disturbed air is subject to random loading. Even so, an aircraft flying into a specified discrete upgust or sidegust presents a perfectly valid stress problem at the design stage. It is a problem of transcendent loading and one reason for its importance is that the competent designer must have more of a feel for his aircraft than a random process analysis can possibly give him.

In a deterministic transient analysis, the approach is to "build up" the functions \(y(t)\) and \(y(t)\) by means of Duhamel integrals used in the "indicial" sense; this requires study of what has come to be called "indicial fluid mechanics". (See refs [2], [12], [13].)
One of the standard cases in the theory of gust loading is based upon the assumption that an aircraft flies straight into a vertical gust. For this, the matrix $F(t)$ must make allowance for the development of the aerodynamic force through geographical location of the aircraft relative to the gust front. This is loosely referred to as the "Klüsner problem". Indicial aerodynamics - and high-performance submarines will probably introduce the subject of indicial hydrodynamics - has so far been confined largely to the determination of Wagner and Klüssner functions and to the application of the results by further use of convolution techniques to the study of response to shaped gusts. It is worth remarking that the application of indicial aerodynamics to gust analysis is by no means the only possibility and that it suggests approaches to accelerated flight, to problems of control input and so forth.

Generally speaking only two types of non-steady flow are studied in connection with the dynamics of aircraft; these are oscillatory aerodynamics and indicial aerodynamics. Now it was first shown by Garrick [15] that these two problems are not unconnected; indeed from a purely mathematical point of view this is obvious. By means of Fourier transform techniques, it is possible to synthesize Wagner and Klüssner functions by means of oscillatory derivatives (i.e., with the contents of the matrices $\mathcal{A}$, $\mathcal{B}$, $\mathcal{S}$) provided that the derivatives are known for all frequencies. It is also possible to proceed straight to shaped-gust analysis by means of oscillatory derivatives without using Buhse techniques. It is therefore arguable that if sufficient is known about oscillatory derivatives for high frequencies, there is no necessity to study indicial functions. (In theory the converse is also true.)

To sum up, then, so far as linear theory is concerned one can approach the functions $\mathcal{F}$ and $\mathcal{F}(t)$ by indicial techniques in which responses to step inputs are employed. Alternatively one can discard this idea of the indicial approach and work solely in terms of sinusoidal oscillations leading to the matrices $\mathcal{A}$, $\mathcal{B}$, $\mathcal{S}$.

Before concluding that it is the latter approach which should be pursued exclusively in the future (since oscillatory derivatives are required for flutter analysis), however, it is as well to recall that "gust design" is not a rigidly defined thing and this obviously throws a good deal of weight on the question of what "gust loading" should be postulated.

The truth is that indicial techniques introduce physically meaningful functions into analysis and it is for this reason that practising engineers would be loth to discard them. Moreover, if linear theory has to be discarded so that neither of the foregoing superposition techniques is admissible, the non-linear problem of a step change of flight condition assumes a special interest in its own right [16].

Concluding Remarks

Oscillations caused by flowing fluids are very common. It is therefore entirely reasonable to expect that this is an identifiable field of investigation that can be surveyed without undue difficulty. There are, however, two major obstacles. In the first place these oscillations are usually so vital that their sheer urgency has effectively prevented a general theory from growing up.

Secondly, there are many forms that these oscillations may take.

In this paper I have selected a particular form of oscillation, namely that which could be described as involving only potential flows. Having done so, I have attempted to write what might be termed an "engineer's appraisal" of this rather daunting aspect of aeronautics. While I naturally hope that this survey represents a fair and lucid statement of present knowledge there can be little doubt that it will be followed by others, as well as by books, on the same subject. For this is a field which must inevitably be made more accessible to those outside aeronautics.
So far as the general subject of this survey is concerned, it is perhaps true that as research it is losing its glamour. The pioneering work has largely been done. My belief is that the work of retrenchment which must now follow is of particular importance technologically. So strongly do I believe this, I suggest that the International Congresses of Acoustics could do much good by publicising the need for this work of interpretation.

This paper relates only to one small part of the title subject. Random loading of structures by turbulent flow is very important. Again the periodic detachment of flow (as in "stall flutter" or in excitations "by vortex shedding") is quite vital in some branches of engineering. Even a survey of what all the technical and scientific implications of aeroelasticity and hydroelasticity are would represent no mean undertaking. In short, this paper should be regarded as a tentative first attempt at providing what the Scientific Committee of this present Congress is undoubtedly right in thinking is something we need.

References


Der Impuls-Schalldtest und seine
raumakustische Beurteilung

Reichardt, Walter
Prof. Dr.-Ing., Direktor des Instituts für
Technische Akustik, Technische Universität Dresden

1. Ziel der Auswertung des Impuls-Schalldtestes für Sprache und Musik

Seit SABINE wenden alle Akustiker eine einfache orientierende
Methode zur akustischen Prüfung eines Raumes an: sie klatschen in die
Hände und hören auf die "Antwort" des Raumes. Man ist sich einig dar-
über, daß diese Methode zwar offensichtliche Fehler, wie Echos,
Überhölligkeit oder auch "zu trocken" schon viel besser anzeigt, als
lediglich Anhören von Sprach- oder Musikproben. Jedoch gestattet es
ein Qualitätsurteil an den verschiedenen Plätzen des Auditoriums
nicht. Die Prüfmethode liefert in dieser Form zu wenig Information.

Die Perfektion des Impuls-Schalldtestes stellt die Aufnahme eines
Reflektogrammes dar. Es entsteht, wenn ein kurzer, knallartiger Im-
puls mit charakteristischem Spektrum und repräsentativer Richtcharak-
teristik an der üblichen Sendestelle erzeugt und die an den einzelnen
Höreinrichtungen der Reflexionen zeitgestaffelt oszillosgra-
phiert werden. Die Auswertung des Reflexogrammes für Sprache ist be-
kannt und ausreichend. Einen ersten Vorschlag von THINL [1], der die
in den ersten 50 ms eintreffenden Schallenergie in Beziehung setzt zur
gesamten an Höreinrichtung einfallenden Energie und dieses Verhältnis
Deutlichkeit nennt, konnte NIEME [2] noch etwas verbessern, indem er
lediglich die über den Nachhallvorgang hinausragenden Spitzen als
Der Impuls-Schaltest und seine raumakustische Beurteilung

Störschallenergie \( W_s \) mit der Nutzschallenergie \( W_n \) zu einem Echograd \( \varepsilon = \frac{W_s}{W_n + W_s} \) in Beziehung bringt, der nun seinerseits wieder über den Echofaktor \( k_E = 1 - 0,3 \varepsilon \) unmittelbar ein Maß für die Beeinträchtigung der Silbenverständlichkeit liefert.

Daß eine so einfache Auswertung für Musik nicht ausreicht, hängt damit zusammen, daß es bei Sprachdarbietungen ausschließlich auf die Verständlichkeit ankommt, während man bei Musik einen Raummitklang sogar bewußt wahrzunehmen wünscht. Der Komponist hat sie in seine Klangschöpfung schon mit einbezogen.

2. Zusammenhang mit der Richtungsverteilung der Schallrückwürfe


In Fällen, in denen der erfahrene Raumakustiker aufgrund der räumlichen Voraussetzungen mit ausreichend erscheinender Sicherheit auf die Richtungsverteilung der Reflexionen schließen kann, ist eine besondere Messung der "Diffusität" entbehrlich. Auch die Fälle der Vorschläge für ein Diffusitätsmaß, das mit dem subjektiven Empfinden Übereinstimmen könnte \([1, 3, 4, 5]\), läßt Skepsis aufkommen. Verschiedene Autoren \([3, 6, 14]\) haben gezeigt, daß diese Diffusitätsmaße...
Der Impuls-Schalltest und seine raumakustische Beurteilung

sich erheblich in den drei Zeitschritten (0-30) ms, (30-50) ms und (50-∞) ms ändern, und zwar in dem Sinn, daß sie sich im Nachhall, (50-∞) ms, zufriedenstellend hohen Werten nähern. Zumindest legt das die Forderung nahe, bei Diffusitätsmessungen für raumakustische Zwecke den Schalleinfall während der ersten 50 ms unberücksichtigt zu lassen. Dort kann (und soll?) Diffusität gar nicht erwartet werden. TISMER kommt daher auch zu dem Schluß, daß die Messung der Richtungs-
diffusität mit Impulsen eine bessere Übereinstimmung mit der subjek-
tiven Beurteilung liefert.

Von der Aufnahme und Auswertung von feingliedrigen "Igeln" nach
[6, 14] ist man zugunsten einer Rundummessung mit Nierenmikrophonen
schon überwiegend abgekommen [3, 7], weil sie zuviel kaum auswertba-
re Information liefert.

3. Die wesentlichen Bestandteile des Reflektogrammes

Die Ausscheidung aller überflüssigen Information ist auch bei
Auswertung der Reflektogramme noch das Kernproblem.

3.1. Das Verhältnis der Anfangsenergie zur Nachhallenergie

Ein wesentlicher erster Schritt war die Erkenntnis, daß das Ver-
hältnis der überwiegend frontal einfallenden Anfangsenergie $E_B$ zur
diffus einfallenden Nachhallenergie $E_R$ das wichtigste Kriterium dar-
stellt und gleich wesentlich ist, wie die Nachhallzeit, mit ihr in
quantitativ angenähbarer Beziehung stehend [9, 17]. Das Ergebnis wur-
de schon in Budapest zusammengefaßt [8] und kann auch in der Form des
Bildes 1 dargestellt werden. Das Nomogramm zeigt für einfache Schall-
felder, die praktisch nur Direktschallenergie $E_B$ und Nachhallenergie
$E_R$ enthalten, daß sich z.B. gleichwertig vertreten können: Bezug-
Hallabstand $E_B = 0$ dB mit der Nachhallzeit $T = 2,0$ s oder $E_B = -4$ dB
mit $T = 1,7$ s oder $E_B = 3$ dB mit $T = 2,2$ s. Alle führen zur gleichen
Räumlichkeitsstufe $R = 7$. Man sieht, daß ein verhältnismäßig enger
Bereich in der Änderung des Hallabstandes schon recht erheblichen Än-
derungen der Nachhallzeit entspricht.

—GF-13—
Diese Erkenntnis wurde auf zwei Wegen gewonnen: Einmal von BERANEK durch Messung des Hallmaßes [10], vgl. auch Bild 3, rechts,
\[ R = 10 \log R_E^{R} \, dB = 10 \log \frac{E_R}{E_D + E_I} \, dB, \quad E_E = E_D + E_I \] (1)
\[ E_R \] nach 50 ms, \[ E_E \] vor 50 ms eintreffende Schallenergie (early energy)
\[ E_I \] bis 50 ms eintreffende "Anfangsreflexionen" (initial reflections)
in verschiedenen Konzertsaalen und Vergleich mit dem subjektiven Urteil an diesen Meßstellen.

Weiterhin aber durch Aufbau synthetischer Schallfelder, die in letzter Zeit von mehreren Autoren [9, 10, 11, 12, 15, 19] zu Untersuchungen der hier behandelten Art verwendet wurden und speziell bei [8, 9, 10, 15] übereinstimmend zu dem genannten Ergebnis führten.

3.2. Hallmaß und Hallabstand

Nun bestehen im einzelnen allerdings noch Abweichungen in den Auswertungsmaßen. Während im vielfach verwendeten Hallabstand [18], vgl. auch Bild 3, links,
\[ H = 10 \log H_E^{R} \, dB = 10 \log \frac{E_D}{E_I + E_R} \, dB, \quad E_H = E_I + E_R \] (2)
die Anfangsreflexionen \( E_I \) zur Hallenergie \( E_H \) geschlagen werden, zählt sie BERANEK beim Hallmaß \( R \) zum Direktschall \( E_D \), vgl. (1). In \( E_I \) sind aber im allgemeinen Wand- und Deckenreflexionen mit enthalten. Es handelt sich weder bei \( R \) noch bei \( H \) um ein reines Verhältnis von frontal und diffus eintreffenem Schall. Ist der Beitrag der Reflexionen innerhalb der ersten 50 ms wichtiger für Steigerung der Räumlichkeit oder der Durchsichtigkeit? Beides ist für Musik von Bedeutung.

3.3. Beitrag der Anfangsreflexionen \( E_I \) zur Räumlichkeit und Durchsichtigkeit

Die zuletzt aufgeworfene Frage wurde im Dresdener Institut des Vortragenden untersucht. Bild 2 zeigt den Einfluß einer similierten

—GP-14—
Wandreflexion W, die seitlich mit einem Winkel von 60° gegen den frontalen Direktschall D um 30 ms verzögert einfiel und im Pegel um \( \Delta L_W \) (\( \Delta L_W = 0 \text{ dB} \) bedeutet \( L_W = L_D \)) geändert wurde (bei gleichgehaltenen Gesamtlautstärke), \( f_+ \) be- deuten die (prozentual aufgetragenen) Urteile von Versuchsperso- nen "räumlicher" oder "qualitativ besser", \( f_- \) die umgekehrten Urteilungen. Es zeigt sich, daß diese Wandreflexion die Räumlichkeit meist erhöht, aber nicht zwangsläufig zugleich die Qualität verbessert.

Untersuchungen ähnlicher Art und Reflexionen anderer Einfallsrichtungen zeigten, daß seitliche Reflexionen am kritischsten Räumlichkeit und Qualität beeinflussen. Das steht im Einklang mit Untersuchungsergebnissen von SCHROEDER e.a. [15] und BERANEK e.a. [10].

Weiterhin wurde eine Minderung der Qualität und Räumlichkeit, auch nahezu konform gehend, bei Beschneidung der tiefen Spektralanteile von \( L_W \) in Übereinstimmung mit beiden Autoren bestätigt, übrigens im gleichen Maße auch beim Nachhall mit \( L_R \).

Das gleichzeitige Vorhandensein anderer Reflexionen mindert den aufgezeigten Einfluß.

Anfangsreflexionen tragen nicht immer zur Steigerung der Räumlichkeit bei. Um nachzuprüfen, ob das Hallmaß nach (1) mit \( E_I + E_D = E_R \) oder der Hallabstand nach (2) mit \( E_I + E_R = E_H \) besser mit der Räumlichkeit korreliert ist, wurde ein synthetisches Schallfeld untersucht, das nach dem in Bild 3 gezeigten Grundschema aufgebaut war. In ihm wurden die Anfangsreflexionen von der Decke mit dem Pegel \( L_D \) (ceiling) und von der Wand mit \( L_W \) um einen variablen Betrag \( \Delta L \) gegenüber der Nachhallenergie mit dem Pegel \( L_R \) oder gegenüber dem Direktschall mit \( L_D \) verschoben. Wird dabei die zu \( L_H \) zusammengefaßte
Der Impuls-Schalldump und seine raumakustische Beurteilung

Energie in $L_0$, $L_w$ und $L_P$ konstant gehalten, so bleibt der Hallabstand $H$ auch konstant. Wird ein solches Schallfeld bis zu sehr großen Änderungen von $\Delta L$ als unverändert empfunden, so darf man schließen, daß der Hallabstand $H$ ein gutes Maß für den Raumeindruck ist. Wird in einem Gegenversuch die in $L_R$ zusammengeräste "early energy" konstant gehalten und bleibt damit $R$ ungeändert, so zeigt weitgehende Unabhängigkeit von $\Delta L$ an, daß das Hallmaß $R$ ein gut geeignetes Beurteilungskriterium ist.

Es zeigte sich, daß $\Delta L$ in beiden Fällen um (10…20) dB schwanken kann, ohne daß 50 % der befragten Personen einen Unterschied bemerkten. Hier- nach erscheinen $H$ und $R$ als etwa gleich gut geeignet.


Das ist so verständlich: Bei Feldern mit $R < -10$ dB ist der Nachhallanteil so unbedeutend, daß schon die Anfangsbestandteile $L_T$ in $L_R$ allein die Räumlichkeit bestimmen. Sie kann dann bei $L_D = L_0 = L_W$ nicht mehr unter einen bestimmten Wert fallen. Andererseits bestimmen bei sehr hälligen Feldern mit $H < -10$ dB bereits die Reflexionen $L_T$ in

---GP-16---
Der Impuls-Schalltest und seine raumakustische Beurteilung

I_R die Räumlichkeit. Deren Bestandteile I_U = I_W wirken aber, wenn sie gleich I_R sind, schon so stark durchsichtigkeits erhöhend, daß die Räumlichkeit nicht über einen oberen Wert wachsen kann.


4. Die Berücksichtigung der Nachhallzeit

Eine Aufteilung von E_I in einen zu E_D und E_R zu zählenden Teil ist notwendig, weil nach dem bisher Bekannten nur über einen Bezugshallabstand E_B = - E_B eine Zuordnung zu einer Räumlichkeitsstufe R möglich ist. Diese kann dann aber die subjektive Wirkung der Nachhallzeit mit einschließen, wenn, soweit nötig, unter Anwendung von Bild 1 auf die Nachhallzeit T = 2 s umgewertet wird. Dabei ist es —GP-17—
gleichgültig, ob die Nachhallzeit bei dem Impuls-Schaltest am Ende des Reflektogrammes abgelesen oder mit anderen klassischen Methoden bestimmt wird.

Die Reflektogramme sollten jedoch auch in diesem abklingenden Teil mindestens in einem Pegelbereich von 20 dB und bis zu 200 ms hin beobachtet werden. Treten dabei konzentrierte Stoßgruppen auf, so wird sowohl die Anwendung des Begriffes Nachhallzeit wie auch die Annahme eines diffusen Einfalles fraglich.


5. Beurteilung der Qualität

Nach 3. und 4. ist eine getrennte Feststellung von \( I_D \) und \( I_I \) notwendig. Sie empfiehlt sich auch deshalb, weil nur mit ihrer Kenntnis eine Umrechnung der beiden bisher bei Auswertung von Reflektogrammen verwendeten Maße \( R \) und \( R \) möglich ist und so Anschluß an die Qualitätsbeurteilungen gewonnen werden kann, die schon in der Literatur bekannt geworden sind [40, 5]. Es wurde geprüft, ob ein optimales Schallfeld durch eine bestimmte Räumlichkeitsstufe oder einen Räumlichkeitsbereich gekennzeichnet werden kann. Bild 5 zeigt für das einfache Schallfeld, dessen Parameter angegeben sind, ein Qualitäts optimum bei der Räumlichkeitsstufe 6, für das erweiterte Schallfeld bei der Räumlichkeitsstufe 5. Das Qualitätssmaß geht auf einen Paarvergleich und eine Auswertung nach Guilford [16] zurück. Der Wert 1 bedeutet, daß die Urteilsabgabe zugunsten dieses Wertes mit der Standardabweichung erfolgt, d.h. von 100 abgegebenen Urteilen fallen 68,3 % zugunsten der zugehörigen Räumlichkeit auf der
Der Impuls-Schalltest und seine raumakustische Beurteilung

Abzisse aus.

Es muß jedoch betont werden, daß dieses Ergebnis nur für das gewählte Motiv und gewisse, dem Beobachterstab gegebenen Richtlinien für die Qualitätsbeurteilung gilt. BERANEK fand bei Auswertung von 5 Konzertsälen [10] ein R zwischen 0 und + 8 dB. Nimmt man R = + 4 dB als das Optimum und setzt man voraus, daß die Hälfte der early energy $E_0$ Direktschall war, so ergibt sich ein optimaler Hallabstand von $H_B = -7.8$ dB und bei durchschnittlich $T = 1.7$ s nach Bild 1 eine optimale Räumlichkeitsstufe von $M = 7.5$. Andererseits ergaben Abschätzungen mit einem kammermusikalischen Motiv einen optimalen Bereich in sehr viel geringeren Räumlichkeitsstufen. Der Stil der Musikkabietung erscheint damit als wichtigste Einflußgröße für die Gütebeurteilung.

5. Zusammenfassung


5.2. Der Bezugs-Hallabstand $H_B$ enthält definitionsgemäß nur Direkt- schallenergie $E_D$ und Nachhallenergie $E_R$. Die Energie der Anfangsreflexionen $E_I$ ist zur Gewinnung von $H_B$ bei sehr harten Schallfeldern ganz zu $E_D$, bei sehr trockenen Schallfeldern ganz zu $E_R$ zu schlagen. Im Übergangsgebiet ist $E_I$ sinnvoll aufzuteilen.

5.3. Anfangsreflexionen tragen zur Räumlichkeit um so mehr bei, je mehr sie seitlich (nicht von Bühnenrückwand, Bühnenrahmen und
Der Impuls-Schalltest und seine raumakustische Beurteilung

Decke) einfallen.

5.4. Nur eine Beschreibung der tiefen Frequenzen in den Rückwürfen
ist für die Räumlichkeit praktisch bedeutungsvoll und ggf. zu
beachten.

5.5. Eine Messung der Diffusität erfolgt, wenn überhaupt nötig, am
besten auch mit Impulsanregung und wird auf den Nachhallteil
beschränkt.

5.6. Der Stil der Musikdarbietung erscheint als wichtigster Einfluß
auf die raumakustische Qualität.

Schriftum

[4] Фирдузев Б.В., Акустические основы вещания. Связиздат, 
Москва, 1960
ersehene demnächst!
[7] Фирдузев Б.В., Акустический журнал, том 11, выпуск 1, 1960
Beiheft 1, S.307
[12] SCHROEDER, M.R. and B.S. ATAL: Computer Simulation... 
(1963) IEEE, Int.Convention Record, Part 7
III. Akust. Konferenz, Budapest 1964, S.380
Speech research of today is an interdisciplinary science involving many specialties such as linguistics, phonetics, physiology, physics, psychology, and electrical engineering.

An intensified research into the structure of speech signals and their production and perception should develop a foundation for man-machine communications, economy of digital speech encoding and new methods and aids for communicating with the hard of hearing and deaf.

Fundamental concepts of speech analysis and synthesis, resonator theory, and theory of perception are reviewed. Of special interest are questions such as the extent to which articulation can be inferred from the speech wave and the evaluation of the various cues in the speech wave that signal a phonetic category.

Earlier studies were static in their scope of establishing correlates to fixed articulations. Considerable effort is today laid on the studies of the temporal dynamics of the speech structure. Special emphasis is laid on the development of models which will enable the prediction of complex and varying speech patterns from the formal composition of the message in a linguistic transcript by general laws of speech production and the particular speaker classification.

Our criteria for a successful analysis are economy of specification and a demand for physiological and psychological reality. Various compromises are possible.
Current Topics of Speech Research

but much work remains before we have anything like a complete knowledge of production and perception.
Some Applications of Ultrasonic Methods for Investigating Dislocations and Point Defects in Crystalline Solid Bodies

Merkulov Lev Grigor'evich
Professor of Leningrad V.I. Ulyanov/Lenin/
Electrotechnical Institute, Leningrad, USSR.

One of the trends in modern physical acoustics is connected with the application of acoustic methods for investigating solid bodies. New interesting possibilities of studying structure and imperfections of crystalline materials arise here.

There exists a close connection between the real crystal structure and its elastic and non-elastic dynamic characteristics. Therefore, the measurements of attenuation or velocity dispersion of sound waves, carried out over a wide range of frequencies and strain amplitudes, can supply extensive information on the movement of dislocations, their interaction with the crystalline lattice and point defects as well as on the diffusion of impurities and vacancies. The methods discussed prove useful for the purposes of continuous observation of rather fine structural changes occurring under different external actions, for instance, under the influence of mechanical loads, ionizing radiations, high and low temperatures.

In developing this trend there has been achieved a considerable progress over the past decade, both in understanding the mechanism of interaction between sound waves and defects of various kinds and in accumulating a great amount of experimental data. Needless to say
that it is quite impossible to report or even enumerate in one paper the problems arising here, so much the more, that a number of reviews, monographs, and particularly, volumes 3 and 4 of Physical Acoustics [1] have been devoted to their detailed consideration. Our paper has more particular task in view — to demonstrate certain new applications of ultrasonic methods for investigating dislocations and point defects.

**Application of Effects of Dislocation Damping for Controlling Impurities Content in Pure Metals**

At present various methods are used for the industrial control of metal purity, such as chemical, spectral, mass-spectrometric, radioactivating, the method of residual resistance. Their practical application involves a number of laborious operations, usually demands the use of costly metal for making specimens and, besides, does not always ensure a sufficiently reliable detection of individual impurities. In this connection the development of non-destructive methods based on the measurements of ultrasonic attenuation is considered to be of interest for the purposes of "microflaw detection" in pure and extra-pure materials [2,3].

The attenuation in a number of metals possessing pronounced plastic properties (aluminium, copper, etc) is to a great extent determined by the dislocation effects, provided the impurity and vacancy concentration is not too high. In megacycle and kilocycle frequency ranges at room temperatures electron-phonon and phonon-phonon mechanisms usually provide a minor contribution to the total attenuation value.

According to Granato's and Lücke's theory [4] the ultrasonic attenuation in the small amplitude region is connected with the vibrations of dislocation loops pinned by point defects. These vibrations take place under the conditions of strong damping \( (\eta / \omega_c \gg 1) \). The at-
Some Applications of Ultrasonic Methods

Attenuation in this case may be represented by the formula corresponding to a simple relaxation process [5]

$$\frac{\alpha}{f} \sim \Lambda L_e^2 \frac{f/f_m}{1+(f/f_m)^2},$$  \hspace{1cm} (I)

where $\Lambda$ is the dislocations density, $L_e$ - the effective length of dislocation loop, $f_m \sim 1/L_e^2$ - relaxation frequency.

For well annealed metals, $L_e \sim 1/\rho_{imp}$ can be assumed to be inversely proportional to the average linear impurity concentration.

Thus, the number of impurity atoms determine attenuation and the value of $f_m$. The measurement of these parameters makes it possible to estimate the amount of impurities in the tested metal. In so far as it is rather difficult to determine the direct analytical connection between the particular impurity concentration and acoustic parameters $\alpha$ and $f_m$, this method is a relative one. For estimating the sensitivity of ultrasonic method it is necessary to compare the results obtained by it with the data of chemical analysis. From formula (I) it is seen that the decrease of impurity concentration results in the growth of $\alpha$, therefore, sensitivity of the ultrasonic method must rise when passing to pure and extra-pure metals, this being a considerable advantage over other physical methods of testing.

The ultrasonic control of impurity content based on the dislo-

Fig.1. Frequency dependence of attenuation in aluminium.

1 - initial material, 99.9997%;
2 - with $5 \times 10^4$ atom% Fe admixture
3 - with $25 \times 10^4$ atom% Fe admixture
cation damping has been developed in our laboratory and is now successfully used in industry. To demonstrate the possibilities of this method we are going to report some results due to one metal only - aluminium.

Fig. 1 shows experimental frequency dependence of attenuation in pure aluminium before the addition of small quantities of iron and after it.

The increase of impurity concentration results in an abrupt decrease of $\lambda$ (at 2 mc the attenuation in specimens 3 and 1 differs by the factor of 30), the relaxation maximum being shifted towards higher frequencies. Similar dependences were also observed by us when Cu, Si, Mg were added. Quantitatively the influence of various impurities on the attenuation is not the same - it is the greatest in the case of Mg and the smallest for Fe [6].

In practice, in many cases it is sufficient to know the total content of main impurities. In industrial control the value of residual resistivity $\rho_0$, measured on special specimens at liquid helium temperature (4.2°K) is usually used as a standard parameter. In this connection it is interesting to compare the values of acoustic attenuation and $\rho_0$

Fig. 2 Dependence of logarithmic decrement ($\nu=62$ ko, $T=290°K$) on residual resistivity for aluminium.

In fig. 2 is seen the logarithmic decrement $\Delta=\alpha \cdot V/f$ as a function of $\rho_0$ for nine specimens of aluminium with total impurity content (Cu and Mg) ranging from $3 \times 10^{-4}$ atom.% to $1.2 \times 10^{-3}$ atom.% (extreme
Some Applications of Ultrasonic Methods

points). It is clearly seen from the data reported that the value of \( \Delta \) can well be used instead of \( \rho_{o} \) as the parameter which characterizes the material purity. In this case a greater sensitivity as compared with the electrical resistivity method can be ensured; approx. 16-fold variation of \( \Delta \) corresponds to 7-fold variation of \( \rho_{o} \).

Metal purity control by the ultrasonic method can be performed not only on specimens but directly on ingots. In Fig.3 are given comparative data of measured values of ultrasonic attenuation and \( \rho_{o} \) along one of aluminium ingots purified by the zone melting. The boundary between the "pure" and "impure" parts is reliably determined in this way.

![Graph showing variation of ultrasonic attenuation](image)

**Fig. 3** Variation of ultrasonic attenuation (pulse echo-method, \( f=1 \) mc) and \( \rho_{o} \) (\( \rho_{o} \) values are measured after the ingot has been cut) along the ingot length; \( x \) is the boundary of the "pure" part satisfying \( \rho_{o} \leq 4 \times 10^{-10} \text{ohm.cm} \).

The measurements appear to take little time and are economical.
which constitutes an important advantage of such method.

When impurity content is increased, the dislocation damping becomes comparable with the losses due to other dissipation mechanisms. Therefore, another method is advisable when the impurity concentration is of the order of 0.01% or greater, which is based on the measurements of post-deformation recovery of attenuation.

In this method the specimen tested undergoes a short-duration pressing load, resulting in a certain variation of its dislocation structure, and the attenuation increases. The dislocations which have been displaced from their initial position, as well as newly generated dislocations (if plastic deformations are used), will gradually become pinned by vacancies migrating towards them, whose number depends on the amount of impurities. By measuring the rate of attenuation recovery one can judge of the material purity [3].

For the case when point defects migration takes place mainly along one of the planes passing through the dislocation line the theory results in the following dependence

\[
\left[\left(\frac{\Delta \alpha_o}{\Delta \alpha_t}\right)^{\frac{t}{\lambda}} - 1\right] = \beta t^{\frac{1}{2}},
\]

where \(\Delta \alpha_o, \Delta \alpha_t\) are variations of attenuation (with respect to undeformed state) at the initial instant of time and instant of time, \(\beta\) is a constant depending on material and temperature.

In fig. 4 one can see the experimental curves of recovery for three aluminium specimens with various impurity content. The investigations indicate that individual impurities affect the recovery rate in a different way; this difference, however, plays a minor part in practical problems of determining the total impurity content.

The examples considered do not certainly, exhaust all the possibilities of acoustic "microflaw detection". Other mechanisms resulting in the attenuation and dispersion of elastic wave, e.g. the re-
Some Applications of Ultrasonic Methods

Laxation effects of interstitial and substitutional atoms, vacancy pairs (Snoek, Zener processes) should also be considered.

Fig. 4 Post-deformation recovery taken after 2-minute application of 30 kg/cm² load, f=ISmc. Impurity total
1. 7x10³ atom.%; 2. 3x10² atom.%; 3. 0.17 atom.%; t = 0 corresponds to the instant of load release.

Extensive experimental studies of acoustic relaxation in high-melting metals, where the detection of low impurity concentrations may also present an important problem, seem to confirm such a possibility [7].


Acoustic magnetic resonance consists in the effect of selective absorption of ultrasonic wave energy by a spin system when the intervals between Zeeman levels correspond to the energy of phonons. The mechanism of this phenomenon represents a reversed effect of spin-lattice (thermal) relaxation, except that in the former single phonon processes are of the main importance. The coupling of propagating elastic wave with spins is established by the modulation of the crystalline electric or magnetic field with ultrasonic oscillations. In accordance with it electric quadrupole transitions which are determined by selection rules \(\Delta m = 1,2\), as well as magnetic transitions \(\Delta m = 1,\)
are excited (m is magnetic quantum number).

Acoustic resonance method opens up interesting possibilities for studying structure imperfections, and dislocations, in particular. The analysis of absorption lines shape may be utilized for this purpose.

The system of hyperfine Zeeman lines is equidistant in an ideal cubic crystal. The distortions due to imperfections result in the appearance of static quadrupole disturbance. If a four-level system for nuclei having the spin I = 3/2 is considered, the transition frequencies satisfying m = 3/2 → 1/2 and m = -1/2 → -3/2 must be shifted above and below the central transition frequency m = 1/2 → -1/2. The resulting spectrum consists of the fundamental line ω₀ and two satellites ω₀ ± Δωₛ. The amount of satellites frequency shift is directly connected with the extent of lattice distortions.

The conventional technique of nuclear magnetic resonance (NMR) for alkaline metals does not make it possible, as a rule, to resolve the given spectrum, i.e. to find Δωₛ. Making use of acoustic magnetic resonance enables us to overcome this difficulty. The method consists in observing the saturation effect of the NMR signal when ultrasonic oscillations are of the double Larmor frequency. The theoretical analysis shows that the acoustic excitation of such a non-equidistant system results in a dissimilar saturation of individual spectral lines[8]. As a result, the NMR signal becomes distorted. The amount of distortion depends on the difference ΔΩ = ωₜₐₜ - ωₙ, where ωₜₐₜ is the angular frequency of introduced acoustic oscillations, ωₙ is the frequency at which the NMR is recorded. Analysing the signal form for different ΔΩ it is possible to determine the frequency shift of the satellites. The respective experimental curves for Na²³ nuclei in NaCl crystal are shown in fig.5.

The shift found corresponds to Δfₛ = Δωₛ/2π = ±1.9 kHz.
Proceeding from the value $\Delta f_3$, the electric field gradient $\langle \phi_{zz} \rangle$ on nuclei, due to the contribution of dislocations, is calculated

$$\Delta f_3 = \pm \frac{eQ_o}{2\hbar} \langle \phi_{zz} \rangle,$$

where $e$ is the electron charge, $h$ - Planck's constant, $Q_o$ - nuclear quadrupole moment. Hence, $\langle \phi_{zz} \rangle = 4.75 \times 10^{11} \text{ CGSE units}$. Assuming that dislocations distribution with density $\gamma$ is random, one can obtain the following expression [8].

$$\langle \phi_{zz} \rangle = 0.188 e G s_{44} \sqrt{\frac{\hbar}{2}} \frac{5 \cdot 10^{-6}}{\gamma \alpha^3} \text{ CGSE units},$$

$a = 2.82 \times 10^{-8} \text{ cm}$ is NaCl lattice constant, $G = 1.3 \times 10^{11} \text{ dynes/cm}^2$ is the shear module, $s_{44} = 2.5 \times 10^4 \text{ dynes}^{-1/2}$ - the component of dynamic gradient tensor (assuming $s_{44}/s_{II} = 4/3$). The numerical solution of equation (4) for the above value of $\langle \phi_{zz} \rangle$ gives $\gamma = 8.7 \times 10^7 \text{ cm}^{-2}$.

**Fig. 5 Saturation of NMR signal on Na$^{23}$ nuclei in NaCl crystal** for different values of $\Delta \Omega$ (1 - without ultrasonic; 2, 3, 4 - with ultrasonic of constant power, $f_o = 7.984 \text{ mc}$).

Dislocations density can also be determined by another independent method from the frequency dependence curve of saturation $q(\Omega)$ which corresponds to the form of ANMR lines [9]. Such dependence is shown in fig.6.
Some Applications of Ultrasonic Methods

As can be seen, experimental points well match the theoretical curve of the type

\[ q(\Omega) = \frac{A}{(\Delta \Omega^2 + \alpha^2)^{3/2}} \]

(5)

here \( \alpha = 0.635 \), \( \Delta \Omega \gamma \), \( \Delta \Omega \gamma \) is ANMR line width at 0.5 level.

Fig.6 Frequency dependence of saturation. Dotted line – experiment; solid line – theoretical calculation.

The curve given in fig. 6 satisfies the values of constants

\[ A = 7.65 \times 10^3 \text{km}^{-1}, \quad \alpha = 9.7 \text{ km}; \text{ hence, the line width will be} \]

\[ \Delta \Omega \gamma / 2\pi = 4.8 \text{ km}. \]

As stated in [10] the total second moment of ANMR line is equal to

\[ \Delta^2 = \Delta_0^2 + \Delta_1^2 \]

(6)

\( \Delta_0^2 \) - the second moment due to dipole magnetic interactions, which can be calculated theoretically and is equal to \( \Delta_0^2 = 1.41 \text{ km}^2 \).

\( \Delta_1^2 \) - the second moment caused by the contribution of dislocations. The value of \( \Delta_1^2 \) is correlated with the electric field gradient by

\[ \Delta_1^2 = \frac{9}{7} \frac{g^2 (I + 1)}{I(I+1)} \left( \frac{e^2 Q_z}{h^2} \right) \left< \Phi_{22} \right> \]

(7)

The total moment found by numerical method from curve in fig.6 corresponds to \( \Delta^2 = 3.82 \text{ km}^2 \), using equations (6),(7) we obtain
Some Applications of Ultrasonic Methods

\[ \langle \phi_2 \rangle = 3.87 \times 10^{11} \text{ GSE units.} \] The corresponding density of dislocations calculated for equation (4) is equal to \[ 7.5 \times 10^7 \text{cm}^{-2}. \] Thus the values of \( \eta \) found by the two methods making use of satellites shift and from ANMR lines width agree quite well.

The determination of dislocations density by the optical method counting the etching pits on different areas of the given specimen surfaces, gave the values of \( \eta \) ranging from \( 10^5 \) to \( 10^8 \text{cm}^{-2}. \) Hence we may conclude that the main contribution to the width of ANMR lines is due to nuclei located in the most deformed parts of the crystal. Therefore the methods discussed above make it possible to estimate the maximum value of dislocations density in the specimen.

Further investigation by the ANMR method is of great interest, for crystals with noncubic symmetry, in particular. In this case dislocations must contribute to satellite lines, as well, therefore, the possibility of obtaining information is greatly extended.

The methods based on the excitation of spin induction signals and echo by powerful ultrasonic pulses may also be promising for studying dislocations. As indicated by the theoretical analysis, in this case it is comparatively simple to separate the contributions of magnetic dipole–dipole, exchange mechanisms and dislocation mechanism, since the amount of spin–phonon interaction is measured here independently by pulse duration for which the induction signal reaches its extreme value. It should be noted, however, that the experimental difficulties on the way of practical realization of pulse technique are rather great, and the relevant problems have been so far worked out only theoretically \([11,12]\).

The utilization of dislocations for direct generation of acoustic magnetic resonance in crystals seems to us to be very interesting. This idea has been recently stated by Kopvillem and Golenishchev–Kutuzov. Several mechanisms of such generation can be indicated.
Some Applications of Ultrasonic Methods

I. Vibratory movements of charged dislocations excite alternating electric field gradient in crystals having nuclei with quadrupole moments.

2. Vibrations of dislocations in a metal cause the redistribution of charge density of conduction electrons surrounding it, which, in turn, results in changing the contact interaction with nuclei.

3. Dislocations in superconductors of the second kind swing the quantum magnetic flux.

Experimental data referring to these completely new effects are not available so far.

REFERENCES

Acoustical Measurements of Chemical Relaxation in Electrolytical Solutions

Konrad Tamm
Professor of Physics
University of Heidelberg
Institute of Applied Physics

The measurement of ultrasonic absorption in a wide frequency range at various temperatures and pressures has proved to be an effective means for studying the structure and the dynamical properties of fluids and especially of electrolyte solutions. For the determination of relaxation times of fast chemical reactions in the range of $10^{-5}$ to $10^{-9}$ sec this method is complementary to other methods, as temperature and pressure step function techniques, electric field dispersion and nuclear magnetic resonance.

The experimental methods have been improved during the last years by extending the absorption measurements to the GHz region and to high concentrations as well as by increasing the accuracy. New results were obtained, e.g. with respect to the ion influence on the solvent and the mechanism of the multistep dissociation, which shall be reported here. They were supplemented by a theoretical treatment of the ligand field stabilization.

Classical Absorption in Pure Liquids

From classical theory an absorption is expected for pure liquids which is given (if heat conduction can be neglected) by the shear viscosity $\eta$ (sound velocity $u$ and density $\rho$) and which increases pro-
portionally to the square of frequency:

\[ \sigma_{\text{class}} = \frac{8\varepsilon^2}{3} \frac{n}{u^2 \rho} \nu^2 \]

In water and in many other pure liquids this predicted frequency dependence was verified experimentally. However, the measured value \( \frac{\sigma_{\text{exp}}}{\nu^2} \) is usually larger than the classical one, in water by a factor which depends only slightly on pressure and temperature (fig.1) and is in the order of 3. The difference \( \sigma_{\text{exp}} - \sigma_{\text{class}} \), which can be ascribed to a volume viscosity, has been explained by Hall [1] as the effect of a very fast "structure relaxation".

**Ion Influence on the Solvent**

If an electrolyte is added to the water the ions form hydration spheres and disturb the structure of the water. In addition, they interact with the ions of opposite charge and may form more or less tightly bound ion pairs or aquo complexes. The magnitude of these effects depends on charge number and radius of the ions.

Generally, in strong 1-1 valent electrolytes the interactions between the ions are small and no chemical relaxation occurs. Such solutions are, therefore, suitable for studying the influence of the ions on the solvent. Indeed, Kurtze [2] found in NaBr solutions a frequency independent value \( \frac{\sigma_{\text{exp}}}{\nu^2} \) which, however, depends on concentration and temperature. Recently it could be shown [3] that this is valid for most of the alkali halides.

In nearly all alkali halide solutions (with the exception of NaF) \( \frac{\sigma_{\text{exp}}}{\nu^2} \) decreases with concentration.

![Fig.1: \( \sigma_{\text{exp}}/\sigma_{\text{class}} \) of H2O vs. temperature and pressure](image)
Acoustical Measurements of Chemical Relaxation

at low concentrations and increases for some of the more soluble ones (LiCl, LiBr, LiI) at very high concentrations to values much greater than in water. In fig.2, as an example, \( \frac{\alpha_{\text{exp}} - \alpha_{\text{H}_2\text{O}}}{\rho^2} = \Delta \alpha/\rho^2 \), the "electrolyte absorption", is plotted versus concentration for LiCl, KBr and NaBr for various temperatures [4]. The dependence for NaBr agrees with that found by Kurtze [2] with the exception of small concentrations.

If the classical part of the absorption is calculated from the shear viscosity a similar dependence on concentration results. However, the variation of the classical absorption is too small -by a factor which was found to be approx. 3- to explain the difference \( \Delta \alpha \). The thereby suggested assumption that generally the absorption of electrolyte solutions can be calculated (as is the case for water) from the classical part by multiplying the latter by a factor \( q \approx 3 \) was confirmed for many electrolytes; for very high concentrations the factor \( q \) is reduced as can be seen for LiCl in fig.3 where \( q \) decreases to approx. 2 at 17 moles/kg \( \text{H}_2\text{O} \). Other electrolytes give similar results. The small dependence of the factor \( q \) on concentration and

--- GP-27 ---
temperature is in accordance to the assumption that shear and volume viscosity have the same origin e.g. breaking of hydrogen bonds [5].

The results found for alkali halides are very useful for the analysis of the absorption measurements of electrolytic solutions showing chemical relaxation.

Chemical Relaxation of Single Step Reactions

That part of absorption which is caused by chemical relaxation, i.e. by a delayed reestablishment of the disturbed chemical equilibrium between ions and the undissociated ion pairs or molecules, can —due to its frequency dependence—

\[
\alpha_{\text{chem}} = A \frac{\omega^2}{1 + \omega^2 \tau^2}
\]

easily be separated from the solvent absorption \( \alpha_S \) if the measurement is extended to frequencies sufficiently above the relaxation frequency \( \nu_m = 1/2\pi \tau \) where \( \alpha \) is reduced to \( \alpha_S \). The development of measuring apparatus for frequencies up to some GHz has been very useful for this purpose. As long as the absorption \( \alpha_S \) of the 'solvent' (influenced by the ions) differs only slightly from that of water \( \alpha_{H_2O} \), i.e. for small concentrations, \( \alpha_{\text{chem}} \) is given by the excess absorption

\[
\alpha_{\text{chem}} = \alpha_{\text{exp}} - \alpha_S \approx \alpha_{\text{exp}} - \alpha_{H_2O} = \Delta \alpha.
\]

If the product \( \Delta \alpha \lambda \) or better the effect of one molecule, the so-called absorption volume \( Q \lambda = \Delta \alpha \lambda / c \lambda \), is plotted versus frequency a simple relaxation curve results as is shown for ammonia [6] as an example in fig.4. The curve is characterized by its maximum \( (Q \lambda)_m \) at the relaxation frequency \( \nu_m = 1/2\pi \tau \), both values depending here on concentration c. This is to be expected for the dissociation reaction occurring in ammonia, which is a weak electrolyte:

\[
\text{NH}_4^+ + \text{OH}^- \overset{k_{12}}{\underset{k_{21}}{\rightleftharpoons}} \text{NH}_3^+\cdot\text{H}_2\text{O}.
\]
Acoustical Measurements of Chemical Relaxation

The time constant $\tau$ of this second order reaction is given by the two rate constants $k_{21}$ and $k_{12}$ and by the concentration $c_A$ and $c_B$ of the ions ($c_A \approx c_B \approx c_1 = \delta c$):

$$\frac{1}{\tau} = k_{21} + k_{12}' \quad \text{where} \quad k_{12}' = k_{12}^0 c_1 = k_{12}^0 \delta \frac{\partial \ln f'}{\partial \ln c} \cdot c.$$

Because of the small equilibrium constant of ammonia $K = k_{21}/k_{12}' = 1.5 \cdot 10^{-5}$ moles/l, it follows $\delta < 1$, $\Pi' \approx 1$, $k_{21} \ll k_{12}'$ and $2\nu_m \approx k_{12}' \cdot 2\sqrt{Kc}$. Therefore $\nu_m$ should be proportional to $\sqrt{c}$, which is fulfilled (s.fig.4). A forward rate constant $k_{12}^0 = 3 \cdot 10^{10} \text{ l/mole sec}$ is obtained, which is of reasonable magnitude for a diffusion controlled process.

For the dissociation of the complex $\text{NH}_3 \cdot \text{H}_2\text{O}$ a rate constant $k_{21} = K/k_{12}^0 \approx 5 \cdot 10^5 \text{ sec}^{-1}$ results.

From $(Q\lambda)_m$ given by the factor $G$, which is a measure of the number of ions and ion-pairs involved in the reaction, and for aqueous solutions—the difference $\Delta V$ of the partial volumes of the states 1 and 2:

$$\frac{(Q\lambda)_m}{\text{cm}^2/\text{mol}} = \frac{\delta(1-\delta)}{N A R T} G (\Delta V)^2 \quad \text{where} \quad G = \frac{\delta(1-\delta)}{2-\delta}$$

one gets $\Delta V = 28 \text{ cm}^3/\text{mole}$ for $\text{NH}_3$, a value being in good agreement with the result of other measurements.

The methods of evaluation shown here for the well-known example of ammonia are generally used, and reaction data of a great number of electrolytic solutions [8,9] including proton transfer [10] in organic substances were obtained. In some cases where intermediate states are

—GP-29—
possible, e.g. in aqueous SO₂
solutions the measured rate
constants allow a decision which
conversion is dominating [11].

If ΔV is small and the dis-
sociation is nearly complete,
very high concentrations have to
be used for the detection of the
relaxation processes. Recently
we found [3] relaxational ab-
sorption for some alkali halides, e.g. CsF (s.fig.5) with a relaxa-
tion frequency of 10 to 20 MHz at concentrations of 2 and 5 moles/l,
which was ascribed to a dissociation process:

(7) \[ \text{Cs}^+ + \text{F}^- \rightarrow \text{CsF} \]

With a degree of association of 5...10 % at 2 m and 20...40 % at 5 m
a partial volume difference ΔV≈0.4 cm³/mole was obtained, which is
somewhat smaller than the value of 1.3 cm³/mole found at lower concen-
trations from ultrasonic vibration potential measurements [12].

**Multistep Dissociation**

As is well-known many of the 2-2 valent electrolytes, e.g. MgSO₄,
show two distinct maxima within the frequency range from 20 kHz to
200 MHz indicating at least two relaxation processes. Eigen and Tamm
[7] ascribed the maxima to sequential steps of a multistep dissoci-
ation or association of the electrolyte:

(8) \[ \text{M}^{2+} \text{aq} + \text{A}^{2-} \text{aq} \xrightleftharpoons[k_{21}]{} \left[ \text{M}^{2+} \text{H}^+ \text{H}^- \text{A}^{2-} \right] \text{aq} \xrightleftharpoons[k_{23}]{} \left[ \text{M}^{2+} \text{H}^+ \text{A}^{2-} \right] \text{aq} \xrightleftharpoons[k_{32}]{} \left( \text{M}^{2+} \text{A}^{2-} \right) \text{aq} \]

In state (1) the ions are free, in state (4) they form ion pairs or
molecules with the ions very close together. The states (2) and (3)
are ion pairs with one or two water molecules between the two ions.
Acoustical Measurements of Chemical Relaxation

In the first step (1-2) of this process an ion-encounter complex is formed. The reaction rate $k_{12}'$, which is given by diffusion, in principle depends (see eq.5) on concentration and $1/\tau_1 = k_{21} + k_{12}'$ should depend on concentration, too. However, this dependence may be nearly compensated for a wide range of concentration by the activity product $\Pi_f^\ell$. The further steps correspond to the stepwise approach of the two ions which is connected with an intersection of their hydration spheres. During this process the sulphate-ion replaces a water molecule of the outer or inner coordination sphere of the metal ion.

As is well-known from the theory of multistep dissociation [6,7,13] the observable relaxation times $1/\tau_j = k_{ki} + k_{ik}'$ are not directly given by the rate constants of the isolated steps 1-2, 2-3, 3-4. They are somewhat larger since the equilibrium of one step in a sequence of steps is influenced by all the faster ones. However, the corrections included in the second terms $k_{ik}'$ do not influence the result very much since the first terms predominate under the given conditions.

In the three corresponding "amounts of relaxation" $(Q\lambda)_{mj}$ (of similar form as eq.6) the participation of more than two states in each equilibrium is of greater influence (especially for $j = \Pi$ and III) since the portions of the faster "admixed" steps contained in the effective partial volumes $\Delta V_j$ are of comparable size as $\Delta V_{ik}$ and can be of opposite sign.

It would have been possible to explain the two distinct absorption maxima of the 2-2 valent electrolytes and their dependence on concentration by a 2-step dissociation at least in principle. However, Eigen and Tamm proposed a 3-step mechanism because it yields a total dissociation constant of the expected magnitude. In this model the "lower" maximum was assigned to equilibrium III (mainly step 3-4),

—GP-31—
the "upper" maximum to equilibrium II (step 2-3). Equilibrium I (mainly step 1-2) was believed not to be measurable because of a small effective partial volume $\Delta V_I$.

New Analysis of the Measurements

The result described above was not very satisfying with respect to step I; therefore, some efforts were made to find a third relaxation maximum. Smithson and Litovitz [14] and Atkinson and his coworkers [15,16] interpreted their measurements in MnSO$_4$ and MgSO$_4$ by assuming an additional maximum at approx. 20 MHz. However, the result of the analysis seems to be influenced by the limitation of the frequency range.

We have tried to improve the old analysis by increasing the exactness of the measurements and extending the frequency range to the GHz region. The results for MnSO$_4$, as an example, are shown in fig.6 for 4 concentrations. At very high concentrations a linear increase of $\eta\lambda$ in the high frequency range, as expected for the ion influence on the solvent, is evident. Its absolute value agrees well with the value predicted from the experience with the 1-1 valent electrolytes. The classical part of the absorption had to be multiplied by a factor between 3 and 2 (decreasing with concentration) to get the best fitting value. Be-
Acoustical Measurements of Chemical Relaxation

cause of the rapidly increasing viscosity the "solvent absorption" predominates at very high concentrations.

The analysis [17,18] shows (lower part of fig.6) that the remaining part of Qλ can be explained by 3 relaxation processes. An optimal superposition of only 2 relaxation curves considerably deviates from the measured values. The assumption of a forth small relaxation process with a strong dependence on concentration, which was supposed [18] before more about the ion influence on the solvent was known, is no longer necessary.

The three relaxation frequencies are listed in table II for a number of electrolytes investigated. The "lower" one depends very much on the cation (for Mn^{2+} it is 10 times larger than for Co^{2+}) and is, therefore, assumed—as before by Eigen and Tamm—to be caused by reaction III, which is mainly the substitution reaction in the inner coordination sphere (step 4-3). The "upper" one, which is practically equal for all the bivalent cations, is now ascribed to "reaction I", the diffusion controlled formation or decay of the encounter complex (step 2-1). The "middle" relaxation frequency with a small but not vanishing dependence on the cation (factor 1.5 between Mn^{2+} and Co^{2+}) corresponds to reaction II, i.e. step 3-2.

The dependence of the three relaxation frequencies on concentration, shown for MnSO_{4} and CoSO_{4} in fig.7, is small, which means a nearly constant value c\varepsilon^{2} II^f. It is very likely that the equilibrium

—GP-33—
Acoustical Measurements of Chemical Relaxation

\[
\begin{array}{|c|c|c|c|c|c|c|c|c|c|}
\hline
\text{MeSO}_4 & 0.5m & 20^\circ C & \text{lat} & \text{1000 at calc.} \\
\hline
\text{State} & S^2_\text{Mn} & S^2_\text{Mg} & K_{ik} \text{ molecules}^{-1} & \Delta V_{ik} \text{ centimeters}^{-1} & K_{ik} \text{ sec}^{-1} \\
\hline
1 & 58 & 56 & 67 & 0.05 & 0.05 & -183 & -180 & 2000 \\
2 & 3 & 0.2 & 4 & 0.021 & 0.021 & +133 & +132 & 370 \\
3 & 20 & 32 & 77 & 1.1 & 1.3 & 34 & 40 & 22 & 8 \\
4 & 17 & 8 & 4 & 0.005 & 0.005 & -81 & -66 & \\
\hline
\text{X234} & 0.005 & 0.005 & 0.0067 & & & & & & \\
\hline
\end{array}
\]

Table I: Data for multistep dissociation of MnSO\textsubscript{4} and MgSO\textsubscript{4}

constants and, therefore, the percentages of ions in each state depend only slightly on concentration. The rapid decrease of the \((QA)_m\) for concentrations higher than 0.5 m leads to the conclusion that \(\Delta V_j\) and the \(\Delta V_{ik}\) also decrease with concentration.

The evaluation of the measured relaxation frequencies by using the above mentioned corrections yields the reaction rates for the dissociation steps \(k_21, k_32, k_43\), which are given in table I. With the assumption of \(k_34\) (from NMR measurements [19]), \(k_{12}\) (calculated from diffusion [7]), and \(K_{ik} = c_i^2/(c - c_i)\) (known from conductivity measurements [20]) the equilibrium constants \(K_{ik}\) and the concentrations \(c_i\) could be calculated. From the measured \((QA)_m\) the \(\Delta V_j\) and with the concentrations the \(\Delta V_{ik}\) of the steps were obtained, which were found to have different signs.

Recent investigations of Purdie and Vincent [21] in the frequency range 15...230 MHz have shown that a similar multistep complex formation process occurs in the sulphates of the trivalent lanthanides (Y, La, Ce, Sm, Eu, Gd, Dy, Yb). Their evaluation of the measurements yielded reaction rates \(k_34 = 2.6 - 10.5 \cdot 10^8\) sec\(^{-1}\) (for the various cations) which were ascribed to the inner-coordination sphere substitution, however, they are remarkably higher than those of the bivalent ions. There is no linear correlation between \(\log k_34\) and the reciprocal cationic crystal radius but a maximum in the middle of the series, which was -since ligand field stabilization should be small-explained by a change in the coordination sphere of the cation.

—GP-34—
Influence of Pressure

Because of the relation $\frac{\partial \ln K_{ik}}{\partial p} = \Delta V_{ik}$, the equilibrium constants, and, therefore, the relative concentrations in the states should vary with pressure. This relation, which was already used by Carnevale and Litovitz [22] to determine the value $\Delta V$ of the dissociation of ammonia from the reduction of the absorption with increasing pressure (in agreement with the values cited above), can be applied to test the values for $K_{ik}$ or $\Delta V_{ik}$ found for the multistep dissociation of $\text{MnSO}_4$. The values expected for 1000 atm are given (in small figures) in table I. The relative concentration of state 4 will be very low and since $\Delta V_{III}$ also will be small, the absorption maximum $(Q_l)_{mIII}$ is expected to decrease strongly with pressure (by a factor of 2 at 1000 atm).

Measurements of the absorption in $\text{MnSO}_4$ at 300 and 500 kHz by Fisher [23] have shown that the "lower" maximum is reduced by a factor of 3 when a pressure of 1000 atm is applied, which agrees quite well with the prediction given in table I. The measured [23] reduction of the electric conductivity of 20% at 1000 atm corresponds at least qualitatively with the decrease of the percentage of the ions. Similar results were found for $\text{MgSO}_4$. Measurements in the region of the other relaxation maxima i.e. at higher frequencies would be of high interest.

<table>
<thead>
<tr>
<th>$a^2$</th>
<th>$20^\circ$</th>
<th>$\nu_{max}$</th>
<th>MHz</th>
<th>$K_{12}$</th>
<th>$K_{23}$</th>
<th>$K_{34}$</th>
<th>$\Delta F_{1}$</th>
<th>$\Delta F_{2}$</th>
<th>$\Delta H_{1}$</th>
<th>$\Delta H_{2}$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\pm 5$ m</td>
<td>(MHz)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(kcal mole$^{-1}$)</td>
<td>(kcal mole$^{-1}$)</td>
<td>(kcal mole$^{-1}$)</td>
<td>(kcal mole$^{-1}$)</td>
</tr>
<tr>
<td>$\text{BeSO}_4$</td>
<td>440</td>
<td>30</td>
<td>0000</td>
<td>-</td>
<td>-</td>
<td>61</td>
<td>19</td>
<td>13</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>$\text{MgSO}_4$</td>
<td>440</td>
<td>70</td>
<td>021</td>
<td>-</td>
<td>0.14</td>
<td>58</td>
<td>92</td>
<td>7</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>$\text{MgSO}_4$</td>
<td>440</td>
<td>30</td>
<td>005</td>
<td>-</td>
<td>-</td>
<td>61</td>
<td>91</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>$\text{MgCr}_2$</td>
<td>440</td>
<td>-</td>
<td>015</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>$\text{VSO}_4$</td>
<td>450</td>
<td>20</td>
<td>-</td>
<td>0.13</td>
<td>53</td>
<td>60</td>
<td>-</td>
<td>-</td>
<td>170</td>
<td></td>
</tr>
<tr>
<td>$\text{CrSO}_4$</td>
<td>600</td>
<td>250</td>
<td>140</td>
<td>0.14</td>
<td>34</td>
<td>53</td>
<td>54</td>
<td>45</td>
<td>45</td>
<td></td>
</tr>
<tr>
<td>$\text{NiSO}_4$</td>
<td>550</td>
<td>100</td>
<td>45</td>
<td>0.21</td>
<td>11</td>
<td>56</td>
<td>74</td>
<td>70</td>
<td>70</td>
<td></td>
</tr>
<tr>
<td>$\text{FeSO}_4$</td>
<td>540</td>
<td>65</td>
<td>10</td>
<td>0.16</td>
<td>25</td>
<td>33</td>
<td>58</td>
<td>80</td>
<td>80</td>
<td></td>
</tr>
<tr>
<td>$\text{CoSO}_4$</td>
<td>480</td>
<td>60</td>
<td>052</td>
<td>0.14</td>
<td>36</td>
<td>58</td>
<td>84</td>
<td>79</td>
<td>79</td>
<td></td>
</tr>
<tr>
<td>$\text{NiSO}_4$</td>
<td>450</td>
<td>40</td>
<td>01</td>
<td>0.16</td>
<td>36</td>
<td>60</td>
<td>106</td>
<td>98</td>
<td>115</td>
<td></td>
</tr>
<tr>
<td>$\text{CuSO}_4$</td>
<td>600</td>
<td>290</td>
<td>140</td>
<td>0.14</td>
<td>29</td>
<td>53</td>
<td>54</td>
<td>53</td>
<td>40</td>
<td></td>
</tr>
<tr>
<td>$\text{ZnSO}_4$</td>
<td>600</td>
<td>140</td>
<td>58</td>
<td>0.20</td>
<td>13</td>
<td>54</td>
<td>59</td>
<td>68</td>
<td>68</td>
<td></td>
</tr>
</tbody>
</table>

Table II: Equilibrium constants and activation energies for multistep dissociation of bivalent metal sulphates.
Acoustical Measurements of Chemical Relaxation

Influence of Dielectric Constant

More information about the dissociation processes can be obtained by changing the dielectric constant of the solvent. Addition of dioxane or ethylalcohol to water shifts the relaxation frequency and relaxation amount considerably [6,24], which is in agreement with an expected change of the reaction rates.

Recently, absorption maxima were found [25] in solutions of alkali halides in pure dimethyl sulphoxide (dielectric constant 46.6) between 2 and 10 MHz, which were ascribed to the formation of a close ion pair (ΔV≈2 cm³/mole) because of the low relaxation frequency.

Ligand Field Stabilization

The activation energies given in table I agree with a sequence of increasing ion radii [26]. However, for the ions of the transition elements large variations exist which correspond to the variations in the reaction rates (see fig.8). Breitschwerdt [27] could show by numerical evaluation of an expanded ligand field theory (see table II, last two columns) that the variations, especially the low values of the activation energy ΔH for Cr²⁺ and Cu²⁺, which cause high values of V_mIII, can be explained by ligand field stabilization. Because of the high values of the activation energies for Ni²⁺ and V²⁺ low relaxation frequencies V_mIII should be expected for these cations. This has been confirmed for Ni²⁺.

Measuring Methods

Ultrasonic absorption in electrolyte solutions can be measured by various methods using progressive or standing waves. Some of the
Acoustical Measurements of Chemical Relaxation

methods have been improved considerably in the last years.

The application of the progressive wave pulse method could be extended to the GHz region [28] by using short delay lines made of crystalline quartz or other piezoelectric materials with lower losses and better coupling factors. These delay lines are excited directly at their surfaces by an electrical resonator field [29].

Very accurate measurements have been made by means of an improved [30] substitution method [31]. (The absorption is measured by continuously replacing -in a fixed path-length- a portion of the liquid to be investigated by a liquid of known absorption.) By using electronic equipment of very high stability absorption coefficients of only 0.01 dB/cm could be measured [32] down to frequencies of some hundred kHz. However, large quantities of liquids are necessary.

A new type of cylindrical resonator or fixed path interferometer for frequencies between 0.5 and 30 MHz has been developed [33] for small quantities of liquids (some ml). Velocity and absorption are derived from resonance frequency and half width of the system which consists of the two quartz plates and the liquid column between them. The accuracy decreases rapidly at frequencies below 0.5 MHz.

For frequencies below 100 kHz acoustical absorption measurements can be made by the spherical resonator method [34] which was improved to high accuracy [35,36] but needs large quantities of liquids and is somewhat complicated.

By using laser light for Brillouin scattering [37] the measurements of dispersion and absorption could be extended to the 10 GHz region.
References

1 Hall, L.: Phys. Rev. 73, 775 (1948)
5 Gierer, A., Wirtz, K.: Z. Naturf. 5a, 270 and 577 (1950)
17 Bechtle, A., Breitschwerdt, K.G., Tamm, K.: to be published
28 Plass, K.G.: Acustica 19, 222 (1967/68)
30 Siegert, H.: Acustica 13, 48 (1963)
32 Uenzelmann, J., Tamm, K.: to be published
33 Eggers, F.: Acustica 19, 323 (1967/68)

Diese Forderungen werden von der schon von SABINE in die Raumakustik eingeführten Nachhallzeit weitgehend erfüllt. Auch heute messen wir in einem Raum als erstes die Nachhallzeit oder wir berechnen sie für einen neu zu bauenden Saal, um festzustellen, ob ein vorliegender Entwurf überhaupt akustisch in Frage kommt und welche akustischen Zusatzmaßnahmen gegebenenfalls erforderlich werden.

Andererseits besteht kein Zweifel darin, daß die Nachhallzeit eines Raumes seine Akustik keineswegs vollständig charakterisiert. Die
Objektive Messungen in der Raumakustik

Erfahrung zeigt, daß es durchaus Räume mit gleicher Nachhallzeit, aber dennoch sehr unterschiedlichen raumakustischen Eigenschaften gibt. Es hat daher nicht an Versuchen gefehlt, der Nachhallzeit ergänzende, objektiv meßbare Parameter an die Seite zu stellen in der Hoffnung, mit ihnen gerade die subjektiven Eindrücke zu berücksichtigen, die von der Nachhallzeit nicht erfaßt werden. Einige davon sind in der nachfolgenden Tabelle zusammengestellt, zusammen mit den Schallfeldeigenschaften, auf die sie sich beziehen:

<table>
<thead>
<tr>
<th>Schallfeldeigenschaft</th>
<th>Kriterium</th>
<th>Urheber</th>
</tr>
</thead>
<tbody>
<tr>
<td>Feinstruktur des</td>
<td>Deutlichkeit</td>
<td>E. MEYER und</td>
</tr>
<tr>
<td>Nachhalls</td>
<td></td>
<td>THIELE, 1953</td>
</tr>
<tr>
<td></td>
<td>Echograd</td>
<td>NIESE, 1956</td>
</tr>
<tr>
<td></td>
<td>Rückwurf-</td>
<td>SCHODDER, 1956</td>
</tr>
<tr>
<td></td>
<td>statistiken</td>
<td></td>
</tr>
<tr>
<td>Feinstruktur der</td>
<td>Transmission</td>
<td>WENTE, 1935</td>
</tr>
<tr>
<td>Frequenzkurve</td>
<td>irregularity</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Frequency irregularity</td>
<td>BOLT u. ROOP,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1950</td>
</tr>
<tr>
<td>Richtungsverteilung</td>
<td>Richtungs-</td>
<td>E. MEYER und</td>
</tr>
<tr>
<td></td>
<td>diffusität</td>
<td>THIELE, 1953</td>
</tr>
<tr>
<td></td>
<td>Kreuzkorrelation in zwei</td>
<td>COOK et al. 1955</td>
</tr>
<tr>
<td></td>
<td>Raumpunkten</td>
<td></td>
</tr>
</tbody>
</table>

Ohne Zweifel kennzeichnen die angegebenen Kriterien wesentliche Schallfeldeigenschaften, und doch kann keine Rede davon sein, daß ihre praktische Bedeutung mit der der Nachhallzeit auch nur entfernt zu vergleichen wäre. Der Grund hierfür liegt sicher nicht allein in dem erforderlichen apparativen Aufwand. Eher ist es darin zu suchen, daß man bei ihrer Aufstellung primär von einem physikalischen Standpunkt...

—GP-40—
Objektive Messungen in der Raumakustik

ausgegangen ist in der Erwartung, daß sich ihre subjektive Aussage-
kraft im Lauf der Zeit und mit wachsender Erfahrung von selbst ergeben
werde, ähnlich wie das bei der Nachhallzeit der Fall war. (Eine gewis-
se Ausnahme hiervon machen die "Deutlichkeit" und der "Echograd", bei
deren Definition der HAAS-Effekt berücksichtigt wurde.) Allerdings
sind in den letzten Jahren eingehende psycho-akustische Untersuchungen
in synthetischen Schallfeldern durchgeführt worden, die manche dieser
Kriterien in einem neuen Licht erscheinen lassen und die zeigen, in
welcher Richtung sich die raumakustische Meßtechnik vielleicht weiter-
entwickeln könnte. Im Folgenden wird daher von solchen objektiven Meß-
verfahren die Rede sein, die durch subjektiv-akustische Befunde unter-
mauert sind.

Beginnen wir mit der Feinstruktur des Nachhalls und der Feinstruk-
tur der Frequenzkurve eines Raumes. Beide hängen bekanntlich über die
Fourier-Transformation aufs engste miteinander zusammen, was ihre ge-
meinsame Behandlung rechtfertigt. Idealer Raumnachhall sollte sich aus
völlig regelloser Schallreflexionen zusammensetzen; dementsprechend
sollte die Frequenzkurve eines Raumes völlig regellos verlaufen. Häufig
enthalten Nachhallvorgänge aber dominierende Reflexionen oder pe-
riodische Reflexionsfolgen, die sich entweder als Echos bzw. Platter-
echos oder aber als Klangfärbung subjektiv bemerkbar machen. Sie sind
in einem entsprechenden Oszillogramm nicht immer klar zu erkennen,
wohl aber aus der Autokorrelationsfunktion des betreffenden Nachhall-
vorgangs. In einem Autokorrelogramm werden nämlich grundsätzlich alle
Zufallskomponenten in einem zentralen Hauptmaximum zusammengefaßt,
während alle anderen Anteile als Seitenkomponenten der Autokorrela-
tionsfunktion in Erscheinung treten. Von besonderem Interesse ist der
Fall zweier ausgeprägter Nebenmaxima, die symmetrisch zum Zentrum lie-
gen. Ihnen entspricht im Frequenzbereich eine periodische "Modulation"
der Leistungsfrequenzkurve.

—GP-41—
Objektive Messungen in der Raumakustik


Dieser gehörbedingte Zusammenhang kann in den Zeitbereich übertragen werden und liefert dann einen Anhaltspunkt für die Beurteilung gemessener Autokorrelationsfunktionen. Unter der Voraussetzung, daß man diese an synthetischen Filtern gewonnenen Ergebnisse auf die Leistungsfrequenzkurven bzw. Autokorrelationsfunktionen von Räumen übertragen darf, kommt man dann zu folgendem Verfahren:

Die experimentell gefundene Autokorrelationsfunktion $\phi_o(\tau)$ des Raumes - gemessen in einem bestimmten Raumpunkt - wird mit einer multiplikativen Bewertungsfunktion $\Phi(\tau)$ versehen, welche den erwähnten Gehöreinfluß berücksichtigt:

$$\phi_s(\tau) = \Phi(\tau) \cdot \phi_o(\tau) \quad (1)$$

Eine hörbare Klangfärbung ist nach den erwähnten Untersuchungen dann zu erwarten, wenn

$$\phi_s(0) < 16 \phi_s(\tau_0) \quad (2)$$
ist. Dabei kennzeichnet $\tau_0$ den Abstand des Nebenmaximums vom Hauptmaximum ($\tau = 0$) der Autokorrelationsfunktion.

Die Bewertungsfunktion $g(\tau)$ nach HILSEN ist in Fig.1 graphisch dargestellt (rechte Ordinate). Das gleiche Bild gestattet auch die Beurteilung der Frage, ob eine wahrnehmbare Klangfärbung auftritt oder nicht: Auf der linken Ordinate ist das gelegentlich als "zeitliche Diffusität" bezeichnete Verhältnis $\Delta = \phi_0(0)/\phi_0(\tau_0)$ abgetragen. Liegt der durch $\Delta$ und den zugehörigen $\tau_0$-Wert gegebene Punkt unterhalb der Kurve, so ist mit einer Klangfärbung zu rechnen, in anderen Fällen nicht.

Als Beispiele sind einige Meßresultate eingetragen, die an verschiedenen Räumen bzw. Raumnachbildungen zur künstlichen Nachhallerezeugung erhalten wurden. Die mit 2, 6 und 25 bezeichneten Kreise beziehen sich auf ein Spiralnachhallgerät, auf dessen beidseitig reflektierend abgeschlossener Spirale entsprechend viele Sperrmassen sitzen, die unregelmäßig angeordnet sind und welche die regelmäßige Reflexionsfolge durch unregelmäßige Partialreflexionen auffüllen. Der Kreis P
Objektive Messungen in der Raumakustik

ist an einer Nachhallplatte gewonnen worden, während die mit I. und H bezeichneten Kreise aus den Autokorrelationsfunktionen eines Laborraums und eines großen Hallraums ermittelt wurden. (Sämtliche Messungen wurden im Frequenzbereich zwischen 900 und 1120 Hz durchgeführt.) Außer den beiden "echten" Räumen ergeben alle Messobjekte hörbare Klangfärbungen, was auch mit der Beobachtung übereinstimmt.


— GP-44 —

Die subjektive Beurteilung solcher Nachhall-Schallfelder führte nun zu dem Ergebnis, daß unser Gehör keine sehr großen Anforderungen an die Gleichmäßigkeit der objektiven Richtungsverteilung stellt. Um ein räumlich wirkendes Schallfeld zu erhalten, genügt es schon, wenn der Schall aus vier oder fünf deutlich verschiedenen, horizontalen Richtungen einfällt. Allerdings ist eine wichtige Zusatzbedingung zu erfüllen: die einzelnen, aus verschiedenen Richtungen eintreffenden Schallanteile dürfen untereinander nicht kohärent sein, sie dürfen also nicht miteinander interferenzfähig sein. Die gegenseitige Kohärenz zweier Signale kann man durch ihren Korrelationskoeffizienten kennzeichnen; dieser muß also wesentlich kleiner als 1 sein. Bei
voller Kohärenz aller Richtungskomponenten scheint aller Schall subjektiv aus einer einzigen, gut lokalisierbaren Richtung zu kommen, so gleichmäßig die objektive Richtungsverteilung auch sein mag.

Im synthetischen Schallfeld kann die gegenseitige Kohärenz durch die Art der Verhüllung gestört werden: Man verwendet bei der Wiederaufnahme des verhallten Signals im Hallraum mehrere, hinreichend weit voneinander entfernte Mikrophone, deren Ausgangsspannungen getrennten Lautsprechern oder Lautsprechergruppen zugeführt werden. Es genügt auch, die einzelnen Lautsprechersignale gegeneinander zu verzögern, wobei die notwendigen Verzögerungszeiten von der Art des Signals abhängen. Der letztere Prozess ist vermutlich auch in realen Räumen für die unvollständige Kohärenz der einzelnen Schallfeldkomponenten und damit für die räumliche Wirkung des Gesamtschalldfelds verantwortlich.

Will man die "subjektive Diffusität" in einem Konzertsaal mess-technisch untersuchen, so genügt es demnach nicht, nur die stationäre Richtungsverteilung mit einem Richtmikrophon auszumessen; man hat viel- mehr auch die Kohärenzverhältnisse zu untersuchen. Im Prinzip wird hierzu die Ausgangsspannung eines Richtmikrophons mit der eines zweiten, in eine andere Richtung weisenden Richtmikrophons korreliert. Allerdings wird der mit einer solchen Korrelationsmessung verbundene zeitliche und apparative Aufwand erst durch die Anwendung des schon früher erwähnten Rückspielverfahrens auf ein tragbares Maß reduziert. Dieses Verfahren sei an Hand von Fig. 2 kurz erläutert:

Der Raum wird über ein passend gewähltes Filter mit einem kurzen Impuls angeregt. Die mit einem Mikrophon \( M_1 \) aufgenommene Impulsantwort wird zunächst auf Magnetband gespeichert. In der zweiten Phase der Mes-
Objektive Messungen in der Raumakustik

... wird die Laufrichtung des Magnetophons und damit das Zeitvorzeichen der Impulsantwort umgekehrt, mit ihr erregt man bei Schalterstellung II den Raum erneut und erhält jetzt am Mikrophon $M_2$ die Kreuzkorrelationsfunktion des Schalldrucks in zwei um die Strecke $x$ entfernten Raumpunkten. (Bei $x = 0$ ergibt sich die Autokorrelationsfunktion.) $M_1$ und $M_2$ können aber auch zwei in verschiedene Richtungen weisende, ortsgleiche Richtmikrophone sein; in diesem Fall müßt man die oben erwähnte Kreuzkorrelationsfunktion zweier verschieden gerichteter Komponenten. Schließlich genügt zu dieser Messung ein einziges Mikrophon, dann kann man zwischen den beiden Meßphasen die Positions- bzw. Richtungsänderung des Mikrophons vornehmen. Die unabhängige Variable der jeweiligen Korrelationsfunktion ist die reale Zeit, man kann das Korrelogramm also auf einem Oszillosgraphen beobachten. Irgendwelche Verzögerungsglieder oder Multiplikatoren werden nicht benötigt.


$$\langle h^2(t) \rangle = \int_t g^2(x)dx = \int_0^t g^2(x)dx - \int_0^t g^2(x)dx$$ \hspace{1cm} (3)

Darin ist $g(t)$ die Impulsantwort des Raumes in einem bestimmten Punkt;
Objektive Messungen in der Raumakustik

\(\langle h^2(t) \rangle\) ist das Ensemblemittel aller Nachhallkurven, die in dem gleichen Raumpunkt bei Rauschverstärkung möglich sind. SCHROEDER hat bereits 1965 Nachhallmessungen nach diesem Verfahren ausgeführt, wobei die durch Gl.(3) vorgeschriebene Integration (erstes Integral) mit einem

![Diagramme](image)

**Fig. 3 Zum Verfahren der "integrierten Impulsantwort"

Digitalrechner vorgenommen wurde. Zur Auswertung mit Analogmitteln eignet sich besonders die zweite Schreibweise der Gl.(3). In Fig. 3 ist eine entsprechende, in Göttingen verwendete Anordnung schematisch wiedergegeben:

Der Raum wird mit einem gefilterten Impuls angeregt; seine Impulsantwort wird mit einem Diodennetzwerk quadriert und mittels eines entsprechend beschalteten Rechenverstärkers integriert. Nach einigen Sekunden wird der geladene Integrationskondensator umgepolt und der ganze Prozess wiederholt sich. Am Ausgang der Schaltung entsteht jetzt die Differenz beider Integrale und damit die gesuchte "mittlere" Nachhallkurve; sie kann mit einem normalen Pegelschreiber aufgezeichnet werden. Die erforderlichen Schaltvorgänge werden von einem (nicht gezeichneten) Schaltwerk übernommen.

In Fig. 4 sind nach diesem Verfahren gemessene Nachhallkurven wiedergegeben. Zur Filterung wurde ein Terzsieb verwendet. Jede Kurve ist doppelt registriert worden, um die ausgezeichnete Reproduzierbarkeit zu demonstrieren. Die noch vorhandenen Unregelmäßigkeiten sind nicht auf Zufälligkeiten der Raumregelung oder auf Überschwingen des Pegelschreibers zurückzuführen, sondern spiegeln echte Schallfeldeigen-
schaften wieder.

Andere Autoren\(^1\) haben zur apparativen Auswertung der Gl. (3) mehr oder weniger abweichende Anordnungen verwendet. Auf jeden Fall erlauben die nach dem Verfahren der "integrierten Impulsantwort" aufgezeichneten Nachhallkurven eine Auswertung, die über die einfache Bestimmung der Nachhallzeit hinausgeht, was von der üblichen Methode kaum gesagt werden kann.

Zusammenfassend ist festzustellen, daß die raumakustische Meßtechnik durch die Entwicklung der letzten Jahre an Zuverlässigkeit und Aussagekraft gewonnen hat. Nach wie vor dürfte die Nachhallzeit die wichtigste und auch anschaulichste Maßgröße sein, doch haben sicherlich auch die auf die zeitliche und richtungsmäßige Schallfeldstruktur bezogenen Parameter ihre reale Bedeutung.

\(^1\) [7], [8]
Literatur


Fourth International Congress on Acoustics, Copenhagen 1962 H 31


Acustica 19 (1967/68)

[3] DAMASKE, P.

Acustica 19 (1967/68) S. 199

[4] BURGTORF, W., WAGENER, B.

Acustica 19 (1967/68) S. 72


Acustica 19 (1967/68) S. 132


J. Acoust. Soc. Amer. 37 (1965) S. 409

[7] BROCH, J.T., JENSEN, V.N.

Brüel u. Kjaer Tech. Rev. No.4, 1966

[8] KURER, R., KURZE, U.

Acustica 19 (1967/68) S. 313
Radiation Induced Nucleation and Ultrasonic Cavitation

D. Sette
Istituto di Fisica, Facolta' di Ingegneria, Universita' di Roma
Gruppo Nazionale di Struttura della Materia del Consiglio Nazionale delle Ricerche, Rome, Italy

1) Introduction

Cavitation in a liquid exposed to a suitable sound field consists in the growth to visibility of tiny bubbles (nuclei) having radii of the order of a few microns.

In studying cavitation in liquids an important distinction must be made between voporous and gaseous cavitations. The first one occurs in degassed liquids; the cavitation nuclei, as well as the cavitation bubbles, are formed only by vapour of the liquid. Bubbles of this kind when collapsing give rise to intense shock waves. Acoustic pressures needed for the onset of such a process are usually very high. On the contrary, gaseous cavitation occurs at a much lower sound pressure. Nuclei and bubbles contain vapours as well as gases initially dissolved in the liquid. The collapse in such bubbles does not reduce the volume to zero. Moreover stabilizing mechanisms may be effective, e.g. if suspended solid particles are in the liquid.

In the present paper we are interested in the nucleation process, i.e. in the creation of microbubbles of micron size. Microbubbles may be introduced into a liquid in various ways: for example, they may be attached to dust impurities, or trapped in microscopic cracks of walls of the container when the tank is filled. Micro-
bubbles may also have been originated by previous cavitation events. Such microbubbles can be eliminated from the liquid, e.g. by pressurization; in such a bubble-free liquid cavitation may be set in again in the presence of ionizing radiation; moreover in the case of normal water, where a stabilization mechanism allows storage of nuclei, a population of nuclei may be restored as a consequence of exposition to ionizing radiation. We wish here to concentrate our attention on such a process of nucleation.

2) Nucleation and ionizing radiation

The effect of ionizing radiation on cavitation was first recognized by Sette\(^{(1)}\) and by Liebermann\(^{(2)}\).

In Sette's experiment the threshold of acoustic cavitation in distilled water held in a large container was measured as a function of time. A remarkable increase was observed by screening the tank with a lead screen of 15 mm thickness, as a consequence of the variation of the ionizing effect of cosmic radiation.

The experiments of Liebermann were performed in pentane and acetone. The cavitation threshold was studied with and without a Po-Be neutron source. Threshold was established at negative pressure of 3.5 bar in pentane and 5.9 bar in acetone in the presence of the neutron source; without the neutron source no cavitation was observed up to 20 bars, the maximum value of acoustic pressure used in the experiments. The liquids were degassed, and vaporous cavitation was observed.

These experiments lead to the conclusion that fast neutrons can nucleate cavitation both in gassed and in degassed liquids. This point of view was successively confirmed by Hahn's experiments\(^{(3)}\) where the tensile strength of liquids was measured with a rotating capillary tube. The fracture time is drastically shortened and the rupture takes place at a lower negative pressure when an ionizing agent is present.

The effects of lead screens and of neutrons was later investigated more deeply by Sette and Wanderlingh\(^{(4)}\). Figure 1 gives
Radiation Induced Nucleation and Ultrasonic Cavitation.

Fig. 1. Cavitation threshold in water: effects of a lead screen and a neutron source.

The results of threshold measurement in distilled water as a function of time, and shows the effect of a 15 mm lead screen and of a 10 mc Ra-Be neutron source introduced in the screened water. It is evident that screening increases the threshold, i.e. decreases the density of nuclei present because the ionizing effect of cosmic radiation is partially cut; the liquid in such a condition becomes sensitive to energetic neutrons, and neutrons from a Ra-Be source are able to produce nuclei in water. The mechanism is that of neutrons displacing atomic nuclei in the liquid and of knocked-on nuclei acting as ionizing particles. Of course proton and oxygen nuclei could be at work in the case of water; protons however do not have a sufficient density of energy deposit during their life in the liquid. This is clearly shown in the experiment of Fig. 2 similar to that of Fig. 1, where however the energy of neutrons entering the liquid has been reduced by a factor of 10. No effect follows the introduction of the source, notwithstanding that protons in both experiments receive from neutrons enough energy to pass during their life in the liquid through the region of maximum energy deposit. Only oxygen recoil nuclei are therefore able to produce cavitation nuclei in water.

The experiments described show also that cavitation nuclei
Radiation Induced Nucleation and Ultrasonic Cavitation.

Fig. 2. Cavitation threshold in water: effects of a lead screen and a neutron source. Energy of neutrons reduced to $\frac{1}{10}$ the value for the experiments of Fig. 1.

have an average life in water of the order of an hour and, moreover that the threshold is a poor parameter for carrying out research on nucleation. In fact, as shown in Fig. 1, the effect of an artificial source on the threshold is of the same order of magnitude as the effect of cosmic neutrons while the neutron flux of artificial source is much more intense than that of cosmic radiation, even if one considers that only neutrons having energy larger than 10 MeV are able to act as nucleating agents.

In order to have a more precise indication of the nucleation process, a statistical procedure has been developed. Such method, which mainly consists in the determination of acoustical pressure at which the first cavitation event occurs in a large number of small samples, allows the experimental determination of the distribution of nuclei population as a function of the acoustical pressure required for their growth to visibility\(^5\).

Fig. 3. Distribution function of cavitation nuclei in normal distilled water.

---GP-54---
Fig. 3 shows typical results of such a measurement. The ordinate gives the number of nuclei able to grow under the acoustical pressure indicated on the abscissa. The method allows detection of very small concentration of nuclei and loses its validity for a concentration such that the probability of finding a nucleus of the given size in the volume under test becomes almost unity.

The difference between concentrations in normal water and in screened water is given in Fig. 4. Positive and negative values indicate nuclei whose creation is respectively inhibited or brought about by the screening. It is evident that screening has two effects: a general decrease of the nuclei population, and an increase of nuclei of a particular size. The possibility of a production of secondary neutrons in the bulk of the screen is suggested by such results.

The difference between normal water and screened water irradiated with Ra-Be neutron source, is shown in Fig. 5. One can see the general decrease produced by the screening, together with an increase due to neutron irradiation. Such an increase refers to the
same kind of nuclei whose concentration is enhanced by the screen
action. The energy of neutrons of a Ra-Be source varies between 0 an
14 MeV, and it seems plausible from various considerations that
nucleation in water is induced by the more energetic neutrons emitted
by the source. This would indicate an energy of the order of 10 MeV
for the secondary neutrons produced by cosmic radiation in lead. This
result is consistent with the nuclear reaction that can take place.

The general decrease in the nuclei population due to
absorption in the screen leads to an effect (increase) on the
threshold which is not compensated by the effect of nuclei created
by secondary neutrons. Therefore the threshold found in screened
water is higher than the one found in water held in a normal
condition (fig.1). The larger production of cavitation nuclei when
the neutron source is present is sufficient to restore the threshold
approximately to the normal value, notwithstanding the screen.

The statistical methods have been used to investigate the
effect of an uranium salt (uranyl nitrate), dissolved in water. The
purpose of this experiments was to observe a possible production of
cavitation nuclei from heavy ionizing fission fragments. The concentrations of
cavitation nuclei in water and in water in which an uranium salt was dissolved\(^6\)
are given in figure 6. The effect of uranium salt is evident. A large nuclei
population is created. The sound pressure at which such nuclei can grow is very
low, as expected for nuclei of larger size than those created by the neutrons.

Many other similar experiments have been performed in various laboratories.
Hahn and Peacock\(^7\) have studied the
Radiation Induced Nucleation and Ultrasonic Cavitation.

effect of a Po-Be neutron source on the cavitation in liquid freon (C Cl F$_3$). They observe the fast light pulse emitted by cavitation bubbles (vapourous cavitation). Cavitation is triggered by the presence of the neutron source. No cavitation appears without a source. In the same experiments the effect of uranium salts on cavitation in methyl acetate has been observed.

R.D. Finch has studied the effect of neutron irradiation on cavitation in partially degassed water$^{(8)}$. In such a case, the cavitation threshold in normal conditions is found at 5 to 7 atm. The results of Finch's experiments are in a good qualitative agreement with those of Sette and Wanderlingh$^{(4)}$. In fig.7 the threshold behaviour in the presence of a neutron beam is given. Fig.8 shows the effect of lead screening. The main difference between these results and those of Sette and Wanderlingh (see for comparison fig.1) is that artificial neutrons give an effect (threshold lowering) without a preventive screening of the container. However degassing of water obviously

---

Fig.7. Cavitation threshold in partially degassed water in presence of a neutron source (R.D. Finch).

Fig.8. Cavitation threshold in partially degassed water: effect of a lead screen (R.D. Finch).

---GP-57---
Radiation Induced Nucleation and Ultrasonic Cavitation.

decreases nuclei population, and such a circumstance makes the sample sensitive to neutrons of sufficient energy.

It is to be noted that the threshold in carefully degassed and cleaned water is ordinarily much higher than that found by Finch. The water used by Finch however, was not filtered and gas remained in it: a stabilizing mechanism operates and explains the observed average life of nuclei.

A statistical study of cavitation has also been performed by Barger. Among other characteristics he studied the effect of a paraffin screening on cavitation rate in water samples prepared in a special way (fig.9). Again a noticeable effect is evident: the rate of occurrence of cavitation events in a given sound field is reduced by a factor of 3 by screening.

More recently a careful investigation on radiation-induced cavitation has been performed by M. Greenspan and C.E. Tschiegg. They have studied the effect of neutrons, α-particles and fission fragments on cavitation in a large number of liquids. Great care was taken to eliminate suspended particles from the liquid. In such conditions the threshold of cavitation is practically unaffected by the amount of dissolved gases, below saturation. Moreover they have observed, even in gassy water, no appreciable life-time of cavitation nuclei produced by neutron irradiation. These results confirm the idea that the presence of suspended particles is a
necessary condition for the stabilization mechanism in water and, in this respect, their results agree with others\(^{(12)}\).

Greenspan and Tschiegg, as well as C. West and C. Howlett\(^{(13)}\) have performed some experiments in neutron irradiated liquids showing conclusively that a cavitation event follows from the action of a single neutron, i.e. it is not the consequence of a cooperative effect of two or more neutrons.

3) Theory of Nucleation

Let us now discuss the mechanism by which a cavitation nucleus may be created by an ionizing particle. A first attempt at explanation has been to apply Seitz's thermal spike theory\(^{(14)}\) which had been developed for the bubble chamber operation. Energetic neutrons produce recoil atomic nuclei which act as nucleating agents in the liquid. While such a theory may be able\(^{(2)}\) to explain results in liquid like pentane and acetone, it fails for water\(^{(4)}\); Seitz' theory requires a value of 47000 MeV/cm as density of energy deposit in order to have nucleation, while the maximum value this parameter can attain for an oxygen recoil nucleus, which produce cavitation nuclei, is about 6000 MeV/cm.

Recently a thermodynamic theory has been proposed for nucleation\(^{(15)}\), which takes into account the statistical existence in a liquid of very tiny microbubbles (embryos) from which a nucleus can originate. It seems able to explain both bubble-chamber operation and the creation of cavitation nuclei. It applies also to gassy water and allows us to explain experiments if due consideration to dissolved gases is given. The conclusion of such a theory for gassy water can be illustrated with reference to fig.10. The full line represents the temperature for which the radius of a microbubble is a critical one, i.e. nuclei grow for \( r > r_c \). Dashed lines give the decrease of temperature of embryos as consequence of their growth; each line refers to a different value of the energy initially furnished by the ionizing particles. Growth stops when the dashed line crosses the
Radiation Induced Nucleation and Ultrasonic Cavitation.

Fig. 10. Thermodynamic theory of nucleation. Temperature of the region overheated by the ionizing particle in water as a function of the embryo dimensions (dashed curves); critical radii and limit of exoenergetic growth in gassed water.

work on this mechanism\(^{(12)}\) has shown the predominant role played by solid impurities.

4. Neutron detection and dosimetry. The action of energetic neutrons in creating cavitation nuclei in liquids leads to the possibility of developing new methods for measurements on high energy neutrons. It is well known how strong are the needs of finding efficient devices in this field. The possibility of neutron detectors based on cavitation was already implicit in the first works on the subject and was put more explicitly later\(^{(16)}\). We wish to indicate here some lines of development.

C. West and co-workers\(^{(17)}\) are trying to develop a single neutron detector. A cylindrical cell driven by a piezoelectric ring is used and one or more pick-ups are placed to detect the shock waves.
Radiation Induced Nucleation and Ultrasonic Cavitation.

produced by collapsing cavities. The cavitation observed is vaporous in degassed tetrachloroethylene. The signal at the microphone produced by the shock waves of the collapsing bubble has a 2 μsec rise time. Some brilliant experiments with a pulsed neutron (74 MeV) source have shown that the acoustic signal follows nucleation within one period of the 20 k Hz sound wave and that the average life of nuclei in the conditions existing in the liquid is less than 2 μsec. The ultimately goal is to have an efficient detector of single high-energy neutrons and, using several microphones, to locate the point at which the shock waves are produced, i.e. the position of the neutron in the cell.

The peculiar properties of gassed water, where a stabilization mechanism allows storage of nuclei, can be used to detect extremely low neutron flux[18], such as the neutron component of cosmic radiation. Moreover, because the energy transfer from neutron to atomic nucleus depends strongly on the mass ratio, and because the ionizing power of the struck nucleus is also mass dependent, a suitable choice of elements present in a water solution makes the liquid sensitive to neutrons of energy higher than a given value. The possibility appears of a spectroscopy of high energy neutrons. Experiments carried on in water solutions having in turn lithium, boron and carbon atoms have in fact shown how the distribution of cavitation nuclei as a function of their radius changes according to the type of atom present.

Greenspan and Tschiegg[11] have also given some preliminary results on the dependence of cavitation threshold in liquids such as alcohols and Freon 113 (C₂Cl₃F₃) on the energy of the incoming neutrons, pointing out the possibility of neutron spectroscopy.
BIBLIOGRAPHY

(1) D. Sette - Sonic cavitation and ionizing radiation - III Meeting of International Commission on Acoustic - Stuttgart 1959.
(5) D. Messino', D. Sette, F. Wanderlingh - IV Meeting of International Commission on Acoustics, Copenhagen, 1962, paper J 34
(9) Private communication.
(14) F. Seitz - Phys. Fluids 1, 2 (1958).
(17) C. West - Nuclear Instr. and Methods 33, 361 (1965);
Psychophysics and the Measurement of Loudness

S. S. Stevens
Professor of Psychophysics
Harvard University

Research in many laboratories has shown that the psychophysical law is a power function. Sensation $\psi$ increases in proportion to the stimulus $\varphi$ raised to a power $\alpha$. Thus $\psi = k(\varphi - \varphi_0)^\alpha$, where $\varphi_0$ is the effective threshold, the value at which the stimulus first becomes effective. The sensation called loudness grows as a power function of sound pressure. The exponent is 0.6, a value that has been recommended by the International Organization for Standardization.

Why do we call it the psychophysical law? Mainly because of its history. In the year 1850 a physicist by the name of Fechner was lying in bed one morning wondering how to connect the inner world of sensory magnitudes with the outer world of stimulus energy. He was thinking about Weber's experiments, which showed how an increment to a stimulus becomes just noticeable when it reaches some fixed percentage of the original stimulus. In acoustics we know Weber's rule as an approximation: the just noticeable difference is approximately a decibel. Fechner seized upon that relation to solve his problem. He decided that he could set up a simple differential equation based on the assumption:

$$\Delta \psi = k \frac{\Delta \varphi}{\varphi}$$

By integrating that equation, Fechner concluded that sensation $\psi$ grows as the logarithm of the stimulus $\varphi$. The psychophysical law, said Fechner, is a logarithmic function.

The logarithmic form of the psychophysical law is still believed in by many people. It is still expounded in many textbooks that deal with sensation, but change is taking place. If the logarithmic law is no longer as highly regarded as it once was, the credit belongs in large measure to the science of acoustics.

Why was it that Fechner's logarithmic law met its first defeat in
acoustics? As I see it, acoustics had two advantages. The first was that fifty years ago acoustics was a backward science. Before the development of the vacuum tube, sound was so difficult to produce and control that a quantitative science was hardly possible. The second advantage was that those who made use of the new electronic tools in the 1930's did not have a long-established tradition to impede them. In order to measure sound, they invented a logarithmic scale called decibels. Next they asked an obvious question: is loudness proportional to decibels, as Fechner's law says it ought to be? The answer appeared to be no. Simple listening tests made it clear that equal decibel differences did not mark off equal differences in loudness. A sound of 50 dB did not seem half as loud as a sound of 100 dB. Actually, instead of sounding half as loud as 100 dB, a sound of 50 dB sounds only about 3 per cent as loud.

Direct experiments were started, beginning about 1930, to determine how the perception of loudness grows with increasing sound pressure. Some of those early experiments were sponsored by commercial companies, because there was a practical engineering need to understand how loudness grows with sound pressure level. Taking into account all the measurements available up to 1937, Harvey Fletcher proposed a loudness function that was concave downward over much of its range (in log-log coordinates). That early loudness function served many useful purposes. In 1936 I had proposed the term some for the unit of loudness, and the loudness function was sometimes called the sone scale.

The early loudness function did not go unchallenged, of course, for it represented a rather radical development. Fletcher's curve was the first empirically based sensation function that commanded wide agreement. Its adherents claimed that it measured the strength of an experience. But many people believed, and some still do, that such a measurement is impossible. I do not want to try to resolve the philosophic argument about the measurability of sensation, but I would remark that we can at least measure what a person says about his experience. For example, we can measure the point at which a person says that one sound seems half as loud as another sound. That method, called fractionation, was a popular procedure in the 1930's and 1940's.

The Power Law

Beginning in 1953, I tried out some other methods for measuring loudness. Since one of the methods, called bisection, showed a curious hysteresis effect, I also tried the method of bisection with visual brightness, where I obtained the same hysteresis effect. That was an interesting outcome, but what seemed even more exciting was the close resemblance between light and sound when the results obtained with several different methods were compared. The method of magnitude estimation, in which the observer tries to assign numbers proportional to the apparent strength of his sensation, gave especially clear results -- results indicating that in both vision and hearing the sensation grows according to a power function of the stimulus intensity. A brief announcement of this psychophysical power law appeared in Science (Vol. 118, 1953, p. 576). I then set about to
find out how general the law really is. Did it hold only for vision and hearing?

By now we have explored about three dozen sensory and perceptual continua. It appears that the same psychophysical power law holds in all sense modalities, so that loudness is by no means a special case. In general, each sense modality has its own exponent. The investigation of other sense modalities has led to the development of another useful method for sensory measurement, called cross-modality matching. If loudness and brightness are both power functions of intensity, what happens if an observer is asked to adjust the level of a sound to make it seem as strong as the brightness of a given light? By matching loudness to brightness for a wide range of intensities, we should obtain a matching function that is also a power function. Not only that, but the exponent of the matching function should be given by the ratio of the exponents for loudness and brightness. Matching experiments between loudness and brightness have been carried out by my former colleague J. C. Stevens who showed, by several procedures, that the predicted results can be readily obtained.

We therefore have a powerful method for validating the loudness scale. As a matter of fact, by means of cross-modality equations, loudness has been matched to at least ten other sensory continua. Some of the results are shown in Fig. 1.

![Diagram](image_url)

Fig. 1 Equal-sensation functions obtained by matches between loudness and criterion stimuli on various other sensory continua. The intercepts of the functions are arbitrary, but the slopes are those determined by the data.

There it can be seen that all the matching functions are power functions (straight lines in log-log coordinates). The slopes of the lines, which determine the matching exponents, accord quite well with the predictions.
Psychophysics and the Measurement of Loudness

based on the exponents of the various functions as determined in separate experiments. The over-all consistency of the total array of matching functions leaves the impression that the question posed in 1953 has been satisfactorily answered, and that the general psychophysical law follows the power-function form.

Significance

What is the significance of the power-function form of the psychophysical law? Although the full significance is far from known, it is interesting to note that the power law can be summed up as a simple invariance: equal stimulus ratios produce equal sensation ratios. This principle of ratio invariance seems to apply to all the sensory systems. An analogous principle holds for many of the functional relations in physics. Indeed, physical laws are often power laws.

Among physical variables, however, the exponents are usually integers or simple fractions, whereas the measured values of sensory exponents show no such simplicity. Some representative exponents are:

- brightness 0.33
- smell (heptane) 0.6
- taste (sucrose) 1.3
- cold (on arm) 1.0
- warm (on arm) 1.5
- vocal effort 1.1
- 60-Hz vibration on finger 1.0
- 250-Hz vibration on finger 0.6
- force of handgrip 1.7
- electric current through finger 3.5

For the origin of the ratio invariance in the response of a sensory system, we must look to the evolutionary history of the organism. In perceiving and reacting to the world, it is advantageous to an animal if the perceived relations among stimuli do not depend on the absolute magnitudes of the stimuli. Fortunately, relations tend toward constancy. Thus the sides of a triangle appear to maintain their proportions when the triangle is moved away from the viewer. Similarly, the ratio between the light and the shaded parts of a photograph seem approximately the same in bright light and in dim light. Relations among the sounds of speech remain much the same whether the speech is soft or highly amplified. The usefulness of these constant relations under widely varying stimulus levels is obvious -- so obvious that we often take them for granted and fail to note their significance. The psychophysical power law makes it possible for us to perceive more or less correctly the rich patterning of the environment despite the enormous ranges of stimulus energy to which we are subjected.

Those sense modalities that are subjected to the widest ranges of energy tend to have low exponents, but that rule does not appear to hold in
Psychophysics and the Measurement of Loudness

reverse. For example, the sense of smell has a fairly low exponent despite the fact that the effective range of odor concentrations is rather limited. When the range of stimulus energy extends over more than a billionfold, as it does in vision and hearing, there is an obvious need for the low exponents that nature has provided. The transducer nonlinearity, evidenced by a low exponent, seems to be nature's way of matching the input from the outside world to the needs of the central nervous system. For example, a billionfold change in sound energy becomes a thousandfold change in apparent loudness. It seems not unreasonable that the nervous system should be able to process changes of a thousandfold, whereas the neural processing of changes exceeding a billionfold would seem beyond possibility. With the help of the nonlinearity of the transducer system it may be possible for the nervous system to perform its role by making only linear transformations.

Calculation of Loudness

Once it had been determined how loudness, in sones, depends on the sound pressure level of a 1000-Hz tone, the question arose whether an equally simple rule might govern the dependence of loudness on the bandwidth of a sound. By combining octave bands of white noise, all bands having been made nearly equally loud to begin with, it was found that the total loudness equals the loudness of the loudest band plus 0.3 times the loudness of all the other bands. This simple summation rule led finally to the development of a practical procedure for using band-level measurements in order to calculate the loudness level of a noise. It is one of the procedures now recommended by the International Organization for Standardization. The same summation rule is also used in the calculation procedure for what is called the perceived noise level in PNdB.

The basic principle that emerges here is that two equally loud bands of noise do not sound twice as loud as a single band heard alone. Instead of a factor of 2, the addition of one octave band to another contiguous octave band increases the total loudness by a factor of only 1.3. Thus the second band seems to add less than a third of its loudness to the combined total. The reason, presumably, is the mutual inhibition, or partial masking of one band by the other.

Loudness under Inhibition

Some of our recent studies at Harvard have been designed to explore further the effects of inhibition on the loudness of sounds. This, of course, is not a new topic, for systematic studies of masking were made several decades ago. It was noted by Steinberg and Gardner that a tone in a white noise sounds less loud than a tone in the quiet, and they proposed a simple linear model to describe the partial inhibition produced by the masking noise. They suggested that the inhibitory effect could be described as the subtraction of a constant loudness from the tone. As further evidence accumulated, however, it became clear that a linear model is not adequate, and that the transformation produced by inhibition belongs to the power group.
Psychophysics and the Measurement of Loudness

of transformations. In other words, the effect of inhibition is to change the exponent of the power function. The exponent appears to increase in proportion to the level of the masking noise.

The same effect occurs in both vision and hearing: a stronger light or a stronger sound depresses the apparent level of a weaker stimulus. The inhibitory effect is asymmetrical in the sense that stronger stimuli inhibit weaker stimuli, but not vice versa. The one-way operation of inhibition can be demonstrated more sharply with visual than with auditory stimuli, because visual stimuli can be separated spatially on the retina, whereas auditory stimuli may overlap one another on the basilar membrane. Nevertheless, power functions whose exponents are made larger by inhibition can be demonstrated in hearing as well as in vision.

![1000 Hz tone in quiet adjusted](diagram)

Fig. 2 Masked loudness functions obtained by adjusting an unmasked tone to match the loudness of a tone inhibited by white noise. The slopes (exponents) were determined by the 30-dB model. The steep part of the recruitment function extends 30 dB above the effective threshold, and the slope increases in proportion to the level of the masking noise.

Figure 2 shows the results of an experiment in which the observers adjusted a tone in the quiet to match the loudness of a tone presented in various levels of white masking noise. The family of slopes (exponents) fitted to the experimental points in Fig. 2 is a family determined by what I have called the 30-dB model. It says that the rapid growth of loudness takes place over a range extending 30 dB above the effective threshold, and that the slope (exponent) of the steepened function grows linearly with the level of the masking noise.

Recruitment

The steep functions in Fig. 2 are often called recruitment curves. Ever since Fowler introduced the concept of recruitment, some 30 years ago, the assumption has been that loudness in a recruiting ear grows rapidly at first and then more slowly as the stimulus increases. Figure 2
suggests, however, that the recruitment produced by masking may result in a function that is a straight line in log-log coordinates. It is not the curved line commonly assumed.

A similar family of straight lines has been observed in the recruitment of some kinds of hearing loss, especially those that result from a widespread cochlear involvement, such as the diffuse degenerative patterns observed in the cochlea following Ménière's disease. The data from hard-of-hearing patients are usually quite variable and the resulting functions are less precise than those shown in Fig. 2. Nevertheless, when Hallpike and Hood examined the recruitment functions for 200 cases of unilateral deafness due to Ménière's disease, they found, as they said, that "all but a very few took the form of a straight line..." What is most interesting in the present context is the fact that the slopes of the recruitment lines tended to follow the pattern predicted by the 30-dB model. On the average, the larger the hearing loss, the steeper the recruitment line, and the loudness in the afflicted ear reached the normal level when the stimulus was about 30 dB above threshold. It appears, therefore, that both Ménière's disease and the inhibitory process produced by masking may alter the exponent of the power function in approximately the same way.

Summary

The new psychophysical law, which replaces Fechner's logarithmic law, states that sensation is related to its stimulus by a power function. Thus in acoustics we find that loudness increases as the 0.6 power of sound pressure, a relation that has now been verified by cross-modality matching experiments. Cross-modality matches have been made between loudness and ten other perceptual continua. The psychophysical power law, which holds in all sense modalities, provides us with a much-needed constancy among perceived ratio relations. In addition, in vision and hearing, where the exponents are less than one, it provides the nonlinearity that is needed for the sensory systems to respond to energy ranges that may exceed a billionfold.

Loudness may be inhibited by a masking stimulus. Mutual inhibition among adjacent noise bands has proved sufficiently regular to make possible a simple procedure for calculating the loudness of complex noises.

When a 1000-Hz tone is inhibited by a wide band of masking noise there occurs an increase in the exponent of the power function. This increased slope of the loudness function is often called recruitment. The recruitment produced by a wide-band masking noise resembles that produced by some kinds of hearing afflictions, such as Ménière's disease. In both cases, the part of the loudness function that is made steeper (in logarithmic coordinates) extends 30 dB above the affected threshold, regardless of the location of the threshold.
References

The following papers are relevant to the foregoing review.


Stevens, S. S. A scale for the measurement of a psychological magnitude, loudness. Psychol. Rev. 43, 405-416 (1936).


COMPUTERS IN ACOUSTICS: 
SYMBIOSIS OF AN OLD SCIENCE AND A NEW TOOL

Manfred R. Schroeder
Director, Acoustics, Speech and Mechanics Research
Bell Telephone Laboratories
Murray Hill, New Jersey, U.S.A.

ABSTRACT

Digital computers are playing an increasing role in acoustics. Their superior power for analysis, modeling and the control of experiments has stimulated much new research. In architectural acoustics and underwater sound, new insights have been gained by digital simulation of wave and ray propagation phenomena. In psychoacoustics and hearing, computer modeling of the mechanical and neural activities of the ear has led to a deeper understanding of auditory perception. In speech research, new concepts of speech production are evaluated by human listeners listening to computer-made speech. Modern digital computers allow the realization of signal processing schemes of great complexity and sophistication – from the coding of speech signals to the simulation and measurement of the acoustic properties of concert halls. Powerful new methods for analyzing and displaying all kinds of physical and subjective data have become feasible and are having a profound influence on our thinking.

But computer technology has also received strong stimulation from acoustics. On-line computation and the use of small “dedicated” computers as research tools was largely pioneered by acoustic researchers. Current research activity in acoustics promises computers capable of understanding spoken instructions, recognizing individual voices, and answering questions in a human-sounding tongue.

—GP-71—
Introduction

During the past decade, some of the most significant advances in acoustics have become possible through digital computers. Many branches of acoustics have benefited: speech, music, physiological and psychological acoustics, architectural acoustics, underwater sound and others.

Acoustics, like other scientific disciplines, has made wide use of computers to perform standard calculations of all kinds. But beyond such routine uses, acoustics has exploited computers in many unforeseen ways, such as the simulation of complex speech processing systems, sound-propagation in concert halls and neural behavior. Other new and fruitful applications include the use of computers in "on-line" fashion as experimental tools in physiological measurements, speech synthesis and psychometric experiments.

The emphasis in this summary paper will be on those applications in which acoustics has pioneered unusual, nonstandard, applications of computers. Illustrative examples are drawn from the following research activities:

- Speech analysis and processing,
- Speech synthesis by rule,
- Automatic speech and speaker recognition,
- Determination of articulatory parameters from acoustic measurements,
- Analysis and synthesis of musical sounds,
- Processing of physiological data and study of neural response patterns by simulation,
- Generation of test stimuli for psychoacoustic experiments,
- Evaluation of subjective quality and preference by multidimensional scaling,
- Simulation of concert hall acoustics,
- Signal generation for acoustic measurements in auditoriums and processing of recorded sound pressure responses,
- Studies of acoustic ray propagation in reverberant enclosures and underwater,
- Monte Carlo simulation of steady-state frequency and space responses in multipath propagation media,
- Determination of the geometrical shape of unknown objects from their acoustic diffraction patterns (The "inverse diffraction" problem).

Speech Processing by Digital Simulation

One of the earliest applications of computers in acoustics occurred in speech processing. Vocoder devices and other speech processing devices had
become so complex that it often took months or even years to implement a new device. When the implementation was finally completed and the result unsatisfactory (as was frequently the case), the question arose "What does the result mean? Is the original idea bad? Or are there, often numerous, approximations in its realization responsible?" Many new and ingenious ideas could not be put to the test because instrumentation would have been too time consuming.

All this was changed by the emergence of large digital computers and the introduction of digital simulation [1]. The device to be simulated if it was not already in digital form, was first represented by a sampled-data equivalent amenable to digital simulation. The bridge between the real world of analog signals and the computer realm of numbers was established by analog-to-digital and digital-to-analog converters which translate analog sample values into binary representations and vice versa [2]. Early programs had to be written in machine language or available assemblers and compilers, but a special block-diagram compiler (BLÖDI" [3]) soon facilitated the task of programming digital simulation of problems representable by sampled-data block diagrams [4,5]. Engineers could learn the use of the BLÖDI-compiler within a few hours. Considerable advances have since been made in speech processing for a variety of practical applications [6]. Spectrum channel vocoders and formant vocoders [7], both pitch and voice-excited [8], were substantially improved and new ideas for syllabic-rate changers [9,10] and the restoration of "helium-speech" of deep-sea divers [11] were successfully tested. New principles for speech processing, based on analytic signal representation [12] and linear predictability [13] of speech signals, were formulated and demonstrated by means of digital simulation.

A new and even more powerful version of the block-diagram compiler (BLÖDI-B) allowed still more complex simulations [14].

The impressive applications of digital simulation in speech processing, aided by block-diagram compilers, soon aroused interest in allied fields, particularly Speech Analysis.

Digital simulation allowed the realization of a long-standing desire of speech researchers: pitch-synchronous (period-by-period) spectrum analysis of speech [15,16]. This led to improved methods for inverse filtering to recover the excitation waveform [17] and to track formant frequencies [18]. Notable gains were made in the extraction of formant frequencies [19] and in the measurement of fundamental frequency by cepstrum [20,21], period histogram [22,23] and other methods [24-27].

—GP-73—
Perceptually significant aperiodicities were found in voiced speech segments [28] by high-resolution spectrograms obtained on the computer [29]. The concept of "spectrum lifting" facilitated spectrum envelope measurement [30,31] and the Fast-Fourier Transform (FFT) [32,33] allowed spectral analysis of high resolution in near real time. Special plotting programs [34] made possible the rendition of half-tone pictures on microfilm plotters, thereby allowing sound spectrograms to be both efficiently computed and plotted by digital machines. New methods for 3-dimensional displays [35] of speech spectra with continuously adjustable resolutions in time and frequency may lead to refined insights into the structure of speech signals and their spectra.

Interesting results were also obtained by computer analysis of the sounds of musical instruments such as the trumpet and the violin [36].

* * *

Although notable successes in automatic word recognition have been achieved with analog devices based on mechanical and optical principles, future progress in Automatic Speech and Speaker Recognition will strongly depend upon computers. Various methods for automatic word recognition have been proposed and tested by digital simulation [37-43]. Automatic voice classification and speaker identification, of great practical potential in future banking operations and limited-access information retrieval, is being explored using spectral data of isolated words [44] and fundamental-frequency contours of connected utterances [27].

In all these studies, the reported "correct" scores are high, but the inventories are still limited to relatively few words and speakers. However, even with such limitations, useful applications can be foreseen in such diverse activities as postal sorting, spacecraft maneuvering [45] and surgery [46].

New methods of analyzing speech spectra (analysis-by-synthesis) [47] combined with deeper insights into the relations between vocal tract shape and acoustic parameters led to Articulatory Analysis Based on Acoustic Measurements.

The formant frequencies of a wide variety of vocal-tract shapes were computed by iterative solution of Webster's Horn Equation and analyzed in terms of articulatory parameters [48]. The "inverse" problem of determining (spatially "bandlimited") vocal tract shapes from formant frequencies [49,50] and lip-impedance functions [51] was also solved. The crucial uniqueness problem ("Is there more than one (bandlimited) shape that could have given rise to the measured formants or impedance singularities?"), for which no mathematical proof is available, was answered
empirically by a near exhaustive, and happily unsuccessful, computer search for multiple solutions [52]. A procedure for calculating vocal tract perturbation functions which, to first order, affect one, and only one, formant frequency for arbitrary vocal-tract shapes was described [53].

While these advances in speech analysis were still in progress, speech synthesis by rule was already attempted by the more ambitious computer types. Speech synthesis by rule [54] (the "talking computer") has been one of the practical goals of speech research for some time. Once it had been shown that natural-sounding speech can indeed be synthesized by electrical networks [55,56] or simulations of such networks, the way to ultimate success seemed open. But it became soon apparent that electromechanical analogies and analog-digital equivalences were only one of many prerequisites for successful automatic speech synthesis. An early speech synthesizer [57] included simple rules for formant transitions and synthesized speech with simulated analogs of the vocal tract. Intelligibility was perfect (if one knew beforehand what was said) and signal quality was amusing (especially in the "singing" mode). Other attempts [58-61] at speech synthesis by rule, using off-line digital computers, achieved various degrees of success.

A quantitative model for coarticulation effects of the articulatory activities of the tongue was developed using a computer-simulated vocal tract [62]. Control rules for vocal-tract shapes and prosodic features for speech synthesis were investigated [63].

Of great importance in the history of speech research was the emergence of a new laboratory tool: the small "dedicated" computer for on-line experimentation [48,64]. The small computer (which has since grown considerably but is still dedicated) allows the manipulation of synthetic parameters and linguistic rule and an immediate subjective evaluation of their auditory effects. The experienced experimenter, using typewriter and graphic input devices to modify the program according to his auditory evaluation, and the computer, with speech synthesizer [65] and oscilloscopes attached, form a tightly coupled loop [66-68]. This most advanced (and expensive) version of trial-and-error methods is uniquely fitted to surmount many of the remaining obstacles on the way to better synthetic speech.

Another early on-line application of note is the "digital spectrum manipulator" [69] which allows the experimenter to generate and examine real-time spectrograms on a cathode-ray tube, make changes in the spectrum and see and hear both the original and modified signals.
A particularly promising approach is speech synthesis by on-line computer based on an articulatory model [70] of the speech process. The computer displays (simplified) outlines of the tongue and lips as well as the corresponding area function and formant frequencies. Articulatory "target" shapes and rules for transition are entered into the program by the experimenter. The almost immediately available acoustic output signal allows rapid subjective evaluation and continual refinement of the rules and the model itself.

One of the advantages of a small owned (or rented) computer which has been paid for (or is being paid for at a fixed rate) is the possibility of processing large volumes of data, a frequent requirement in speech research, at low cost. The saving can be as much as a factor of 10 compared to large central computers whose extensive arsenal of hardware and software is often not needed but still has to be paid for. Thus, once "free" access to a small machine is obtained, many problems can be run more economically on an existing on-line facility. A realistic model of the human vocal cords [71] and their interaction with the vocal tract was thus successfully tested and refined on a small Honeywell DDP-516 computer using a powerful General Electric GE-645 only for compiling and debugging the (Fortran IV) program. The same basic program was used to imitate the sound of a Flügelhorn [72] in which the vibrating lips of the trumpet player play the role of the vocal cords and the horn that of a (fixed-shape) vocal tract.

* * *

Programs and compilers [73] for more Avant-Garde Sounds

(music? [74]) permitting the simultaneous synthesis of many "instruments" [75] were adapted to small on-line computers to permit the electronic composer to hear and modify the results of his inspiration almost instantly [76], much like the conventional composer behind his analogue equipment (piano). With wide availability of this new tool to exercise people's acoustic imagination and perception, recreational composing may come to occupy a significant portion of future generations' leisure time which, no doubt, will be plentiful as a result of increasing automation.

* * *

The success of on-line computing in speech research and sound synthesis led to the use of small computers for real-time experimentation in Psychoacoustics.

As is well known, the collection of psychometric data is a time-consuming occupation, to say nothing of the difficulties due to boredom-induced shifting of perceptual thresholds and time-varying decision

--- GP-76 ---
criteria. If each new stimulus or stimulus pair (or triplet) could be chosen in such a way that it extracts a maximum amount of information from the subject, a considerable speed-up in acquisition of psychometric data could be achieved. This is precisely what an on-line digital computer allows one to do. In determining, for example, the 50%-point of a psychometric function, the computer can select that stimulus for presentation which is nearest to the expected mean as computed from past responses. The tracking of time-varying means [77] is likewise facilitated. Other parameters of psychometric functions can be similarly estimated [78].

In some tests, the computer can generate the test stimuli in real time. In others, the stimuli have to be prerecorded [79]. In either case, signals and noises of high reproducibility can be obtained [80]. With an on-line computer, it is not even necessary to prerecord all possible stimulus combinations. It suffices to record each test stimulus only once (or twice). A computer-controlled tape deck goes to the proper position on the stimulus tape for the presentation of the next stimulus. The computer knows the position of the tape from a shaft-encoder coupled to the tape transport and a timing signal on a second tape track. This system has been successfully employed in a study of speech quality resulting from replacing the vocal excitation function by simplified waveforms and constraining the variability of successive fundamental periods [81].

One of the most remarkable advances made possible by computers is Multi-Dimensional Scaling of perceptual and other subjective data [82,83]. These scaling methods allow either linear or arbitrary monotonic transformation of the subject's data which may be in the form of subjective preferences or similarity judgments. Geometrical representations of the transformed data are calculated by the computer and the so constructed "subjective spaces" are plotted, two dimensions at a time, by microfilm plotters or oscilloscopes. On-line computers in conjunction with 3-dimensional display facilities allow the simultaneous representation of more than 2 dimensions and rotation of the display for more effective evaluation [84].

The perceptual aspects of a variety of acoustic stimuli including vowel sounds, tone-ringer signals [85] and speakerphone systems [86] have been explored by this powerful approach. In many cases a clear correlation between the physical parameters of the signal and its perceptual attributes could be established - a correlation that was obscure in the untransformed data.

--- GP-77 ---
On-line computation has also been brought to bear on Physiological Acoustics.

In fact, the average response computer [87], used for accumulating neural response patterns of cats [88], was probably the first instance of on-line computing with a computer as an experimental tool. The response of single neurons and neural nets [89,90] has been simulated off-line on digital computers. A model for the first-order auditory fibers has been shown [91] to match closely the results of physiological measurements on cats [92].

These and other results have demonstrated that computer methods are ideally suited to study neural behavior and to analyze complex neurophysiological data. Of particular import for speech production (and perhaps perception) are the investigations using electromyographic signals from various organs participating in the articulatory process [93].

Our understanding of the peripheral hearing process was significantly advanced by a mathematical model of the basilar membrane [94,95]. Digital simulations of the model for a variety of acoustic stimuli were instrumental in interpreting psychoacoustic experiments on monaural and binaural hearing. As an ultimate tutorial refinement; 3-dimensional motion pictures of waves traveling along the basilar membrane were made on the computer [96].

---

Another complex problem which calls for model studies by simulation is the propagation of sound waves in reverberant enclosures. With the advent of large digital computers, classic scale-modeling techniques, using light rays and ultrasonic waves, were supplemented by Digital Simulation of Concert Hall Acoustics.

This work originated with the digital simulation of "allpass" filters and "comb" filters to create natural-sounding artificial reverberation [97] and pseudo-stereophonic effects [98] without spectral distortion of the audio signal. This approach to adding 3-dimensional realism and an illusion of space has been exploited to enhance computer-generated electronic sounds [99]. Since the computer-simulated acoustics can be changed easily by program, illusions of auditoriums rapidly varying in size and reverberation characteristics are feasible.

Another astonishing percept made possible by computer processing of auditory signals is a kind of "super-stereophony." Preprocessed sound signals, emanating from two closely-spaced (±22.5°) loudspeakers in front of a listener, combine at his ears in a manner that closely resembles the sound pressure waves from a sound source at 90° (or some other, arbitrary, angle) off to the side [100]. The effect is astounding, particularly when the listener turns his head in an attempt to locate the nonexistent...

---

GP-78
third sound source - at which point the unique perception vanishes because the computer program is based on the measured sound diffraction of a listener's head oriented in the forward direction.

The possibility of digitally simulating realistic reverberation processes, together with the ability to add discrete arrivals (reflections off the stage, sidewalls and ceiling) with proper spectrum and directionality, allows sound transmission in concert halls and auditoriums to be simulated on the computer [101]. Reverberation-free speech or music signals are fed into the computer which is programmed to add echoes and reverberation. The computer produces two or more audio output signals. These are radiated from loudspeakers in an anechoic chamber to create, at a listener's ears, sound fields which perceptually resemble those in an actual hall. This method is the digital equivalent of creating complex sound fields by means of electronic analog devices (delay lines and reverberators) and a multiplicity of loudspeakers suspended in an anechoic chamber [102].

The two most important applications of this simulation method are:

(1) auditory evaluation of new concert hall designs before construction and
(2) psychoacoustic studies of auditory perception in reverberant spaces. In the second application, the method can be used for instantaneous comparison of different reverberation processes acting on identical sound signals. The two reverberation processes can differ in any of a number of physical parameters (spectral composition of the direct sound, delay and direction of an early echo, direct-to-reverberant energy ratio, reverberation time as a function of frequency, shape of the decay curve, etc.). Needless to say, the degree of flexibility of computer simulation is impossible to achieve by listening in real concert halls. Owing to digital simulation, a better understanding of the complex relationship between the physical parameters of a reverberation process and its perceptual attributes has already been attained in a few cases.

---

One of the important applications of digital simulation of reverberation processes has been a study of Subjective Reverberation Time.

A variety of nonexponential reverberation processes were simulated and compared with exponential decays in paired comparison tests [103]. Based on the results of these tests, a new criterion for objective reverberation time (best straight-line fit to the initial portion of the decay) was formulated which matches the subjective reverberation time surprisingly well (see Fig. 1, next page). How misleading the frequently used criterion, based on the decay between 5 dB and 35 dB, can be, is shown.
in Fig. 1 for one extreme case (×) with a subjective reverberation time of about 1.1 sec and an "objective" reverberation exceeding 2.5 sec!

Fig. 1: Ordinate: subjective reverberation times found in subjective comparison tests of computer-made nonexponential reverberation processes with exponential decays. Abscissa: physical reverberation time based on a computed best straight-line fit to the initial 160-msec portion of the decay. The cross (×) shows the "objective" reverberation time for one extreme case based on an older but widely-used definition of reverberation time.

Computers have also quietly revolutionized methods of Acoustic Measurements in Auditoriums.

Such measurements used to be confined to what was possible by available, often inadequate, instruments. But our knowledge about human perception in reverberant sound fields is steadily increasing and it is of utmost importance to measure what (and in a manner that corresponds to how) we perceive - using both appropriate test signals and subjectively meaningful evaluation procedures. Both of these tasks are greatly facilitated by digital computers.

The computer is ideally suited to act as a signal generator producing precisely tailored test signals to reveal acoustic deficiencies both from a physical and a perceptual viewpoint. Such computer-generated test signals (e.g., short tone-bursts with a third-octave bandwidth) are radiated from a loudspeaker on the stage and the resulting sound pressure responses are recorded at selected positions throughout the hall [104]. These responses are fed back into the computer for bandpass filtering (to minimize the influence of ambient noise on the results) and envelope detection. Integrals over selected time and frequency intervals and summations over various spatial positions are computed. The results of these computations are tabulated or plotted by means of a microfilm plotter attached to the computer. Subsidiary programs fit straight lines to decay curves, compute direct-to-reverberant energy ratios, "clarity" [105] and other quantities that are of subjective significance or that will aid in pinpointing physical defects. As new knowledge is accumulated, in part through comparison of different halls evaluated in this manner,
programs can be modified or expanded to incorporate new insights into auditory perception in reverberant sound fields.

Computer-aided measurements in a large concert hall (Philharmonic Hall, New York) [106] revealed a considerable attenuation at low frequencies of the energy reflected from the suspended ceiling ("clouds") and confirmed earlier measurements on a scale model of the ceiling [107]. An unexpected lack of low frequencies in the direct sound has subsequently been explained as a "seat" effect [108,109]. The combination of these two effects led to a perceivable weakness of the low-frequency response on the main floor in Philharmonic Hall and a near inaudibility of the cello during tutti passages [110].

Of particular importance in concert hall acoustics is the accurate Measurement of Reverberation Time.

A new method [111] for obtaining the ensemble average of many decay curves and a correspondingly more accurate value of reverberation time was first applied using the computer method just described. The ensemble average decay is obtained by integrating the squared envelope of a single tone-burst response. The great variability of individual noise decay curves, which result from the randomness of the exciting noise signal, is avoided. As a result, significant double slopes in the decay are clearly revealed by the ensemble average method.

Computers have led to both a more meaningful subjective criterion for reverberation time and a more accurate and reproducible method of measuring it. But computers are also playing a part in refining the Theory of Reverberation and the Dependence of Reverberation Time on Room Shape.

Existing formulas relating reverberation time to surface absorption depend on the total volume of the enclosure and the total surface area and absorption coefficient of the absorber.

Fig. 2: Ordinate: reverberation time according to the "Eyring" formula for 32 different room shapes. Abscissa: reverberation time found by ray tracing on the computer. The discrepancies depend on room shape and exceed 50% in 12 of the 32 cases.
They do not account for the shape of the enclosure and the distribution of the absorber—contrary to experimental evidence. This problem has recently been attacked by ray tracing studies on a digital computer [112]. Figure 2 shows one of the results of this study, a comparison between values of reverberation time according to the widely used "Eyring" formula [113] and reverberation times found by ray tracing on the computer. The discrepancies exceed 50% (!) in 12 out of the 32 different (2-dimensional) room shapes examined. Similarly, large errors were found even for diffuse wall reflections and scattered placement of the absorbers. It appears that, until these discrepancies are more fully understood, the only way to derive a reliable formula for a given hall shape and absorber distribution is by ray tracing on the computer.

An early computer attempt [114] at obtaining values for the "mean-free path" and collision rates of sound rays with different walls was regrettably marred by a mistake in the averaging over the solid angle[115].

Ray tracing and wavefront construction on the computer have also been used to study

**Sound Propagation in the Ocean.**

Slow-motion pictures of sound propagation in the deep ocean have been produced by a microfilm plotter in conjunction with a large computer programmed to repeatedly apply Snells' law of wave refraction [116]. Sensitive signal processing schemes, such as matched filtering [117] and "volume focussing" [118] applied to the outputs of hydrophone arrays have likewise been studied by digital simulation [119], permitting a rapid and reliable evaluation of their potential.

**Other wave propagation phenomenon which have been illuminated by computer simulation are the**

**Random Wave Fields**

resulting from wave interference in multipath propagation media. Some quantities of such wave fields can be computed analytically (for example, the expected density of interference maxima in space or frequency) [120] but others are beyond today's methods of analysis. An example of such a quantity is the expected level of the highest interference maximum in a given frequency range. This quantity, apart from being of interest in certain detection tasks, determines the stability of public address systems in auditoriums [121]. To find this quantity, a computer was programmed to generate frequency response curves of reverberant rooms with appropriate statistical properties. The desired quantity was then determined by the computer using a Monte Carlo method [122]. Many other
quantities pertaining to random wave fields have been calculated by Monte
Carlo methods, such as correlation functions of spatial and frequency re-
sponse curves [123] of random wave fields and the effect of frequency
averaging on the fluctuations of frequency response curves [124].

One of the most recent applications of computers in acoustics is
Image Reconstruction with Sound Waves.

It is well known that the distribution of a field quantity obeying
a linear wave equation inside a specified volume can be reconstructed
from its values on the boundary. Since one or more field quantities
change abruptly at the surface of an object whose impedance differs from
that of the surrounding medium (e.g., a rigid body in air or water, or an
air "bubble" enclosed in a solid or liquid), the shape of the object can
be computed from the (sound) field scattered by it and measured on some
closed surface surrounding the object. This computation is a process not
unlike wavefront reconstruction by holography. In fact, the holographic
process can be viewed as an analog computer for wavefront reconstruction
with one important difference: With sound waves, one can record linear
complex amplitudes of the scattered field, as opposed to merely some non-
linear function of intensity in the case of photographic film. Thus, for
example, in the acoustical case, one can use a wide-spectrum or impulsive
"illumination" of the object - as opposed to the essentially monochromat-
ic illumination in optical holography - thereby adding another dimension
to acoustic image reconstruction. To utilize the phase information
available in the acoustic case, the computer is essential.

Figure 3 illustrates the method of acoustic image reconstruction.
The object is a letter "H" (left). The center of Fig. 3 shows the mag-
nitude of the scattered sound field at some distance from the object. The
right-hand part of Fig. 3 shows the computer-reconstructed H, for the
case of a sound wavelength equal to about 1/6 of the dimensions of the
object [125]. Both the scattered sound field and the reconstructed image

Fig. 3 Acoustic image reconstruction

GP-83
were computed as complex amplitudes. Figure 3 shows only the magnitude rounded to one (black) and zero (white). Potential applications of this method are nondestructive materials testing and geometrical shape mination of inaccessible objects.

* * *

This paper has been about the past and the present of computers in acoustics—almost precisely one decade of growth and evolution. What will the future bring? Nobody can be certain. We have only glimpses of things that are already underway like the proposal to use an on-line computer to control X-ray microbeams to limit radiation in articulatory experiments [126]. There are bound to be many other new and exciting developments.

* * *

REFERENCES

- A comprehensive listing of current speech research activities in Japan can be found in J. Acoust. Soc. Am. 39, 428 (1966).

—GP—84—
COMPUTERS IN ACOUSTICS

[45] Astronauts apparently need to supplement manual controls by voice-controlled maneuvering capabilities at crucial junctures on missions (1967).
[46] Surgeons, during complicated operations, often need additional information which we like to call for from a machine by voice.
[74] H. Scherchen.

—GP-86—
COMPUTERS IN ACOUSTICS

[110] G. Szell (personal communication).
[126] O. Fujimura (work in progress, personal communication).

The Perception of Timbre

J.F. Schouten
Director
Institute for Perception Research
Insulindelaan 2, Eindhoven
The Netherlands

Introduction

When describing a sound as it presents itself to us in auditory perception we are able to characterize it by various perceptual attributes. It may sound loud or weak, high or low, long or short. These attributes: loudness, pitch and duration are the easiest to ascertain in the overall impression of any sound. For all other qualities we have scarcely more at our disposal than the one and all embracing term: timbre. A very vague way of bringing all other unresolved attributes under one general heading.

This is an extremely disappointing state of affairs. The tone of the violin, for instance, may sound loud, high and long. But these easily perceptible attributes are the very ones which are the least invariant with respect to the particular sound of a violin. The vague heading "timbre", though, is precisely the one which covers those invariant acoustic properties which make us recognize the violin.

In this lecture I shall try and arrive at a more detailed analysis of this elusive concept of timbre.

Perception and Identification

Before doing so, let me start by demonstrating a general property of human perception. Listen to the following sounds (demonstration). What is your immediate response?

Without any conscious analysis you recognized: a violin - a crying baby - a piano - footsteps in gravel - the rolling thunder -
a street car - the cuckoo amongst other birds.

In terms of immediate recognition the sounds provided you with a sort of overall image, in particular a visual one, of the objective sources causing those particular sounds. This indeed is the essential function of the totality of our sensory and perceptual equipment.

The physical world around us provides us with information through our sense organs. Our brain, through life long experience, interprets this information in terms of that physical world around us. So, we live in our very own world of perception, in a world of our making, which, by accumulated testing of our reactions, gradually evolves into an almost unbelievably consistent image of that physical outer world. Indeed you were correct: it was a violin first, then a crying baby etc. You recognized right. That is, you guessed right, because you heard the sounds before and learned to guess right. This goes to show that our auditory perception is cued to the recognition of familiar sounds.

The following fragment of music will, most probably, convey much more meaning to the inhabitants of these islands than to us, foreigners. (Fig. 1) Indeed it is the sound of the "Biwa", the Japanese lute (demonstration).

---

Fig. 1: Spectrogram of the Sound of the Biwa (Japanese Lute)
The next instance will probably not convey any direct meaning to any of you (demonstration). In fact what you have just heard was the rumbling of human intestines (Fig. 2).

Now listen to this sound (demonstration). You will recognize a large bell and perhaps visualize its shape and size and even the way in which it is being struck. But let me play the record faster and faster. The bell will seem smaller and smaller and finally you will discover that the sound you heard initially was just a very slow playback of the recording of this tiny table bell (demonstration).

When listening to this piece of music (demonstration) you will guess that it was produced by an organ, but I played you another elementary acoustic trick by inverting the direction of playback. (Fig. 3) Hearing it the right way round it is easy to recognize that the original sound stemmed from a piano (demonstration).

Let us now leave the cognitive aspects of overall auditory perception and turn to our powers of perceptual analysis. That is, not to identify a sound source but rather to cue our hearing system so as to analyse the various perceptual attributes, in particular timbre.
Fig. 3: The Piano sounds like an Organ
Above: Reversed playback (organlike)
Below: Straight playback (piano)

Five major parameters of timbre

In most textbooks timbre is defined as the overtone structure
or the envelope of the spectrum of the physical sound. This definition
is hopelessly insufficient, as I hope to prove by demonstrating
that timbre can be expressed in terms of at least five major para-
eters which I shall now try to unravel.

First let us state that we can roughly distinguish between
tonal sounds and noiselike sounds. Subjectively the former have a
well defined pitch and the latter have not.

Acoustically there is a roughly corresponding scale ranging
from perfect periodicity through imperfect periodicity to random
noise. Compare singing to hissing.

To cover all noise under one heading "random noise" is a great
oversimplification. The variety of random noises is rather large
depending upon the particular form of randomizing. Let me give you
one typical example.

When electrical engineers speak about noise they mean one partic-
ular type of noise which can be imitated by random pulses occurring
at a sufficiently fast rate. The noise is "white", that is the spectrum is uniform, if the pulses are sufficiently narrow. The noise is "coloured" if the spectral envelope is non uniform.

Let us produce these fast narrow random pulses. If now we decrease the rate of the pulses we observe that, although the acoustic specification of the noise remains white, the sound impression changes drastically (demonstration) from the original hiss through rustling and fizzling to the random ticking, such as we obtain e.g. from a Geiger counter. So, even normally randomized pulses are not only characterized by their spectral envelope but also by the rate at which they appear. We term this noise "rustle noise" and characterize it by the rustle time $\Theta$, the average interval between the pulses. When listening to rustle noise filtered by bandpass filters (demonstration) it is not difficult to hear similarity with familiar sounds, for which, moreover, we have a great many onomatopoeic verbs at our disposal (Fig. 4).

So our first parameter of timbre is given by its location in the range between tonal and noiselike character.
Fig. 5: Spectrogram of the diphthong "I"
Notice the formant glides (hills on the right) 
and the intonation (ridge on the left)

The second parameter is the familiar spectral envelope. Tonal phonemes, like e.g. the vowels, sung with the same loudness, at equal pitch and with the same duration are strongly dissimilar due to different formants, meaning the overtone structure. Similarly the noise-like phonemes (the fricatives). Umpteen examples can be taken from a large variety of sound sources.

Whereas the above two parameters are sufficient to describe stationary sounds they do not suffice for the description of dynamic properties inherent in most, if not all, natural sounds.

This leads to the third parameter of timbre: the time envelope. Any sound, when analysed physically, has a rise time, a duration and a decay time. The piano tone, in being produced by striking, has a very short rise time and a long decay time. The organ sound has a gradual rise. These properties made you interpret the reversed piano music as a sound resembling the organ.

Here we come up against the following curious phenomenon. Our analysing powers of auditory perception are insufficient to follow the rise, duration and decay, and yet our hearing mechanism is of
sufficient accuracy to distinguish one form of time envelope from another.

Interference with the time envelope, e.g. by introducing abrupt onsets make us hear a click preceding a pure tone and a /t/, /p/ and /k/ before an /s/, /f/ and /χ/ respectively.

Thus the /t/, /p/ and /k/ manifest themselves as separate percepts, that is as separate perceptual entities, whilst acoustically they are no more than the sudden onsets of the /s/, /f/ and /χ/ respectively.

We now come to the fourth major parameter of timbre being that of change. Many sounds, be it in nature, in music or in speech, can not be described sufficiently as a source of given tonality tailored by a spectral and a time envelope.

First, the spectral envelope may change drastically during the production of the sound. In speech we call the particular vowel a diphthong, like the English word "I" (demonstration). This is called a formant glide (Fig. 5).

Secondly, the fundamental frequency may change and affect timbre. We know that in the intonation of speech the pitch changes constantly "from one vowel to another" (_______ or _______). But within each vowel, and within a great many tonal sounds for that matter, acoustical analysis shows the phenomenon of micro-intonation viz. a slight up and down movement in fundamental frequency (demonstration). Again we fail to perceive this micro-intonation in terms of a pitch change and yet its introduction in synthetic speech shows a marked improvement in naturalness. This micro-intonation is also a characteristic feature in the recognition of the Biwa.

The fifth and last major parameter of timbre is the acoustic prefix. A great many sounds start off with a sound quite different from the ensuing tone body. This is particularly pronounced in organ pipes where the prefix may either consist of just noise, or of overtones which either appear earlier than the others or disappear at all after some 50 ms (Fig. 6).

Again our perceptual analysis is too slow to follow these changes and yet the onset or prefix produces a marked difference in timbre (demonstration). Similar observations may be made with regard to the acoustic suffix.
Fig. 6: Sonagrams of three organ pipes
Notice the acoustic prefix, the short initial sound, different from the ensuing steady vibration

Summing up, the elusive attributes of timbre can be considered to be determined by at least five major acoustic parameters:

1. The range between tonal and noiselike character
2. The spectral envelope
3. The time envelope in terms of rise, duration and decay
4. The changes both of spectral envelope (formant glide) and fundamental frequency (micro-intonation)
5. The prefix, an onset of a sound quite dissimilar to the ensuing lasting vibration.

I drew your attention to the fact that our overall perception leads to an immediate identification of the physical object producing a particular sound, without the necessity of conscious subjective analysis. Evidently we should never be able to perform these subtle distinctions, unless the peripheral auditory mechanism provides us with the required power of analysis.

By cueing our ear some of these distinctions can easily be perceived, such as loudness, pitch and total duration.

In timbre we find acoustic parameters some of which, like the tonal or noiselike character, can be perceived consciously with
Fig. 7: Sonagrams of the Cuckoo and of the Cork


Comparative ease. Others, though, like the minute changes in spectral and time envelope, and in particular the acoustic prefix, may prove hard to bring to conscious analytical perception.

The cuckoo and the cork

Let me finish by giving you two examples from everyday life to show you how poorly we use our powers of perceptual analysis.

The first is the sound of the cuckoo normally described as two tones, the second being a minor or a major third below the first.

One day, in the country, when listening to repeated cuckoo cries I heard to my astonishment (and immense deception as a scientist who believes himself to be a keen observer) that the first tone of the cuckoo seems to have a definite upward movement in pitch (Fig. 7).

The spectral recording bears out that there is even more to it than this, in that the frequency goes up and down, a typical case of micro-intonation, whereas the frequency of the second tone is constant throughout (demonstration).

My second example, I am ashamed to admit, I learned from a parrot. We all know how to imitate the popping of a cork pulled from
a bottle (demonstration).

The parrot of a friend of mine, though, when it sees its master approaching the wine cabinet, anticipates what is going to happen, produces the popping sound yet preceded by a distinct squeak. Indeed, this squeak can be heard in most cases (demonstration) showing clearly on the sonagram and is evidently due to the friction between cork and glass prior to the influx of air into the bottle producing the plop (Fig. 7).

Conclusion

Isn't it remarkable, ladies and gentlemen, how fantastic our capabilities are in identifying any sounds we are familiar with, how unbelievably acute we are in interpreting that vague term called timbre in distinguishing the most subtle varieties of those familiar sounds, and yet, how poor we are, when keying up our perceptual analysing powers, to trace the particular perceptual aspects which enable us to make these distinctions.

The author is greatly indebted to Messrs. B. Lopes Cardozo, R.J. Ritsma, J. 't Hart, I.H. Slis and Th. A. de Jong from the laboratory and to Prof. A. Cohen from the University of Utrecht for their co-operation in preparing this lecture.
A

PHYSIOLOGICAL
and
PSYCHOLOGICAL ACOUSTICS

CONTENTS

A–2

A–2–1 Propriétés biophysiques particulières des osselets
de Mammifères aquatiques et marines
Rene Guy Busnel, Dominique Giraud-Sauveur A–1

A–2–7 A model for the spatio-temporal discharge patterns
of neurons in the auditory nervous system
Kengo Ohgushi, Hisao Sakai 5

A–2–8 Prediction of temporary threshold shift following
exposure to noise having arbitrary spectrum and
temporal characteristics
Takeo Yamamoto, Hikaru Shoji, Koichi Takagi 9

A–2–9 Growth of temporary threshold shift from
impulse-noise exposure
David C. Hodge, R. B. McCommons 13

A–2–10 New graphs on noise induced hearing loss
Jan van den Eijk, W. Passchier-Vermeer 17

A–3

A–3–4 A model describing temporal effects in loudness
and threshold
Eberhard Zwicker A–21

A–3–5 Binaural loudness summation as a function of
bandwidth
Bertram Scharf 25
A-3-6  On the difference between localization and laterazalization
Toshio Sone, M. Ebata, T. Nimura .......... A-29

A-3-7  Binaural fusion of tone bursts different in frequency
Masanao Ebata, Toshio Sone, T. Nimura .......... -33

A-3-8  Masking and its clinical use for differential diagnosis of perceptive deafness
Isamu Ebihara .......... -37

A-3-9  Lateralization of high-frequency complex stimuli
Hisao Sakai, Tsuneo Inoue .......... -41

A-3-10  Ein Beitrag zur Theorie des Vorwaerts-ruckwaerts-eindruckes beim Hoeren
Jens Blauert .......... -45

A-3-11  Auditory perception and blind guidance
Derek Rowell, Leslie Kay .......... -49

A-3-12  Pattern of the noise images and the binaural summation of loudness for the different interaural correlation of noise
Nikolai Andreevich Dubrovsky,
Rozaliya Isaevna Chernyak .......... -53

A-4

A-4-1  Control of psychophysical experiments with an on-line digital computer
Cecil Harold Coker, Joseph Lindley Hall,
Arron Edward Rosenberg .......... A-57

A-4-2  Vocal compensation for change of distance
Richard M. Warren .......... -61

A-4-3  Possibilities for the objective measurement of the loudness level of speech
O. Juhl Pedersen, S. Boel Pedersen .......... -65

A-4-5  Zur Messung des Effektivwertpegels und der Lautstarke von Logatomen
Otto Brosse, Rolf Wehrmann .......... 69
A-4-6 Some proposals on realizing the ideal hearing aid
*Mitsuhiko Katakura, Tomio Yoshida,
M. Okamoto

A-4-7 Identification of synthetic vowels in perceptive deafness
*Masao Tomisawa, Tetsuya Shitara, Tsunemasa
Sato, Ichiro Kirikae, Hiroya Fujisaki

A-4-8 Measurements on speech-audiometers
*Horsi-Gunther Diestel

A-5 Psychological scales for evaluation of vibrations
*Toshisuke Miwa, Yoshiharu Yonekawa

A-5-2 TACTPHONE as an aid for deaf
*Hisayoshi Suzuki, Ryuzemon Kagami,
Tazimi Takahashi

A-5-3 Formulation of spatial summation and spatial masking effects of vibratory sensation of fingers
*Akira Watanabe, Kouichi Abe,
Hisayoshi Suzuki, Ken'iti Kido

A-5-4 Recherches sur la propriété de quasi-stationnarité
parrapport à l'oreille humaine
*Hulina Maria Ruyfert

A-5-5 Pitch: Result of an auto-correlation process
*Roelof J. Ritsma

A-5-7 Pitch of filtered noise
*Andrew Rakowski

A-5-9 Tonal differentiel limen of the speech transmission
system containing single dip in frequency-response
*Ryoichi Matsuda

A-5-10 Descriptive adjectives on the listening of reproduced sounds
*Yutaka Enomoto, Tomio Yoshida

A-5-11 Factor analytical research of tone colour
*Otoichi Kitamura, Seiichiro Namba,
Rimpei Matsumoto
<table>
<thead>
<tr>
<th>Page</th>
<th>Title</th>
<th>Authors</th>
<th>Pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>A-5-12</td>
<td>A new theory of psychometric estimation concerning the sound quality of voice</td>
<td><em>Shuji Gotoh, Tomio Yoshida</em></td>
<td>A-121</td>
</tr>
<tr>
<td>A-5-13</td>
<td>Additivity in consonance and its calculation method</td>
<td><em>Akio Kameoka, Mamoru Kurtyagawa</em></td>
<td>125</td>
</tr>
<tr>
<td>A-5-14</td>
<td>Design of reproduced sound quality by ESP method</td>
<td><em>Rikuo Miyagawa, Takeshi Nakayama, Tanetoshi Miura</em></td>
<td>129</td>
</tr>
<tr>
<td>A-5-15</td>
<td>Über das kontextabhängige auditive Perzeptionsverhalten bei nicht-periodischen akustischen Schwingungsstrukturverläufen</td>
<td><em>Volker Nösselt</em></td>
<td>133</td>
</tr>
</tbody>
</table>
Propriétés biophysiques particulières des osselets de Mammifères aquatiques et marins.
R.-G. Ausnal et D. Giraud-Sauveur
Laboratoire de Physiologie Acoustique
C.N.R.Z.
78 - Jouy-en-Josas

Introduction

Les Cétacés Odontocètes possèdent des propriétés auditives particulières en rapport, notamment, avec leur capacité d'écholocation et leur perception qui s'étend très largement dans le domaine ultra-sonore, ces animaux analysent, en particulier, les spectres de bruit de leurs clics.

Nous nous sommes proposés de voir si, au niveau des osselets, ces éléments relais du système auditif n'avaient pas de propriétés biophysiques particulières, en les comparant avec celles d'autres os du crâne et avec celles des os et osselets d'autres mammifères marins ou amphibies qui n'utilisent pas l'écholocation.

Huit espèces d'Odontocètes ont été étudiées : Delphinus delphis L., Globicephala melas Traill, Phocaena phocaena L., Physalia physalus L., Stenella griseocephala L., Stenella bredanensis Lesson, Tursiops truncatus Montagu, Ziphius cavirostris Cuvier, ainsi que des Mammifères terrestres, Canis familiaris L.,
Cavia sp. L., amphibies d'eau douce Castor canadians Kuhl, Hippopotamus amphibius L., et amphibies marins Phoca vitulina L., Arctocephalus norvegicus Cuvier, pour permettre une comparaison.

Vitesses de propagation sonore

On fait la mesure par référence à celle d’un matériau connu (quartz) avec un transducteur de 5 ou 8 MHz relié à un générateur d’impulsions, selon la technique décrite par GARNIER (1).
Densité

Les mesures ont été faites par la méthode de flottation utilisée par Pernell et Dreher [3].

Dans toutes les espèces étudiées, la densité de l'os du crâne est de l'ordre de 2,10. Par contre, les osselets ont une densité plus grande que celle de l'os du crâne et ceux des Cétacés Odontocètes ont nettement la densité maximum (2,65) comparée à celle des Mammifères terrestres, amphibiens (2,25) et même aux Cétacés Mysticètes (2,45). (Fig. 1).

Il y a une différence de l'ordre de 4 % entre les différents osselets ainsi qu'entre la bulle tympanique et le marteau ; celle-ci ne s'observe pas chez les Odontocètes.

Graphique 1  Graphique 2

Fig. 1 - Comparaison des valeurs de la densité de fragments d'os du crâne (graph. 1) et de celle de marteau (graph. 2) chez les espèces suivantes :

G.R 1 Espèces ne faisant pas d'écho-location : 1 aérienne, 2 amphibiens ;
G.R 2 Espèces faisant peut-être de l'écho-location : 3 amphibiens, 4 aquatiques ;
G.R 3 Espèces faisant de l'écho-location : 5 aquatiques, 6 aériennes.

Minéralisation

Le dosage du calcium au spectrophotomètre de flamme donne des valeurs maxima chez les Cétacés Odontocètes qui atteignent 30 % du poids de l'os pour le marteau, contre 20 % pour l'os du crâne.

Chez les Phoques et les Otaries, les valeurs sont de 22 % pour le marteau et l'os du crâne ; ces valeurs sont les mêmes chez le Castor, le Cobaye et le Chien.

Le pourcentage de cendre est de 85 % pour le marteau des Odontocètes, contre — A-2 —
62 % pour l'os du crâne, alors que pour les autres espèces, ces valeurs sont de
70 % pour le marteau contre 60 % pour l'os du crâne.

**Microradiographie**

On a mesuré comparativement l'irradiation aux Rayons X des lames minces d'osselets
(100 μ) sous une intensité (10 mA) et une distance (20 cm) constantes, en faisant
varier à la fois le voltage et le temps nécessaire pour obtenir avec les osselets
des différentes espèces des clichés de noircissement égal.

On obtient les valeurs suivantes :

<table>
<thead>
<tr>
<th>Animal</th>
<th>Voltage lv</th>
<th>Temps mm</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chærous avunculus</td>
<td>13</td>
<td>15</td>
</tr>
<tr>
<td>Phoca</td>
<td>15</td>
<td>15</td>
</tr>
<tr>
<td>Canis</td>
<td>15</td>
<td>17</td>
</tr>
<tr>
<td>Delphinus</td>
<td>15</td>
<td>17</td>
</tr>
</tbody>
</table>

**Radiocristallographie**

La méthode utilisée est celle des poudres, dite de Debye-Scherrer.

Les diagrammes de diffraction montrent que les osselets minéraux des osselets
d'Odontocètes, donnant des raies fines et bien individualisées, sont mieux cristal-
lisés que ceux des os craniens de ces espèces, à raies floues et larges, ainsi que
les os de la plupart des osselets des animaux des autres groupes à raies floues et larges
egalement (fig. 2).

Ces diagrammes caractérisant la nature chimique des os de l'os, analogues
daus toutes les espèces et qui est du phosphate tricalcique hydraté.

**Fig. 2 - Diagrammes de diffraction.**

1. Témoin : Phosphate tricalcique hydraté.

1. Canis : crâne (a) marteau (b) - 2. Halichoerus : crâne (a) marteau (b)

3. D. delphis : crâne (a) marteau (b)
Conclusion

Les osselets des Cétacés Odontocètes étudiés possèdent des propriétés biophysiques particulières, d’abord vis-à-vis des os crâniens de ces espèces, mais aussi les os et osselets des autres Mammifères marins ou amphibies ou aériens.

Ils sont caractérisés par leur densité, une vitesse de propagation, un taux de cendre, une opacité aux rayons X, considérablement plus élevés que tous les autres groupes de Mammifères (les osselets des Cétacés Mysticètes n’ont pas nettement les caractères exceptionnels de ceux des Odontocètes, et ils en diffèrent notamment par la densité, la teneur en cendre).

Ces propriétés spéciales sont l’apanage des espèces qui utilisent des systèmes sonars actifs en milieu marin et sont probablement en relation d’une part avec les très hautes fréquences des signaux et avec la vitesse de répétition des impulsions d’autre part.

Références

1. Genuer (M.L.) - Appareil à mesurer les vitesses de propagation des ultra-sons.


2. Giraud-Sauvour (D.) et Bricard (M.C.) - Différence entre les diagrammes radio-cristallographiques des osselets de l’oreille moyenne et des os du crâne chez les Mammifères aquatiques et marins caractérisant les espèces qui utilisent l’écho-location.


3. Parnell (J.E.) et Dreher (J.J.) - Estimates of auditory frequency response limits as a function of mammalian ossicular density.

   Rapport n° 17090, 1963, du 14 octobre, de la Lockheed California Company, contract 3680 (00).

A model for the spatio-temporal discharge patterns of neurons in the auditory nervous system

Kengo Ongushi and Hisao Sakai
Broadcasting Science Research Laboratories,
Japan Broadcasting Corporation, Tokyo, Japan

Introduction

A model was constructed with the aid of basic physiological data in order to get an insight into the spatio-temporal discharge patterns of neurons at each level when a tone stimulus is received by the auditory system.

Model

Circuit models of the basilar membrane and the hair cell were constructed. The neuron and neural networks, however, were simulated on a digital computer. Fig.1 is a diagrammatic representation of a model of the auditory nervous system.

An electrical network was designed to make the voltage at each section correspond to the displacement along the basilar membrane. The network has 140 sections, each with the configuration shown in Fig.2. A hair cell model was composed of a diode and resistors. The output voltage of each of the 140 sections of the basilar membrane was rectified by

Fig.1 Diagram of the auditory nervous system
A model of the auditory nervous system

the hair cell model.

Concerning the lateral inhibition among nerve fibers of primary neurons, forward inhibition as proposed in a previous paper \(^1\) \(^2\) has been adopted. We assumed that the inhibitory coupling function among nerve fibers had a Gaussian distribution, with a standard deviation of 1.5mm. From the measured output voltage of each hair cell and this coupling function, the internal potential of each primary neuron can be computed. If the internal potential attains a threshold level, it causes the primary neuron to discharge an impulse of about 1msec. The neuron model is absolutely refractory for about 1msec. The relative refractory state starts at the cessation of the impulse. This exponential recovery toward its resting threshold has a time constant of 2.7msec.

The secondary neuron model has 11 excitatory and 37 inhibitory synapses. The coupling function of this neural net was also assumed to have a Gaussian distribution, with standard deviations of 0.3mm for excitatory and of 1.5mm for inhibitory coupling. The EPSP (excitatory postsynaptic potential) of a secondary neuron is built up by output impulses from primary neurons which have reached it through excitatory synapses. The IPSP is built up by primary neuron impulses through inhibitory synapses. There is a simple temporal and spatial summation of synaptically produced PSPs. The time course of EPSP was assumed \(^3\) to be synaptic delay of 0.38msec; time to summit of 1.14msec; time constant decay of 3.8msec. The time course of IPSP was assumed to be a mirror image of EPSP \(^3\). If the PSP of a secondary neuron reaches the threshold level, it discharges an output impulse. The refractory time

---A-6---
A model of the auditory nervous system

of a secondary neuron is the same as that of a primary neuron.

Experiment

An experiment was made with this model applying a 400Hz tone burst of 10msec duration to the basilar membrane model.

Results are shown in Fig.3-6. Fig.3 shows the temporal variation of the output voltage of each hair cell. Here darker color indicates higher voltage. The travelling waves along the basilar membrane can be observed.

Fig.4 shows the internal potential of primary neurons resulting from being sharpened by lateral inhibition. Fig.5 and 6 show the output impulses of primary and secondary neurons respectively.

In Fig.5, the primary neurons discharged impulses at intervals of about 2.5msec or 5.0msec, and their latencies were different for different neuron. These properties of the model agree with physiological data from squirrel monkeys and cats(4)(5).

On the other hand, each secondary neuron discharged only one or two impulses as shown in Fig.6 because of the effect of inhibition in the neuron. It is known physiologically too that higher order neurons discharged fewer impulses than peripheral neurons.

References

A model of the auditory nervous system

Fig. 3 The output voltage pattern of the hair cell models for a 400Hz tone burst of 10msec

Fig. 4 The internal potential of the primary neuron models for 400Hz tone burst of 10msec

Fig. 5 Spatio-temporal discharge pattern of the primary neuron models

Fig. 6 Spatio-temporal discharge pattern of the secondary neuron models
Prediction of Temporary Threshold Shift Following Exposure to
Noise Having Arbitrary Spectrum and Temporal Characteristics.

Takeo Yamamoto, M.D.
Professor of the Faculty of Engineering, Kyoto University

Hikaru Shoji, M.D.
Professor of the Faculty of Engineering, Kyoto University

Koichi Takagi
Assistant of the Faculty of Engineering, Kyoto University

Introduction

For the past ten years or so, many attempts have been made to
clarify the relation between temporary threshold shift (TTS) and the
noise exposure. However, there still remain unsolved such important
problems as follows. 1. How to correlate noise spectrum to TTS.
2. How to predict TTS following exposure to non-steady state noise.

To give the answers to these questions, we have undertaken the investi-
gation of TTS using the various types of exposure noise.

1) Critical band for TTS

From the results of experiments with high pass and low pass noises,
it was confirmed that the basic notion of the critical band could be
applied to the deafening effects of noise.\(^{1}\) To determine the center
frequency and the width of the critical band for TTS, the data obtained
by twenty minutes exposure to (I) noises having linear spectrum and
(II) noises of narrow band (1/6 octave band) were analyzed by the
following method.

a) Calculation of the center frequency of the critical band.

TTS at frequency \(F\) is assumed to be expressed as
\[ TTS_F = aX + b \cdots (1), \]
where \(a\) and \(b\) are the constants that depend on exposure time and test
frequency, \( x \) is the critical band level and is expressed as \( x = S(f_c) + 10 \log_{10} \Delta f \) ..... (2), where \( S(f_c) \) indicates the spectrum level at the center frequency \( f_c \) of the critical band, and \( \Delta f \) is the critical band width. In case of noise whose spectrum is a linear function of \( \log_{10} f_c \), \( S(f_c) \) is given by \( S(f_c) = a \log_{10} f_c + \beta \) ..... (3), \( a \) : spectrum slope \( \text{db/oct} \), \( \beta \) : intercept \( \text{db} \). From equations (1), (2) and (3), \( \text{TTS}_p = a \left( \log_{10} f_c + \beta - L \right) = a \left[ S(f_c) - L \right] \) ..... (4), where \( L = -10 \log_{10} \Delta f + b/a \). Equation (4) implies that \( \text{TTS}_p \) becomes a linear function of spectrum level at the center frequency of the critical band. The values of \( a, f_c \), and \( L \) can be calculated by the least square method;

\[
\Delta = \frac{\sum (y_i - a(\log_{10} f_c + \beta - L))^2}{\sum \Delta_i^2} = 0, \frac{\Delta}{\frac{\Delta}{\Delta f}} = 0, \frac{\Delta}{\frac{\Delta}{\Delta c}} = 0,
\]

where \( y_i \) is TTS produced by noise whose spectrum is \( a \log_{10} f + \beta \).

One of these results is shown in Fig. 1.

b) Calculation of the critical bandwidth.

The values of TTS produced by noise (I) and noise (II) are expressed by notation \( Y \) and \( y \) as:

\[
y = a(S_1 + 10 \log_{10} \Delta f) + b, \quad y = a(S_2 + 10 \log_{10} \Delta f) + b,
\]

where \( S_1 \) represents the spectrum level of noise (I) at the center frequency of the critical band and \( S_2 \), the spectrum level of noise (II) having bandwidth \( \Delta f \) in the critical band. Therefore

\[
10 \log_{10} \Delta f = 10 \log_{10} \Delta f + S_2 - S_1 + (Y - y)/a.
\]

From the equation above, the
Prediction of TTS

critical bandwidth in db was calculated for test frequencies 15, 2, 3, 4, 6, and 8kc (Table 1).

<table>
<thead>
<tr>
<th>Center frequency (cps)</th>
<th>1.5kc</th>
<th>2kc</th>
<th>3kc</th>
<th>4kc</th>
<th>6kc</th>
<th>8kc</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1010</td>
<td>1400</td>
<td>2620</td>
<td>3040</td>
<td>3840</td>
<td>4950</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>95% confidence limit (cps)</th>
<th>upper</th>
<th>1140</th>
<th>1440</th>
<th>2710</th>
<th>3260</th>
<th>4800</th>
<th>5980</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>lower</td>
<td>880</td>
<td>1360</td>
<td>2570</td>
<td>2900</td>
<td>3430</td>
<td>4450</td>
</tr>
</tbody>
</table>

| Bandwidth in db and its 95% confidence limit | 24. ±1.3| 26±0.8| 29.7±0.9| 30.5±0.8| 29.8±1.0| 33.3±0.8|

2) TTS produced by steady state noise.

The empirical equations which give TTS as a function of exposure time and the spectrum level at a center frequency of the critical band were derived from the data obtained by the experiments in which white noise and octave band noise were exposed at a various level for durations from 5 minutes to 8 hours. These are given in Table 2.

3) TTS produced by non-steady state noise.

After the various examination of TTS produced by non-steady state noise, we adopted the following method. The pattern of exposure noise in Fig. 2 can be expressed, by unit step function \( U(t) \) \( t \leq 0 \) as \( S_1[U(t) - U(t-T_1)] + S_2[U(t-T_1) - U(t-T_1-T_2)] + \cdots + S_n[U(t-T_1-T_2-\cdots-T_{n-1}) - U(t-T_1-T_2-\cdots-T_1)] + \cdots \), where \( S_1 \) is the level of the exposure noise. When we regard this as the input, then the output is expressed as \( f_{S_1}(t) = U(t-T_1)f_{S_2}(t-T_1) + U(t-T_1)f_{S_2}(t-T_1) - U(t-T_1-T_2)f_{S_2}(t-T_1-T_2) + \cdots + U(t-T_1-T_2-\cdots-T_1)f_{S_1}(t-T_1-T_2-\cdots-T_1) + \cdots \).

\[ S : \text{Spectrum level at center frequency of critical band.} \]
\[ T : \text{Exposure time in minutes.} \]

<table>
<thead>
<tr>
<th>Test frequency (cps)</th>
<th>Empirical equation</th>
</tr>
</thead>
<tbody>
<tr>
<td>8000</td>
<td>0.98(5-44.1)log_{10}T - 0.11S + 8.4</td>
</tr>
<tr>
<td>6000</td>
<td>0.88(5-33.1)log_{10}T - 0.33S + 17.6</td>
</tr>
<tr>
<td>4000</td>
<td>1.36(5-41.9)log_{10}T - 0.45S + 18.7</td>
</tr>
<tr>
<td>3000</td>
<td>0.85(5-374)log_{10}T + 0.27S + 6.4</td>
</tr>
<tr>
<td>2000</td>
<td>0.25(5+193)log_{10}T + 0.75S - 5.0</td>
</tr>
</tbody>
</table>

---

- A-11 -
Prediction of TTS

$f_{S_1}(t)$ denotes TTS produced by steady state noise at a level $S_1$, and is a linear function of logarithm of exposure time. The applicability of this method was confirmed from the results of the experiments in which six types of noises were used. Some results are given in Fig. 3. In order to make sure that the method is applicable for the different type of exposure, we re-examined the data from which on fraction rule\(^2\) exposure-equivalent rule\(^3\), and other rules\(^4\) had been derived by Ward et al. The good agreement was observed between the data and the predicted values. It was concluded that the method is available in predicting TTS produced by non-steady state noise.

Fig. 3. Comparison of TTS between observed and predicted values.

References
Growth of Temporary Threshold Shift from Impulse-Noise Exposure

David C. Hodge and R. Bruce McCommons

Research Psychologists, Behavioral Research Laboratory
U. S. Army Human Engineering Laboratories,
Aberdeen Proving Ground, Maryland 21005 USA

Introduction

One of the significant procedural problems in investigating the effects of impulse noise on hearing is that the range of susceptibility to impulse-noise-induced temporary threshold shift (INITTS) is far greater than is the case with steady noise. This makes it necessary to select exposure conditions very carefully. If the conditions are too severe there is considerable risk of either (a) permanently damaging the subject's hearing or, at the very least, (b) the recovery time will be lengthy (3). On the other hand, if the number of impulses in an exposure is selected to protect the most susceptible subjects, the least susceptible subjects will demonstrate either zero TTS or, as we have found in a number of instances, negative TTS (2). Exposing all subjects in an experiment to the same number of impulses is advantageous when the purpose is to establish the distribution of TTS resulting from a given set of noise parameters. However, when the purpose is to study more general features of the hearing mechanism's response to noise exposure, or when the effects of several conditions are to be compared, it would be more beneficial to obtain a measurable, positive, TTS from each subject. To accomplish this end the amount of noise exposure would have to be tailored for each subject individually; then the dependent variable would be the amount of exposure required to reach a set criterion level of TTS.

This paper relates experiments in which a procedure was developed for observing the growth of INITTS from zero to a criterion level. The general approach was somewhat akin
to that employed by Elwood, et al. (1), in their method for separating subjects into categories of susceptibility.

Procedure

The subjects were first trained to give reliable thresholds with a discrete-frequency Bekesy audiometer. They were then exposed to groups of impulses produced by a .30 cal. small arm. (A representative impulse waveform is shown in Fig. 1). The peak level of each impulse was 155 dB re 0.0002 μbar and the A-duration was 0.35 msec (free-field measurements; subject absent). During the exposure the subject was seated in a position such that his left ear was oriented at normal incidence to the oncoming shock wave. After each group of impulses the subject's thresholds at 2, 4, and 6 kHz were measured. This alternation of noise exposure and threshold measurement was continued until the subject's left ear demonstrated 15 dB TTS₂ at one or more of the test frequencies. When the criterion of 15 dB TTS₂ was reached, no further noise exposure was administered. Post-exposure audiometry varied according to the purpose of the experiment. Sometimes complete audiograms were taken on both ears. In other cases only the left ear was tested, repeatedly, in order to establish recovery functions for INITTS.

Experiment #1

32 subjects were exposed to groups of 10 impulses. A maximum of 20 such 10-impulse groups was administered; thus the least susceptible subjects were exposed to no more than 200 impulses. A monaural exposure was used, the subject being required to wear half of a noise-attenuating earmuff during the exposure.

The distribution of INITTS growth rates (in terms of numbers of 10-impulse groups required to cause criterion TTS) is shown in Fig. 2. 56% of the subjects reached criterion after exposure to only three groups of impulses. We concluded that this particular procedure
would not be suitable for use in a future study of, for example, higher peak levels, since the higher levels might be expected to cause more rapid TTS growth. If TTS grew more rapidly than it did in the present case it would not be possible to discriminate among the growth rates.

Another question which arose during this first investigation was whether there would be any difference in TTS growth rate in the test ear when binaural, rather than monaural, exposure was used.

To test such a possibility, 19 of the subjects were given binaural as well as monaural exposures. The distributions of INITTTS growth rates are shown in Fig. 3. The two distributions are remarkably similar. The mean number of impulses to criterion was 41.1 for the binaural exposure and 45.3 for the monaural exposure. The difference between these means was not significant. Four subjects demonstrated identical growth rates under the two conditions; 8 subjects’ TTS grew more rapidly under the binaural condition; the remaining 7 subjects’ TTS grew more slowly under the same condition. Hence, we concluded that, while there might be individual differences in rates of growth under the two conditions, there were no consistent or obvious trends favoring either approach.
Experiment #2

To better define the initial portion of the TTS growth function, 31 subjects were given binaural exposures to the same peak level and duration used in the first study. The number of impulses in the groups was changed: the first four groups contained 5 impulses each; the next four, 10; the next four, 15; the last four, 20. The distribution of growth rates is shown in Fig. 4. Compared to Fig. 2 and Fig. 3, the distribution is skewed to the right. The number of subjects reaching criterion in the first few groups has been reduced and a more nearly rectangular distribution of growth rates results. From this we concluded that the procedure used in Experiment #2 was more nearly acceptable for our purposes.

References


Introduction

The several existing "damage risk criteria" are based on laboratory investigations on temporary threshold shifts (TTS) and an assumed relation between temporary and permanent threshold shifts (PTS). Unfortunately very little is known in fact about this relation: only at 4000 Hz it is reasonably established. Therefore, we decided to investigate whether an upper limit for safe noise levels could be established on the basis of measured hearing levels of labourers who have worked in noise for years. Data have been collected from the available literature as well as from measurements by members of our working group. The data relate to 20 groups of employees; all in all about 4600 people.

Noise and noise exposure

We limited the investigation to steady-state broadband noise and to exposures for 8 hours a day and at least 5 days a week. The noise is characterised by a noise rating number (NR), which is established by comparing the octave band levels with midfrequencies 500, 1000 and 2000 Hz of the noise with a set of noise rating curves [1]. The curves are numbered according to their dB-levels at 1000 Hz. In this paper the NR of a noise is the number of the NR-curve just not surpassed by any of the octave band levels with midfrequencies 500, 1000 and 2000 Hz.

All the results in this paper are formulated in NR-terms. They can be used however with reasonable accuracy for sound levels in dB(A) too. The number

1) Publication no. 302 of the Res.Inst. for Public Health Engng.TNO. The investigation forms part of the research program of the Working Group "Relation between noise and hearing loss" of the Research Committee on Occupational Health TNO. (TNO = Netherlands Organisation for Applied Scientific Research)
obtained by adding 5 to the NR value is numerically equal to the corresponding sound level in dB(A).

**Median noise-induced hearing loss**

To simplify the problem we leave out first the influence of individual susceptibility and confine ourselves to the median of the hearing levels, which we consider to be influenced by:
- the NR of the noise
- the total exposure time.

We define the median of the noise-induced hearing losses (\( D_{50\%} \)) at a given frequency of a group of employees exposed to noise for a given time, as the noise-induced increase of the median hearing level at that frequency, i.e. the difference between the median hearing level (\( L_{e,50\%} \)) of the group exposed and the median hearing level (\( L_{n,50\%} \)) at the same frequency of a group of people not exposed to noise, but having the same mean age as the group exposed. In Fig. 1 for a number of frequencies and for an exposure time of 10 years, \( D_{50\%} \) is given as a function of NR.

We see that for a 10 years exposition \( D_{50\%} \) is a maximum at 4000 Hz, although the difference with 3000 Hz is small. Furthermore at frequencies above 4000 Hz as well as below 4000 Hz, \( D_{50\%} \) is smaller the larger the difference between the frequency considered and 4000 Hz. This is true for each NR. Further we see that even for NR = 75 there is some median hearing loss at the frequencies 3000, 4000 and 6000 Hz.

For an exposure time longer than 10 years we could expect of course higher values of \( D_{50\%} \), especially for the higher noise ratings. Fig. 2 gives the results for an exposure time of 40 years. By comparing Fig. 2 with Fig. 1 we see that \( D_{50\%} \) does not increase in the same way for all frequencies. For the frequencies 500, 1000 and 3000 Hz it turns out that per frequency for all NR’s considered, \( D_{50\%} \) increases after 10 years of exposure with the same percentage of \( D_{50\%} \), which is due to exposure for 10 years. This percentage, however, is different for different frequencies. For 4000 Hz there is no increase and for the frequencies 6000 Hz and 8000 Hz the situation is more complicated; for NR’s up to 92 there is no increase either, whereas for NR’s higher than 92 there is an increase which is larger, the higher the NR. In the Table the different percentages are given.

<table>
<thead>
<tr>
<th>Table</th>
</tr>
</thead>
<tbody>
<tr>
<td>Increase per year of ( D_{50%} ) for exposure times longer than 10 years, in percentage of ( D_{50%} ) due to exposure for 10 years</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Frequency (Hz)</th>
<th>500 Hz</th>
<th>1000 Hz</th>
<th>2000 Hz</th>
<th>3000 Hz</th>
<th>4000 Hz</th>
<th>6000 Hz</th>
<th>8000 Hz</th>
</tr>
</thead>
<tbody>
<tr>
<td>Percentage</td>
<td>2%</td>
<td>2.5%</td>
<td>10%</td>
<td>1%</td>
<td>0%</td>
<td>0.3(NR&gt;92)</td>
<td>0.3(NR&gt;92)%</td>
</tr>
<tr>
<td></td>
<td>for NR&gt;92</td>
<td>for NR&gt;92</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
New Graphs on Noise-Induced Hearing Loss

Remarkable is that at 4000 Hz there is no increase at all after 10 years of exposure, whereas at 2000 Hz the increase is large: 10% per year! Our analysis showed that $D_{50\%}$ at 2000 Hz is a linearly increasing function of exposure time from the very beginning of exposure.

Spread due to individual differences in susceptibility

Up till now we considered only the median noise-induced hearing losses, i.e. the increase, due to exposure to noise, of the median hearing level of people exposed to noise. Although this gives important information about the damage caused by noise, we should not base on it an upper limit of safe noise, since we do not have information yet about the hearing losses that are less favourable than the median values. Therefore, instead of the noise-induced increase ($D_{50\%}$) of the median hearing level, we shall now consider the noise-induced increase ($D_{75\%}$) of the hearing level, not exceeded in 75% of the people exposed. $D_{75\%}$ is of course equal to the difference between the hearing level ($L_{e,75\%}$), not exceeded by 75% of the people exposed, and the hearing level ($L_{h,75\%}$), not exceeded by 75% of non-exposed people with the same mean age as the people exposed. We call $D_{75\%}$ noise-induced hearing loss not exceeded by 75% of the employees.

Our analysis showed that for exposure times of at least 10 years, $D_{75\%} - D_{50\%}$ is independent of exposure time for all frequencies considered. This difference, however, depends upon the ER. The results are presented in Figure 3:

$D_{75\%} - D_{50\%}$ is increasing with ER at 2000, 3000, 6000 and 8000 Hz, decreasing at 4000 Hz and stays at 500 and 1000 Hz. Furthermore the differences between $D_{75\%}$ and $D_{50\%}$ are rather small. This means that the increases, due to exposure to noise, of the hearing levels not exceeded by 50% and 75% of the people are about the same. This does not mean, of course, that the hearing levels of people exposed all are the same, but only that the spread in hearing levels of people exposed are about the same as that of non-exposed people.

Our next step was to consider 90% of the people exposed. If we define $D_{90\%}$ analogously to $D_{75\%}$ and $D_{50\%}$, we have a good indication that the difference between $D_{90\%}$ is about twice as large as the difference between $D_{75\%}$ and $D_{50\%}$. We shall investigate this more thoroughly.

A full report [2] of the investigation is available at the Institute for Public Health Engineering THO, P.O. Box 214, Delft, The Netherlands. It seems to us that, as soon as the data in this report are completed with sufficient data on $D_{90\%}$ it is justified to decide on this basis which sound levels or NR's of steady-state broadband noise are acceptable in view of hearing damage.
New Graphs on Noise-Induced Hearing Loss


---

Fig. 1

Fig. 2

Fig. 3

---

A - 20 ---
A model describing temporal effects in loudness and threshold

Eberhard ZWICKER

Professor und Direktor des Instituts für Elektroakustik
der Technischen Hochschule München

Temporal effects in loudness especially the increase of loudness as a function of duration can be approximated by the simple assumption that sound intensity may be transferred through a RC-network using the peak value for further calculation. The value of the time constant given by different authors varies from 23 ms /1/ to 100 ms /2/ and 200 ms /3/. There is a tendency to the longer time constant but there is also the argument /4/ that the time constant describing temporal effects in loudness can not be longer than 35 ms because just noticeable degree of amplitude modulation as function of frequency leads to a time constant of 15 ms /5/. For backward masking the situation is similar. Although there is no good conformity between the decay of backward masking as function of delay time and the approximation of the decay of a RC-network /6/, the time constant of 5 ms necessary to get agreement in delay for short delay times, is very different to the 100 ms-time-constant in loudness.

Shortening the time constant for loudness to get better agreement between the different time constants describing temporal effects in hearing processes, is simple but inconsistent. It may be much better to give up on the simple conception /4/ of using one single time constant to describe the very complicated dynamic functions of the hearing process, like loudness, threshold, forward masking or just noticeable difference in degree of amplitude modulation. A somewhat more complicated conception which is able to describe the temporal effects of various sensations is needed.
A model describing temporal effects in loudness and threshold

A first step in this direction is the model shown in fig. 1.

![Diagram](image)

Fig. 1. Function model.

The sound pressure $p(t)$ is transferred to a filterbank which is responsible for the selectivity of the ear as it is described /6, 7/ by the excitation $E$ as a function of tonalness $z$. As an approximation of the very many outputs of the bank there are only 24 in regard to the 24 critical bands. Within or connected to the bank are rectifiers and square law transformers, so that at a certain output of the bank the excitation $E^*(t)$ is available. The excitation $E^*(t)$ is transferred to an RC-network with a time constant $\tau_1$ of only several ms integrating only over a relative short time. The signal is further transformed through a power function with an exponent $0.25$ of the specific loudness /7/ and lead through a proportional differential transfer function to an other RC-network with a longer time constant $\tau_2$ of some ten ms.

The output of this last RC-network gives the specific loudness $N_v^I(t)$ within the $v^{th}$ channel. $N_v^I(t)$ is a provisional value and may be processed into threshold by forming $N_v^I$ or into loudness $N(t) = \sum_1^2 N_v^I(t)$ by summing $N_v^I(t)$ over all channels. Hence loudness $N(t)$ is a function of time, the loudness of a short impulse will be the maximum value $\text{N}_{\text{max}}$ (within about 100 ms).
A model describing temporal effects in loudness and threshold

Now \( E(t) \) is transformed through the different steps into \( N'(t) \) may be shown in fig. 2 using rectangular impulses of different duration. The time functions have been taken from a model with the values \( \tau_1 = 35 \) ms. The time constant \( \tau_2 \) of the proportional differential element was 8 ms with the relation \( b = 0.5 \) between the peak value and the steady state value for a step function.

The step from \( E(t) \) to \( E^*(t) \) through a normal RC-network is rather simple. On the other hand the nonlinear transfer from \( E^*(t) \) to \( N''(t) \) produces not only a much quicker rise of the time function and thereby an increment of the peak value of shorter pulses but also a 4 times prolonged decay time in regard to the exponent \( \frac{1}{4} \) of the power function. The proportional differential element produces an overshoot at the beginning. The time constant \( \tau_2 \) is chosen in such a way that the decay function remains positive. The final RC-network leads to the specific loudness \( N'(t) \) which is shown in the lowest part of fig. 2. The impulse reaches the asymptote only for durations longer than 100 ms. For shorter impulses the peak value becomes smaller. It should be pointed out that the peak is reached for a time after the end of the original impulse. Coming from short durations, the peak value of \( N'(t) \) increases for about a factor of 2 if duration is increased for a factor of 10. This means that the loudness level increases for about 10 phon, if
A model describing temporal effects in loudness and threshold

duration is increased for a decade (for example from 4 ms to 40 ms). This is in good agreement with most of the published results.

Measurements on the model simulated with analog computers have shown, that the temporal effects in loudness - as dependence of duration for single impulses and dependence of repetition rate for pulses as well as forward throttling of loudness - can be described with a model of the described kind. Using the same model, temporal effects in forward masking can be explained quantitatively, too. Even backward masking can be discussed.

The results show that a single RC-network is an approximation just too simple to interpret the very different temporal effects in hearing, while the described model - still a very first approximation - clarifies how even short time constants of RC-networks in cooperation with nonlinear transformation may account for the different temporal effects measured in loudness and threshold respectively.

Literature

/1/ NIESE, H., Hochfrequenztechnik und Elektroakustik 68 (1959), 143
/2/ ZWICKER, E., Acustica 17 (1966), 11 - 22
/3/ STEVENS, J.C. and HALL, J.W., Perception and psychophysics 1 (1966), 319
/4/ REICHARDT, W., Acustica 15 (1965), 236
/5/ ZWICKER, E., Acustica 6 (1956), 365-381
/6/ ZWICKER, E. und FELDTKELLER, R., Das Ohr als Nachrichtenempfänger. Verlag Hirzel Stuttgart
/7/ ZWICKER, E. und SCHARF, B., Psych.Rev. 72 (1965), 3-26

— A-24 —
Binaural Loudness Summation as a Function of Bandwidth

Bertram Scharf

Associate Professor in the Department of Psychology,
Northeastern University, Boston

A sound led to both ears is louder than the same sound led to one ear. To make the binaural sound and the monaural sound equally loud, the monaural sound must be set to a higher sound pressure level than the binaural sound. The intensity difference between the two sounds depends upon their absolute sound pressure level. Reynolds and Stevens (1960) reported that for a white noise the intensity difference increased from about 4 dB at 30 dB SPL to more than 10 dB at 100 dB SPL. Because Fletcher and Munson (1933) had earlier reported similar results for pure tones, Reynolds and Stevens naturally assumed that binaural loudness summation is independent of bandwidth and type of sound. However, in measurements of dichotic loudness summation for pure tones (Scharf, 1967), I consistently measured less binaural loudness summation, i.e., a smaller intensity difference between a monaural and binaural tone, at high sound pressure levels than at moderate levels. Forsolt and Irwin (1967) reported the same finding, also for pure tones. The results of Causse and Chavasse (1942) tended in this direction. The only important differences between the measurements of Reynolds and Stevens and my own were in the type of stimulus. I therefore undertook to measure, in the same experiment and on the same subjects, binaural loudness summation for pure tones and for bands of noise, both narrow and broad.

Four different stimuli were used: a 1-kHz pure tone, a band of white noise with a bandwidth equal to one critical band (924 to 1091 Hz), an octave band of white noise (690 to 1400 Hz), and a broad band of white noise (high cutoff at 7000 Hz). Sounds were presented through a pair of dynamic earphones (TH-39) mounted in sponge Neoprene cushions (MX-41/AR). A monaural sound was presented through a single earphone to the left ear of half the subjects and to the right ear of the other half. The monaural and binaural sounds alternated, each lasting 750 msec with a silent interval of 600 msec between them. In half the trials the subject adjusted
the intensity of the monaural sound to make it as loud as the binaural sound. In the other half of the trials he adjusted the binaural sound. The standard sound, either the monaural or binaural, was presented at each of 7 sound pressure levels from 20 to 110 dB. Binaural sounds were presented in phase at the two ears so that the sound was heard in the center of the head. All but one of the 12 subjects were college students about 20 years old.

Our measure of binaural loudness summation is the number of decibels by which the sound pressure level of the monaural sound exceeded the level of the binaural sound when the two sounds were judged equally loud. Figure 1 shows how this difference varied as a function of the sound pressure level of the binaural sound. Each point is the median of 24 loudness matches, 12 matches in which the monaural sound was adjusted and 12 in which the binaural was adjusted. For the broad-band noise, loudness summation increases monotonically as a function of sound pressure level. For the pure tone and the critical band of noise, loudness summation is maximum around 40 dB SPL and then changes very little between 40 and 110 dB SPL. (At all but one level, the pure tone and the critical band gave the same result which is combined in a single circle, half white and half black.) An octave band of noise fails between the narrow critical band of noise and the broad band of noise. The inter-subject variability for these data was fairly small; the inter-quartile range varied between 3 and 4 dB. These results suggest that binaural loudness summation varies in an orderly fashion as a function of bandwidth. At low sound pressure levels binaural loudness summation appears to be greater for narrow-band stimuli and at high levels for broad-band stimuli.

A second series of measurements with 3-tone complexes, generated by three oscillators, confirmed the conclusion that binaural loudness summation depends on bandwidth. The procedure was the same as described above. The middle frequency of the complex was 2 kHz. The lower and higher frequencies were set to give overall frequency separations, or "bandwidths," from 150 Hz, less than the critical bandwidth, to 4200 Hz. Once again, for narrow-band stimuli binaural summation was maximum around 40 dB SPL, whereas for broad-band stimuli it increased with sound pressure level. However, at low levels the results were not clear-cut, probably because differences in loudness among the equally intense components of the complex.
Binaural Loudness Summation as a Function of Bandwidth

It is difficult to know why the present results with narrow-band stimuli differ so clearly and consistently from the results of Fletcher and Munson (1933). Since all but 2 of my 12 subjects gave results of the same form as the average curves in Figure 1, it is unlikely that we can ascribe the discrepancy to subject differences, although Fletcher and Munson did not state how many subjects they used. Probably the discrepancy arises from procedural differences. Fletcher and Munson used a method of constant stimuli with the comparison 1-kHz tone always coming second. I used the method of adjustment, as did Reynolds and Stevens (1960) and Porsolt and Irwin (1967). Guasse and Chevasse (1942) did not describe their procedure.

These data pose two closely related questions about the form of the monaural and binaural loudness functions (some functions) and the amount of binaural loudness summation, measured in sones. In this short paper, it is possible only to touch upon these points. The non-monotonic relation between monaural and binaural loudness summation as given in Figure 1 for a 1-kHz tone implies that both monaural and binaural loudness cannot be simple power functions of sound pressure. If the binaural tone function has an invariant exponent of 0.6 (above 30 dB SPL), then the monaural function must have a variable exponent. However, the amount of variability is likely to be too small to detect by the direct psychophysical procedures upon which the tone function is based. Since the broad-band data do agree with those of Reynolds and Stevens (1960), they fit their conclusion that the binaural loudness function for broad-band noise has a larger exponent than the monaural function. However, their assumption that the same rule applies to pure tones and that, indeed, the same loudness functions hold for a 1-kHz tone and for white noise is contrary not only to the present data but also to data from many experiments in which broad-band stimuli and narrow-band stimuli or pure tones have been matched in loudness (see Zwicker & Scharf, 1965).

The second question is just how much louder is a binaural sound than a monaural sound. We can use the present results to answer this question provided we can convert the intensity differences measured in decibels to loudness differences measured in sones. To make the conversion, we must know the form of the loudness function, and, given the small range of differences, we must know it rather precisely, but the precision of the loudness function is uncertain. Nevertheless, it has been well established that the loudness of a pure tone grows more rapidly at low levels than at higher levels (Hellman & Zwilocki, 1961). So although the difference between the monaural and binaural tone is only 5 dB at 20 dB SPL, the corresponding difference in loudness may be larger there than at higher levels where the intensity difference is greater than 5 dB. It has also been shown that at low levels the loudness of broad-band stimuli grows even more rapidly than the
loudness of a pure tone (Scharf, 1959). So the 3-dB difference for the white noise at 20 dB SPL does not necessarily mean that binaural loudness summation is smaller for white noise than for a pure tone at that level. Preliminary calculations based on a large variety of data suggest that below 100 dB SPL the differences among the sounds in Figure 1 largely disappear when binaural summation is measured in sones instead of decibels. However, owing to the uncertainties in these calculations, I conclude by noting simply that the amount of binaural loudness summation, measured as an intensity difference, depends upon both sound pressure level and bandwidth.

1. Research supported by a grant from the National Institute of Neurological Diseases and Blindness, U. S. Public Health Service.

References


Forsløt, K. D. & Irwin, R. J. Binaural summation in loudness of two tones as a function of their bandwidth. Amer. J. Psychol., 1967, 80, 384-390.


On the Difference between Localization and Lateralization

Toshio Sone
Associate Professor of Faculty of Engineering, Tohoku University

Masanao Ebata
Assistant of Faculty of Engineering, Tohoku University

Tadamoto Nimura
Professor of Faculty of Engineering, Tohoku University

Introduction

Sounds coming to the ears from external sources are heard as "out there", while sounds coming through earphones are heard as inside the head. Spatial aspect of the former is perceived as "localization" and that of the latter as "lateralization". What factor causes the difference between both perceptions in spatial aspect is discussed here.

Factors which exist in field-listening, but are absent in earphones-listening are as follows: (1) Diffraction and scattering of sounds by head and pinnae, i.e. the directivity of ears. (2) Effect of the head movements on the perception of sound location. (3) Sound input through bone conduction. (4) Visual information about external sound sources.

On the contrary, factors which are added in earphones-listening, though absent in field-listening, are as follows: (5) Abnormal feeling caused by setting the earphones on ears. (6) Bone-conducted sound transmitted from earphones to the skull.

These factors are investigated here one after another.

Procedure and Results

(1) Effect of Diffraction and Scattering of Sounds by means of the Head

An artificial head is placed in the sound field and two microphones are set up on its both ears. When a sound from those microphones is heard through binaural earphones, it is heard as inside the head. This factor, therefore, makes little contribution to the difference between localization and lateralization.

(2) Effect of the Head Movements and Visual Information

Both ears of a subject sitting in the sound field are covered with the ear protector(sound isolation 35-40dB) in order to exclude the air-
borne sound, and sounds through microphones attached to it are led to his both ears. When a level of the sound out of earphones is of the same order with that in the sound field (reference level), the sound is heard as "out there". When it is 20dB above the reference level, however, the sound is heard as inside the head even if the head is moved and the source is stared at. These factors, therefore, make little contribution to the difference between localization and lateralization.

(3) Effect of Bone Conduction (I)

In order to investigate what role the bone-conducted sound plays in field-listening, the spatial aspect in the case that bone-conducted sound is given to a subject from sound field is compared with that in the case without bone-conducted sound.

Procedure: Both ears of a subject being covered with the ear protector, the sounds from microphones standing at the sides of the ears are led to them through earphones. Fig.1 shows the block diagram for the experiment. Band limited noise (600-4800Hz), speech and musical sounds are used as source sounds. If the switch is turned to the position "Sw.1", bone-conducted sound is not presented. If the switch is turned to the position "Sw.2", outputs of the taperecorder are radiated into sound field through a loudspeaker, and the sounds are led to a subject's ears through earphones after two microphones. Intensity of sound through earphones is adjusted with Att.1 and Att.2 to be equal to that in the external field. This intensity level is taken as the reference level. The experimenter sets an attenuation level for Sw.1 to previously arranged level, and then, a subject is required to adjust

--- A-30 ---
an intensity for Sw.2 as equal to that for Sw.1 and to record the spa-
tial aspects for Sw.1 and Sw.2.
Results: Fig.2 shows some of the results. The followings are clarified
in this experiment. (a) For 65dB SPL of sound pressure in the field and
sound level of 0dB re. 65dB SPL through earphones, a sound is lateralized
at the back of the head for Sw.1 and is heard as "out there" for Sw.2.
(b) For earphones' level higher than the reference level, an image shifts
to the upside of the head for Sw.2, ultimately to the back of the head,
as earphones' level increases. (c) For earphones' level lower than the
reference level, sound is heard as coming from a distant position in
front of a subject and localization is indistinct for Sw.2. (d) If the
head is placed behind the line between two microphones, a sound image is
lateralized at the back of the head for Sw.2 and a contribution of the
bone-conducted sound from the field to spatial aspect of sound disap-
ppears. It is meant, therefore, that the time difference between the
bone-conducted sound through frontal area of the head and sound entering
the ear canals plays an important role in the perception of spatial
aspect.
(4) Effect of Bone Conduction (II)
Now the bone-conducted sound is given to a subject through a bone
conduction receiver (BCR) put on his forehead. Effects of the time dif-
ference between the bone-conducted sound and one through earphones, and
the level difference between both sounds on spatial aspect of sound are
investigated. Source sounds are a band limited noise, speech and musical
sounds as before. The results are as follows; (a) A lateralized image
shifts forward when the bone-conducted sound leads the earphones' sound
about 0.3msec. (b) But if the time difference is about 0.15msec, the
sound is heard as coming from the upside of the head. (c) If it is about
10msec, an image for the bone-conducted sound is separated from that for
the earphones' sound. (d) So that an image may shift forward clearly,
it is required that the bone-conducted sound should be 10-20dB louder
than the reference level. From these results, it is clarified that the
bone-conducted sound through a BCR is effective in shifting a sound
image forward (toward forehead). However, sound is not heard as "out
there".
(5) Effect of Bone Conduction (III)
When a stereophonic sound is given to a subject through two BCR put
on left and right side of his forehead, a lateral expanse of sound is
produced and the sound is heard more natural than in the case of a single
receiver. But the sound is not heard as "out there".
(6) Effect of Bone Conduction on Decision of Front or Rear
Localization and Lateralization

If a bone-conducted sound transmitted to the skull through our forehead plays an important role in our deciding a sound as coming from the front, the decision as to whether the sound comes from the front or the rear of us is expected to be inaccurate when we listen to sound in the sound field of less intensity than the bone conduction threshold. To test this point, sounds, 1/2 oct. of band limited noise, with intensities of 20, 40, 60 and 80dB SPL are presented to a subject in random order through one of six loudspeakers placed in front and rear of him, and the proportion of correct judgments is obtained as to which direction the sound comes from, the front or the rear. Fig. 3 shows the result that the decision about front or rear is inaccurate for a sound of less intensity. In this experiment a subject is prohibited from moving his head.

(7) Effect of Reverberation

If reverberant sounds, conversations among several persons in a normal room, are presented to a subject in another room through binaural earphones and a BCR, he can appreciate not only speakers' lateral disposition but also distances from sound sources, and the spatial aspect approaches to that for localization.

(8) Effect of Putting Earphones on Ears

Even if earphones are removed and two loudspeakers are set near to both ears, spatial aspect of sound is the same as that for earphones.

Discussion

As a result of the investigation into several factors which are suspected to have possibilities of contributing to the difference between localization and lateralization, a bone conduction is considered to be the most possible factor. There remains the further question as to the fact that the intensity of bone-conducted sound must be 10-20dB louder for receiver-listening than for field-listening in order to shift a sound image forward. As to this point it is required to examine whether the difference between vibration of the skull excited through a bone conduction receiver and that given from a sound field is important or not.

--- A-32 ---
Binaural Fusion of Tone Bursts Different in Frequency

Masanao Ebata
Assistant of Faculty of Engineering, Tohoku University

Toshio Sone
Associate Professor of Faculty of Engineering, Tohoku University

Tadamoto Nimura
Professor of Faculty of Engineering, Tohoku University

Introduction

Even if the wave forms of sounds given to our both ears are fairly different from each other, they can be fused into a sound image as far as the durations of those sounds are short enough and the time difference between them are small, within a certain limit. When the different continuous tones are given to each of our both ears respectively, a fused image is not obtained and both sounds are perceived as separate sounds, one at the left ear and the other at the right, even if the frequencies of both tones are apart only a bit from each other. These phenomena whether a fused image is obtained or not arises from the interaction between both ears, and it is considered to be caused by less binaural interaction that sounds given to both ears are perceived as separate sound images.

The relation between interaural difference in frequency components of sounds given to each of our both ears and their binaural fusion into a sound image is investigated here.

Apparatus and Procedure

Apparatus for the experiment is shown in Fig. 1. Sine waves out of two oscillators are turned into tone bursts of the same duration through
gate circuits, and they are led to each of a subject's ears respectively. Frequency of a tone from one oscillator is fixed at 1000Hz and that from the other oscillator is variable. If the frequency of the tone generated by the latter approaches to that of the former, tone bursts led to both ears can be fused into an image, while if the difference between two tones in frequency is large enough, they are perceived as separate images. The minimum frequency difference between both tones over which they are perceived as separate images is obtained by the limiting method. Tone bursts with durations of 20, 50, 100 and 200msec, along with 4msec of rise-fall time, and continuous tones are used in this experiment as stimulus sounds. Subjects are 5 young females and 3 young males with normal hearing acuity.

Results

Fig.2 shows the results. The abscissa in the figure is the duration of tone bursts and the ordinate shows the minimum frequency difference between two tones given to each of both ears over which they are perceived as separate images. It is seen from the results that the minimum frequency difference between them becomes gradually small and approaches to a constant value as the duration increases. The fact that the minimum frequency difference become gradually small along with an increase of durations of tones should be attributed to the reason why the frequency analysis is sharpened as the durations become long. It is
meant by the fact that the minimum frequency difference approaches to a constant value along with an increase of durations of tones, that the stimulus sounds are observed in our auditory system through a limited time window.

Discussion

Binaural fusion of different sounds into a sound image is considered to arise from the binaural interaction. Now we propose a model of the mechanism for binaural fusion as shown in Fig.3. In this model sounds entering both ears get to the process for frequency analysis in the first place. Spectra for both sounds are compared with each other in the next place. Then, if both spectra are similar to each other, a fused image can be obtained, while if the similarity between them is comparatively small, those sounds are separately perceived.

In order to explain a fusion mechanism for sounds different in frequency components, now an auditory step for frequency analysis is simplified to a frequency analyzer with a finite length of time window. Short time spectrum for a sound at a given instant of time \( t \) after the sound input which entered the analyzer is expressed as follows:

\[
F(w, t) = \int_{-\infty}^{t} f(, t - \lambda) e^{-j\omega \lambda} d\lambda
\]

Fig.2 Minimum frequency difference between two tone bursts given to both ears over which they are perceived as separate images

(Diagnosis)

Comparison between both spectra

(Frequency Analysis)

Left Ear Right Ear

Fig.3 Model for the binaural fusion
Binaural Fusion

where
\[ f(t) = (1 - e^{-\frac{t}{\tau}}) \cos \omega_c t \]
\[ \mathcal{A}(t) = t^2 e^{-\alpha t} \]
and \( f(t) \) represents the stimulus sound which exponentially builds up and approaches to a constant value, and \( \mathcal{A}(t) \) shows the time window. Spectra for various durations of sounds are calculated according to this formula.

Next problem is in what way the spectra for sounds given to both ears are compared with each other and the difference between them is detected.

As an estimation of the similarity between both spectra, cross-correlation and total spectral difference between both sounds are calculated. For cross-correlation it is studied whether there is the minimum value of cross-correlation independent of the durations of tone bursts over which a fused image is obtained. For the latter it is studied, in the same way, whether there is the maximum value independent of the durations within which a fused image is obtained.

Fig. 4 shows these two kinds of estimations for sounds in the maximum frequency difference within which a fused image is obtained versus durations of tone bursts. As seen from the figure, the curve for the total spectral difference is nearly flat independently of the durations, while the curve for cross-correlation greatly depends on the durations of tone bursts.

A binaural fusion of sounds different in frequency components is considered, therefore, to arise from the decision about the total spectral difference between both sounds.
Masking and Its Clinical Use for
Differential Diagnosis of Perceptive Deafness
Isamu Ebihara, M.D.
Associate Professor of the Institute
for the Deaf, School of Medicine,
Tokyo Medical and Dental University

Introduction

Various investigators have studied of the behaviour of masking
and also have employed tone bursts as experimental procedures in
regard to temporal summation effects of the ear and different kinds
of masking. In general, when hearing thresholds are measured by tone
bursts, it is appeared that there is a tendency at thresholds deter-
mination in which variance of thresholds becomes smaller than done
by continuous tones, and also it is found that quantitative relat-
ions between sound stimuli and their hearing responses can be stud-
ied more in detail. I believe, that is why tone burst train is em-
ployed for more close examinations in hearing.

In order to make full use of masking for the differential diag-
nosis of perceptive deafness, I have studied various kinds of mask-
ing phenomena, say, masking due to different noises mixed with sig-
nal tones, masking level difference phenomenon (MLD) and contra-
lateral remote masking (CRM) by using several combinations of burst
train in experiments since about ten years ago. From these investi-
gations mentioned above, it would be clear that some of masking phe-
nomena are able to use in clinical audiometry. Hearing test proce-
dures newly developed could be classified to the following three

--- A-37 ---
Masking and Its Clinical Use for Differential Diagnosis

groups; 1) those which are in use of critical bandwidth level as monaural hearing test. 2) those which are in use of MLD as binaural hearing test. 3) those which are in use of CRM as binaural hearing test.

The findings presented here might be conveniently divided into two parts, basic investigation and use of masking for clinical audiometry. Needless to say, the purpose of basic investigations is to obtain a combination of tone bursts to be useful for hearing test in daily clinic, and studies of clinical utilization of masking aimed at exploring a relation between each of masking phenomena and its physiological backgrounds acted on.

Procedure and Subjects
Subjects: Three groups are used in this study; 1) normal hearing population in which each subject has an air and bone conduction loss no greater than 10dB as determined by routine pure tone audiometry and demonstrated a type I Bekesy audiogram. Ages of the subjects ranged from 25 to 35 years. 2) impaired hearing population with inner ear lesion who demonstrated a type II Bekesy audiogram. 3) the other population who consisted of impaired hearing of conductive origin with surgically removed middle ear muscles and impaired hearing with central auditory pathway lesion at different portions.

Apparatus: A Multi-Dimensional Audiometer 301B (Nagashima Medical Instrument Co., Ltd., Japan) connected with dynamic receivers TR-33 (Toshiba, Japan) was used to conduct pure tone audiometry and monaural or binaural masking threshold measures in case of both basic investigation and clinical utilization of masking. A Multi-Dimensional Audiometer 301B(so-called MD-Audiometer) developed by the author equipped with pure tone oscillator, white noise generator, narrow band noise generator such as about critical bandwidth noise, electronic switch sets with interval timer, phase shifter, mixing circuit through which signal with signal, noise with noise and signal with
noise can be mixed, and attenuators amplifiers used in both monaurally and binaurally hearing test.

The masked threshold measures were collected in a two-room sound treated recording suite.
Procedure: On the basis of previous works on signal detection problem, tone burst trains as shown in Fig. 1 were used. In these burst trains, it is of importance that time delay of more than 250 msec to 500msec between the onset of the masker and the onset of the signal keeps exactly, because signal detectability gets higher with —out a warning light. Due to these presentation of sound, fluctuation of thresholds gets smaller and drops within almost range of 4dB.

Masked thresholds were measured for each subject at 250, 500, 1000, 2000, 4000, and 8000cps in a white noise whose frequency spectrum was 100 to 10000cps and in a narrow band noise whose frequency spectrum was about critical bandwidth. These thresholds were measured for signal durations of 17, 267 and 467msec calculated on DALLOS’s report. Most of all bursts have the both rise and decay time of 25msec. There was a 3000msec repetition time between each signal. The masked thresholds were measured at spectrum level ranged from -15 to 60dB, which correspond to 5dB to 90dB overall SPL.

The method of constant stimuli was used. Measurements were made for each subject in two or three listening sessions. The listening sessions separated by at least a 24 hour period. Each subject was
instructed to listen for the signal tone in a background of noise or another masker in different experimental conditions.

Results and Conclusions:
The results presented here are only remarkable findings and also are arranged on the basis of sound presentations as shown in Fig. 1 for basic investigations. Those are the following:
1) in case of monaural masking; (a) critical bandwidth level (CBL) in dB obtained by tone bursts whose durations are shorter than 200msec gets smaller than by continuous tones, and is almost the same as Fletcher's. (b) there is not the difference of CBL between normal hearing subjects and inner ear impairment. (c) masking effects of white noise (WN) upon hearing threshold shifts were about 5dB smaller than of narrow band noise (NB) in all tested frequency range, but when the same noise as the testing ear is presented to opposite ear, the difference of masking effects between them completely disappears so that CBL gets the smallest or lost in frequency range lower than 500cps.
2) in case of MLD; super-imposed pip(SIP) was employed in binaural hearing. (a) the values of MLD in pure tone get larger in case of NB used as masker than in (WN), and also get larger in shorter duration signal than longer (b) the interaural time difference threshold derived from MLD is about 0.5msec which value is agreed with physiological findings of the tectal auditory neurons in normal subjects (c) the discrimination threshold of interaural time difference of sound lateralization is less than 0.05msec in normal subjects
3) in case of CRM; sound stimuli as shown in Fig. 1-3A and -3B were employed (a) in which procedure, CRM take place at time when the spectrum level of masker in contralateral ear reached at the level of hearing threshold (SPL) of ear listening masker sound, even if the intensity of masker is faint (b) in the procedures as shown in Fig. 1-3A, masked thresholds of CRM increase as the intensity of masker increases, but the rate of increase is different among kinds of masker, namely, pure tone, WN and NB, closely observed, the rate of increase is almost less than 4dB for each 0.01dB increment of masker (c) under the condition as shown in Fig. 1-3B, when signal duration of 17msec is used threshold shifts of CRM remarkably increase in frequency range of lower than 1000cps as the intensity of masker increases only in case of using NB especially, in presenting pure tone to right ear and noise to left ear, the rate of increase is attractive and is not remarkable in vice versa. in white noise used as masker, these findings are barely observed. In case of this, the difference of threshold shifts between NB and WN get larger until above 20dB in normal subjects.

On the basis of the results described above, the procedures already utilized for differential diagnosis of perceptual deafness are as follows: 1)-(a), -(b) used for retrocochlear lesion, 2)-(a) and 1)-(c) for auditory cortex, 2)-(b) for brain stem, 2)-(c) for auditory cortex, 3)-(a), (b) for brain stem and 3)-(c) for dorminancy of brain hemisphere.

Acknowledgement.
I acknowledge with gratitude the readiness and patience of Drs. Y.Hiramatsu, Y.Hashimoto, H.Tsuji, K.Fujimori and M.Motosugi who make it possible to carry out all the experiment. I am also indebted to Ms. F. Funaki and Mr. H. Fujimoto who generously donated their time to assist me in the preparation of this manuscript.
Lateralization of High-Frequency Complex Stimuli
Hisse Sakai and Tsuneo Inoue
Broadcasting Science Research Laboratories,
Japan Broadcasting Corporation, Tokyo, Japan

Introduction

Studies on the lateralization have shown that interaural phase
difference of the pure tone can not be detected at frequencies above
about 1,300 Hz. For complex stimuli, however, the temporal information
processing for lateralization might be different, as contrasted to the
case of the single high-frequency tone(1)(2).

In the present study, experiments are made on the lateralization
of the amplitude modulated tone of high frequency due to the inter-
aural phase difference of the envelope, and processing of temporal
information in the auditory system is discussed with respect to inter-
aural phase difference.

To determine the
degree of lateralization,
complex stimuli and 1,000
Hz tone were presented
alternatively to subjects,
and they were instructed
to match the image of
1,000 Hz tone to that of

Fig. 1. Block diagram of the apparatus.
the complex stimulus by changing the interaural intensity difference of the tone. Accordingly, the degree of lateraliztion is expressed in interaural intensity difference of the 1,000 Hz tone.

**Lateralization of modulated tone containing only high frequencies**

The lateralization of the 3,000 Hz tone, modulated at frequency of 200 Hz, is shown in Fig.2. The lateralizing effect is maximal at 90° of interaural phase difference of the envelope and decreases progressively as phase difference is changed from 90° to 0° or 180°. It is similar to the case of the pure tone at low frequency.

![Fig.2. Lateralization of modulated tone of high frequency.](image)

**Lateralization of the modulated tone of low frequency**

In the case of the 500 Hz tone modulated at 200 Hz, three tone images with different pitches corresponding to the carrier, lower sideband and upper sideband frequency were observed separately. Fig.3 shows the lateralizing effect for each of them as the function of the phase difference of the envelope. In this case, one of images corresponding to two sideband components was lateralized in opposition to the other as shown in Fig.3, presumably due to the different sign in equations expressing the amplitude modulation.
Lateralization of High-Frequency Complex Stimuli

**Influence of modulation frequency and modulation depth**

Fig. 4 shows the lateralizing effect of the modulated tone of high frequency due to phase difference of 90°. Influence of modulation frequency was observed up to 1,300 Hz, with the tendency of the lateralizing effect being decreased as the modulation frequency was increased. The depth of modulation also had influence upon the lateralizing effect. On the contrary, influence of the difference of carrier frequencies was not observed as shown in Fig. 5.

**Influence of high frequency masking noise**

It has been suggested (3) that the neural information for the lateralization of brief sounds came largely from the basal turn of the cochlea. In order to check if this assumption is applicable to the lateralization of modulated tone, masking experiment was made. High frequency noise was added monaurally to the modulated tone with interaural difference of 90°.

Fig. 6 shows that it hardly changes whether the masking noise is presented or not. Therefore, it seems that the basal turn does not contribute essentially to the processing of the temporal information.

**Discussion**
Lateralization of High-Frequency Complex Stimuli

Above experiments show that the interaural difference of slow intensity variation is an important factor to the lateralization even when the modulated tone contains only high frequency components.

As the lateralization of modulated tone of high frequency is not affected by the masking narrow-band noise centered at the modulation frequency, we consider that nervous signals from the part of the peripheral system corresponding to the modulation frequency do not contribute to the lateralization in this case.\(^{(4)}\)

Hence the temporal information for lateralization of high-frequency complex may be transmitted in the form of the temporal variation of nervous pulse train at the part of the peripheral system corresponding to high frequency.

(3) B.H. Deatherage, Binaural interaction of clicks of different frequency content. J. Acoust. Soc. Am. 33, 139-145 (1961)
EIN BEITRAG ZUR THEORIE DES
VORWAERTS - RUECKWAERTS - EINDRUCKES BEIM HOEREN

Jens Blauert
Institut für elektrische Nachrichtentechnik
der Technischen Hochschule Aachen

EINLEITUNG


Für diese Erklärung spricht u.a., daß sich für schmalbandige Geräusche und Dauertöne die Koinzidenz der Richtungen von Schallquelle und Hörereignis statistisch nicht signifikant nachweisen läßt. Die im Folgenden beschriebenen Messungen zeigen jedoch, daß sich auch im Falle der Beschallung mit schmalbandigen Geräuschen Aussagen über die zu erwartenden Richtungen der Hörereignisse machen lassen.

VERSUCHSANORDNUNG

20 normalhörende, männliche Versuchspersonen im Alter von 25 - 30 Jahren wurden mit fixiertem Kopf in einem reflexionsarmen Raum so beschallt, daß die an den beiden Ohren anliegenden Signale identisch waren (BILD 1).

Nachdem in Vorversuchen festgestellt worden war, daß die Hörereignisse der Versuchspersonen sämtlich etwa in der Medianebene des Kopfes und in keinem Falle mehr als 15° unterhalb der Horizontalebene auftraten, wurde zur Beurteilung der Richtung der Hörereignisse die in Bild 2 gezeigte Nominalskala mit den Bereichen "v", "o" und "h" vereinbart. Die Versuchspersonen schätzten, in welchem der Bereiche sich ihr Hörereignis jeweils befand.

ERGEBNISSE

In den Bildern 3, 4 und 5 sind für den Fall der Beschallung von vorne und hinten die Häufigkeiten der einzelnen Beurteilungen in Abhängigkeit von der Frequenz und vom Pegel dargestellt. Für die anderen in Bild 1 gezeigten Beschallungsarten ergaben sich analoge, lediglich im Falle der Kopfhörer stärker streuende Ergebnisse, so daß hier auf eine vollständige Darstellung verzichtet werden kann.

Es zeigt sich eine deutliche Frequenzabhängigkeit der Häufigkeiten der einzelnen Beurteilungen. Unabhängig von der Vorspielrichtung gibt es Frequenzbereiche in denen die Hörereignisse vorwiegend im Bereich "v", andere, in denen sie vorwiegend in den Bereichen "o" bzw. "h" auftreten.

Als Maß für die Unabhängigkeit der Beurteilungen von der Vorspielrichtung mag gelten, daß in 79 % der Fälle beim Vorspielen von vorne die gleiche Beurteilung erfolgte wie beim Vorspielen von hinten mit gleicher Frequenz, wobei zudem bei den Fällen ungleicher Beurteilung alle möglichen Varianten etwa gleich häufig.
BILD 3
Häufigkeit der Beurteilung "v" in Abhängigkeit von Schallpegel und Frequenz.
20 Versuchspersonen wurden pro Terz je einmal von vorne und einmal von hinten beurteilt.

BILD 4
Häufigkeit der Beurteilung "o" in Abhängigkeit von Schallpegel und Frequenz.
20 Versuchspersonen wurden pro Terz je einmal von vorne und einmal von hinten beurteilt.

BILD 5
Häufigkeit der Beurteilung "h" in Abhängigkeit von Schallpegel und Frequenz.
20 Versuchspersonen wurden pro Terz je einmal von vorne und einmal von hinten beurteilt.
vorlagen. Eine Abhängigkeit der Beurteilungen vom Pegel scheint lediglich im Bereich von etwa 5 - 8 kHz vorzuliegen. Dort geht mit steigendem Pegel die Häufigkeit der Beurteilung "h" zugunsten der Beurteilung "o" zurück.

In BILD 6 sind die Ergebnisse über die vier gemessenen Pegel gemittelt und zusammenfassend dargestellt worden. Betrachtet man die jeweils häufigste Beurteilung in einem Frequenzbereich als Maß für eine bevorzugte Lokalisiertheit der Hörereignisse in dem entsprechenden Segment der Mediänebene, so ergeben sich die Richtungen der Hörereignisse, die in BILD 6 oben eingezeichneten Vorzugsbereiche.

Die erhaltenen Ergebnisse legen den Versuch nahe, auch breitbandigere Hörereignisse, wie z.B. Musik oder Sprache, unabhängig von Standort des Lautsprechers in gewünschter Richtung entstehen zu lassen, indem man die jeweils entsprechenden Vorzugsbereiche gegenüber den anderen Frequenzbereichen anhebt. Durch Anhebung der entsprechenden Vorzugsbereiche um 30 dB gelang es dem Verfasser z. B. vorne abgestrahlte Musik hinten erklingen zu lassen und umgekehrt. In ähnlicher Weise erklärt sich die verstärkte "Präsenz", die sich im Tonstudio bei Anhebung der Frequenzen um 3 - 5 kHz zeigt.


LITERATURHINWEISE


Burger, J. F. 1955 Front-back discrimination of the hearing system Acutea 8, S. 301 - 302

Kietz, H. 1953 Das räumliche hören Acutea 5, S. 79 - 85

Klensch, H. 1949 Die Lokalisation des Schalles im Raum Die Naturwissenschaften 36, S. 145 - 146

Tarnoczy, Th. 1958 Über den Vorwärts- Rückwärts-Eindruck Acutea 6, S. 343


Wallach, H. 1946 The role of head movements and vestibular and visual cues in sound localisation Journal of exp. Psych 27, S. 339 - 360

— A - 48 —
Auditory Perception and Blind Guidance.

D. Rowell

Electrical Engineering Department, University of Canterbury, Christchurch, New Zealand.

L. Kay

Professor and Head of Electrical Engineering Department, University of Canterbury, Christchurch, New Zealand.

---

Introduction

This paper examines some of the problems involved in utilising the auditory system as an alternative sensory input through which the environment may be perceived by a blind person. The discussion is confined to the use of an echolocation system, with binaural presentation of the spatial information. In defining the optimum form of the signals at the man-machine interface, it is shown that the natural properties of the human auditory system place severe restrictions upon the form of signal coding that can be used for the perceptual display of more than a single object.

The Perception of Multiple Objects

'Selective attention (the "cocktail-party" effect) (1) enables a person to concentrate on a single source in a field of independent sound sources. When the sounds from two or more sources are correlated, however, this ability disappears. Most echolocation systems utilise a temporal analysis to resolve separate objects at different ranges within the environment. The received signal is the sum of returns from all objects, and range differences are coded as time delays between components of the signal. However, the auditory system is not capable of assimilating information in this form.
The precedence effect in audition is well known (2). This temporal inhibition of all but the first of a rapid succession of stimuli is useful in communication where reverberation is a problem, but it prevents the use of the auditory system for analysis of temporal range information. The small delays between components of the received signal from the immediate environment are such that only the nearest object will be perceived. The perception of all other objects will be inhibited by the auditory system.

In a simple experiment five loudspeakers, all emitting the same sound were arranged in front of a blindfolded subject. His task was to state the number and identify the direction of the loudspeakers he could perceive. Clicks, pulsed tone and white noise were used. All subjects reported that they could only perceive a single source, which was invariably the nearest one. When asked to move blindfolded through the field of loudspeakers subjects were confused, and reported that a slight movement served to change the apparent direction of the loudspeaker perceived.

This experiment served to simulate the perceptual effect of a simple echolocation system in a multiple target situation. It demonstrated that a mobility aid based upon this form of sensory input coding is not suitable for coupling to the auditory system, and that some other signal coding must be used.

The only way to overcome this effect is to present the range information in the frequency domain and to allow the frequency analysis function of the auditory system to resolve the range differences between objects. In other words, the range of objects must be characterized by the frequency of the audio signals, which must therefore be of narrow bandwidth and hence long duration, instead of the short duration signals of the temporal analysis system. Multiple objects will then be represented by a complex chord, and the resolution possible will depend upon the ability to resolve the components of this
Auditory Perception and Blind Guidance.

The linear F.M. mobility aids for the blind (3) (4) present the range information in this manner. The frequency of the audio signals is directly proportional to the range of reflecting objects. This form of output coding not only allows resolution of two or more objects but also gives a readily interpreted measure of the range of these objects.

**Auditory Localization**

The restrictions imposed upon the structure of the stimuli in the above section restricts the possible usage of cues for binaural localization of the auditory image. It is well known that wide bandwidth signals are more easily localised than tones, (especially with regard to front-back discrimination and the determination of elevation). Recent hypotheses have asserted that additional cues in wide bandwidth signals may be provided by diffraction effects around the pinnae (5) (6). With narrow bandwidth sounds the localization must suffer.

For true localization of the auditory image with dichotic presentation, the audio signals should contain the same cues that would result in normal listening, if the reflecting object itself was making the sounds. Although adaptation may allow the use of unnatural cues, these must not be in conflict with the mobility cues derived through the normal auditory system.

In normal hearing the path length differences when a source is displaced from the median plane give rise to the binaural cue of interaural time (phase) difference. If, as explained above, the frequency of the audio signal is processed to be proportional to the propagation path length, any path differences between the binaural receptors must cause a further cue of interaural frequency difference. There will therefore be no constant phase relationship between the binaural stimuli. Phase differences will not be available to the
Auditory Perception and Blind Guidance.

localization mechanism. The predominant localization cue becomes inter-aural amplitude difference at all frequencies. The cues in a binaural system of this type are therefore unnatural but it has been found in experiments using a prototype aid that binaural fusion is not destroyed by the frequency different. It appears that envelope similarity is sufficient to maintain, fusion with a small interaural frequency difference.

In static situations, with the head fixed, the auditory image is lateralized, that is perceived only with "sidedness," and originating from within the head. It has been found by the authors however that with head movement and freedom to move in the environment, the image is easily projected into space, and with some experience with the form of range coding, perceived in its true position.

Object Characterization

Objects can only be characterized, in static situations, by differences in the form of the audio signals, caused by differences in reflecting characteristics of the objects within the environment. Wide bandwidth echolocation signals must be used, so that the characterization information may be impressed as spectral differences in the returned signals. With the form of range coding specified, these differences should be expressed in the audio signals as different envelope forms upon the tonal pulses. The ability to discriminate between objects will depend upon the ability to perceive and recognize the form of this modulation.

References

Pattern of the noise images and the binaural summation of loudness for the different interaural correlation of noise

R.I. Cheruvyak and M.A. Dubrovsky
Acoustics Institute, Academy of Sciences of the USSR, Moscow

In binaural hearing through earphones the acoustic images are perceived inside the head in a volume which may be designated the "subjective" acoustic space. The pattern of the noise images depends on the interaural crosscorrelation of noise fed into two earphones [1,2]. In our work we intended to explore the influence of the regular changes in the interaural noise correlation $r$ on the pattern of the local estimates of sensation in some projection of the subjective acoustic space and also on the summed sensation of loudness.

In the first experiment the subjects were asked to imagine a plane which is a crosssection of the subjective acoustic space in the frontal direction through the crown of the head and the ears. The following convention in presenting such a plane was adopted: a semicircle with a radius 10 cm was drawn on a sheet of paper (Fig.1). The semicircle was divided into small squares. The subjects projected the local estimates of their sensation on that semicircle. First of all the subjects had to determine the position of the maximal sensation and to designate to the corresponding square the number "6". The sensation in the other squares was compared to the maximal one.
Pattern of the noise images

Fig. 1. For details see text.

The results of the experiment showed that (1) the interaurally independent noise produces two images placed near the corresponding ears. The images are clearly localized and become diffuse in a proximity of their margins; (2) the completely correlated noise produces one image in the center of the subjective space. The image also becomes diffuse near its margin. The mixing of the correlated and the independent noises does not change the shape and the position any of three independent images. However the distribution of the local estimates of the noise pattern are changed with a variation of the noise cross-correlation. These distributions for the values of $r$ 0.0, 0.4, 0.85 and 1.0 are shown in Fig. 1. The small squares with the estimates ranged from "4" to "6" are shaded by the narrowly displaced lines; the squares with the estimates ranged from "2" to "4" are shaded by the widely displaced lines and the squares with the estimates ranged from "0" to "2" have no shading at all.

In the second experiment the subject had to compare the loudness of the binaurally presented noise for the different values of the
interaural crosscorrelation $r$. The results are shown in Fig. 2. The horizontal axis represents the values of $r$ and the vertical axis represents the ratio $P/P_0$, where $P$ is a sound pressure of the noise with a variable value of $r$ and $P_0$ is a sound pressure of the noise with $r = 1$. For the given set of measurements the value of $P_0$ was kept constant. In Fig. 2 the different marks indicate the data pertaining to three subjects. One can easily see that the noise with a value of $r \neq 1$ is equally loud to the noise with $r = 1$ when $P = P_0$.

Consequently our data show that the binaural summation of loudness does not depend upon the interaural correlation of the noise. The simple algebraic manipulation allows to gain that in the case of the coherent summation of the noises from the left ear and from the right ear the binaural loudness of noise $L$ must depend on $r$ the following way:

$$L = L(\sqrt{2} \cdot P \cdot \sqrt{1+r})$$

(1)

Here we have taken into consideration the formulas for the determination of $r^2[3]$. The expression (1) implies that for the coherent summation the loudness of the unity correlated noise shall be equal to the loudness of the partly correlated noise provided that the
sound pressure $P$ at both ears will be changed with $r$ as follows:

$$\frac{P}{P_0} = \sqrt{2}/\sqrt{1+r}$$  \hspace{1cm} (2)

Likewise the calculation for the case of incoherent summation

/which we prefer to call "statistical" summation/shows that instead

of the expression (2) we have got the following expression

$$\frac{P}{P_0} = 1$$  \hspace{1cm} (3)

The solid curve in Fig. 2 mirrors the expression (2); the dashed

curve in Fig. 2 mirrors the expression (3). A comparison of the cal-

culated and the experimental data indicates that there is actually

the statistical summation of the coherent components of the noise

from the left and from the right ear.

This findings can be described quite well by the following model.

The temporal interaural relations are coded by the spatial relations

at low level of the auditory system. These spatial relations are

reflected in the pattern of noise images described at the beginning

of our paper. Due to a loss of the fine temporal structure after such

a coding the way of the summation turns to be essentially the same

regardless of the interaural crosscorrelation at the input of the

auditory system.

References

1. J.C.R. Licklider. Influence of Interaural Phase Relations upon

the Masking of Speech by White Noise. J. Acoust. Soc. Amer.,

20, N. 2, (1948).

2. H.C. Blodgett, W.A. Wilbanks and L.A. Jeffress. Effect of Large In-
teraural Time Differences upon the Judgement of Sideness.

3. L.A. Jeffress and D.E. Robinson. Formulas for the Coefficient of
Interaural Correlation for Noise. J. Acoust. Soc. Amer.,
Control of Psychophysical Experiments with an On-Line Digital Computer

C. H. Coker, J. L. Hall and A. E. Rosenberg

Bell Telephone Laboratories, Incorporated

A moderate-size high-speed digital computer is being operated on-line to supervise psychophysical experiments. The computer performs the usual tasks of a human operator—presenting stimuli and recording stimuli and responses. In addition, it performs tasks that would be impossible for a human operator. It can generate stimuli and control complex equipment at high speeds and its computational ability makes possible the use of sophisticated procedures for on-line analysis of subject responses. This paper describes the facility and two types of experiments that are being performed.

The computer, a Honeywell DDP-516, has a memory of 8000 16-bit words and a cycle time of .96 microseconds (see figure). Add and subtract times are 1.92 μsec; multiply and divide times are 5.84 and 11.52 μsec. It is equipped with a typewriter, paper tape punch and reader. A magnetic tape drive and a high-speed disc storage unit have been installed locally. The disc has a capacity of 394,000 words, a transfer rate of 196,000 words/sec, and maximum access delay of 33 msec. An oscilloscope display is driven from a track on the disc.

Digital and analog channels provide communication with subjects and external equipment. Two 16-bit digital output channels switch stimulus generation equipment and signal subjects. Two 16-bit digital input channels sense the status of equipment and subject responses. A 160-kHz 15-bit analog-to-digital converter and 40-channel multiplexer are provided for analog signal inputs. Digital-to-analog converters are also provided; one of these drives a 16-channel analog multiplexer with hold circuits. Three real-time clocks are provided. Up to four
analogue stimulus outputs can be obtained from 14-bit converters with combined sampling rates up to 200 kHz and controllable relative time delays resolvable to 1 µsec.

The computer is provided with a Fortran IV compiler, and Fortran-callable subroutines have been written for all equipment. As a result, the task of writing programs to implement experiments is considerably facilitated.

One use of this facility concerns general methods in the determination of psychometric functions - functions relating some aspect of a stimulus to the probability of a given response by the subject. Since determination of these functions involves the exploration of regions of uncertainty in subject response, a central problem for experimental efficiency is to keep the stimulus within the range where uncertainty does exist. This requires, however, some estimate of the result of the experiment before the experiment is concluded.

One approach to this problem is the class of sequential procedures in which one or several prior responses determines the direction of change in the stimulus in some manner so as to converge to the region
of uncertainty. We have developed a procedure, in which the computer is essential, that obtains a maximum likelihood estimate of the psychometric function after each response and uses this estimate to determine the placement of the next stimulus.

The maximum-likelihood estimate of the psychometric function is determined as follows: First, the experimenter must assume some parametric form for the psychometric function - cumulative normal from zero to one, for example. The program selects a set of response-curve parameters and computes the probability that the stimuli that have been presented would have elicited the obtained responses. Then it selects another set of values and computes another probability. The search is continued in a systematic way until the parameters of the psychometric function are found for which the probability of the actual responses is a maximum. The resulting curve is taken as an estimate of the actual psychometric function and is used to determine the placement of the next stimulus.

It is apparent that if the experimenter assumes an incorrect form of the psychometric function, this procedure will give incorrect results. However, this procedure allows the experimenter a great deal of freedom in selecting the form. In addition, the experimenter is free to give whatever weighting he chooses to previous responses. If the subject's performance is known to be stable, the maximum-likelihood estimate can be based on all previous responses. If the subject's performance is suspected to be unstable, the estimate can be based on only the more recent responses. Changes in the subject's performance will then be tracked more rapidly.

The amount of computation involved is such that this procedure would be completely impossible without an on-line computer.

The facility has also been applied to the subjective evaluation of speech. Most synthesis or processing of speech at Bell Laboratories is done by large-scale computer off-line. This presents a problem for the subjective evaluation of such material since tests must be completely recorded. In an A-B comparison test, the number of comparisons increases as the square of the number of samples in the repertoire. With just a moderate number of samples in the repertoire such a test can be quite long and cumbersome, both to prepare and to run. One can conceive of tests in which the number of comparisons can be minimized by allowing the sequence of comparisons to be a function of the history.
of the subject's responses. With a prerecorded test this strategy is, of course, impossible, since the sequence must be fixed in advance.

A system has been devised using the DDP-516 computer that facilitates the preparation and running of A-B comparison tests and provides for sequential strategies. The repertoire of speech samples is stored so that specified samples may be easily fetched and reproduced under computer control. The storage medium is ordinary magnetic tape. This medium provides much more storage capacity than the disc, and faster access than digital tape. The system utilizes two ordinary tape reproducers (AMPEX AG-350) each with an identical recording of the sample repertoire. The beginning of each sample is marked by a tone pulse on the second track of the tape. A shaft encoder coupled to the shaft of each supply reel encodes the angular position of the reel into a binary word that can be fed into the computer. By sweeping through each reel and sensing the marker pulses it is possible to store an address for each sample on the tape. Tape motion is controlled by relays at the computer output. Given the address of a sample and taking into account the dynamics of the tape motion, the system can fetch the sample and reproduce it for a listener. Fine positioning of the sample is achieved by sensing the marker pulse. Two tape reproducers are used to eliminate waiting time for fetching a sample. While a sample is being reproduced on one machine the next sample is being fetched on the other machine. The only constraint is that the longest time required to fetch a sample must be less than the sample duration plus the duration of adjacent preset interstimulus intervals.

We have described two applications of an on-line computer to the control of psychophysical experiments. The computer facilitates the control of experiments and makes it possible to modify experiments as a function of subjects' responses.
Vocal Compensation for Change of Distance

Richard M. Warren

Professor of Psychology at the
University of Wisconsin-Milwaukee

Introduction and Purpose

This experiment was designed to determine if individuals can control the intensity of self-generated sound so as to compensate for changes in distance from themselves to a target microphone. Of course, it is not necessary that S be aware of amiable to describe the physical principle governing the change of intensity with distance. It is the type of perceptual and motor skill which Polanyi has called "tacit knowledge." If S functioning as a sound source could compensate accurately for the effect of distance upon intensity, then his sound pressure level should vary inversely as the distance to the target, and the corresponding sound energy inversely as the square of the distance. Thus, for halving distance, S if accurate, would decrease intensity by 6 db (corresponding to 1/2 sound pressure level and 1/4 sound energy).

The existence of this ability to compensate for intensity on the basis of distance has some interest if considered as an isolated phenomenon, but the main reason for undertaking this experiment is to help resolve a controversy concerning the relation between spatial localization and subjective loudness of sounds.

The physical correlate theory of sensory intensity states that judgments of subjective magnitudes are based upon experience with physical magnitudes associated with changes in stimulation (Warren, 1958). Subjective relative loudness judgments
of sounds, whether externally generated or self-produced, are considered to be based upon familiarity with the manner in which intensity varies with distance from the source.

In agreement with this theory, it was demonstrated by Warren, Sersen, & Pores (1958) and confirmed by Stevens & Guirao (1962) that, with externally generated sounds, judgments of half loudness were quantitatively equivalent to estimates of the attenuation produced by doubling distance from listener to sound source. Further, if care was taken to minimize cues indicating the fixed position of the loudspeaker used to deliver sounds at different intensities, then 6 db was selected correctly for doubling distance and also for half loudness (Warren, Sersen, & Pores, 1958).

I have extended this theory to judgments of "autophonic" level of self-generated sounds. Lane, Catania, & Stevens (1961) found that Ss' judgments of half autophonic level of the vowel [a] corresponded to a decrease of 6 db (which would agree with the physical correlate theory for loudness). However, they concluded that despite instructions to judge the sound they produced, Ss based their responses upon "muscular effort," since the same function was obtained when masking noise (delivered through headphones) obscured Ss' hearing of their own vocal productions. I questioned their conclusion that hearing was not employed in estimating the level of one's own vocal production, and had Ss judge half loudness for three different types of self-generated sounds (the vowel [a], the unvoiced consonant [s], and a note produced by blowing on a pitch pipe) each involving a different mode of production and different sets of muscles (Warren, 1962). As anticipated, no significant difference existed between the attenuations produced for half subjective loudness with any of these sounds. The values obtained in this study are given in Table 1.

I suggested that while muscular effort might serve as a surrogate if hearing were blocked, autophonic judgments were basically judgments of the level of sound produced, and that half autophonic level was equivalent to estimates of the intensity needed to project one's voice for half the distance. In order to verify this theoretical equivalence, the present experiment employed the same sounds used for
half autophonic level (Warren, 1962), but instructions were changed from the subjective judgment of sensory magnitude to estimates of the physical effect of distance upon self-generated sounds.

Procedure

Forty students taking the introductory psychology course served as Ss. They were informed by printed instructions that the purpose of the experiment was to determine how accurately they could compensate for the effect of distance upon the intensity of sounds which they would produce. They were seated at the end of a long table, and saw a microphone alongside a tape measure which extended from where they were seated. They were informed that the distance to the microphone was ten feet. They were to produce a sound, holding it steady for a few seconds. Then they were to pause, and take a breath while the experimenter moved the microphone to a position at a point five feet from them. Then they were to produce the same sound at a different intensity, so that the sound reaching the microphone would be the same as when it was in the further position. In other words, they were to compensate for the change in distance from ten to five feet.

The sound was actually monitored by a Bruel and Kjaer sound level meter located six feet from S. Half Ss started with [4] and half with [5], each S then producing the second speech sound. All Ss produced as their last sound the note of F above middle C on the same pitch pipe used by Warren (1962). Three successful productions of paired distance estimates (i.e. the level held within ±1 db of a central value for 3 sec, both with the target at 10 ft and at 5 ft) were obtained for each of the three sounds. In order to minimize the influence of ambient noise in the testing room (a converted class room) an octave band filter was used in conjunction with the sound level meter (16 db/octave fall off) with the center frequency set at 1000 Hz for [4] and 2000 Hz for the other sounds.

Results and Discussion

The results obtained for intensity compensation for halving distance are shown in Table 1. Due to the skewed nature of the response distributions, medians are the statistic of choice.
### TABLE 1

<table>
<thead>
<tr>
<th></th>
<th>Half Distance Compensation Median</th>
<th>Half Distance Compensation Mean</th>
<th>Half Loudness Median</th>
<th>Half Loudness Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vowel [æ]</td>
<td>7.1</td>
<td>7.4</td>
<td>7.0</td>
<td>7.3</td>
</tr>
<tr>
<td>Unvoiced consonant [j]</td>
<td>5.0</td>
<td>5.8</td>
<td>7.0</td>
<td>7.4</td>
</tr>
<tr>
<td>Pitch pipe note</td>
<td>8.3</td>
<td>8.3</td>
<td>8.0</td>
<td>8.0</td>
</tr>
</tbody>
</table>

*Data from Warren (1962)*

It can be seen that values for half distance compensation were close to the correct value of 6 db for all three sounds. A $X^2$ test showed that differences from 6 db were not significant for any of the sounds. Table 1 also shows the values reported earlier (Warren, 1962) for half loudness of self-generated sounds. These values also show no significant differences from 6 db using $X^2$.

The close agreement of half subjective levels with estimates of the physical change produced by halving distance is further evidence for the physical correlate theory of sensory intensity. Together with other evidence (see Warren, 1963 for a review), this experiment indicates that individuals can estimate accurately the change in intensity produced by changing distance, whether while acting as passive listener or sound generator, and that loudness judgments, whether involving self-generated or external sounds, are based upon the physical correlate of distance.

### References


### Acknowledgment

The work reported here was supported by the National Institutes of Health, Grant Number NB 05998.

---

---
Possibilities for the Objective Measurement of the Loudness Level of Speech.

Pedersen, O. Juhl, and Pedersen, S. Boel.

The Acoustics Laboratory, Technical University, Lyngby, Denmark.

Some preliminary investigations on the possibilities of using the indication of a number of different "instruments" as a measure of relative loudness level of speech are reported.

Such measurements are important in some branches of acoustics, e.g. in connection with audiometry and intelligibility measurements. In many cases a knowledge of the absolute value of the loudness level is, however, not necessary as the purpose of the measurements is often to make it possible to keep the loudness level or loudness of different speech samples approximately constant.

"Loudness level" measurements are often carried out using the VU-meter [1] or the "peak"-level meter [2] available at most broadcasting and recording organisations for programme level monitoring, even if these instruments are developed for overload protection and not for loudness level measurements.

The speech samples. It was assumed that the loudness level of speech is judged not only from the intensity of the phonemes but also from the intervals between words and syllables. For that reason two texts (in Danish) were chosen.

Text A was a short passage (15 words) from a political essay read fluently.

Text B was a short dialogue (18 words) from a play read with three pauses.

Both texts were recorded by three male speakers with voices subjectively judged as different in pitch and quality but within "normal" range. The duration of the presentations was 4,8 - 6,8 s.

Measurements of loudness level. Each of the six speech samples was presented via headphones to listeners who had to compare the loudness of the speech with that of a reference signal (1/3 octave noise centered at 1 kHz). Each speech sample was presented alternating with 3,5 s noise reference signals adjusted to a sound pressure level of 73 dB/2·10^{-5}N/m^2. (Pauses between the noise and the speech 0.5 seconds). Each speech sample was presented at six levels 1 dB apart, approximately
centered on the equal loudness point of the listener. The order of presentation of the levels was chosen at random and each level was offered three times during a single measurement. The listeners had to decide whether the speech signal or the noise was the loudest. They were allowed to listen to the signals as long as they wanted; the total time for a single measurement (18 decisions) was 10-15 minutes. From the decisions of each subject the attenuation of a speech sample necessary to give equal number of judgments "speech louder than noise" and "noise louder than speech" was taken to express the subjects assessment of a loudness level of 73 phons for that sample. 25 subjects, age 19-26 years, with normal hearing were used.

The results showed a spread of attenuator settings over a 20 dB range for each sample, but the samples were rated according to loudness in the same order by nearly all subjects. The probability density function of the attenuator settings apparently showed a small secondary maximum about 10 dB below the main maximum. This seems to indicate that a small number of listeners judges the loudness level of the samples some 10 dB below the judgment of the majority.

It was realized that the resulting standard deviation of the equalization of the levels of the speech samples would be smaller, if the lower judgments were excluded, but no satisfactory apriori reasons for an exclusion could be found. So all results were pooled to give a mean value and a standard deviation of the attenuator settings for each speech sample. Using these mean attenuator settings for the equalization of the samples each sample was adjusted to 73 phons with a 95% confidence interval of ± 2 phons.

The instruments used for the measurements on the levelled speech samples:

1) VU-meter; basically a moving coil instrument preceded by a full wave rectifier the output of which is proportional to the 1.2nd power of the input voltage. The averaging time of the instrument is 100-300 ms (for the instrument used in the measurements described here 165 ms). Calibrated in VU's in the range ± 3 to -20.

2) "Peak"-level meter; a light spot meter preceded by a logarithmic amplifier and a full wave rectifier. The averaging time is much lower than the "decay" time. For the instrument in question 3 ms and 1.5 s respectively. Calibrated in dB; scale approximately linear over + 5 to -40 dB.

The VU- and "peak"-level meters were read by two skilled observers.

3) Cathode ray oscilloscope preceded by a rectifier and possibly an integrating RC network. The traces were recorded photographically.

4) AFL-meter. An instrument specially developed to measure the average peak level of speech. The reading of the instrument is derived from a measurement of the probability density function of the logarithm of those speech amplitudes exceeding a certain threshold [7].

5) Bruel & Kjaer level recorder, type 2305.

Objective measurements

Using the mean value of the attenuator settings found as described above the
relative levels of the six speech samples were measured in the following ways:

1. With a VU-meter using: a) readings of greatest deflections
   b) readings of the average of peak deflections
   (in both cases after exclusion of a few "deflections of unusual amplitude")

2. With a "peak"-level meter read as above

3. With a cathode ray oscilloscope preceded by
   a) a linear full wave rectifier; readings of
      1) maximum amplitude
      2) average amplitude
   b) a square law rectifier; reading average amplitude
   c) a square law rectifier and RC network with
      1) $\tau = 35$ ms, German proposal for an impulse sound
         level meter, DIN 45 633
      2) $\tau = 100$ ms, proposal by Zwicker and Feldtkeller [5].
      In both cases maximal and average amplitudes were measured.

4. With an APL-meter [4].

5. With a level recorder with the settings:
   a) Rectifier response: RMS, writing speed: 25 mm/s, 50 dB potm..
      Readings of average levels
   b) Rectifier response: EMG, writing speed: 1000 mm/s, 10 dB potm..
      Readings of peak levels

Averaging time according to [6]: a: 5 s, b: 5 ms.

Results. For each type of measurements the difference $\Delta$ between the highest and lowest level measured on the subjectively balanced samples is calculated and listed in Table 1 in an order corresponding to increasing magnitude. (The values are rounded off to the nearest full dB (or VU)). Some of the results are also shown in fig. 1.

Best results are obtained with the peak indicating oscilloscope and the "peak"-level meter. The VU-meter shows a relatively poor response as do most of the instruments tested having a large integrating time. It must, however be borne in mind, that these results are based on a limited number of measurements of a preliminary nature only.

It looks, as if the majority of listeners judges the loudness of the speech samples from the peak value of the sound pressure level. As mentioned above a small group judged the loudness approximately 10 phons lower which in fact corresponds fairly well to the mean level difference found between peak and average sound pressure levels of the samples.

The authors acknowledge Mr. P.E.Lyregaard, M.Sc. who made most of the measurements described.
### Table I.

<table>
<thead>
<tr>
<th>Rating No.</th>
<th>Instrument</th>
<th>Deflection depending on</th>
<th>Reading</th>
<th>( \Delta )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Oscilloscope</td>
<td>(</td>
<td>p</td>
<td>)</td>
</tr>
<tr>
<td>2</td>
<td>&quot;Peak&quot; level meter</td>
<td>(</td>
<td>p</td>
<td>, t_a = 3 ms )</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Average peak</td>
<td>5 -</td>
</tr>
<tr>
<td>3</td>
<td>Oscilloscope</td>
<td>( p^2 )</td>
<td>Average</td>
<td>6 -</td>
</tr>
<tr>
<td></td>
<td></td>
<td>( p^2, t = 100 ms )</td>
<td>&quot;-&quot;</td>
<td>6 -</td>
</tr>
<tr>
<td>4</td>
<td>Oscilloscope</td>
<td>( p^2, t_a = 35 ms )</td>
<td>Peak</td>
<td>7 -</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Average</td>
<td>7 -</td>
</tr>
<tr>
<td></td>
<td>Level recorder</td>
<td>( p^2, t = 5s )</td>
<td>&quot;-&quot;</td>
<td>7 -</td>
</tr>
<tr>
<td>5</td>
<td>APL-meter</td>
<td></td>
<td></td>
<td>9 -</td>
</tr>
<tr>
<td>6</td>
<td>VU-meter</td>
<td>( p^{1,2}, t_a = 165 ms )</td>
<td>Average peak</td>
<td>9 VU</td>
</tr>
<tr>
<td></td>
<td>Level recorder</td>
<td>( p^2, t_a = 5 ms )</td>
<td>Peak</td>
<td>9 dB</td>
</tr>
<tr>
<td>7</td>
<td>VU-meter</td>
<td>( p^{1,2}, t_a = 165 ms )</td>
<td>&quot;-&quot;</td>
<td>11 VU</td>
</tr>
</tbody>
</table>

### Dev. f. av. val.

<table>
<thead>
<tr>
<th>dB</th>
</tr>
</thead>
<tbody>
<tr>
<td>+4</td>
</tr>
<tr>
<td>0</td>
</tr>
<tr>
<td>-4</td>
</tr>
</tbody>
</table>

**Fig. 1.** Deviations from av. readings on levelled speech samples.


Zur Messung des Effektivwertpegel und der Lautstärke von Logatomen

Brose, O. und R. Wehmann
Forschungsinstitut des Fermmeldetechnischen Zentralamtes des Deutschen Bundespost

1. Einleitung


2. Mögliche Verfahren zur Pegelmessung


Die Messung des Spannungspegels bereitet insofern Schwierigkeiten-


3. Pegelmessung mit Hilfe eines Pegelschreibers

Die Amplituden der Pegelschreiberaufzeichnungen sind abhängig von der Schreibgeschwindigkeit des Schreibstiftes. Nach der IEC-Empfehlung 179 soll die Schreibgeschwindigkeit eines Pegelschreibers zur Aufzeichnung des Effektivwertes 100 bis 160 dB/s betragen. Die Maximalwerte der so aufgezeichneten Amplituden liefern eine Verteilung $h$ der maximalen Effektivwertpegel $L_{\max}$ (Bild 1). Der Mittelwert der Verteilung ist der mittlere maximale Effektivwertpegel.


Die Häufigkeitssumme $H$ ist keine Gerade, d.h. die Pegel sind nicht normalverteilt, wie man es hätte erwarten können, da der Sprecher alle Logatome mit annähernd gleicher Lautstärke, genauer gesagt Sprechenergie, auf das Tonband aufgesprochen hat. Der Verlauf von $H$ deutet auf ein Mischkollektiv.
Zur Messung des Effektivwertpegel und der Lautstärke von Logatomen

hin. Die Erklärung hierfür liegt darin, daß die Maximalwerte der Logatomepegel im wesentlichen durch die Intensität der Vokale bestimmt werden und daß die verschiedenen Vokale, auch wenn sie mit gleicher Energie gesprochen werden, verschiedene Pegel haben. Eine getrennte Auswertung der Logatome mit den Vokalen i, u, e, o und a ergibt die Verteilung h(i), h(u), h(e), h(o) und h(a). Die Teilkollektive mit den Vokalen e, o und a sind jetzt normal verteilt. Den höchsten Pegel haben die Logatome mit dem Vokal a, dann folgen die Logatome mit o, e, u und i. Diese Reihenfolge haben auch die von verschieden Autoren angegebenen Pegel gesprochener Vokale. Die Teilkollektive der leisern Vokale u und i sind nicht normalverteilt, da hier die stimmhaften Konsonanten häufig schon den maximalen Logatomepegel bestimmen und so wiederum ein Mischkollektiv entsteht.

Die maximalen Pegel der Logatome erstrecken sich über einen Bereich von 67 bis 87 dB re $2\cdot10^{-4}$/ubar. Der mittlere maximale Effektivwertpegel aller Logatome beträgt 78,5 dB re $2\cdot10^{-4}$/ubar.

4. Pegelmessung mit Hilfe des Pegelhäufigkeitszählers

Schaltet man an den Pegelschreiber einen Pegelhäufigkeitszähler an, der die aufgezeichneten Amplituden mit einer möglichst hohen Abtastfrequenz abtastet und die Abtastergebnisse in Zählwerken speichert, die den verschiedenen Amplitudenstufen zugeordnet sind, so erhält man eine Pegelverteilung, die nicht nur die Maximalwerte sondern den gesamten Pegelverlauf der Logatome berücksichtigt. Wir haben diese Messung mit einem Häufigkeitszähler mit 10 Pegelklassen a 5 dB und einer Abtastfrequenz von 10 Hz (das sind etwa 10 Abtastimpulse pro Logatom) durchgeführt. Es ergaben sich Pegelwerte im Bereich von 53 bis 86 dB re $2\cdot10^{-4}$/ubar, der Mittelwert lag bei 64 dB re $2\cdot10^{-4}$/ubar, also um 14,5 dB unter dem mittleren maximalen Effektivwertpegel.


Die Frage, ob der aus den Pegelschreibungsaufzeichnungen resultierende mittlere maximale Effektivwertpegel als Effektivwertpegel der
Zur Messung des Effektivwertpegel und der Lautstärke von Logatomen

Logatome betrachtet werden kann oder ob auch die kleinen Logatomepegel - wie bei der Messung mit dem Häufigkeitszähler - zu berücksichtigen sind, wird durch die im folgenden beschriebenen Messungen der Logatomlautstärke beantwortet.

5. Messung der Logatomlautstärke

Zur Bestimmung der Lautstärke verglichen 11 Versuchspersonen, die mit Silbenverständlichkeitsevaluationen vertraut waren, in 3 Messreihen die Logatome mit einem in der Amplitude regelbaren 1 kHz-Ton, der mit Hilfe eines Sprachdetektors während der Pause zwischen zwei Logatomen über Kopfhörer dargeboten wurde. Der 1 kHz-Ton war der konstant gehaltenen Logatomlautstärke anzulegen. Die Logatomlautbreite wurde wiederum auf dem Fernsprechband von 300 bis 3400 Hz begrenzt.


Damit ist der aus den Pegelschreiberaufzeichnungen ermittelte mittlere maximale Effektivwertpegel als Effektivwertpegel der Logatome anzusehen. Der mit dem Pegelhäufigkeitszähler gewonnene, um 14,5 dB niedrigere Pegel, der auch die Nebenmaxima der Logatome berücksichtigt, kann nicht als Effektivwert bezeichnet werden.
Some Proposals on Realizing the Ideal Hearing Aid

Mitsuhiro Katakura and Tomio Yoshida
Matsushita Communication Industrial Co., Ltd.
Yokohama, Japan
Michinari Okamoto
Kanto Rosai Hospital
Kawasaki, Japan

1. Introduction

Since the main purpose of the hearing aid is to make the wearer understand conversation clearly and comfortably, it seems to be a first step on realizing the ideal hearing aid to clarify the relation between articulation comfortableness as the psychological properties aroused in the wearer and frequency response, amplification as the physical properties of the hearing aid. This sort of research has been carried out by various researchers, but the fixed opinions have not been found out yet, especially for the perceptive deaf.

Two experiments were carried out based on these considerations.

2. Experiment

(a) Apparatus

The schematic diagram is shown in Fig. 1. The speech sound reproduced by a tape recorder was received in a dynamic receiver through the transmission system. The over-all frequency response of this system was in ±2dB at the range of 200 - 7000Hz. The frequency equalizer indicated with dotted line in Fig. 1 was provided only in experiment I.

(b) Method

Experiment I: The relation between articulation score and frequency response was studied according to the audiogram (American Standard Specification) of examinees. 43 perceptive deafs were tested as examinees.

Articulation score was measured by the articulation test with 50 syllables
Some Proposals on Realizing the Ideal Hearing Aid

through "flat system" and "mirror system". Flat system means the transmission system without the frequency equalizer in Fig. 1. Mirror system does that with the frequency equalizer which has the frequency response to supplement hearing loss. Improvement of articulation score caused by adopting mirror system in place of flat system is called "mirror effect". Fig. 2 shows some examples of audiogram and corresponding equalizer frequency response.

Experiment 2: The relation between comfortableness and sound volume was tested. Comfortableness and uncomfortableness were measured based on examinees' judgements on a short sentence presented by an ascending method. 29 perceptive deafs were tested here.

3. Results and discussion

(Experiment 1)

Fig. 3 shows mirror effect as classified by hearing loss which was calculated by Japanese Standard Method \( a + 2b + c/4 \). o mark indicates the articulation score obtained through the flat system, x mark does through the mirror system. The dark column of the histogram shows the mean value of the articulation scores through the mirror system, and the white one does through the flat system. Articulation score dotted here is the maximum one. Examples without mirror effect are not illustrated here. The ratio between examples with mirror effect and the ones without mirror effect comes to 20/43, 23/43 respectively which appear to be nearly equal. But, further
Some Proposals on Realizing the Ideal Hearing Aid

analysis tells us that 16 examples remain after eliminating two unnecessary factors: one is examples having a level shaped audiogram in which no mirror effect was seen, the other is examples having higher articulation score over 80% in the flat system which are considered not to make troubles for the normal daily conversation. 15 out of 16 examples showed mirror effect. Besides most examples could obtain articulation score over 60% through the mirror system. In the flat system, the score was widely scattered at the range of 40 - 80%.

(Experiment 2)

The relation among speech sound levels which provide comfortableness uncomfortableness and the maximum articulation score is denoted in Fig. 4 by three straight lines which mean regression lines calculated by the method of least squares. According to Fig. 4, it seems to be proper to discuss the relation through classification of hearing loss into two ranges.

(a) 0 - 40dB: In near-normal or mild deaf, sound level providing comfortableness is nearly equal to that providing the maximum articulation.

(b) 40 - 80dB: In case of serious or severe deaf, it is found out that in order to get the maximum articulation, the sound level must exceed the sound level of comfortableness, especially for severe deaf, occasionally rather uncomfortableness.

Fig. 3 The mirror effect as classified by hearing loss

Fig. 4 The relation among sound levels providing comfortableness (MCL), uncomfortableness (UCL), and maximum articulation score (MAL). I mark indicates 95% confidence interval.
4. Conclusion

(1) Mirror effect was recognized through the transmission system with wide frequency range.

(2) Most deafs with articulation score ranging 40 - 80% in the transmission system with over all frequency response (200 - 7000Hz) within ±2dB showed mirror effect.

(3) Most deafs obtained over 60% articulation score through the mirror system.

(4) Speech sound level providing comfortableness was nearly equal on an average to that providing the maximum articulation score for the near normal, mild deafs with hearing loss under 40dB.

(5) For the serious or severe deafs, with hearing loss over 40dB, the maximum articulation score was obtained when it exceeded the level of comfortableness on an average. Especially, the severe deaf over 60dB hearing loss, it exceeded over uncomfortable level.

5. Reference

(1) M. Katakura, T. Yoshida, M. Okamoto, Reports of the '66 Autumn Meeting of the A.S.J.

(2) M. Katakura, T. Yoshida, M. Okamoto, Reports of the '67 Spring Meeting of the A.S.J.

(3) H. Davis et al., "Hearing Aid".
Identification of Synthetic Vowels in Perceptive Deafness
Masao Tomisawa, Ichiro Kirikae, Tetsuya Shitara and Tsunemasa Sato
Department of Otorhinolaryngology, Faculty of Medicine,
University of Tokyo
Hiroya Fujisaki
Faculty of Engineering, University of Tokyo

Introduction: Interpretation of test results of conventional speech audiometry is often difficult since the test sounds are recorded human voice, the acoustic structure of which can not be controlled. The present authors employed synthetic vowel sounds in speech audiometry in order to clarify the mechanism of speech perception in perceptive deafness. Synthetic vowel sounds are ideal for this purpose because their acoustic properties can be exactly specified and easily controlled.

Stimuli: In this study, synthetic vowel sounds were generated by a terminal analog speech synthesizer with controlled variables of F0, F1, F2, F3, F4 and the higher pole correction. Figure 1 shows the block diagram of the synthesizer. The first and the second formants were used as variable parameters, while the third and the fourth formants were held constant at 2,700 Hz and at 3,500 Hz, respectively. The fundamental frequency was set at 140 Hz.

Fig. 1—Simplified diagram of the test instrumentation.
Kr4: Higher pole correction, R: Radiation characteristics, F1-4: Formant circuits.
Identification of Synthetic Vowels in Perceptive Deafness

Table 1.—The first and the second formant frequencies used for the synthetic stimuli.

<table>
<thead>
<tr>
<th>F1</th>
<th>200</th>
<th>250</th>
<th>300</th>
<th>350</th>
<th>400</th>
<th>450</th>
<th>500</th>
<th>575</th>
<th>650</th>
<th>750</th>
<th>850</th>
</tr>
</thead>
<tbody>
<tr>
<td>F2</td>
<td>750</td>
<td>850</td>
<td>1000</td>
<td>1200</td>
<td>1400</td>
<td>1650</td>
<td>1900</td>
<td>2200</td>
<td>2550</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

A set of 87 synthetic vowel stimuli was chosen by the combination of F1 and F2 variables shown in Table 1. The duration of the test sound was 800 msec, and the rise time constant for pitch and intensity was approximately 20 msec. These sounds were recorded on tape in random order and were presented to subjects through a single earphone in a sound proof room. The intensity of sound stimuli was kept at a level over the threshold appropriate for each subject. The subjects were forced to identify each test sound as one of the five Japanese vowels, /a/, /i/, /u/, /e/ and /o/.

Subjects: Forty cases of bilateral perceptive deafness were selected for the present study. Diagnosis of perceptive deafness had been made by means of conventional audiometry prior to the experiment. These cases ranged in age from 11 to 70. For the purpose of comparison, a control group consisting of ten female students with normal hearing was also examined.

Results: It is revealed that subjects of the control group show agreement over 85% in identification for 51 out of 87 test sounds.

The response contours for these 51 sounds are given in F1 - F2 diagram as five areas corresponding to the Japanese vowels as shown in Fig. 2.

In cases of perceptive deafness, however, identification patterns for these 51 test sounds are often found to deviate from the normal pattern. Based on the pattern of
Fig. 5—F1-F2 diagrams and pure tone audiograms in cases of perceptive deafness. 
(a): Type O, (b): Type 1, (c): Type 2 and (d): Type 5.

--- A-79 ---
Table 2---Comparison of articulation scores of normal and impaired groups for synthetic vowels and natural speech syllables.

<table>
<thead>
<tr>
<th>Number of subjects</th>
<th>Articulation scores</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>51 synthetic vowels</td>
</tr>
<tr>
<td>Normal</td>
<td>10</td>
</tr>
<tr>
<td>Type 0</td>
<td>15</td>
</tr>
<tr>
<td>Type 1</td>
<td>8</td>
</tr>
<tr>
<td>Type 2</td>
<td>10</td>
</tr>
<tr>
<td>Type 3</td>
<td>7</td>
</tr>
</tbody>
</table>

Identification for the 51 test sounds, these cases of perceptive deafness can be classified into the following four types:

Type 0: This minor group shows similar pattern to the control group (Fig. 3-a).

Type 1: This group shows poor discrimination for vowels differing in F1, but show good discrimination for differences in F2 (Fig. 3-b).

Type 2: Contrary to Type 1, subjects of this type show poor discrimination for F2 (Fig. 3-c).

Type 3: This group shows generally poor identification and the boundaries between vowels on F1 - F2 diagram are not definite (Fig. 3-d).

In general, there exists a close relationship between the above classification and the results of pure tone audiometry. In Type 0, disagreement in identification of test sounds is usually less than 14%. The maximum threshold elevation in pure tone audiometry among these cases of this type is found to be 30 dB at 250 Hz, 40 dB at 500 Hz, 50 dB at 1,000 Hz and 60 dB at 2,000 Hz, respectively. In most cases Type 1 corresponds to low tone hearing loss in pure tone audiometry, while Type 2 corresponds to high tone loss in hearing. In Type 3, subjects usually show far advanced hearing impairment over the entire frequency range. There are some exceptional cases who show no appreciable correlation between the results of pure tone audiometry and the type of identification of synthetic sounds noted above.
Measurements on Speech Audiometers

Horst Gunther Diestel
Leitender Direktor und Professor
bei der Physikalisch-Technischen Bundesanstalt
Braunschweig, Germany

---

Introduction

Speech-audiometers are devices for reproducing speech material at known intensities for making tests of articulation or intelligibility. To get comparable results when different audiometers are used it is necessary to standardize the test material, the manner of presentation, and the characteristics of the audiometers. The elements of a speech-audiometer are usually a record (magnetic tape or phonograph disk), an amplifier, a meter for monitoring the output of the amplifier, an attenuator, earphones, and a loudspeaker.

For the elaboration of standard specifications a couple of speech-audiometers were tested at the Physikalisch-Technische Bundesanstalt in Braunschweig. At first their frequency curves using loudspeakers and earphones were measured. Secondly, the levels of the test words were measured with different instruments and the readings compared with relative loudness estimations of these words. Finally, the reference levels of the 50% articulation scores for both numbers and monosyllabic words were determined for some audiometers.
Frequency curves

The following types of audiometers were tested: ATLAS-ELEKTRONIK, Bremen, Type EM 69 and Type EM 58; BEOTON, Essen, Type V; HANSATON, Hamburg, Type EB 450/V; PHONAK, Stuttgart, Type 215; and SIEMENS, Erlangen, Type "Audiator".

The frequency curves at a suitable setting of the intensity attenuator were obtained by giving at the input sinusoidal reference signals of test tapes (see German Standard DIN 45 513; 1966) or test discs (see DIN 45 541; 1966). The corresponding sound levels generated by the loudspeaker were measured in a non-reflecting surrounding by means of a calibrated microphone. The sound levels generated by the (binaurally used) earphones were characterized by the SPL of a plane progressive wave which generated the same loudness in the listeners' (free) ears as the earphone did. The wave reached the person from ahead and had the same frequency as the earphone.

The frequency ranges, presented in table I, are defined according the rules for hearing aids: The average SPL at the frequencies 500 Hz, 1000 Hz and 2000 Hz is set to be 0 dB, and the cut off frequencies are given by the -10 dB points of the frequency curves.

<table>
<thead>
<tr>
<th>Audiometers</th>
<th>frequency range using earphones</th>
<th>frequency range using loudspeaker</th>
</tr>
</thead>
<tbody>
<tr>
<td>Atlas EM 69</td>
<td>200 Hz ... 3 500 Hz</td>
<td>100 Hz ... 8 000 Hz</td>
</tr>
<tr>
<td>Atlas EM 58</td>
<td>140 Hz ... 10 000 Hz</td>
<td>125 Hz ... 10 000 Hz</td>
</tr>
<tr>
<td>Beoton V</td>
<td>100 Hz ... 10 000 Hz</td>
<td>125 Hz ... 18 000 Hz</td>
</tr>
<tr>
<td>Hansaton</td>
<td>180 Hz ... 6 300 Hz</td>
<td>200 Hz ... 6 500 Hz</td>
</tr>
<tr>
<td>Phonak 215</td>
<td>200 Hz ... 3 500 Hz</td>
<td>180 Hz ... 12 500 Hz</td>
</tr>
<tr>
<td>Audiator</td>
<td>100 Hz ... 8 000 Hz</td>
<td>110 Hz ... 12 500 Hz</td>
</tr>
</tbody>
</table>
It can be seen that the average frequency range of the mentioned audiometers reaches from 150 Hz to 9000 Hz. The types of earphones calibrated in terms of free field SPL were Beyer, Heilbronn, Type DT 48 with flat and with circumaural rubber o assim; MIEFONBAU (MB ELECTRONIC), Schweizingen, Type K 600; TELEPHONICS, Huntington (USA), Type TDH 39; HOLMBERG, Berlin. Type HOLMO 100 BC-1. All types were calibrated using the same attenuator setting of the audiometer on a NBS-Type 9A coupler too, so that the differences between free-field SPL and the SPL in the coupler are known for these types. These differences are needed for a simple objective calibration of the audiometers.

**Level of speech**

The test material consisted of numbers, i.e. polysyllabic words, and of monosyllabic words (see DIN 45 621: 1961). The levels of the component syllables (words) were measured by means of different instruments, i.e. impulse sound level meter (rms time constant 35 ms), Sound level meter ("fast"), VU-meter, "Aussteuernsmesser" (see DIN 45 406: 1966, quasi-peak meter, time constant 10 ms) and the hp-loudness analyser. The relative readings of the words were compared with the relative loudness of these words. The impulse sound level meter had the highest correlation coefficient ( .52) between reading and loudness estimation.

It is proposed therefore to define the average level of speech as the average value of the maxima of the levels of the component syllables, the maxima being measured by an impulse sound level meter.

**Calibration of speech audiometers by articulation tests**

The accuracy of calibrations by articulation tests were studied on some audiometers in a quiet surrounding. At different settings of the intensity attenuator the percentage articulation scores for unknown numbers and monosyllabic words were measured by ten normal hearing persons having one ear—
phone on one ear. From the articulation versus intensity curves the average setting for 50% articulation is interpolated.

For this setting, the above defined average speech level was measured in terms of voltage at the terminals of the earphone by an impulse sound level meter. As average response between voltage and equivalent free-field SPL of the earphone the average value of the response at the frequencies 500 Hz, 1000 Hz, and 2000 Hz was taken. The results are presented in table II.

<table>
<thead>
<tr>
<th>Free-field SPL for 50% articulation (monaural hearing)</th>
<th>Atlas EM 69</th>
<th>Atlas EM 58</th>
<th>Phonak</th>
<th>Boston V</th>
</tr>
</thead>
<tbody>
<tr>
<td>numbers</td>
<td>23</td>
<td>21</td>
<td>22</td>
<td>18</td>
</tr>
<tr>
<td>monosyllables</td>
<td>36</td>
<td>34</td>
<td>37</td>
<td>33</td>
</tr>
</tbody>
</table>

From table II can be seen that in a quiet surrounding the average 50% reference level (monaural hearing) for numbers is 21 dB (equivalent free field SPL) whereas the level for monosyllabic words is 35 dB re 20 μW/m².

Considering these measurements it can be assumed that the accuracy of calibrations by articulation tests is about ±3 dB. Audiometers with a higher cut off frequency seem to have a smaller reference level than others.

The author gratefully acknowledges the co-operation of Dr. H. Mrass and Dr. K. Brinkmann and the help of the members of the staff who carried out the measurements.
Psychological Scales for Evaluation of Vibrations

Toshisuke Miwa & Yoshiharu Yonemura

Staffs of National Institute of Industrial Health

Introduction

For rating vibrations in residential, industrial, and traffic areas with respect to their annoying qualities such as interference with rest, working efficiency and social activities, psychological scales for vibration were determined by the same procedures as psycho-acoustics. New vibration tables of electro-dynamic type in vertical and horizontal directions were constructed for this purpose. These tables could be vibrated between 0.5 and 300 Hz by a transistorized power amplifier without an output transformer. The vibration sensation for whole body was measured on ten male subjects sitting or standing on the vibration table. The hand sensation was measured at the similar posture as handling the vibration tools. Vibration greatness level (VGL) corresponding to loudness level (phon) and vibration greatness (VG) to loudness (sone) were proposed.

Measurement of Vibration Greatness Level \(^1\),\(^2\)

Equal sensation levels were obtained by comparing various vibrations in the range from 0.5 to 300 Hz with that at 20 Hz selected as standard. The experiment was carried out by the method of paired comparisons and levels of the vibrations were changed in descending and ascending series by the tester. Threshold levels of the vibrations were also observed by the same method as described above. There were no differences of equal sensation
Psychological Scales for Evaluation of Vibrations

between sitting and standing for whole body vibration and also between the vertical and the horizontal vibrations for hand (Fig.1 and 2). We have defined vibration greatness level as the vibration acceleration level (VAL) at 20 Hz sensationaly equalized with the vibration of the other frequency. The VAL value was indicated as $20 \log_{10}(a/a_{ref})$, $a$: rms acceleration value $(g)$, $a_{ref}: 10^{-5}g$. The frequency characteristics of human sensation may be explained as velocity type between 7 and 100 Hz for whole body and hand as shown in Fig.1 and 2.

![Fig.1 Vibration greatness level for whole body.](image1)

![Fig.2 Vibration greatness level for hand.](image2)

**Measurement of Vibration Greatness**

A scale of vibration greatness was obtained by the study on whole body and hand in the vertical and horizontal vibrations. This scale has true ratio properties, and numbers on this scale having a given ratio refer to vibrations whose sensation has that ratio. The corrected ratio method devised by Garner for loudness measurement was used in our experiment. This method consisted of two experiments which were based on fractionation and equisection judgment. The unit value of VG was decided as 40 VGL (1 VG = 40 VGL). The experiments were carried out mainly at 20 Hz. The effect of frequency on vibration greatness was examined at 5, 30 and 60 Hz. The vibration greatness for both of whole body and hand was shown by following two Eqs.
Psychological Scales for Evaluation of Vibrations

\[ \log VG = 0.030 \cdot VGL - 1.20 , \quad \text{below 1 VG.} \]
\[ \log VG = 0.023 \cdot VGL - 0.92 , \quad \text{above 1 VG.} \]

Dependency on frequency was not observed.

**Evaluation of Compound Sinusoidal Vibrations**

The vibration greatnes level of the artificial compound vibrations composed of several sinusoidal vibrations at the center frequencies of the octave bands between 1 and 250 Hz were measured by comparison with the standard frequency of 20 Hz. Combinations of amplitudes and frequencies of their components were changed variously. The level of the standard vibration was varied by the tester by the same procedure as in the measurement of the vibration greatnes level. Conversion of observed VGL value to VG one was calculated by Eq.1. Effects of frequency intervals, levels and numbers of components in compound vibrations on vibration greatnes were examined. On whole body and hand for vertical and horizontal vibrations, total vibration greatnes \( \left( V_{GT} \right) \) could be estimated from each VG value of components \( \left( V_{G1} \right) \) and the maximum value \( \left( V_{GM} \right) \) using equation proposed by S.S.Stevens \(^{6}\).

\[ V_{GT} = V_{GM} + 0.3 \left( \sum_{i} V_{G1} - V_{GM} \right) \]

**Conversion of the Horizontal Vibration to the Vertical One**

On whole body vibration, the sensation of the horizontal was equated to that of the vertical at the same frequency between 5 and 300 Hz. The sensation of the vertical vibration above 5 Hz was stronger than that of the horizontal by 10 dB for sitting and by 15 dB for standing. The horizontal vibrations at each frequency from 0.5 to 200 Hz were equated in sensation with the vertical vibration of 20 Hz at 30 dB VAL. The contours obtained in this experiment and the curves of the equal sensation for horizontal vibrations (Fig.1) were quite similar with each other except that VAL value at 20 Hz as the standard was lower 10 dB for sitting and 15 dB for standing than the standard level in the equal sensation measurement. While, the results of similar experiments with hands for vertical and horizontal vibrations showed no difference between them. This reason may be explained by the fact that the difference of the threshold of hand between vertical and horizontal vibrations.
was not observed. On the other hand, the difference of threshold of whole body between vertical and horizontal vibrations was about 10 dB.

Measurement of Unpleasant and Tolerance Limit Levels

To obtain the relation between the VGL value and the human emotion such as unpleasant and untolerable, the sinusoidal vibration was impressed to the subjects from low to high level at 6 steps (1 step: 5 dB) for 5 min. at each step. The subject was required to answer the emotional response to the given vibration at each step, unpleasant or untolerable or not. These results were plotted on an ogive by each frequency, and tolerance limit and unpleasant levels were chosen at 50% point on them. The emotional contours for whole body vibration in vertical and horizontal directions and for hand vibration were almost similar with Fig.1 and 2 respectively. These values at 20 Hz were shown in Table 1.

<table>
<thead>
<tr>
<th>Tolerance limit</th>
<th>Whole body, vertical</th>
<th>Whole body, horizontal</th>
<th>Hand</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unpleasant</td>
<td>55</td>
<td>50</td>
<td>65</td>
</tr>
<tr>
<td></td>
<td>40</td>
<td>50</td>
<td>50</td>
</tr>
</tbody>
</table>

Table 1. Tolerance limit and unpleasant levels at 20 Hz (VGL).

Conclusion

When the vibration affecting whole body or hand is analyzed by the octave band filter in the frequency range of 0.5-300 Hz, each VGL value of its components can be obtained from Fig.1 or 2. The vibration greatness can be calculated by Eq.1. Total VGL value of compound vibration is obtained by summing up each VGL value by Eq.2 and VGL value of this vibration can be determined by using Eq.1 again. Then, the vibration effect on human body can be assessed from comparison of the obtained VGL value with Table 1. Moreover, if the vibration has two components of vertical and horizontal directions, its horizontal component can be equated in sensation to the vertical one and the VGL value of total vibration can be determined by Eq.1 and 2.

References

TACTPHONE as an Aid for the Deaf

Hisayoshi Suzuki
Assistant Professor of Research Institute of
Electrical Communication, Tohoku University

Ryuzaemon Kagami
Licensed Teacher of Yamagata School for the Deaf

Tajimi Takahashi
Licensed Teacher of Yamagata School for the Deaf

Introduction

TACTPHONE is one means for supplying cutaneous information of the short time amplitude spectrum of speech. This has been made as a teaching instrument to use in combination with lip reading at Yamagata School for the Deaf.

We have tried to train the deaf children to recognize Japanese speech sound by lip reading supplemented by tactile information with TACTPHONE. The report mentioned below is an abstract obtained from the trial use of about six months.

The Outline of TACTPHONE

TACTPHONE is composed of a filter bank analyzer, similar to that used in a vocoder, modulators and vibrators, as shown schematically in Fig.1. Ten contiguous band pass filters divide the frequency range 160 to 6600 Hz into ten channels. Table 1 shows the frequency range of the channels, which have equal band width in mel scale. Their outputs are rectified and smoothed to obtain values of the short time spectrum at ten frequency bands. The ten time varying voltages are used to amplitude modulated individual rectangular carriers of 200 Hz. The analyzing channel of lowest frequency is led to the small finger of the left hand, and the channel of highest frequency is led to the
small finger of the right hand.

**Discrimination and Identification of Japanese Speech Sound**

Three boys and two girls joined experiments as the subjects. They were in the 6th grade and aged 13. Their hearing losses (H.L) were 83 -- 96 dB even in their better ears, while their intelligence quotients (I.Q) were 75 -- 112.

After 6 months of training and practice, which were about 40 minutes for a day, all the subjects were able to make sound discrimination comparable to, and sometimes better than, that achieved in only lip reading.

In order to estimate an effect of TACTPHONE, an experiment to let small children discriminate and identify speech sound were carried out. 11 kinds of tests were given. For example, a test was to distinguish an item such as a syllable, a word or a phrase, from the others. The items selected were difficult to be distinguished from one another in lip reading.

The half of those tests were performed after some practices showing the lists of items in the question to the subjects, and the other half were performed without such practice nor showing the lists.

The results of discrimination of vowels were fairly good in each case where the vowels were pronounced separately or presented in a word. There were no difference of performance between the case aided by TACTPHONE and with only lip reading. Pairs of [a]-[1], [a]-[u] and [u]-[e] were almost completely discriminated. The pairs [i]-[u] and [u]-[o] were 72% and 76% in discrimination rate.

For the discrimination test of consonants, 65 pairs of words were used. A word in a pair was composed of the same series except just one phoneme which was to be tested. For example, "maku" (to wind) was paired with "naku" (to cry). Since those two words have the same accent pattern, a discrimination between [maku] and [naku] must be made in terms of the phonetical difference between [n] and [m].

---A-90---
In our experiments, word pairs which had to be discriminated in terms of [i], [ni], [ki], [ʃi] and [tʃi] were 25 -- 30% in the lip reading and 52 -- 55% in the case aided by TACTPHONE. Since those phonemes were pronounced with narrow mouth opening, it was more difficult to discriminate the difference of their place of articulation in lip reading than the difference of cutaneous stimulation of short time spectrum presented by TACTPHONE. In the case in which discrimination had to be made in terms of [a] and [ha] or [ka] and [ga], the words were pronounced with relatively wide mouth opening. As their place of articulation were relatively inner part of mouth cavity, the appearances were almost the same and consequently, it was profitable to use TACTPHONE. For example, the words such as "kagaku" (science), "agaru" (to rise), "kakaru" (to anchor) and "hakaru" (to measure) were discriminated about 30% by only lip reading and about 43% aided by TACTPHONE. In the case of discrimination of bilabials, [p], [b] and [m], those consonants were pronounced with the same place of articulation and could not be discriminated by only lip reading. For example, the words such as "patto" (suddenly), "batto" (a bat), "manto" (a cape) are discriminated 25% by lip reading and 45% aided by TACTPHONE.

In the articulation test for numerals, rate of correct response was 88% in lip reading and 97% aided by TACTPHONE. Most of confusions occurred between 1 and 2, which were pronounced in Japanese [itʃi] and [ni:] respectively.

**Conclusion**

From the observation of our experiments it is concluded that the vibratory stimulus on the fingers of subjects supplies the gross features of spectrum and amplitude of speech sounds. As the time patterns such as duration and rhythm of speech must have been clearly perceived with TACTPHONE, the type of consonant, the length of vowel and the number of syllables in words are identified much better in the case aided by TACTPHONE than in the case by only lip reading. In the
combined use of lip reading and TACTPHONE, the cutaneous information about the type of consonants and the number of syllables improved the recognition rate without detracting from lip reading.

From an educational point of view, it is important to note that the children have a feeling of attendance at the sound field, and they are apt to feel a mental unity with their teacher when they perceive vibrations of sound. Theoretically, a potential advantage of a tactual hearing aid lies in the similarities between the tactual and the auditory sensations to the time-intensity patterns of speech sounds. At present, there are at least two or more problems to be solved. One of them is to measure the capacity of information of tactile organs and the second is the problem of how to extract the essential quantities which provide the tactual signals from input speech sound. Our research efforts are now being directed toward those problems.*

Introduction

In order to construct an ideal tactual communication system for hard of hearing, we must investigate the human ability when he receives information by means of the vibratory sensation. In this report, the theoretical procedures to formulate spatial masking and summation effects, that are related to the ability, have been verified by conducting the experiments.

Spatial masking effect in the case where a lateral finger is stimulated

When two fingers were stimulated simultaneously, we assumed that the interaction should be as follows:

\[ \Psi_1 - \xi, \Psi_m \geq 0 \quad, \quad \Psi_1'' = \Psi_1 - \xi, \Psi_m < 0 \quad, \quad \Psi_m' = 0 \]

In the formula mentioned above: \( \Psi_1 \) (or \( \Psi_m \)); loudness in the case where only finger 1 (or \( m \)) is stimulated in the intensity \( L_1 \) (or \( L_m \)). \( \Psi_1'' \); loudness of the finger 1 in the case where the fingers 1 and \( m \) are stimulated simultaneously in the intensity \( L_1 \) and \( L_m \) respectively. \( \xi \); constant which represents the degree of the spatial masking.

If a subject adjusts the stimulus of the finger 1 to the threshold intensity \( L_{11} \) under the condition where the finger \( m \) is stimulated in the intensity \( L_m \), the following equation can be obtained from eq.(\( \text{eq.1} \)).
Spatial Summation and Spatial masking effects.

$$\psi'_i = \psi_i - x_i \psi_m = 0$$

$$\therefore \frac{\psi_i}{\psi_m} = \varepsilon_i = \text{const.} \quad \ominus$$

where $$\psi_i$$ is loudness in the case where the finger $$i$$ is stimulated alone in the intensity $$I_{si}$$.

Loudness function in the vibratory sensation could be represented by

$$\psi = k (I^n - I_0) \quad \ominus$$

In eq. (2) $$I_0$$; intensity of threshold, $$n$$; constant ($$0.4 \sim 0.5$$). $$k$$; constant which depends upon the choice of unit.

Using eq. (5), $$\psi_{si}$$ and $$\psi_m$$ in eq. (3) can be represented as follows:

$$\psi'_i = k_m (I^n_{\psi_m} - I_0) \quad \ominus$$, $$\psi'_i = k_s (I^n_{\psi_s} - I_0) \quad \ominus$$

In eq. (5) and (6) $$I_{\psi_m}$$ and $$I_{\psi_s}$$, threshold intensity which is influenced by the stimulus duration and the antecedent stimulus.

$$I_{\psi_i}$$; threshold of the finger $$i$$ in the case where the finger $$m$$ is stimulated in the intensity $$I_m$$.

Under the various conditions (the stimulus duration and the antecedent stimulus), the relations between $$I_{\psi_i}$$ and $$I_m$$ were measured by the method of adjustment, and the ratios ($$\psi_{si}/\psi_m$$) were calculated according to eq. (6) and (7). An example of the results is shown in Fig. 1.
Spatial Summation and Spatial masking effects.

The ratio is constant approximately, and constant for 4 subjects was as follows:

\[ \frac{\psi_{oi} \psi_m}{\psi_m} = \xi_i = 0.13 \sim 0.20 \] \( \text{eq.1} \)

From eqs. 6, 7 and 8, the spatial masking is formulated as

\[ M_m = 20 \log_{10} \frac{L_m}{L_o} = 20 \log_{10} \left[ \frac{L_m}{L_o} \right] \frac{L_m}{L_o} + 1 \] \( \text{eq.2} \)

therefore threshold shift \( M_m \) is represented by

\[ p_m = 20 \log_{10} \frac{L_m}{L_o} = M_m + 20 \log_{10} \frac{L_m}{L_o} \] \( \text{eq.3} \)

where \( L_m \) is specific threshold in the finger \( i \).

The comparison of the theoretical values (solid line) in eq.2 with the measured values (plots) by the experiments is shown in Fig.4.

Spatial summation effects

After the total loudness in the case where the fingers \( m \) and \( r \) were stimulated simultaneously was measured by the method of magnitude estimation, the relation between the total loudness (\( \psi_{mr} \)) and the loudness of individual finger (\( \psi_r \) or \( \psi_m \)) was examined for two following hypotheses.

(I) Hypothesis of orthogonal vector sum:

\[ \psi_{mr} = (\psi_r^2 + \psi_m^2)^{\frac{1}{2}} \] \( \text{eq.4} \)

(II) Hypothesis of linear sum in considering the influence of the spatial masking:

\[ \psi_{mr} = (\psi_r - \psi_m) + (\psi_m - \psi_{rm}) \] \( \text{eq.5} \)

An example of the comparison of the empirical values with the theoretical values by the hypothesis (I) and (II) is shown in Fig.3. The hypothesis (II) is more accurate than hypothesis (I).

Spatial masking effect in the case where two lateral fingers are stimulated.

When fingers \( m, r \) and \( i \) are stimulated simultaneously, the loudness of the finger \( i \) (\( \psi'_i \)) can be represented by replacing \( \psi_m \) and \( \xi_1 \) in eqs.1 and 2 with \( \psi_{rm} \) and \( \xi_2 \), respectively. And the expansion
Spatial Summation and Spatial masking effects.

of eq. 3 requires the below mentioned eq.:

\[ \frac{\psi_i}{\psi_m} = \xi_{2} = \text{const.} \quad (12) \]

In the same notations as eqs. 5 and 6,

\[ \begin{align*}
\psi_i &= k_i (L_{i} - L_{m}) \quad (5), \quad \psi_m = k_m (L_{m} - L_{r}) \quad (6) \\
\psi_{01} &= k_1 (L_{01} - L_{0}) \quad (6)
\end{align*} \]

Under the various conditions (the various combinations of the fingers m and r, and the intensity levels Lm and Lr), the threshold of the finger 1 were measured by the method of adjustment. From those results, eq. 12 was verified.

(Fig.1.). Therefore the spatial masking is represented as follows:

if \( \psi_i - \psi_m \geq 0 \) and \( \psi_i - \psi_r \geq 0 \),

\[ \mu_{1}^{(0)} = \frac{20 \log \{ (\xi_{1} \mu_{1}^{(0)} (L_{i} - L_{m}) + (\xi_{r} \mu_{1}^{(0)} (L_{i} - L_{r}) + 1 \} \quad (9) \]

and if \( \psi_i - \psi_m \geq 0 \) and \( \psi_i - \psi_r < 0 \),

\[ \mu_{1}^{(0)} = \mu_{1}^{(0)} \quad (9) \]

The comparison of the theoretical values (solid line) in eq. 19 with the measured values (plots) by the experiments is shown in Fig. 4.

Conclusion

If the loudness of vibratory sensation is defined in the ratio scale, the following conclusions can be obtained.

1. Spatial masking effects can be represented as the simple interaction in the loudness.

2. Spatial summation effects can be represented as the linear sum of loudness in considering the spatial masking.

3. As the spatial masking effect is independent of the combination of fingers, we can analogize that the neural networks are symmetrical among the fingers.
Recherches sur la propriété de quasi-stationnarité par rapport à l'oreille humaine

Halina Maria Kiffert
Doc. Dr ès sciences, Chaire d'Acoustique et de la Théorie de Vibrations
Université A. Mickiewicz - Poznań
Pologne

La méthode d'évaluation des propriétés et de détermination des paramètres d'un processus acoustique donné, basée sur l'analyse spectrale détaillée, a amené une évolution de la notion de stationnarité d'un processus acoustique. La possibilité d'observer les changements de la forme d'un spectre dynamique (notamment, du spectre évolutif ou instantané, ou des autres spectres obtenus selon le même principe/1/), qui ont lieu pendant la durée du processus même, a justifié l'introduction de la notion de stationnarité spectrale. La méthode a en outre permis d'établir certaines limites de tolérance pour les changements du spectre instantané, limites admettant de considérer le processus comme à peu près stationnaire (quasi-stationnaire). En outre, puisque la forme du spectre dépend des paramètres de l'analyseur, on a attribué à cette notion un caractère relatif qui la met en rapport avec le dispositif analyzant /2/. Ainsi, la stationnarité est en quelque sorte une notion de limite, car la non-stationnarité peut être amortie par le dispositif. La constatation qu'un processus est quasi-stationnaire, basée sur une évaluation des changements d'une grandeur spectrale, conduit à déterminer son temps de quasi-stationnarité par rapport au dispositif utilisé. Ce temps représente une grandeur-limite pour le temps de la mémoire du dispositif analyseur. Si ce dernier est excessivement long, il y a tendance à produire la grandeur spectrale sous forme de sa valeur moyenne, en y effaçant les détails, tandis qu'un temps de la mémoire trop court ne donne pas /à cause de fréquentes oscillations/ d'informations précises et bien établies ni ne rend compte de la composition spectrale du processus et des changements dans le temps de celle-ci.

La notion de stationnarité spectrale doit être considérée tout à fait indépendamment de celle de périodicité d'un processus. En effet, elle apparaît dans les transitoires de l'acoustique de la parole, de la musique, du bruit, des salles, etc.

Ici, il sera avantageux de choisir comme objet de nos considérations le simple
transitoire final d'une composante spectrale, obtenue avec un filtre à courbe de transmission rectangulaire, pour une vibration périodique soudainement interrompue, dont l'enveloppe est du type $\sin \frac{\pi t}{T}$. Le problème consiste à déterminer si et dans quelles conditions un transitoire ainsi défini pourra-t-il être approximativement stationnaire par rapport à une fonction de mémoire, convenablement choisie, d'un analyseur. Compte tenu de la relation entre la forme du spectre instantané et la fonction de la mémoire, le problème équivalent résidera dans le choix et la définition du temps optimal pour cette fonction garantissant que les conditions précédemment exprimées pour la quasi-stationnarité spectrale seront remplies.

D'autre part, en basant sur le dualisme entre la présence dans un processus/ de la propriété appelée quasi-stationnarité et le choix des paramètres d'un dispositif analyzant/, on pourra vice versa essayer de trouver les processus qui, par rapport au dispositif donné, s'avéreront approximativement stationnaires, c'est-à-dire quasi-stationnaires spectralement. Cette approche devient particulièrement intéressante lorsqu'on veut définir une telle propriété par rapport à un analyseur aux paramètres constants.

Or, l'organe de l'ouïe, malgré les essais nombreux d'établir un modèle de son fonctionnement, apparaît toujours comme un analyseur à paramètres constants très peu connu. Vis-à-vis du problème que nous venons d'énoncer, l'oreille présente des difficultés toutes spéciales, dues surtout au fait que, en tant que récepteur de son, elle agit comme un analyseur /et même de haute précision/ mais que simultanément elle effectue une synthèse, dont le résultat est que nous percevons chaque son, autant complexe qu'il soit, comme un tout.

Malgré cela, il paraît utile d'élargir la notion de quasi-stationnarité en y annexant certains éléments et résultats nouveaux, afin de la rendre applicable dans le domaine psycho-acoustique. Une étape dans cette direction consisterait à introduire les notions de stationnarité et de quasi-stationnarité psychologique.

L'impossibilité dans laquelle nous nous trouvons de faire des observations directes sur les résultats de l'analyse effectuée par l'oreille aux différentes étapes /sur la membrane basilaire elle-même, et après transformation des variations de pression en des séries d'impulsions transmises par les fibres nerveuses/ exclue toute évaluation des changements du spectre de différents processus acoustiques.

Toutefois, la synthèse /en tant que le réciproque de l'analyse/ effectuée dans les centres supérieurs du système nerveux, et dont les résultats sont perçevables à l'observateur comme objets d'étude, fournit les informations qui peuvent être considérées comme équivalentes pour l'évaluation du spectre et comme résultats des changements subis par celui-ci. Ce sont ces informations à partir desquelles on pourra attribuer à un processus donné la propriété dite quasi-stationnarité considérée, dans le présent cas, selon des critères psychologiques.
Recherches sur la propriété de quasi-stationnarité

Ici encore, il sera avantageux de choisir comme point de départ le transitoire final d’une vibration sinusoidale subitement interrompue. Une telle allure de la composante spectrale dans son transitoire final peut être considérée comme l’effet d’un processus généralisé de modulation de l’amplitude de la vibration sinusoidale. Cette généralisation consiste aussi bien en l’utilisation d’un facteur non typique de la forme \( \frac{\sin x}{x} \) qu’en l’hypothèse d’une valeur-limite du coefficient de modulation \( n = 1 \).

Indépendamment de l’étude du processus de modulation et des seuils d’audibilité de celle-ci, nous avons fait des recherches /5/ sur l’audibilité des changements de l’enveloppe du transitoire final, produit artificiellement, d’un processus monochromatique, enveloppe donnée analytiquement par une fonction du type \( \frac{\sin x}{x} \).

Les résultats de ces recherches peuvent servir de point de départ pour vérifier la conception théorique ayant comme but de déterminer les processus acoustiques que l’oreille estime être approximativement stationnaires dans le sens psychologique.

On a pu établir que l’apercuevabilité par l’observateur de deux maxima consécutifs dans l’enveloppe du transitoire de forme \( \frac{\sin x}{x} \) a lieu ou non selon l’intervalles-temps entre ces maxima, et on a trouvé une valeur-limite de celui-ci séparant la région où l’oreille traitait l’enveloppe comme monotomique de celle où il y avait audibilité de deux maxima distincts. Caractéristiquement, la valeur trouvée pour cet intervalle entre les deux maxima ne dépend ni de la fréquence de la vibration qui réplique l’enveloppe ni du niveau d’intensité du son; elle est toutefois liée à un intervalle d’incertitude /6/ de prise de décision par l’observateur.

Ces résultats permettent de considérer cette valeur-limite de l’intervalle-temps entre maxima comme une grandeur caractéristique pour l’évaluation par l’oreille des changements d’enveloppe des processus acoustiques. En outre, on peut considérer cette valeur-limite, qui est de 250 msec, comme temps de quasi-stationnarité pour les processus de cette espèce par rapport au temps de la mémoire de l’oreille, dans la mesure où le notion de quasi-stationnarité a pu être appliquée au fonctionnement de l’organe de l’oreille en tant que récepteur de sons.

Une comparaison détaillée avec les travaux sur la modulation d’amplitude et de fréquence aboutit à une bonne concordance des résultats. En effet, dans le processus de modulation, l’intervalle de sensibilité maximale de l’oreille est aussi indépendant de la fréquence portante et de l’intensité du son, et a lieu pour environ 4 Hz. A admettre, par généralisation, que l’on peut attribuer une valeur correspondant à un hémicycle à l’intervalle des maxima de l’enveloppe précédemment défini, on obtient que la fréquence qui se trouve coordonnée à un intervalle de temps de 250 msec est égale à environ 2 Hz.

En outre, la nécessité s’impose de faire une comparaison entre la notion de temps de quasi-stationnarité psychologique de l’oreille introduite plus haut et celle de la constante de temps de l’oreille. Les différentes manières de définir cette grandeur.
Recherches sur la propriété de quasi-stationnarité

dans la littérature conduisent à des valeurs entre 35 et 500 Hz selon le cas. La valeur de 250 Hz obtenue ici pour le temps de quasi-stationnarité se situe dans ces limites; elle pourra être considérée comme représentant une notion à partir de la constante de temps de l'oreille, notion rapportée aux observations des changements de l'enveloppe des transitoires. Bien que l'exemple que nous venons de considérer soit un transitoire de forme bien déterminée, il est pourtant représentatif d'un groupe de phénomènes aux paramètres convenablement variables et correspond à des processus appartenant à la réalité de la physique.

L'introduction de la notion de temps de quasi-stationnarité psychologique, et la spécification de processus acoustiques que l'oreille traite comme quasi-stationnaires /dans le sens antérieurement défini/ à temps de quasi-stationnarité de 250 msec, présente de l'importance pour un cercle de problèmes bien déterminés. En particulier, cette notion appliquée dans des recherches ultérieures pourra servir à trouver à quel point l'oreille est-elle capable de distinguer des changements d'enveloppe de transitoires dans un domaine plus généralisé, ou si au contraire, en produisant des valeurs moyennes, l'oreille n'efface certains détails porteurs d'informations supplémentaires sur les différents paramètres acoustiques et la manière dont ils dépendent les uns des autres.

Bibliographie:
PITCH: RESULT OF AN AUTO-CORRELATION PROCESS
Roelof J. Ritama
Instituut voor Perceptie Onderzoek,
Eindhoven, Holland.

In a previous paper 1) it has been shown in using stimuli consisting of the sum of a low-frequency band of a unipolar pulse train with fundamental frequency \( f_0 \) and a high-frequency band of a unipolar pulse train with fundamental frequency \( f_0 + \Delta f \), that a rather limited spectral region is utilized by the ear in producing a well-defined pitch perception. For fundamental frequencies in the range of 100 to 400 Hz and for sensation levels of the entire signal up to at least 50 dB, the pitch derived from the frequency band consisting of the third, fourth and fifth harmonics tends to dominate as long as the amplitude of this band exceeds a minimum absolute level of about 10 dB above threshold.

We now like to introduce the concept of dominance for broad-spectrum signals in general. Fig. 1 illustrates the concept. The waveform of an acoustical signal is given at the eardrum, the stapes and various points on the basilar membrane. So pitch information is available at a large part of the membrane the ear uses only the information at one specific point. This point depends on subjects and is 3-5 times the pitch value.

Due to this concept of dominance the physical parameters determining pitch might be the same as already found for A(mplitude) M(odulated) sinusoids. This type of signals with narrow frequency spectrum produces a "ready made" envelope at the corresponding place on the basilar membrane and so we can be sure that the displacement waveform at that place is about the same as the acoustical waveform. For AM-sinusoids it has been shown that to a first approximation the
time distance between two positive peaks in the fine structure of
the waveform near two successive crests in the time envelope is an
adequate parameter of the pitch perceived 2). Experiments on the
pitch behaviour of a quasi FM-sinusoid (derived from an AM-sinusoid
by shifting the phase of the center frequency over 90°) have estab-
lished this. The experiments have ruled out any theory based on the
mean frequency of the power spectrum of the signal or on subharmo-
nics of the strongest component in the signal 3).

The displacement waveform of the basilar membrane produced by
a broad-spectrum signal at the place which is maximally sensitive to
the frequency $f_0$ is similar to the waveform of the acoustical signal
passing through an adequate bandpass filter with center frequency $f_0$.
The response of an ideal bandpass filter with center frequency $f_0$
and bandwidth $\Delta f$ to a Dirac pulse can be shown 4) to be equal to
$$\delta_0(t) = \frac{2}{\pi t} \sin \pi \Delta f t \cos 2 \pi f_0 t$$
In the same way we find for the response to a 90°-phase shifted
Dirac pulse
$$\delta_{90}(t) = \frac{-2}{\pi t} \sin \pi \Delta f t \sin 2 \pi f_0 t$$
For $t = 0$ the envelope $\frac{2}{\pi t} \sin \pi A t$ is at maximum, thus, in this vicinity the fine structure has its greatest peaks. In particular, $\delta_0(t)$ has its major positive peak for $t = 0$; $\delta_{90}(t)$ for $t = -1/4 f_0$; $\delta_{270}(t)$ for $t = 1/4 f_0$; and $\delta_{180}(t)$ has two major peaks, one for $t = -1/2 f_0$ and the other for $t = 1/2 f_0$.

Now, if pitch ($P_x$) due to the interaction of a Dirac pulse and a $x^2$-phase shifted Dirac pulse (pulse distance $\tau$), is indeed correlated with the reciprocal value of the time distance between the major positive peaks of the fine structure, the pitch behaviour for that place on the basilar membrane which corresponds to $f_0$, must satisfy the following relations:

$$P_0 = \frac{1}{\tau}; \quad P_{90} = \frac{1}{(\tau - 1/4f_0)}; \quad P_{180} = \frac{1}{(\tau^2 + 1/2f_0)}; \quad P_{270} = \frac{1}{(\tau - 1/4f_0)}$$

These relations have been tested for various types of broad-spectrum signals.

a. The non-filtered periodic alternating-polarity pulse train.

Two subjects made pitch measurements for pulse rates from 60 up to 250 pps. From the results the existence of two distinct pitches could be established. The pitch values found follow the relation for $P_{180}$ rather closely.

In a second experiment we tried to prove more directly the concept of dominance: the pitch is derived from the displacement waveforms in one restricted spatial region of the basilar membrane. Therefore the non-filtered periodic alternating-polarity pulse train was mixed with either high-pass or low-pass masking noise.

Two subjects did pitch measurements for various values of the cutoff-frequency of the masking noise. The results show that the perception of the low-pitch, $1/(\tau - 1/2f_0)$, and the high-pitch, $1/(\tau - 1/2f_0)$, did not change as long as the frequency $f_0$ was not masked by the noise. Moreover, if the frequency $f_0$ was masked by the noise, both pitches were shifted. In the case of high-pass noise the frequency $f_0$ shifted to a lower value near the cutoff-frequency of the masking noise; in the case of low-pass noise the frequency $f_0$ shifted in the opposite direction.

b. The interaction of a sound with its repetition.

When a sound and the repetition of the same sound after a delay $\tau$ are presented together, a pitch is evoked corresponding to the reciprocal value $1/\tau$. The sound may be noise or merely a pulse (repeated periodically or at random).
These phenomena have been described by Fourcin 5) and by Thurlow and Small 6). Bilsen 7) made control measurements and similarly found pitch values corresponding to $1/\tau$ in a range up to 800 Hz. After adding the repetition with a negative sign (a phase shift of 180°) a bivalent pitch was perceived, viz. $P_{180} = 1.14/\tau$ and $0.88/\tau$. By shifting the phase of all the frequency components of the repeated sound by $\pm 90°$ pitches were heard lying between the values for $P_0$ and $P_{180}$, viz. $P_{90} = 1.07/\tau$ and $P_{270} = 0.94/\tau$. Thus far this pitch behaviour could not be explained simply in terms of time or frequency analysis. But in using the concept of dominance a value for the dominant frequency $f_0$ has been found of

$$f_0/P_0 = 3.9 \pm 0.2$$

This value is in accordance with the results from the test with periodic pulses.

Independently of these measurements the concept of dominance was checked in the same way as for periodic unipolar pulse trains. In the experiments a test signal was used consisting of the sum of a low-frequency band and a high-frequency band of a noise added to its repetition after a delay $\tau_1$ and a center frequency band of the noise signal added to its repetition after a delay $\tau_2$. The 3 dB-points of the center band were situated at 715 and 1400 Hz. It was found that for a pitch range of 150 - 300 Hz this frequency band tends to dominate the pitch sensation as long as its amplitude exceeds a minimum absolute level of about 10 dB above threshold irrespectively of the level of the other frequency bands in the signal (up to 50 dB).

These experiments suggest that the concept of dominance is not only valid for the residue pitch but can also be extended to effects known as repetition pitch, time separation pitch, time difference tone and sweep tone.

References

Pitch of Filtered Noise

A. Rakowski
Assistant Professor
of the Academy of Music, Warsaw

Introduction

Listening to white noise filtered with the aid of low-pass or high-pass filters evokes the sensation of pitch, very closely correlated with the cutoff frequency of the filter. Bekesy /1/ found that an octave band of noise elicited the sensation of two pitches corresponding to both half-power frequencies of the noise band 400 and 800 Hz. Small and Daniloff /2/ reported that cutoff frequencies of two low-pass bands of noise may be subjectively tuned in octave intervals. According to them, when high-pass bands of noise are used, the possibility of tuning is limited to filter cutoff frequencies higher than 600 Hz. Rakowski and Rossypal /3/ compared the accuracy of tuning the frequency of a pure tone to the cutoff frequency of a low-pass noise in various frequency regions and for two different sensation levels. The filters used had very steep slopes of frequency response beyond the cutoff frequencies.

The aim of present study was to investigate the effect of varying the slope of skirts of the noise band upon clarity of its pitch sensation.

Experimental Procedure

The Gaussian noise was filtered with the aid of low-pass and
high-pass filters. The slope of filter characteristics beyond the
cutoff half power frequency equaled to 150, 50 and 15 dB per octave.
The cutoff frequencies of the filters were 200, 500, 1000, 2000 and
5000 Hz. These values were purposely varied within ±12 % for filters
with different frequency response in order to avoid harmonic relations.

Samples of noise were
recorded on magnetic tape
with frequency response flat
within 40 – 16000 Hz. The
result of spectral analysis
of a typical recording is
shown in Fig. 1. The ana-
lysis was taken with the use
of Radiometer FRA-1 heterodyne analyser and Bruel & Kjaer 2305 level
recorder.

Loops of tape with recorded samples of noise were replayed from
a taperecorder and signal fed to a two position switch operated by
the subject. To this switch sinusoidal signal from oscillator was
also delivered. The subject was seated in a sound treated booth and
listened binaurally to the signals through Beyer DT-48 earphones.
The frequency response of the earphones was corrected to listening
in free field conditions and, including corrector, was flat in
frequency region 40 – 16000 Hz. The subject's task was to effect
a pitch match between the noise and the pure tone from oscillator.
He was permitted to listen
alternatively to noise and
tone and adjust the frequen-
cy of the oscillator with
dial covered, until he con-
sidered that he had
accomplished a pitch match.
Pitch of Filtered Noise

The value of adjusted frequency was read by the experimenter from a digital counter.

The loudness level 40 phons was held constant for a tone independently on the frequency, due to a corrector whose frequency response corresponded to the appropriate isophonic curve of the ear. The loudness level of all other stimuli was 80 phons. A schematic block diagram of the apparatus used in the study is shown in Fig. 2.

Results

It has been previously stated /3/ that subjects in the experiments involving judgement of pitch should be very carefully selected and properly trained if a serious vagueness of the results is to be avoided. Accordingly, the present experiment was carried out in two parts. In the first part 32 audiologically tested young musicians, men and women, listened during 3 sessions to "steep skirt" noise bands being instructed to match the pitch of a tone to the most prominent component of noise. 16 subjects who had been making the most consistent judgements participated in the second part of experiment and each of them performed various pitch matches in 6 sessions. Only the results of this part of the experiment were taken into account.

The average number of single results per stimulus was 60.

In Fig. 3 are presented frequency distributions of the pitch matches at filter cutoff frequencies 200, 1000, and 5000 Hz /a,b - low-pass noise, c,d - high-pass noise, a,c - 150 dB/oct, b,d - 15 dB/oct/. The results are grouped in 1/6 octave class intervals.

Conclusions

The sensation of pitch for low-pass and high-pass noise is well established in general, except for broad band high-pass noise, and corresponds to the cutoff frequency of a filter. Octave and fifth errors are observed. The accuracy of the pitch judgement decreases for extreme low and high frequencies. The increase in steepness of noise band skirts improves the accuracy of the pitch judgements.
but at steepness 15 dB/oct judgement may still be made with considerable consistency. Pitch of low-pass noise may be evaluated more accurately than that of high-pass noise.

The phenomenon described may effect in raising subjective distortions in sound transmitted through electroacoustic channel in which filters are used.

References

Tonal Differential Limen of the Speech Transmission System containing Single Dip in Frequency-response

Ryōichi Matsuda

Electrical Communication Laboratory, N.T.T.,
Tokyo, Japan

Introduction

In the case of electroacoustic transducers such as speakers, microphones and telephone receivers and transmitters, peaks and dips are often seen in their frequency-response. For the design of these transducers, it is important to evaluate the effects of these irregularities in frequency-response on the transmission quality.

In this paper, just detectable depth of dip of a speech transmission system containing single dip in frequency-response as shown in Fig. 1 is discussed in analytical and experimental standpoints.

Analysis

Speech signals, from the point view of its amplitude fluctuations, is interrupted by its silent interval with some build-up or decay time constant, and speech signals have also a fluctuation in its frequencies.

--- A-109 ---
of harmonic components. Then for our analysis, speech signals are approximately expressed by a group of sine-waves (complex tone) as

$$
F(t) = A \sum_k e^{\frac{jt}{h_0}} \sin(2\pi k h(t) + \phi_k) \\
= A \sum_k \sin(2\pi k h(t) + \phi_k)
$$

where \( h(t) \) is fundamental frequency of the complex tone and it is considered as a function of time due to the fluctuation.

As regard to frequency fluctuations, only one sinusoidal harmonic component is considered for simplicity. And dynamic frequency-response of the system for the signal, whose frequency is proportionally varied with time, is analyzed and effective depth of dip is obtained as shown in Fig. 2. In Fig. 2 \( D_s \) is the dip depth illustrated in Fig. 1 and \( D_h \) is the effective dip depth where the frequency of input signal varies in fluctuational speed of \( h \) Hz/sec. From Fig. 2 it is seen that effective dip depth \( D_h \) for the frequency fluctuated signal decreases compared with \( D_s \).

As regard to amplitude fluctuations, the analysis on the envelope of output signal of the system for the input signal of Eq. 1 is carried out, but in this case \( h \) is assumed to have no frequency fluctuations. As the results of this analysis, the intensity \( I_t \) and equivalent duration \( T \) of transient tone caused by the amplitude fluctuation are obtained as follows.

$$
I_t = \frac{I_o \left( \Delta\Omega \cdot Q / (1 + \Delta\Omega \cdot Q) \right)^2}{2\pi f_0 (1 + \Delta\Omega \cdot Q)}
$$

$$
T = Q / 2\pi f_0 (1 + \Delta\Omega \cdot Q)
$$

where \( I_o \) is intensity of one component of the complex tone of Eq. 1,
Tonal Differential Limen

\( f_0 \) is anti-resonant frequency of the system, and
\[
\gamma(\frac{1}{\delta}) = \left( \frac{1}{1 - \frac{1}{\delta/\alpha}} - 1 \right) e^{-\frac{\delta}{\delta/\alpha}} - \frac{1}{1 - \frac{1}{\delta/\alpha}} \delta^{\frac{3}{2}} \alpha
\]
\[
= 2e^{-\frac{1}{\delta/\alpha}} \delta = \alpha
\]

where \( \alpha = \pi f_c(\Delta \Omega + 1/Q) \).

Normalized representation of \( I_t \) and \( T \) is shown in Fig. 3, where it is seen that the transient tone will decrease as the time constant \( 1/\delta \) of the input signal decay increases.

On the other hand, the author has already reported a method to predict the hearing threshold of transient tone succeeded to the complex tone. Applying this result to this case here, we can also predict the hearing threshold of the transient tone of Eq. 2 and Eq. 3. The predicted minimum dip depth of the system where the transient tone is just detectable is shown in Fig. 4 as solid lines.

**Measurement of Tonal Differential Limen**

Listening tests are performed using conversational speech sound and synthesized complex tone comprising 14 pure tones of equal amplitude in a harmonic relation. The fundamental frequency of the synthesized tone is 240 Hz simulating average fundamental frequency of the female voice. Four adult females with normal hearing were engaged in these tests. The tonal differential limen is defined as the dip depth in which there is 50% probability of aurally distinguishing the
Tonal Differential Limen

output of the system with dip characteristic from that of the system with flat characteristic. For the latter system the hearing sound pressure level is 64 dB relative $2 \times 10^{-4}$ microbar.

An example of the test results is shown in Fig. 4. In Fig. 4 circles indicate the just detectable dip depth for conversational speech as input signal, and dotted line indicate that for the synthesized tone without any fluctuations. In Fig. 4 it can be seen that in the range of large $\Delta \Omega$ just detectable dip depth for speech has almost same value to that for the synthesized tone, and in the range of small $\Delta \Omega$ the predicted dip depth mentioned above agrees well with the measured value for speech sound.

Conclusion

In the range of wideband dips the sharper the dip the more difficult to distinguish it, and in this range the fluctuations of speech signal has almost no effect on the tonal differential limen. But in the range of narrow dips, dips can be rather easily distinguished because of the transient tone generated in the anti-resonant system due to the amplitude fluctuation of speech sound. It is noticeable that there exists a range where any dips can not be detectable.


Fig. 4 Tonal differential limen

---

-A-112---
Descriptive Adjectives on the Listening of Reproduced Sounds

Yutaka Enomoto and Tomio Yoshida
Matsushita Communication Industrial Co., Ltd.
Yokohama, Japan

Introduction

The customer using an audio apparatus expresses usually sound qualities by means of not physical data but descriptive terms. The salesman receives information and interprets it with his own judgement, and communicates customers' opinions to engineers or designers with descriptive terms.

As an adjective has slightly different meaning for each person, everyone can not communicate exactly his opinion each other for sound qualities. It is difficult to standardize a position of descriptive adjectives in the semantic space. However, we can communicate exactly information by clarifying the deposition of descriptive adjective. Discussions were made on the difference of adjectives in the semantic space by means of factor analysis in which ordinary people, engineers and psychologists use.

Method and Procedure

First of all, about 200 descriptive adjectives were collected from magazines, books and various papers. One term was adopted from similar terms and technical terms were omitted for ordinary subjects. Therefore, 33 terms were presented to the subjects.

As stimuli four reproduced sounds (piano, symphony, male-voice and noise) which were received from ordinary recording disks were used. The subjects
### Descriptive Adjectives on the Listening of Reproduced Sounds

<table>
<thead>
<tr>
<th>No.</th>
<th>descriptive adjectives</th>
<th>O Group</th>
<th>F Group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>I</td>
<td>II</td>
</tr>
<tr>
<td>1</td>
<td>o o hiku yōna</td>
<td>echoic</td>
<td>-.29</td>
</tr>
<tr>
<td>2</td>
<td>zarasara shita</td>
<td>rough</td>
<td>-.34</td>
</tr>
<tr>
<td>3</td>
<td>fukami no aru</td>
<td>deep</td>
<td>.48</td>
</tr>
<tr>
<td>4</td>
<td>nebai</td>
<td>viscid</td>
<td>.11</td>
</tr>
<tr>
<td>5</td>
<td>yaseta</td>
<td>thin</td>
<td>-.42</td>
</tr>
<tr>
<td>6</td>
<td>kurai</td>
<td>dark</td>
<td>-.28</td>
</tr>
<tr>
<td>7</td>
<td>baranu no toretarare</td>
<td>well-balanced</td>
<td>.71</td>
</tr>
<tr>
<td>8</td>
<td>katai</td>
<td>hard</td>
<td>-.41</td>
</tr>
<tr>
<td>9</td>
<td>sōkaina</td>
<td>refresh</td>
<td>.46</td>
</tr>
<tr>
<td>10</td>
<td>rittaikan no aru</td>
<td>presence</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>yawarakai</td>
<td>soft</td>
<td>.40</td>
</tr>
<tr>
<td>12</td>
<td>sōsōshii</td>
<td>noisy</td>
<td>-.31</td>
</tr>
<tr>
<td>13</td>
<td>namesakana</td>
<td>smooth</td>
<td>.72</td>
</tr>
<tr>
<td>14</td>
<td>tsuya no aru</td>
<td>glossy</td>
<td>.79</td>
</tr>
<tr>
<td>15</td>
<td>heibon na</td>
<td>common</td>
<td>.04</td>
</tr>
<tr>
<td>16</td>
<td>hinjaku na</td>
<td>poor</td>
<td>-.42</td>
</tr>
<tr>
<td>17</td>
<td>usupperana</td>
<td>thin</td>
<td>-.33</td>
</tr>
<tr>
<td>18</td>
<td>sensai na</td>
<td>delicate</td>
<td>.11</td>
</tr>
<tr>
<td>19</td>
<td>gōka na</td>
<td>gorgeous</td>
<td>.37</td>
</tr>
<tr>
<td>20</td>
<td>kinsoku sei no</td>
<td>metallic</td>
<td>-.07</td>
</tr>
<tr>
<td>21</td>
<td>nobi no aru</td>
<td>extended</td>
<td>.58</td>
</tr>
<tr>
<td>22</td>
<td>marumi no aru</td>
<td>rounded</td>
<td>.49</td>
</tr>
<tr>
<td>23</td>
<td>hori no aru</td>
<td>chiseled</td>
<td>.44</td>
</tr>
<tr>
<td>24</td>
<td>hade na</td>
<td>gay</td>
<td>.21</td>
</tr>
<tr>
<td>25</td>
<td>chikarasuyoi</td>
<td>powerful</td>
<td>.26</td>
</tr>
<tr>
<td>26</td>
<td>kusunda</td>
<td>dull</td>
<td>-.51</td>
</tr>
<tr>
<td>27</td>
<td>yutaka na</td>
<td>rich</td>
<td>.62</td>
</tr>
<tr>
<td>28</td>
<td>akarui</td>
<td>bright</td>
<td>.51</td>
</tr>
<tr>
<td>29</td>
<td>keikai na</td>
<td>light</td>
<td>.53</td>
</tr>
<tr>
<td>30</td>
<td>tokeatta</td>
<td>blended</td>
<td>.51</td>
</tr>
<tr>
<td>31</td>
<td>sukkiri shita</td>
<td>clear</td>
<td>.47</td>
</tr>
<tr>
<td>32</td>
<td>ikiliki shita</td>
<td>vivid</td>
<td>.51</td>
</tr>
<tr>
<td>33</td>
<td>surudoi</td>
<td>sharp</td>
<td>.20</td>
</tr>
</tbody>
</table>

| Ratio of variance contributed by factors (%) | 47 | 35 | 18 | 50 | 27 |

---

*Table 1: Factor Matrix*
### Descriptive Adjectives on the Listening of Reproduced Sounds

<table>
<thead>
<tr>
<th>III</th>
<th>h^2</th>
<th>I</th>
<th>II</th>
<th>III</th>
<th>h^2</th>
<th>No. persons (O Group), 8 audio engineers (E group) with a few years experience and 5 psychologists (P group) with an experience of a few years in the experiment of semantic analysis.</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.60</td>
<td>0.130</td>
<td>0.10</td>
<td>0.31</td>
<td>-0.44</td>
<td>0.307</td>
<td>1</td>
</tr>
<tr>
<td>0.43</td>
<td>0.637</td>
<td>-0.55</td>
<td>0.62</td>
<td>0.10</td>
<td>0.714</td>
<td>2</td>
</tr>
<tr>
<td>0.21</td>
<td>0.661</td>
<td>-0.79</td>
<td>0.16</td>
<td>-0.27</td>
<td>0.736</td>
<td>3</td>
</tr>
<tr>
<td>0.09</td>
<td>0.536</td>
<td>0.27</td>
<td>-0.10</td>
<td>-0.47</td>
<td>0.311</td>
<td>4</td>
</tr>
<tr>
<td>0.66</td>
<td>0.666</td>
<td>-0.70</td>
<td>-0.05</td>
<td>0.35</td>
<td>0.624</td>
<td>5</td>
</tr>
<tr>
<td>0.12</td>
<td>0.380</td>
<td>0.37</td>
<td>-0.18</td>
<td>0.23</td>
<td>0.228</td>
<td>6</td>
</tr>
<tr>
<td>0.11</td>
<td>0.567</td>
<td>-0.79</td>
<td>-0.14</td>
<td>0.13</td>
<td>0.670</td>
<td>7</td>
</tr>
<tr>
<td>0.62</td>
<td>0.639</td>
<td>-0.30</td>
<td>0.52</td>
<td>0.26</td>
<td>0.435</td>
<td>8</td>
</tr>
<tr>
<td>0.41</td>
<td>0.491</td>
<td>0.72</td>
<td>0.12</td>
<td>0.31</td>
<td>0.643</td>
<td>9</td>
</tr>
<tr>
<td>0.17</td>
<td>0.399</td>
<td>0.58</td>
<td>-0.47</td>
<td>-0.27</td>
<td>0.650</td>
<td>10</td>
</tr>
<tr>
<td>0.07</td>
<td>0.639</td>
<td>-0.58</td>
<td>-0.48</td>
<td>-0.27</td>
<td>0.656</td>
<td>11</td>
</tr>
<tr>
<td>0.43</td>
<td>0.485</td>
<td>0.47</td>
<td>0.71</td>
<td>0.04</td>
<td>0.745</td>
<td>12</td>
</tr>
<tr>
<td>0.16</td>
<td>0.537</td>
<td>0.65</td>
<td>-0.56</td>
<td>-0.23</td>
<td>0.804</td>
<td>13</td>
</tr>
<tr>
<td>0.18</td>
<td>0.664</td>
<td>0.76</td>
<td>-0.36</td>
<td>-0.16</td>
<td>0.747</td>
<td>14</td>
</tr>
<tr>
<td>0.33</td>
<td>0.259</td>
<td>-0.37</td>
<td>-0.37</td>
<td>0.03</td>
<td>0.285</td>
<td>15</td>
</tr>
<tr>
<td>0.47</td>
<td>0.699</td>
<td>-0.69</td>
<td>-0.10</td>
<td>0.18</td>
<td>0.522</td>
<td>16</td>
</tr>
<tr>
<td>0.62</td>
<td>0.635</td>
<td>-0.68</td>
<td>-0.19</td>
<td>0.28</td>
<td>0.592</td>
<td>17</td>
</tr>
<tr>
<td>0.54</td>
<td>0.563</td>
<td>0.57</td>
<td>-0.19</td>
<td>0.28</td>
<td>0.445</td>
<td>18</td>
</tr>
<tr>
<td>0.27</td>
<td>0.289</td>
<td>0.56</td>
<td>0.55</td>
<td>0.08</td>
<td>0.630</td>
<td>19</td>
</tr>
<tr>
<td>0.61</td>
<td>0.591</td>
<td>0.33</td>
<td>0.61</td>
<td>0.37</td>
<td>0.625</td>
<td>20</td>
</tr>
<tr>
<td>0.05</td>
<td>0.629</td>
<td>0.77</td>
<td>0.18</td>
<td>-0.06</td>
<td>0.635</td>
<td>21</td>
</tr>
<tr>
<td>0.12</td>
<td>0.728</td>
<td>0.75</td>
<td>-0.29</td>
<td>-0.21</td>
<td>0.738</td>
<td>22</td>
</tr>
<tr>
<td>0.09</td>
<td>0.570</td>
<td>0.69</td>
<td>-0.10</td>
<td>-0.05</td>
<td>0.493</td>
<td>23</td>
</tr>
<tr>
<td>0.46</td>
<td>0.377</td>
<td>-0.25</td>
<td>0.56</td>
<td>0.23</td>
<td>0.445</td>
<td>24</td>
</tr>
<tr>
<td>0.10</td>
<td>0.307</td>
<td>0.56</td>
<td>0.72</td>
<td>-0.10</td>
<td>0.859</td>
<td>25</td>
</tr>
<tr>
<td>0.15</td>
<td>0.537</td>
<td>-0.58</td>
<td>0.22</td>
<td>-0.21</td>
<td>0.431</td>
<td>26</td>
</tr>
<tr>
<td>0.11</td>
<td>0.692</td>
<td>0.89</td>
<td>0.04</td>
<td>-0.14</td>
<td>0.830</td>
<td>27</td>
</tr>
<tr>
<td>0.35</td>
<td>0.495</td>
<td>0.68</td>
<td>0.09</td>
<td>0.49</td>
<td>0.715</td>
<td>28</td>
</tr>
<tr>
<td>0.48</td>
<td>0.566</td>
<td>0.60</td>
<td>0.07</td>
<td>0.56</td>
<td>0.685</td>
<td>29</td>
</tr>
<tr>
<td>0.31</td>
<td>0.732</td>
<td>0.71</td>
<td>0.37</td>
<td>0.04</td>
<td>0.653</td>
<td>30</td>
</tr>
<tr>
<td>0.19</td>
<td>0.631</td>
<td>0.73</td>
<td>-0.07</td>
<td>0.30</td>
<td>0.639</td>
<td>31</td>
</tr>
<tr>
<td>0.30</td>
<td>0.395</td>
<td>0.80</td>
<td>0.42</td>
<td>0.06</td>
<td>0.824</td>
<td>32</td>
</tr>
<tr>
<td>0.35</td>
<td>0.394</td>
<td>0.40</td>
<td>0.56</td>
<td>0.39</td>
<td>0.637</td>
<td>33</td>
</tr>
</tbody>
</table>

---

**Results and Conclusions**

The factor matrix is presented in Table 1. From the mean squared factor loadings at the bottom columns in Table 1, it was found out that factor I extracted from the above analysis can be accounted for.
Descriptive Adjectives on the Listening of Reproduced Sounds

approximately 50% of the total variance in judgements. Adjectives of "well-balanced (baransu no toreta)" and "rich (yutakana)" used in the all groups occupied almost the same position in the semantic space. In comparison with the other groups, there were found out in E group "noisy (sōsōshii)" and "gay (hade na)" characterised on the semantic space, in P group "thin (usupera na)" and "chiseled (hori no aru)". "Echoic (o o hiku yōna)" scattered at different positions in the space in each group.

There are descriptive adjectives that we can use to communicate the opinion for sound qualities or not use in same meaning. Therefore, descriptive adjectives that are placed at same position in semantic space should be use to communicate the information for sound qualities.

Reference

Factor Analytical Research of Tone Colour

Otoichi Kitamura
The Institute of Scientific and Industrial Research,
Osaka University

Seiichiro Namba
The college of General Education, Osaka University

Himpei Matsumoto
Audio Research, Central Research Laboratory,
Matsushita Electric Industrial Co., Ltd.

Introduction

The timbre has usually been defined as the peculiarity by which one is able to distinguish a sound from other sounds and it depends upon the structure of partial tone of the sound.

The expression, "tone colour" used in this paper, does not only mean the peculiarity of sound but it means the whole tonal experience when we hear a sound. In other words, the tone colour means the impression which is made from the pattern of pitch and loudness in the total sensation of tone heard.

The impressions which are made by all sounds, can be described by various kinds of language. It is necessary to obtain the required and sufficient number of descriptive adjectives of tone colour, for the purpose of describing the whole range of tonal experience.

Solomon did the study in an attempt to derive a limited number of descriptive adjectives with which to characterize passive sonar sounds. But there was no attempt to make a study of the tone colour of all kinds of sound, namely music, voice and noise. Then through the factor analysis, 21 studies were carried out to research the number of factors of tone colour of all kinds of sound. The methodology employed utilized the Semantic Differential Method developed by Osgood.

Three orthogonal factors were extracted in each study and the universality of the factors of tone colour was confirmed.
Factor Analytical Research of Tone Colour

Procedures

As many descriptive adjectives of tone colour as possible were collected from a number of books and magazines of music and acoustics, and from acoustical engineers, musicians and general civilians.

The sound stimuli were selected from music, noise and human voice. The music and noise were selected from discs, and the human voice was selected from readings by radio announcers.

Ordinary, duration of each stimulus was about 5 seconds.

The subjects estimated on a seven-point scale, each scale being defined by a polar adjective, (uni-polar), or a pair of polar-opposite adjectives, (bi-polar), at every presentation of a stimulus.

To confirm the universality of the factors of tone colour, 21 experiments were made by changing stimuli, scales, subjects and listening rooms. In experiment No. 1, 138 seven-point uni-polar scales were used and in experiment No. 2, 64 seven-point bi-polar scales were used. The intercorrelations between scales, (calculated over both subjects and sounds), were factor analyzed by the Thurstone complete centroid method. In the following experiments, (No.3, No. 4 etc.), the scales being defined by a pair of polar-opposite adjectives were constructed using these results of factor analysis.

Results and discussion

The final rotated factor matrix is presented in Table II. This is the result of the first half of these studies. Three principal orthogonal factors of tone colour were extracted: the factor of beauty, power and metallic impression. The summation of variance of these three factors was more than 80% throughout all experiments.

As the result, the tone colour can be indicated by a point in the "semantic space" constructed of the typical scales of these three factors.

The experiments were made by changing the degree of nonlinear distortion, the frequency range of reproduced sound and the pressure level of reproduced sound.

These sounds were ranged on the above-mentioned typical scales. Meaningful relationships were found between ranks on certain scales and the physical properties of these sounds.

### Table I List of subjects and stimuli

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>no.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>Unipolar 138</td>
<td>Unipolar 64 (pair of negative words)</td>
<td>Bipolar 10 (pair of antonyms)</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

#### Scale

<table>
<thead>
<tr>
<th></th>
<th>age</th>
<th>Tot.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>20</td>
<td>30</td>
</tr>
<tr>
<td></td>
<td>40</td>
<td>Tot.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

#### Subjects

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

#### Listening room

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>1</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

#### Kind of sound

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

#### Date

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### 1. Scale

<table>
<thead>
<tr>
<th></th>
<th>Bipolar 10 (pair of antonyms)</th>
<th>Bipolar 11 (pair of antonyms)</th>
<th>Bipolar (pair of antonyms 10, pair of negative words 3)</th>
<th>Bipolar 10 (pair of antonyms)</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>30</td>
<td>40</td>
<td>50</td>
<td>60</td>
</tr>
<tr>
<td>30</td>
<td>40</td>
<td>50</td>
<td>60</td>
<td>70</td>
</tr>
<tr>
<td>50</td>
<td>60</td>
<td>70</td>
<td>80</td>
<td>90</td>
</tr>
<tr>
<td>60</td>
<td>70</td>
<td>80</td>
<td>90</td>
<td>100</td>
</tr>
</tbody>
</table>

---

### 4. Kind of sound

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

---

### 5. Date

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

---

### 6. Bipolar 10 (pair of antonyms)

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

---

### 7. Bipolar 11 (pair of antonyms)

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

---

### 8. Bipolar (pair of antonyms 10, pair of negative words 3)

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

---

### 9. Bipolar 10 (pair of antonyms)

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

---

### Notes

- Factor Analytical Research of Tone Colour
Table II: Final rotated factor matrix

<table>
<thead>
<tr>
<th>Descriptive Adjectives</th>
<th>Factor 1</th>
<th>Factor 2</th>
<th>Factor 3</th>
<th>Factor 4</th>
<th>Factor 5</th>
<th>Factor 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 beautiful</td>
<td>0.82</td>
<td>-0.08</td>
<td>-0.14</td>
<td>0.61</td>
<td>-0.53</td>
<td>-0.69</td>
</tr>
<tr>
<td>2 ugly</td>
<td>-5.7</td>
<td>1.12</td>
<td>1.4</td>
<td>3.3</td>
<td>-5.2</td>
<td>1.5</td>
</tr>
<tr>
<td>3 clean</td>
<td>5.3</td>
<td>-2.4</td>
<td>1.1</td>
<td>-1.4</td>
<td>3.9</td>
<td>0.4</td>
</tr>
<tr>
<td>4 dirty</td>
<td>-3.4</td>
<td>3.1</td>
<td>2.0</td>
<td>2.5</td>
<td>-3.4</td>
<td>3.1</td>
</tr>
<tr>
<td>5 balanced</td>
<td>-0.7</td>
<td>-2.5</td>
<td>0.1</td>
<td>3.2</td>
<td>1.1</td>
<td>-1.9</td>
</tr>
<tr>
<td>6 unbalanced</td>
<td>-3.1</td>
<td>0.5</td>
<td>1.3</td>
<td>0.4</td>
<td>-2.2</td>
<td>1.2</td>
</tr>
<tr>
<td>7 powerful</td>
<td>0.19</td>
<td>0.71</td>
<td>0.01</td>
<td>-1.2</td>
<td>2.1</td>
<td>1.1</td>
</tr>
<tr>
<td>8 feeble</td>
<td>-1.4</td>
<td>-2.0</td>
<td>-3.2</td>
<td>-0.6</td>
<td>-0.5</td>
<td>-0.8</td>
</tr>
<tr>
<td>9 rich</td>
<td>-0.3</td>
<td>-4.4</td>
<td>0.55</td>
<td>-0.5</td>
<td>-0.8</td>
<td>3.2</td>
</tr>
<tr>
<td>10 poor</td>
<td>-0.3</td>
<td>-4.4</td>
<td>0.55</td>
<td>-0.5</td>
<td>-0.8</td>
<td>3.2</td>
</tr>
<tr>
<td>11 deep</td>
<td>-0.3</td>
<td>-4.4</td>
<td>0.55</td>
<td>-0.5</td>
<td>-0.8</td>
<td>3.2</td>
</tr>
<tr>
<td>12 metallic</td>
<td>-0.3</td>
<td>-4.4</td>
<td>0.55</td>
<td>-0.5</td>
<td>-0.8</td>
<td>3.2</td>
</tr>
<tr>
<td>13 bright</td>
<td>0.33</td>
<td>0.55</td>
<td>0.01</td>
<td>4.6</td>
<td>-1.9</td>
<td>0.9</td>
</tr>
<tr>
<td>14 dark</td>
<td>0.68</td>
<td>-0.40</td>
<td>0.09</td>
<td>3.0</td>
<td>-2.0</td>
<td>-2.1</td>
</tr>
<tr>
<td>15 hard</td>
<td>-0.39</td>
<td>0.24</td>
<td>0.40</td>
<td>-2.3</td>
<td>0.4</td>
<td>4.0</td>
</tr>
<tr>
<td>16 soft</td>
<td>0.92</td>
<td>-0.01</td>
<td>-0.05</td>
<td>0.96</td>
<td>-0.08</td>
<td>0.95</td>
</tr>
<tr>
<td>17 ugly</td>
<td>0.91</td>
<td>-0.12</td>
<td>-0.06</td>
<td>0.95</td>
<td>-0.07</td>
<td>0.71</td>
</tr>
<tr>
<td>18 clean</td>
<td>0.79</td>
<td>0.26</td>
<td>0.05</td>
<td>-0.20</td>
<td>0.04</td>
<td>0.38</td>
</tr>
<tr>
<td>19 dirty</td>
<td>0.60</td>
<td>0.27</td>
<td>-0.41</td>
<td>-0.09</td>
<td>0.12</td>
<td>0.74</td>
</tr>
<tr>
<td>20 balanced</td>
<td>0.34</td>
<td>0.72</td>
<td>-0.13</td>
<td>0.22</td>
<td>0.08</td>
<td>0.01</td>
</tr>
<tr>
<td>21 unbalanced</td>
<td>0.21</td>
<td>0.74</td>
<td>-0.37</td>
<td>0.11</td>
<td>0.04</td>
<td>0.10</td>
</tr>
<tr>
<td>22 powerful</td>
<td>-0.01</td>
<td>-0.19</td>
<td>0.64</td>
<td>0.14</td>
<td>0.15</td>
<td>0.40</td>
</tr>
<tr>
<td>23 feeble</td>
<td>0.24</td>
<td>-0.04</td>
<td>-0.73</td>
<td>-0.11</td>
<td>0.20</td>
<td>0.54</td>
</tr>
</tbody>
</table>
A New Theory of Psychometric Estimation concerning the Sound Quality of Voice

Shuji Gotoh and Tomio Yoshida
Matsushita Communication Industrial Co., Ltd.
Yokohama, Japan

Introduction

As a method of evaluating the quality of speech transmission systems, there is the articulation theory which was established by H. Fletcher (1) aiming at the intelligibility. It is the author's purpose to establish a method evaluating a reproducing system in view of the similarity of reproduced sound.

In this report, it is described to calculate the quality of a sound system by observing the frequency response and transmission level.

Here, we define that the change of sound quality $\Delta q$ is difference of similarity between the reproduced sound made through a high pass filter system and the reproduced sound made through the system of the ideal flat frequency response.

The sound quality is scaled by subjective method and is represented numerically with the psychological scale. The scaling method is used the method of paired comparisons (based on the Thurstone's Case III).

Experiments & Consideration

Fig. 1 shows the change of sound quality reproduced through the high pass filter and the low pass filter systems at the time when the cut-off frequencies were changed. Under this experiment, the transmission level was supplemented properly.
in order to keep the whole loudness constant, even if the cut-off frequencies were changed. The curve I indicates the change of sound quality in the high pass filter system and the curve II indicates that of the low pass filter system respectively.

By the limitation of the experimental conditions, the dotted curve III which is drawn symmetrically against the curve II at the point of Fo will be assumed to be an extended part of the curve I. The whole change of sound quality (Thurstone's Scale q) over the all frequency band is normalized in value Q ranging from 1 to 0.

The slope of these curve I & III \( |\Delta Q/\Delta f| \) may represent the volume of sound quality elements per one cycle. Fig. 2 shows \( |\Delta Q/\Delta f| \).

This curve is similar to the curve of the loudness function \( G_4 \) (defined by Fletcher and measured by Yamaguchi in Japanese), not only in its figure but also in its numerical value. This curve (we named here the frequency importance function of the quality) means the contribution to the sound quality of each frequency.

The sound quality being varied with the change of the level , on which is the lowest frequency band (Fig. 3 hatched part) of the transfer band in the high pass
filter systems, was obtained under an experiment.

The loudness \( N \) corresponding to \( \alpha \) being changed is shown with the longitudinal axis in Fig. 3.

The result of this experiment is shown with dots in Fig. 4. The line in Fig. 4 shows the value which was calculated according to a theory described below.

Let's consider the frequency band divided into small bands. Its width is indicated as \( \Delta f_i \) and the sound quality element included in \( \Delta f_i \) is indicated as \( \Delta Q_i \) (\( i = 1, 2, \ldots, n \)). The summation of the quality elements included in each band is supposed to constitute the whole sound quality. The quality element of each band is a function of band sensation level \( Z \) and is represented as \( \Delta Q_i (1 - k_i(Z)) \). \( \Delta Q_i \) is the maximum value in \( i \)-th band, and at the optimum transmission level of \( Z_0 \), it is obtained as \( k_i(Z_0) = 0 \). Therefore, the total sound quality \( Q \) is given by

\[
Q = 1 - \sum_{i=1}^{n} k_i(Z) \Delta Q_i
\]  

(1)

On the other hand, since \( |\partial Q/\partial f| \) is the density of the quality \( \Delta Q_i \) can be represented as follows:

\[
\Delta Q_i = \left| \frac{\partial Q}{\partial f} \right| \Delta f_i
\]  

(2)

As \( k_i(Z) \) can be determined according to the frequency, it may be expressed by the function of frequency \( f \) , \( k(f, Z) \), instead of affixing \( i \). Tending \( \Delta f_i \) to zero, \( Q \) becomes as follows:

\[
Q = 1 - \int_{0}^{\infty} k(f, Z) \left| \frac{\partial Q}{\partial f} \right| df
\]  

(3)
Next, let's assume that the sound quality change refers to the product of the change of loudness $\Delta N$ by $|\partial Q/\partial f|$ at the position where the loudness change occurs, because the sound quality change is one of the phenomena on the basilar membrane being caused by the transmission level and frequency response change.

The scale of sound quality change $\Delta q(\alpha)$ is obtained from the equation (4) with $\Delta N$, $x$ (position from a helicotrema percent of total nerve endings)\(^{(1)}\) and $|\partial Q/\partial f|$ as follows:

$$\Delta q(\alpha) = a \int_{0}^{100} \Delta N_x(\alpha) \left| \frac{\partial Q}{\partial f} \right| dx = a \int_{0}^{\infty} \Delta N_f(Z) \left| \frac{\partial Q}{\partial f} \right| F(f) df \ldots \ldots (4)$$

But "a" is a proportional constant. $\Delta N_x(\alpha) = N(Zo) - N(Z)$ is the loudness change corresponding to the change of level $\alpha$ in $\Delta f_i$.

$\alpha = Zo - Z$ is the change of level, $F(f)$ is the function to convert $x$ to $f$.

After all, the equation (3) becomes as follows:

$$Q = 1 - a \int_{0}^{\infty} \Delta N_f(Z) \left| \frac{\partial Q}{\partial f} \right| F(f) df \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots (5)$$

Substituting the experimental condition for the equation (5), the calculated result fits accurately to the result of the experiment, as is shown in Fig. 4.

**Conclusion**

It can be analysed that how to influence the change of the frequency characteristics and transmission level upon the quality of the voice. Then, it is found that the quality change is depended on the loudness change $\Delta N$ and the frequency importance function of the quality $|\partial Q/\partial f|$ at the position where the loudness change occurs. The formula to calculate the change of sound quality reproduced by the high pass filter system was obtained by an equation which is a function of frequency ($f$) and transmission level ($Z$), i.e.

$$Q = 1 - a \int_{0}^{\infty} \Delta N_f(Z) \left| \frac{\partial Q}{\partial f} \right| F(f) df$$

This equation fits accurately to the result of the experiment.

\(^{(1)}\) H. Fletcher: Speech and Hearing in Communication
ADDITIVITY IN CONSONANCE AND ITS CALCULATION METHOD

Akio KAMEOKA and Mamoru KURIYAGAWA
Central Research Laboratory
Tokyo Shibaura (Toshiba) Electric Co., Ltd.
1 Komukai Toshiba-cho, Kawasaki, Japan

Introduction
Any conception is not established yet for addition of such prothetic sensation as consonance, which does not apparently correspond with physical value. The consonance is here defined as an auditory sensation of "clearness" and the dissonance, "turbidity". It is an important attribute of complex tones as well as the loudness or pitch, and plays an important role in determining pleasantness of musical tones and it also has close relation with the nonlinear distortion of audio instruments. The purpose in this paper is to establish a theory of dissonance calculation for static multi-component tones from dissonance values for two-partial tones experimentally determined.

The Power Law and Its Extension
Some preliminary experiments confirmed the consonance sensation is additive toward the direction of dissonance, and the dissonance belongs to a prothetic continuum. Prothetic continua generally follows the Power Law proposed by Stevens, which express psychological response \( R \) as a function of physical intensity \( I \) as follows:

\[
R = k I^n
\]

where \( k \) is a constant depending on the scale unit of \( R_s \) and the exponent \( n \) is a constant consistent with the sensation. In case of the dissonance, however, its corresponding physical values are not obvious. In order to extend the Power Law to such prothetic sensation, the authors introduced a new concept of "Physiological Power" in Fig. 1, which is assumed to have additivity and to exist at or near the final stage of physiological perception processings, presumably in a form of impulse train. The auditory perception process is divided into two, the first represent Physiological Process which converts the physical stimuli into the physiological power \( P \), and the second, Psychological Process which converts \( P \) into the psychological response \( R \). For all prothetic continua, the Power Law is assumed by extension to

Fig. 1 Perception model for prothetic continua in Class I and II
Additivity in Consonance and its Calculation Method

hold in the Process II between P and R as follows:

\[ R = \phi_2(P) = k_2 P^{n_2} \]  \hspace{1cm} (2)

Prothetic continuum in Class I which is apparently a function of physical intensity I such as loudness or brightness, is well known to follow the Stevens' Power Law. Consequently, the Process I in those continua is expressed by a power function as follows:

\[ P = \phi_1(I) = k_1 I^{n_1} \]  \hspace{1cm} (3)

Then, the Power Law is re-written as:

\[ R = k_2 P^{n_2} = k_2 (k_1 I^{n_1})^{n_2} = k_2 k_1^{n_2} I^{n_1 n_2} = k I^n \]  \hspace{1cm} (4)

\[ k = k_2 k_1^{n_2} \quad \text{and} \quad n = n_1 n_2 \]  \hspace{1cm} (5)

where \( k_1 \), \( n_1 \) and \( k_2 \), \( n_2 \) are constants for Process I and II respectively.

On the other hand, for the prothetic continuum in Class II which is functions of several physical parameters as the dissonance sensation, the function \( \phi_2 \) of Process I is usually more complex and is unknown. Fig. 1 shows the perception model for prothetic continua in Class I and II.

Sone Scale and \( \lambda \) Scale Interpreted by the Extended Power Law

In loudness calculation of multi-component tones, the sone scale adds directly as a linear summation \( S_s = \sum S_i \), while the \( \lambda \) scale of Garner\(^5\) adds as a square summation \( \lambda_\lambda = (\sum \lambda_i^2)^{1/2} \), where \( S_i \) and \( \lambda_i \) are total loudness, and \( S_i \) and \( \lambda_i \) are loudness of each components in their respective scales. The fact that both of those methods are fairly successful, indicates \( S_s \) and \( \lambda_\lambda \) are additive values and have a same dimension with and be proportional to the physiological power of loudness \( L_{pi} \), which is independent of the method of scaling. Therefore, the exponents \( n_s \) and \( n_\lambda \) in Process II for sone and \( \lambda \) scale are turned out 1.0 and 0.5 respectively. Accordingly, the sone and \( \lambda \) scales are expressed as a function of \( L_{pi} \) as follows:

\[ S_s = k_2 s L_{pi}^{1.0} \quad \text{and} \quad \lambda_\lambda = k_2 \lambda L_{pi}^{0.5} \]  \hspace{1cm} (6)

According to Stevens' measurement, \( n_s = 0.3 \), then \( n_{1s} = n_s / n_2^s = 0.3 / 0.5 = 0.3 \), which is close to 0.132 measured by Garner, Later, Stevens\(^6\), Hellman and Zwisloski\(^7\), and Lochner and Burger\(^8\) independently measured the exponent for sone scale as 0.27. The exponent for \( \lambda \) scale is similarly estimated as \( n_\lambda = 0.27 x 0.5 = 0.135 \), which agrees very good with the Garner's result of 0.132. This supports the model in Fig. 1, for the sone and \( \lambda \) scales are clearly interpolated as shown in Fig. 2. It may be concluded that...
the ratio judgements adopted for the scale takes 1.0 and the combination of ratio and equi-section judgements adopted for the scale takes 0.5 as the exponent in Process II.

**Dissonance Perception Model**

The basic thought above mentioned is applied to dissonance perception as shown in Fig. 3. All the components comprising a complex tone are first decomposed into all the combinations of two components in the Comparison Process (1), which outputs the physiological power of dissonance or "Dissonance Power" \( D_{p2i} \) here called. Put in parallel to (1) is the Noise Detection Process (2) which gives dissonance power \( D_{pn} \) of the external and internal noises. The next is the Addition Process (3) where the dissonance power linearly adds. The total dissonance power \( D_{pt} \) is finally transformed in the Transformation Process (4) into the total dissonance \( D_m \) of the complex tone, which is expressed in the absolute dissonance scale\(^9\) (AD).

**Procedure of Dissonance Calculation**

The above dissonance perception model automatically shows the calculation procedures as depicted in Fig. 4. First, decompose all the components into two-component pairs and calculate absolute dissonance \( D_{2i} \) for each pair with empirical equations\(^2\) which were already determined for every possible two-partial tones. Fig. 5, for example, shows a typical consonance curve for two-partial tones with an equal sound pressure level of 57 dB(SPL). The dissonance value \( D_{2i} \) is then transformed once into its corresponding dissonance power \( D'_{p2i} \) by the extended Power Law. After subtracting the dissonance power of noise \( D_{pn} \), the real dissonance powers \( D_{p2i} (= D'_{p2i} - D_{pn}) \) and \( D_{pn} \) are added, then the total dissonance power \( D_{pt} \) is again retransformed into the absolute dissonance \( D_m \), as follows:

\[
D_m = k_2 \left( D_{pt} \right)^{\beta_2} = k_2 \left( \sum_{i=1}^{M} D_{p2i} + D_{pn} \right)^{\beta_2}
\]  

(7)

where the exponent \( \beta_2 \) is a constant consistent with the dissonance. The result of calculation for two experiments are shown in Fig. 6 and 7, together with the experimental results scaled by the method of incomplete paired comparisons after scale conversion to AD. The subjects are men with average age of about 23, who are taking a regular course of training as audio-consultants. The exponent was decided in experiment I so as to minimize the discrepancy between the calculation and the experimental results. The same value was applied to experiment II, in which the samples include more components and covers wider frequency range than the first. The results of calculation agrees very good with the experiments.

---

**Fig. 4** Dissonance calculation procedures for static multi-component tones

**Fig. 5** Consonance characteristics of two-partial tones with equal sound pressure of 57 dB(SPL)
Conclusions

Introducing a new concept of "Physiological Power" at or near the final stage of physiological perception processing, the Power Law was extended to the prothetic continua in Class II, which is not simply a function of physical intensity, but of several physical parameters. The perception process of prothetic continuum was divided into two steps: Physiological Process (I) and Psychological Process (II). The Power Law was assumed to hold in the Process I. As an example of prothetic continuum in Class I which is a simple function of physical intensity, the loudness sensation was discussed. The difference between the sone scale and λ scale was clearly explained by the model. Next, the extended Power Law was applied to construct a perception model of dissonance in Class II. And a dissonance calculation method was developed depending on the model. The calculation showed good agreements with the experiments. This also suggests, as the first example do, the above conception on the prothetic continua will be applicable to other prothetic sensations both in Class I and II.

Acknowledgements

The authors are indebted to Dr. Tarow Inadō, Dr. Soichiro Kuroki and Dr. Hisao Sakai for their advices and discussions.

Fig. 6 Comparison of theoretical calculation with the experiment (1) for 11 complex tones. The ordinate represent \( D_m \) in AD.

References

4) S. S. Stevens, Psychol. Rev., 64 153-181 (1957)
6) S. S. Stevens, Psychom. 26 35-47 March (1961)

Fig. 6 Similar comparison of calculation with the experiment (2) for 14 complex tones
Introduction

In a previous paper, we discussed the basic concepts and methodology for a new method of designing sound quality for reproducing systems. The present report, partly overlapping the previous one, deals with the basic concepts of this method and some examples of application to the design of stereo phonographs.

Model of the process of sound quality evaluation

We assumed a model of the process of sound quality evaluation as shown in Fig. 1. In this model, the psychological process in a listener is divided into two, namely the elemental sensory process and the synthetic emotional one. The former is considered to be invariant from person to person or from era to era, and therefore can be related consistently and reliably to the physical characteristics of the transmitting system. On the other hand, the latter varies depending on the individual and era involved.

In the elemental sensory process, multidimensional elemental sensations $D_i (i=1,2,...,n)$ are formed by listening to reproduced sounds. $D_i$ can be expressed as functions of $t_j$, the physical parameter of transmitting system $T$, and of $s_i$, the physical component of program source $S$ contributing to $D_i$. Namely,

$$D_i = \Phi_i (t_j, s_i)$$  \hspace{1cm} (1)

It is assumed that "preference" (synthetic emotional response) $R$ for reproduced
Design of reproduced sound quality by ESP method

sound can be expressed, as shown by Eq. (2), by the sum of products of \( D_i \) and the
weighting factor \( w_i \) which is determined by the need of the listener.

\[
R = \sum w_i D_i
\]  

(2)

It is considered that \( w_i \) varies depending upon not only listener, \( L \), but also musical
content of program source, \( S \), and era, \( A \). In other words,

\[
w_i = \Psi (L, S, A)
\]  

(3)

In accordance with the consideration given above, we first quantified the relationship between various physical characteristics and \( D_i \) using the multidimensional scaling method,\(^{(2)}\) and made clear that the sensation of reproduced sounds can be
expressed sufficiently by five or so dimensional elemental sensations.\(^{(1)}\) Next, we
proved experimentally, as described later, that between \( R \) and \( D_i \) there exists the
relationship as shown by Eq. (2). This method was named "ESP method" because the
evaluation and design of sound quality are based on the above-mentioned relationship
among Emotional response, Sensation and Physical characteristics.

Quantification of synthetic emotional response

In order to study the validity of the linear model (Eq. (2)), some experiments
were made. First,\(^{1)}\) two dimensional elemental sensory scales, \( D_1 \) and \( D_3 \), and
preference scale \( R \) were obtained as shown in Fig. 2. Calculating multiple linear
regression coefficients of \( R \) to \( D_i \), we obtained \( w_i \). \( \bar{R} \), the estimated value of \( R \), was
obtained by the sum of products of \( w_i \) and \( D_i \). \( \bar{R} \) is shown in Fig. 2 together with \( R 
and \( D_i \). As observed in this figure, there is a relatively good coincidence between
\( R \) and \( \bar{R} \). Similar experiments\(^{3)}\) were conducted by increasing the number of \( D_i \) up to
five. Thus, it was observed that the coincidence between \( R \) and \( \bar{R} \) was vastly improved.

Now, we assumed that \( w_i \) is a function of only \( L \), \( S \) and \( A \), but does not vary
with \( t_j \) as shown by Eq. (3). Therefore, once \( w_i \) for certain \( L \) and \( S \) is obtained,
as far as \( L \) and \( S \) are fixed, \( R \) for various \( t_j \) can be estimated by the curves of \( \phi_i \).
The estimated value \( \bar{R} \) should coincide with \( R \) for such \( t_j \). Some experiments\(^{3)}\) were
made to prove these assumptions. An example of the results is shown in Fig. 3. It
is evident from the figure that \( R \) and \( \bar{R} \) coincide with each other very well, and that
\( w_i \) is independent of \( t_j \). It is believed that this fact gives another piece of

- A - 130 -
Design of reproduced sound quality by RSP method

evidence of the validity of our model. Furthermore, some experiments were made by
the multidimensional quantification method, to classify the listeners into several
groups in which preference for sound quality is homogenous. As a result, there were
found few differences in $D_1$ in spite of a great many in $R$ among the listener groups
that differ from each other in preference, and the validity of our model was proved
again. Finally, "Designing chart of sound quality" was prepared by the
numerous data obtained in our studies. Using this, $\bar{D}_1$ and $\bar{R}$, the estimated values
of $D_1$ and $R$ respectively, were calculated for a certain number of stereo phonographs
from their physical characteristics. At the same time, $D_1$ and $R$ were measured for
these sets. Some of the calculated and the measured values are shown in Fig. 4 and
Fig. 5. As evident in these figures, the calculated and the measured values coin-
cide with each other very well as to both $D_1$ and $R$

Thus, with the designing chart, it has become possible to estimate the degree of
preference for sound quality of a transmitting system by measuring its physical
characteristics, and vice versa. However, because $\omega_1$ varies depending upon $L$, $S$
and $A$, it is necessary from now on to classify $L$ and $S$ effectively, and to obtain $\omega_1$
in each combination of classified $L$ and $S$ at an interval of several years. Studies
on these programs are now in progress.

Conclusion

By the sound quality evaluation method given above, it has become possible to
design a sound reproducing system with a sound quality most preferable to the
customer. This method was named "RSP method" and has been practically applied to
design acoustic products.

We wish to express our profound gratitude to Prof. T. Itoo of Waseda Univ.,
to Dr. C. Hayashi of Inst. Statist. Mathematics, to Prof. M. Omiya of Ochanomizu
Women's Univ. and to Dr. S. Kuroki of Fundamental Research Lab. of the Japan Broad-
casting Corp., who kindly gave us valuable suggestions and many opportunities for
discussion.
Design of reproduced sound quality by ESP method

Fig. 1. Model of the process of sound quality evaluation

Fig. 2. Relation between preference scale R and elemental sensory scales D_i

Fig. 3. Comparison of estimated scale value of preference R and measured scale value R

Fig. 4. Comparison of calculated scale value D_i and measured scale value D_i for stereo phonograph sets

Fig. 5. Comparison of calculated value of preference R and measured scale value R for stereo phonograph sets

References

1) T. Koshikawa, T. Nakayama, R. Miyagawa: "On the designing method of reproduced sound quality..." 5th ICA, M67


Über das kontextabhängige auditive Perzeptionsverhalten bei nicht-periodischen akustischen Schwingungsstrukturverläufen

Volker Nössel

z.Zt. Institut voor Psychoakoestiek en Elektronische Muziek
Rijksuniversiteit - Gent


Um als Basis hierfür einen realistischen Einblick in die komplexe Verarbeitungsweise des auditiven Perzeptionsmechanismus bei nicht-periodischen akustischen Schwingungsstrukturen zu erhalten, wurden folgende Hörsstitute durchgeführt (4)(6):

Der auditive Stimulus bestand aus einem Muster zweier Impulse (tIB) mit bestimmter Zeitpunktierung (tIS).

Diese fixen Muster (FRM) traten im stochastischen Zeitabstand (ZTA) repetierend auf.

$tIB$ und $tIS$ des FRM konnten in 10 Stufen variiert werden.

Es wurden Intervallschritt-Hörsstitute (4)(6) durchgeführt, die bei aleatorischer Werte-Programmierung folgendes Resultat ergaben:

Bei akust. Elementevariation (SKI nach SKII)

( Typ $\alpha$: $t_b < t_a$ / Typ $\beta$: $t_a < t_b$ )

können in der auditiven Tonhöhen-Bewegungs-
Tendenz-Bestimmung (TBTB) (steigend: $+$; fallend: $-$ )

TBTB-Inversionen als Response auftreten.

Durch Zusammenfügung von $\alpha$- mit Equivalen-
ter $\beta$-Variation zu $\alpha/\beta$-Einheiten ergeben sich in der Darstellung als auditives Resultat folgende Möglichkeiten:

<table>
<thead>
<tr>
<th>TBTB bei Variation: $\alpha$</th>
<th>$\beta$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1) positive Reversibilität</td>
<td>$+$</td>
</tr>
<tr>
<td>2) negative</td>
<td>$-$</td>
</tr>
<tr>
<td>3) positive Irreversibilität</td>
<td>$+$</td>
</tr>
<tr>
<td>4) negative</td>
<td>$-$</td>
</tr>
</tbody>
</table>
Durch Parallelssetzung der IS-Ebenen läßt sich eine TTB-VFerumlkehrung zeigen, aus der abschließend die Bestimmung der akust. Kondition für ein reizadäquates TTB-Verhalten abgeleitet werden kann.

Die relativen Tonhöhenpositionen der TTB lassen sich durch die Maximal-t-Werteprünge [(TETW: IB/IS; IB+IS (symmetr. oder asymmetr.))] äquivalent mit der akust. Skala (t = Z (t2IB+IS) perzeptiv in reizadäquater topologischer Darstellung zeigen. (Die Lmns der akust. Skala sind: Tmax: 3x10 / Tmin.: 3x1)

Trotz des adäquaten Verlaufes von akust. und auditiver Skala, kann man die Resultate nicht ohne weiteres traditionell auf Amplituden-Rhythmus-Pitch (ARP)-Ebene als "time-separation-pitch" (TSP) nach McClellan (1) interpretieren.

Man betrachte den folgenden Fall, wobei TSP = t; wäre: perzeptiv dürfte hierbei keine auditive Differenz eintreten. Da jedoch eine Perceptonedifferenz eintritt, wird somit eine TSP-Interpretation gegenstandlos.

Frühere eigene Experimente (2)(3) (mittels "spontanen dichotischen Schwebungseffektes") ergaben als Resultat, daß die perzeptive Konjunktion von IB+IS als TSP-Fall nur in der relativen Refraktärzeit eintrat, während davor eine perzeptive Disjunktion von IB und IS erfolgte.

Abgesehen davon ist eine periodizistische Interpretation im Bereich der Aperiodizität wenig sinnvoll.


Wenn somit eine bistabile Verhaltensweise des Perzeptionsmechanismus in diesem Fall angenommen werden kann, so wird hierfür eine Präferenzentscheidung durch eine Polarisierungsfunktion des Perzeptionsmechanismus stattfinden müssen.

Diese Polarisierungsfunktion der Präferenzentscheidung des Perzeptionsmechanismus kann reaktionsfähig werden: a) durch akustisch dominierende Informations-Variationselemente und wäre abhängig von:


Um einen weiteren Einblick in die Verarbeitungsweise von IB oder IS der PRM durch den Perzeptionsmechanismus zu erhalten, wurden Hörteile mittels der Methode des "spontanen dichotischen Schwebungseffektes" (2)X wobei eine Synchronisationsmöglichkeit des Stimulus mit einem Sinuston bei komplexer Verarbeitung des Nervensystems eintreten kann) durchgeführt.
Über das kontextabhängige auditive Perzeptionsverhalten

Im Hörtestprogramm war der Frequenzbereich von 100Hz-1100Hz des Vergleichsinstimulums über 92 Frequenzen (mit durchschnittlicher Distanz von 10Hz) gleichmäßig aufgelöst sowie die Frequenzwerte im Programm algoritmus verteilte.

Aus der Gesamttmasse der Versuchsergebnisse wurden folgende systematische Elementevarianzreihen zusammengestellt:

Aus den Bildern wird ersichtlich, daß auf der akust. Ebene folgende Zeitsegmente perzeptiv informationsbestimmend sein können:

a) \( t_{IB} \)
   - partielle disjunktive
b) \( t_{IS} \)
   - " "
c) \( t_{IB+IS} \)
   - partielle konjunktive
   \( \{ \text{PRM-Verarbeitung} \} \)
d) \( t_{2IB+IS} \)
   - totale "
e) nicht identifizierbar: X

Die Testergebnisse mittels des "spontanen dichot. Schwebungseffektes" zeigen, daß der Perzeptionsmechanismus bei einer mehrelementigen akust. Information in diversen Modi ansprechbar sein kann, (d.i. mehr als ein Verarbeitungsmodus) d.h. in gewissen Fällen ist keine einzelne Entscheidung möglich durch mehrfache Synchronisationspunkte.

Hieraus ist auch 1) die aufgetretene Kontextabhängigkeit,
2) die TBTB-Inversionen, 3) die Reversibilität/Irreversibilität (+ oder -) identischer Elemente, die in den Intervallschritt-Testprogrammen auftreten, erklärbar.

Die Resultate ergeben, daß bei Aperiodizität eine perzeptive Verarbeitung akust. Zeit/Amplitudensegmente der PRM durch das periphere Gehörorgan möglich ist.

Derartiges ist jedoch nur bei einem Verarbeitungsmodus mit diphasischen Spike-Auslösungs punkten akustischer Schwingungskurven im Cortischen Organ denkbar. (7)

Die Volley-Theory (9) dürfte hierfür keine Beschreibungsmöglichkeiten bieten, da sie an die beschränkenden Grundbedingungen starrer monophasischer, statt variabler diphasischer Spike-Auslösungs punkte gebunden ist.
Über das kontextabhängige auditive Perzeptionsverhalten

Außerdem ist die Volley-Theory nur für den Periodicity-Pitch-Bereich konzipiert.

Daß die Volley-Theory im Bereich der Aperiodizität nicht mehr anwendbar ist, geht auch aus der Tatsache hervor, daß z.B. IS/Zeitsegmente < 1ms nicht mehr disjunktiv perzeptiert werden; wäre das Volley-Prinzip funktionsfähig, so dürften die Refraktärzeiten in diesem Fall keinen Einfluß ausüben.

Wenn \( t_{IS} \) bei \( t_{IH} \) (200µs) noch disjunktiv (2) perzeptiert wird, dann wird daraus geschlossen, daß \( t_{IH} \) (200µs) zeitlich noch mittels zweier Spike-Auslösungen durch die Cochlea-Transformation verarbeitet werden kann (\( t_{IS} \) wäre sonst nicht disjunktiv perzeptierbar); d.h. ein 'off'-Spike (P1) muß noch dem 'on'-Spike (P2) folgen können.

Dabei ist jedoch die \( t_{IS} \) (wenn \( t_{IS} > 1 \text{ms} \) bei \( t_{IS} = 200 \mu \text{s} \) audativ bestimmt.

Aus der Tatsache, daß hierbei die \( t_{IS} \) (wenn \( t_{IS} > 1 \text{ms} \) bei \( t_{IS} = 200 \mu \text{s} \) ist) audativ bestimmt ist, wird geschlossen, daß zwar eine lineare und adäquate 1:1 Übertragung durch die Transformation in der Cochlea erfolgen kann, B) daß jedoch andererseits bei der weiteren Verarbeitung im komplexen Nervensystem die Refraktärzeit eine selektive Funktion ausübt, so daß als auditives Endresultat in diesem Fall die \( t_{IS} \) bestimmt wird.

Somit wären A) und B) als zwei autonome Verarbeitungssysteme anzusehen.

Hieraus wird, entgegen der Volley-Theory, eine refraktärzeitabhängige akust. Signalverarbeitung im komplexen Nervensystem bei Aperiodizität konkludiert.

Aus den Resultaten ergibt sich noch folgendes:

1) Wenn die \( t_{IS} \) und \( t_{IS} \) informationstragend sind, dann ist der Begriff Frequenz auf der traditionellen Amplituden-Rhythmus-Pitch(ARP)-Ebene (vorallem im Bereich der Aperiodizität) für eine Beschreibung nicht mehr als ausreichend anzusehen.

2) Außerordentliche Ergebnisse geschlossen, daß unter der ARP-Ebene noch eine Perzeptionschicht einer Amplituden/Zeit-Quantum-Perzeption vorhanden sein muß.

(2) Nössel, V.: Über die Tonhöhenempfindung des Menschen bei nicht-period. Schallvorgängen (Forsch. u. Fortschr., 1965, 9, Pg. 263).