Foreword

This volume contains invited papers given at the Eighth International Congress on Acoustics held at Imperial College, London in July 1974. The contributed papers have appeared in two separate volumes containing single page contributions from over seven hundred delegates.

The main theme of this Eighth Congress was Environmental Acoustics. The opening address and the afternoon invited lectures were related to various aspects of this topic.
Invited Lectures

presented at

Eighth International Congress on Acoustics,

London, 1974

Under the auspices of the International Commission on Acoustics (ICA) of the International Union of Pure and Applied Physics (IUPAP) with the support of UNESCO.

Sponsored by the Institute of Acoustics, The Institute of Physics and the Noise Advisory Council of the U.K.

© 1974 The Contributors

Printed in Great Britain by Goldcrest Press, Trowbridge, Wiltshire.
LIST OF INVITED LECTURES

J.E. Ffowcs-Williams (UK) Physics of Sound Creation 1
L. Cremer (BRD) Building Acoustics 11
Z. Maekawa (Japan) Environmental Sound Propagation 29
M.R. Schroeder (BRD) Hearing Research – Nonlinear Mechanics and Neural Transduction 47
I.P. Goliamina (USSR) Ultrasonic Surgery 63
L. Filipczyński (Poland) Ultrasonic Medical Diagnostic Methods 71
J.K. Johnsen (Norway) Models of the Underwater Acoustic Channel 87
J.H. Collins (UK) Acoustic Surface Waves and Systems Applications 107
R.G. Busnel (France) L'Intelligibilite et Certaines Donnees Psychophysilogiques Sous-estimes 123

The Lectures listed below were not received in time for presentation in this volume. It is hoped, however, to include them in a supplementary issue which will be distributed to all delegates after the Congress.

Opening Address by R.H. Bolt on 'Environmental Acoustics'

Invited Lectures:

H.E. von Gierke (USA) Noise - How much is too much? The Measurement and Rating of Environmental Noise with respect to its Effect on Man

F. Ingerslev (Denmark) Strategies for Controlling and Planning Against Transportation Noise
The dictionary gives the meaning of 'source' in physics as 'a point or centre from which a fluid or current flows'. Evidently the term dates from 1878, introduced to the subject by Helmholtz. As long as the source is concentrated on a point, the only point at which

$\mathbf{Q} \neq 0$, such a definition is perfectly adequate in acoustics even though no fluid or current need necessarily originate there. The source point is the centre at which mass may be created, from which waves of potential or pressure or dilatation or energy (or sound) emerge, where work is done on the fluid, which is the common focus of all wave crests and is the only point at which physical laws governing the equilibrium motions of a continuum breakdown. On any measure the source is the cause, origin, sponsor and birthplace of the radiated sound. It is the only peculiar spot; it is a singular point.

That source can generate fields of great character, for it can be subdivided into an extensive class of source types. Though confined to a point, it may none-the-less contain distinct elements between which an interesting interference is possible. Such a source is termed 'multipole', and the degree to which self-cancellation occurs between the fields of separate elements determines the 'multipole order'. The higher the order the more sensitive is the field to changes in the time scale on which the source varies, and the more directional structure can be imparted to the radiation field. But with this multipole generalisation is associated the first distinct ambiguity, for there is absolutely no difference between the field radiated by the simple point source of strength density

$\frac{2}{2} \mathbf{Q}$ and the isotropic point quadrupole of strength density $C_{ij} \mathbf{Q}$. This is also true when $\mathbf{Q}$ represents a multipole distribution, so the subdivision into multipoles, even point multipoles contains a significant freedom of choice, and must be accomplished in the manner judged most convenient to the job at hand. There is no unique prescription that labels the source type. In my view this does not detract from the value of such labels; on the contrary, it is usually possible to arrange the source definition in such a way that the source type conveys at once the most interesting features of the field. But a failure to stress the essential non-uniqueness of practical sources often leads students to untold confusion. That is a theme I will revert to during this lecture.

The possibility of confusion is easily appreciated when considering the degree of variety that can be contained in the fields of a point multipole system. The variety is enormous for it can represent the distant sound field of any bounded source system. Furthermore those equivalent multipoles can be positioned at any selected point within that source. Of course the equivalent multipole strength depends on the point at which the multipole is assumed to exist.
EIGHTH INTERNATIONAL CONGRESS
ON ACOUSTICS, LONDON 1974

SOURCES OF SOUND

But there is a far more interesting indeterminacy associated with the
fact that the sound of the real source is identical with that of the
infinitely many possible multipole series that represent it. That
means that it is impossible to determine from a distance the precise
nature of the source field! It means also that it is impossible to
determine from the source free field alone the location of the sound
sources.

This point can be emphasised by reminding you that Kirchhoff's
theorem provides a means of generating with a surface distribution of
monopole and dipole sources an identical exterior field to that
radiated by any source within the bounding surface. From the sound in
the auditorium one could not distinguish the difference between the
Halle orchestra playing on the stage and an equivalent source distribu-
tion on a surface separating the orchestra from the audience! This
principle has already found its place in optics where the hologram
regenerates the light waves of the real event. No doubt acoustic
applications will follow - indeed some are already accomplished. We
can look forward to the day when acoustic reproduction techniques are
perfect in the sense that the reconstituted field in the vicinity of
the listener will be in every way identical with the original.

But this duality of real and reconstituted fields means that
source location procedures based on acoustic telescopes are incapable
of pinpointing with any degree of uniqueness the actual 'source'!
Rayleigh and Helmholtz were concerned with the elegant structure of
singularities capable of generating what invariably seemed to be
pleasant sounds. A possible source was as good as the next for this
purpose - in fact the simpler the better! Times have changed. Our
modern concern is to locate the 'cause' of chaotic sound, noise, with
a view to eliminating it. It is typical of nature's cunning that
though that noise is all too evident its source is an ingeniously
guarded secret! So much so that to enquire into its origin can often
be a naïve question to which there are infinitely many equally true
and equally irrelevant answers!

Once this point is appreciated the study of sources becomes very
much more interesting, for one must exercise an art and choose the
most satisfactory source according to some criterion of elegance.
The study is particularly interesting when scattering takes place for
that process converts one type of dynamical motion to another, and
then the origin of sound, the region from which it seems to radiate,
is often far removed from its cause. The term 'source' is then mean-
ingless pending a definite prescription by which it can be identified.

For example, consider a large elastic plate driven at some near-
central point by a randomly excited vibrator. If the plate is large
enough that vibratory waves take longer to travel to the plate
boundary and return to the driving point with the news of the plate's
scale than the time scale of the excitation, then the energy flows
from the driver at a rate independent of the plate geometry. In
general the energy radiated as sound waves centred on the driving point is negligible compared to that imparted to plate vibration, but that also ultimately ends up as sound. It does so because the vibratory waves are reflected at the plate boundaries, and lose energy at each reflection by a sound producing edge scattering process. In this case then, though all the energy flows from the vibrator to the sound, the process is indirect, the sound being generated at the remote edges of the plate, edges whose presence eventually brings about the total conversion of the driving power into sound but which like the sound have no influence on the magnitude of that power.

Bubbles play a similar scattering role in underwater sound. A bubble being much more compliant than water is easily driven into vigorous motion by an unsteady pressure field, and that motion often induces a powerful sound wave, centred on the bubble. But the bubble cannot be the means by which that sound is energised because the bubble undergoes a purely passive driven response, yet the field centred on the bubble is often very much more powerful than that centred on the source of energy. To be specific suppose a rod of radius \( r_0 \) is in transverse vibration which radiates sound of wavelength \( \lambda \). If a bubble of radius \( a \) is now positioned close to the rod, the field is supplemented by a wave centred on the bubble, its strength being determined by the bubble response to the local unsteady field. This scattered sound is more energetic than the direct sound by a factor \( a^2 \lambda^2 / r_0^4 \), which can easily amount to several tens of decibels. Here again the obvious 'source' of sound is the vigorously responding bubble - but the source of energy lies elsewhere, in fact in the work done on the fluid by the vibrating rod, which has to work harder when the bubble's presence increases the resistive component of the force between the rod and fluid.

Neither need scattering involve vibration of the secondary surfaces. In 1868 Stokes observed that the sound became much stronger when the blade of a large knife obstructed the motion local to the plane of a tuning fork. Some fifty years later the essential analysis of that problem was performed by Sommerfeld. It is a relatively straightforward task to show through that analysis that the presence of the motionless screen increases the sound energy radiated from a vibrating cylinder by a factor \( \lambda / r_0 \), \( \lambda \) again being the wavelength of the sound and \( r_0 \) the separation between the cylinder and the edge of that screen. Again this factor can be very large so that once more the dominant part of the sound field is centred on the rigid surface from which no energy could possibly be extracted. The clear geometric origin of the field is different from its energising centre. Either could reasonably be defined as the source, but neither claims unique recognition as that source.

All these examples concern motions clearly recognisable as sound,
governed by the equation $\Box^2 \phi = 0$ in the source free volume through which it propagates. The strength of the field is formally found by matching conditions at a physical boundary, the pressure and/or normal component of particle velocity being set equal to their counterparts on that physical boundary. The source of sound need not necessarily be uncovered in this process which merely determines the acoustic motion consistent with boundary conditions that are themselves determined by the source but need not be that source.

The situation is more straightforward when no boundaries exist. Then if $\Box^2 \phi = 0$, $\phi = 0$ is the only solution satisfying the radiation condition, i.e. $\Box^2 \phi = 0$ guarantees that no sound, and therefore no sources exist. So it is convenient to define the degree to which $\Box^2 \phi \neq 0$ as the source $Q$.

$$\Box^2 \phi = Q.$$  

(1)

$Q$ is termed the simple source strength density. But this definition is actually a matter of choice. Outside the source, $Q = 0$ so that $\Box^2 Q$ is also zero and we could solve for $\phi + Q$ instead of $\phi$.

$$\Box^2 (\phi + Q) = Q + \Box^2 Q$$

might then be considered as the governing equation, $Q + \Box^2 Q$ being another possible source distribution that generates the same exterior field $\phi$. It is impossible to tell from the source free part of the field which of $Q$ or $Q + \Box^2 Q$ or indeed any one of a large variety of different sources actually generated the sound. Again one must refrain from enquiring too deeply into a sound's origin.

The character in the field of a complicated source distribution results from destructive interference between various source elements. Sometimes this interference is complete, there being just as many negative sources, or sinks, as there are positive ones. The source distribution is then said to be multipole. Formally, if the simple source strength is zero, then

$$\int_{\text{volume}} Q \text{d}v = 0$$

and $Q = \nabla \cdot F$ or $\frac{3F_1}{\partial x}$, $F$ being termed the dipole strength. The dipole elements may also be arranged in a self-cancelling array; the source distribution is then quadrupole, and so on.

The source order is immediately recognisable by the order of the divergence operator in the source distribution, an $n$th order source being
sources of sound

$$\square^2 \phi = \mathcal{Q} = \frac{\delta^n \psi_{ij...mn}}{\delta x_i \delta x_j ... \delta x_m \delta x_n}.$$  

It is not always important that the source structure be recognised. It is really immaterial if every nearby source is opposed exactly by distant elements mimicking their motion in antiphase, though of course at points equidistant from the two there will be a substantial interference, the field being somewhere increased and elsewhere weakened. But if the opposing elements are close together, their sound being heard at every point in antiphase, then the source system is an extremely ineffective generator of sound. The phase changes arising from path length differences from the various source elements are much smaller than a complete cycle when the separation between source elements is much smaller than a wavelength. Then we say that the source distribution is compact, and it is extremely helpful that the source structure, or multipole order, be recognised. The interference between such elements is entirely destructive and the multipole combination is relatively quiet, the more so the higher the multipole order. The possibility of destroying the delicate destructive balance between source elements is quite obvious, and that is precisely the mechanism by which the passive blade of a large knife placed near the tine of tuning fork results in powerful amplification of the sound generated by the fork which constitutes a quadrupole array.

Since it is only the residue of imperfect cancellation that escapes as sound from a multipole it is obvious that high order multipoles will respond vigorously to factors affecting that cancellation, the main one being the compactness ratio that measures the source dimension on a wavelength scale. Any reduction in the wavelength due to an increased frequency, or to a lowering of the wave speed, or due to the Doppler contraction if the source is moving, results therefore in a rapidly increasing acoustic output. A lower bound on the degree of such effects is established once the source type is known, and this is very important in practice because many source fields are so complicated that one cannot expect to know more about them than global diagnostic features. This seems to be especially true in the more powerful acoustic sources; turbulence for example in the exhaust of civil aircraft is right now generating some tens of megawatts of acoustic power. Now nothing significant is known about turbulence, nor in my view is it ever likely to be known, so that the most effective description of its sound-producing properties is that which makes most use of what scanty information we have such as the source order, its scale and characteristic frequency.

The precise, though optional, definition of $\square^2 \phi$ as the source can be formally extended to a bounded source-free domain by choosing
a Heaviside operator, \( H \), equal to unity in the wave field and zero elsewhere and simply multiplying it with the wave equation \( \Box^2 \phi = 0 \) and then taking the Heaviside function inside the wave operator, viz.

\[
H \Box^2 \phi = 0
\]

\[
\Box^2 H\phi = -c^2 \frac{\partial}{\partial x_i} \left\{ \frac{\partial H}{\partial x_i} \phi \right\} - c^2 \frac{\partial}{\partial x_i} \frac{\partial H}{\partial x_i} \frac{\partial \phi}{\partial x_i} = Q, \quad \text{say.}
\]

Now this equation is valid over all space, and since \( H\phi \) equals \( \phi \) in the wave field, the source of \( H\phi \) might well be regarded as the source of \( \phi \). This source consists of two parts, both concentrated on the surface that bounds the wave field, the only region where \( H \) is not constant and its derivative non-zero. The first part is a surface dipole contribution, the dipole strength being determined by the boundary value of \( \phi \). The second term is a surface monopole the monopole strength being determined by the gradient of \( \phi \). This classification is of course the one familiar from Kirchhoff's theorem.

This approach is valid even when \( \Box^2 \phi \) is defined everywhere. We could still select a function \( H \) to be unity in a chosen source-free part of that space and still represent the sources of \( H\phi \) as the monopole and dipole boundary distributions. Then we are creating an analogy. We recognise that though \( \Box^2 \phi \) may be perfectly well defined everywhere, and we could call its value \( Q \), the source distribution, the effect of that source is precisely the same in the wave field as is the effect of the equivalent but quite different surface distribution \( \tilde{Q} \). It might be much easier to determine the field due to \( \tilde{Q} \) than that due to \( Q \), in which case the analogy would be helpful. But sometimes \( \tilde{Q} \) is no simpler to determine than \( Q \) itself, and then the analogy is useless.

Actually we are creating an analogy by writing \( \Box^2 \phi = Q \), when in so doing we give the impression that \( Q \) will be specified as a prerequisite to the determination of the field. \( \phi \) may not be part of a wave motion at all. It may be governed by some quite different equation

\[
P\phi = S, \quad \text{say with } S \text{ specified independently of } \phi.
\]

We could choose to write this as a driven wave equation in an acoustic analogy, though a hopelessly sterile one, as follows:

\[
\Box^2 \phi = Q, \quad \text{with } Q = \Box^2 \phi + P\phi - S.
\]

The field \( \phi \) that we have chosen not to be a wave field is indeed the same as the \( \phi \) that is driven by the source distribution \( Q \) in a perfect wave field. The two problems are analogous - but it will rarely be the case that \( Q \) can be specified independently of \( \phi \), the essential requirement for the analogy to prove useful.

The definition of the source can from this example be seen to be not only non-unique, but also irrelevant unless there is some
additional expectation that the source can be specified before the solution is known! When this is the case, the acoustic analogy can be both elegant and useful and one such analogy finds widespread application in the acoustics of moving fluids.

Lighthill chose the density, $\rho$, as the central field quantity in his theory of aerodynamic sound because that choice naturally highlights the essential role of fluid compressibility in distinguishing between unsteady pressures confined to the vicinity of a flow and the waves that escape as sound. $\Box^2 \rho$ is zero in the sound field and $\Box^2 \rho$ is conveniently defined therefore as the source, expressible through the exact equations of fluid motion as a quadrupole distribution $\frac{\partial^2 \mathbf{T}_{ij}}{\partial x_i \partial x_j}$. Given a description of $\mathbf{T}_{ij}$, the field $\rho$ is determined; the aerodynamic sound equation,

$$\Box^2 \rho = \frac{\partial^2 \mathbf{T}_{ij}}{\partial x_i \partial x_j},$$

looks deceptively simple. But the expectation of knowing much about $\mathbf{T}_{ij}$ is extremely small, for the equation is actually nothing more than a disguised version of the Navier-Stokes equation the known exact solutions to which can be counted on the fingers of one hand, and a few appear wavelike. It is the generality of Lighthill's theory that impresses most: the remarkable fact that the density perturbation in any fluid motion, however complicated, is exactly the same as that generated in the analogous problem where a quadrupole distribution $\mathbf{T}_{ij}$ radiates into a perfect acoustic medium at rest. But the generality is also its weakness for most practical problems are too complicated for detailed modelling and the analogy is powerless pending a further prescription for its use and a guide to the right procedure for approximation. Lighthill provided that too in a scheme that exploits the fact that in many important situations the scale of the source flow is much smaller than the acoustic wavelength it generates. The known properties of compact quadrupole fields then led him directly to one of the most celebrated equations of modern acoustics, the eighth power law relating sound energy to the source velocity scale. In unconstrained compact fluid motion sound can be regarded as the by-product of those non-linear and inhomogeneous stresses that remain unbalanced by the linear accelerations and compressions of a purely acoustic motion.

The acoustic effects of compact foreign bodies readily follow. When the quadrupole is interrupted by a boundary, which may isolate from the fluid a part of the quadrupole and leave behind a much more efficient dipole or monopole residue, the sound level is vastly increased. The dipole strength is the force imbalance between the body and the fluid due to the abrupt ending of the stress field, and
the monopole is the mass displaced by volumetric changes. It follows immediately that because of superior radiation efficiency in the compact limit, only the most elementary source need be considered. The dipole structure of compact aerodynamic surfaces with the attendant sixth power of velocity law is well documented as is the monopole air bubble in an unsteady water flow whose resonant vibration generates a loud and musical note.

In all these compact cases the stress tensor $T_{ij}$ can be specified unambiguously in advance of the sound and the acoustic analogy provides an exact procedure for defining a possible source field and for locating the source region with an accuracy much better than the acoustic wavelength.

But the loudest sources are the biggest sources which generally fall outside the scope of the compact form of the acoustic analogy. For this the source motion must have a velocity scale at least comparable with the acoustic velocity and $T_{ij}$ cannot be specified independently of the sound field it prescribes. Some other scheme is then needed to define the approximate modelling of the source field. One important case rests on the limit that at sufficiently low amplitude the acoustic elements in $T_{ij}$ change only slowly and can therefore be approximated by their linear form. This gives rise to the continuum scattering equations where the interaction of sound with unsteady flow, or the interaction of sound with sound, generates distinct disembodied sources of secondary waves. The analogy provides in these cases an efficient means of describing the secondary fields and can cope equally easily with the propagation of sound through a slowly evolving inhomogeneous medium and with the parametric array where the non-linear terms in a high frequency beam form a useful source of secondary waves. The analogy describes these sound fields in complete generality and any non-uniqueness of the source field is of little or no concern.

Sources near a non-compact scattering surface induce an extensive distribution of linear surface terms acting as non-compact monopoles and dipoles. These usually account for a far greater radiation than does the source itself, though they need provide none of the field's energy. The analogy whereby $D_0^{2p}$ is termed the source emerges then as pure formalism and in fact can be dealt with formally by solving an elementary case to generate the exact Green's function for the problem. The high radiation efficiency of quadrupoles, or flow, near a sharp edged plate was worked out in this way and the exact result that the surface is neither a monopole nor a dipole but something in between was a surprising but now well verified outcome of that analysis. The field scattered by a sharp edge can usefully be thought of as originating in a 'one and a half pole'. The sound energy generated by a lifting wing of span much greater than the wavelength can be similarly treated, again the result implying a fractional order source, and again the problem must be worked out quite formally. The interesting
modern source studies do seem more formal and the results less easily
generalised to a wider class of problems.

The interaction of sound with flow that is extensive on the wave-
length scale is a good example wherein is found a whole spectrum of
different behaviours. Ray theory can be used at high frequency to
show that sound propagates through slowly varying (on a wavelength
scale) flow with its energy conserved, the details of its propagation
path being determined by the flow. That case conforms to the acoustic
analogy also, but it is clear that the equivalent quadrupole distribu-
tion accounting for high frequency refraction is entirely determined
by the sound and cannot possibly be known in advance of it; the
analogy is then purely formalistic. If the flow is thin in one direc-
tion and extensive in the others, energy can be exchanged between the
sound and the flow, and the acoustic analogy's equivalent sources has
more meaning. But in fact those cases are again ones in which the
sources cannot be known in advance, so that the analogy provides no
basis for evaluating the field. Also, sometimes the disturbance may
originate inside the flow and disturb it into its natural unsteady
motions which may themselves be highly coupled to sound. An unstable
jet flow is a case in point; there small disturbances at the nozzle
exit can excite the jet into unsteady growing oscillations that
generate far more sound than the initial triggering disturbance.

Again these cases have, to date, all been worked out on the basis of
evaluating the linear motion and its attendant field directly, the
issue of where the sources exist being by-passed in the process. In
these problems the sound field is established as the asymptotic form
of the disturbance which, in the flow, may not be in the least bit
like a sound wave but is coupled to it, the sound evolving slowly out
of the basically hydrodynamic motion. In this case there seems little
point in seeking a prescription to define a source in some unique
or preferred way. In such a mixed situation it must be possible to
choose an even greater variety of satisfactory equivalent source
fields than in the purely acoustic case - where as we have seen the
choice is in any event infinite! The basic difficulty is that the
source and sound fields fail to have an independent existence whenever
the flow is grossly non-compact.

Even the effect of uniform flow on sound can sometimes prove
surprising. A compact vibrating body not only radiates sound whose
wavelength is influenced by flow through the Doppler effect, it also
radiates sound more effectively. Even in the absence of friction,
there is a steady drag force between the flow and the body, and the
work done by the flow in overcoming this drag, passes directly into
the acoustic field whose energy increases in consequence.

Compliant surfaces can similarly cause the conversion of mean
flow energy into sound by a relatively indirect coupling where the
source process is quite elusively hidden. Uniform flow over compliant
boundaries, of the type used to absorb sound, can be unstable and the
surface may flutter in a slowly travelling but growing surface wave. That wave may be scattered into sound very effectively by surface inhomogeneities, or by an edge. The noise of the flapping flag is a good example. Problems such as these are only beginning to be posed in a tractable manner and though they obviously generate sound effectively enough the detailed modelling of the source process is yet to be accomplished.

Finally I would like to make some observations regarding the 'acoustic telescope' methods of source location that are becoming increasingly common in noise control work and which have of course for a long time been a familiar item of the underwater sonar principle. I have been concentrating so far on the technical task of establishing the mechanics of a source process. In doing so I have emphasised the essential non-uniqueness of source definitions, and how, though the acoustic analogy can specify the field completely, it is not possible to go backwards and from the sound field determine the source process unambiguously. Despite this there is no doubt in my mind that source locating principles will find widespread and useful application. By continued use an operator will learn the idiosyncrasies of the scheme and what confidence level can be placed in its results. This is true of sonar, where the possibility of decoy signals is obvious but can be countered effectively through operating experience. Also there is no important risk of everyday visual images being actually generated by a hologram - though the possibility is there in principle. What the likelihood is of distinguishing between important and trivial acoustic sources in noise control work is still to be determined and the problem must be greater in situations where flow is inextricably intermingled with sound, as it must be in non-compact situations. But eventually such techniques will I'm sure yield sufficiently unambiguous indication of source activity that they will provide worthwhile clues as to the area most likely to contain the most source suppression potential. That will be a most worthwhile contribution.
EIGHTH INTERNATIONAL CONGRESS ON ACOUSTICS, LONDON 1974

STAND DES SCHALLSCHUTZES IM WOHNUNGSBAU
(NOISE ABATEMENT IN DWELLING HOUSES)

L. Cremer  Institut für Technische Akustik,
Technische Universität Berlin (BRD)

Mögen auch – glücklicherweise – die im Wohnungsbau auf-
tretenden Schallpegel weit unter denen liegen, denen der
Mensch in Verkehrsmitteln oder in Fabriken tagsüber aus-
gesetzt ist, er benötigt zwischendurch, insbesondere
nachts, Oasen der Ruhe in seinem Heim. Auch und gerade
für den Schallschutz gilt das englische Sprichwort
"my home is my castle" – a castle protecting against
noise.

Ähnlich wie Henry quatre jedem Bauern seines Landes
sein sonntägliches Huhn im Topf versprach, sichern die
Deutschen Baunormen jedem Bürger einen Pegel unter
30 dB(A) in seiner Wohnung zu. Das bedeutet, daß er durch
ausreichende Decken und innere Wände geschützt sein muß
gegen den Luftschall, den Nachbarn über oder neben ihm
erzeugen, aber auch durch ausreichende Außenwände gegen
Luftschall von außen, ferner gegen unmittelbare Körperschallanregung im Haus, zu denen namentlich der Trittschall gehört, und schließlich gegen die Geräusche, die
alle Wasserleitungen erzeugen.

1. Luftschallschutz gegen den Nachbar

Über die physikalischen Vorgänge, die bei der Luftschall-
übertragung durch Decken und Wände zu beachten sind, hat
im Rahmen der ICA zuletzt K. Gösele auf dem Stuttgarter
Kongreß im Jahre 1959 referiert /1/, über die Bewertung
der Ergebnisse und Zulassungsbedingungen O. Brandt 1962
in Kopenhagen /2/.

Nicht nur, weil seither einige neue Erkenntnisse hinge-
zechenden Tagungsteilnehmern eine Übersicht zu geben, muß ich die
wesentlichsten, hierbei geltenden Richtlinien, bzw. die
getroffenen Vereinbarungen kurz wiederholen.

a) Das Meßverfahren

Das durch die 1960 verabschiedete ISO-Recommendation 140
international vereinbarte Meßverfahren schließt sich
eng an die im Wohnungsbau auftretende Situation an. Zwei
zimmergroße Räume, die über – oder nebeneinander-liegen,
werden durch das Meßobjekt akustisch getrennt. In dem
einen Raum 1 wird mit Hilfe eines Lautsprechers Luft-
scall erzeugt, und zwar terzbreites Rauschen, wobei die
Termmittenfrequenz bei der Darstellung der Ergebnisse
die Abszisse darstellt. Man beschränkt in der Bauakustik
den Frequenzbereich auf 100 bis 3200 Hz. Die untere
Grenze ist zunächst durch die Größe der Meßräume und bei
dickeren Trennobjekten auch durch deren Größe gegeben. Schon bei 100 Hz sind sie eigentlich zu klein, um noch mit einer statistischen Schallverteilung auf alle Eigen- schwingungen - anders ausgedrückt auf alle Ausbreitungsrichtungen - rechnen zu können, wie es die Theorie verlangt, die der Auswertung zugrunde liegt. Gilt aber diese Theorie nicht mehr, so erhält man wie schon HECKL und SEIFFERT /3/ am eindimensionalen Fall und erst kürzlich MULHOLAND und LYON /4/ am räumlichen Problem ausführlich darlegten, Ergebnisse, die so von den jeweiligen Abmessungen abhängen, daß ihnen keine allgemeine Aussage zukommt. Glücklicherweise läßt nach tiefen Frequenzen unsere Ohrenempfindlichkeit nach, so daß diese Beschränkung auch vom subjektiven Standpunkt leidlich gerecht- fertigt ist.

Bei hohen Frequenzen ergibt sich die Beschränkung einmal daraus, daß die erzeugte Schallenergie je Terz geringer wird, mehr aber noch daraus, daß die Dämmung hoher Frequenzen keine Schwierigkeiten bietet. Das liegt an dem allgemeinen Ähnlichkeitsgesetz, daß jede Objektdicke im Verhältnis zur Wellenlänge zu bewerten ist, daß sie umso wirkungsvoller wird, je höher die Frequenz ist.

Die spezielle Wahl von 3200 Hz geht darauf zurück, daß man ursprünglich die Frequenz 100..200..400..3200 Hz als Hauptreihe der Oktavschritte ansah. Nachdem die Hauptreihen auf aber 125, 250,... 2000, 4000 gelegt wurde - um die Lautstärke-Vergleichsfrequenz 1000 Hz einzubeziehen -, wäre es zweckmäßig, die Terzen bei 2500 Hz abzubrechen. Man erhält dann den gleichen insgesamt überdeckten Frequenzbereich, gleichgültig ob man in Terzen von 100 bis 2500 Hz, oder in Oktaven von 125 bis 2000 Hz mißt.

Der Schalldruckpegel des Terzrauschens wird nun im Senderaum als $L_1$ und in dem dann durch das Messobjekt getrennten Empfangsraum als $L_2$ gemessen. Man erhält als primäres Meßergebnis eine Pegeldifferenz

$$ D = L_1 - L_2 $$

(1)

Die gemessene Pegeldifferenz ist aber nicht nur vom Messobjekt abhängig, sondern auch davon, welche schallabsorbierenden Elemente im Empfangsraum sind, was durch die sogenannte äquivalente Absorptionsfläche $A$ erfaßt wird. Würde man die im Senderraum erzeugte Schalleistung als gegebene Größe ansehen, so würden die äquivalenten Absorptionsflächen beider Räume $A_1$ und $A_2$ von Bedeutung sein. Da man aber den Pegel $L_1$ als gegeben ansieht, z.B. annimmt, daß jeder sein Fernsehgerät auf die sogenannte, nie genau definierte Zimmerlautstärke einstellt, korrigiert man die Pegeldifferenz nur in Bezug auf $A_2$ (1.F. = $A$), indem man ihr ein Korrekturglied anfügt und
so eine Normpegeldifferenz definiert:

\[ D_N = L_1 - L_2 + (10 \log \frac{1}{A}) \text{ dB} \quad (2) \]

Gewisse Meinungsverschiedenheiten bestehen hinsichtlich der zweckmäßigen Wahl von \( A_0 \). Es sei auf diese nicht näher eingegangen, zumal es bei Normierungen mehr darauf ankommt, daß man sich einigt, als worauf. (Nur sollte man Übereinkünfte nicht bei jeder Neuausgabe ändern, weil es den Anwender in seinem Vertrauen auf die Fachexperten unsicher macht.)

Wichtiger ist, daß \( D_N \) keine Kenngröße für die spezifischen Merkmale eines Objektes wie Zahl der Schalen, ihre Dicke usw. darstellt, sondern daß \( D_N \) umso kleiner wird, je größer die Trennfläche \( S \) ist. Eine solche Kenngröße würde im Idealfall der Transmissionsgrad darstellen, d.h. das Verhältnis von durchgelassenen zu auflaufender Leistung

\[ \mathcal{L} = \frac{P_2}{P_1} \quad (3) \]

bzw. das aus ihm sich ergebende Schalldämmmaß

\[ R = (10 \log_{10} \frac{1}{S}) \text{ dB} \quad (4) \]

und dieses erhält man - jedenfalls unter idealen Bedingungen -, indem man der Pegeldifferenz eine Korrektur anfügt, die \( A \) und \( S \) enthält:

\[ R = L_1 - L_2 + (10 \log_{10} \frac{S}{A}) \text{ dB} \quad (5) \]

Hier herrscht Einigkeit, daß dies die Größe ist, die man im Labor - jedenfalls bei der systematischen Entwicklung neuer Wand- und Deckentypen - zu bestimmen sich bemüht.

Dagegen gehen die Meinungen auseinander, ob man auch bei Messungen am Bau mehr an \( R \) oder an \( D_N \) interessiert ist. Wenn sich beide Größen nur durch so leicht bestimmte Daten wie \( S \) und evtl. \( V \) (Volumen des Empfangsraumes) unterscheiden, könnte jeder das Ergebnis "nach seiner Fasson" angeben. Leider lehrt die Erfahrung bei
guten Trennobjekten, daß der Schallübertritt zwischen zwei Räumen nicht nur infolge direkter Anregung und direkter Abstrahlung des Trennobjektes erfolgt (in Fig. 1, Weg Dd), sondern auch infolge der Anregung eines flankierenden Bauteils im Senderraum und entsprechender Abstrahlung im Empfangerraum (in Fig. 1 Weg Pf), sowie über Eck, wobei die Anregung über das flankierende Element, die Abstrahlung von dem trennenden erfolgen kann (Weg Pd) oder umgekehrt (Weg Df). Die in den Empfangsraum ein- dringende Leistung $P_2$ setzt sich somit zusammen aus:

$$P_2 = P_{Dd} + P_{Df} + P_{Pd} + P_{Pf}$$

(6)

worauf in der Neufassung von R 140 hingewiesen ist. (In Erkenntnis der Tatsache, daß die Definition von R nur mit dem Weg Pd rechnet, gibt man R dort, wo man mit Flanken- oder auch sonstigen Nebenwegen rechnen muß, also insbesondere am Bau einen Beistrich; R').

Jedenfalls sollte man sich davon überzeugen, ob und wie sehr solche Flankenübertragungen beteiligt sind. Die einfachste Kontrolle bietet die unmittelbare Abtastung der Schnelle der den Empfangsraum begrenzenden Wände, der Decke und des Bodens. Haben sie alle eine gewisse Dicke und damit Steife, kann man daraus unmittelbar auf die Anteile ($P_{Dd} + P_{Pd}$) einerseits und ($P_{Df} + P_{Pf}$) andererseits schließen. Sind einzelne Elemente sehr biege- weich, so ist ihre Abstrahlung schlechter als die der biegesteifen Schalen. Man überschätzt daher ihren Anteil, kann aber sagen, daß er sicher kleiner ist, als es dem Vergleich der Schnellen entspricht.

Der andere Weg, der Flankenwege schrittweise zu unterbinden gestattet, besteht gerade im Vorsetzen solcher biegeweicher Vorsatzschalen (Fig. 1 unten).

Dies ist zudem eine für Prüfstände empfehlenswerte Aufbauweise, wenn man dort möglichst die Wand oder Decke "an sich" prüfen will.

Man hat ursprünglich einmal geglaubt, daß man statt dessen auch das Trennobjekt vor den Flankenwänden durch so genannte schallweiche Zwischenlagen isolieren kann, stellte aber fest, daß dabei das ermittelte Schalldämm- maß nicht größer, sondern kleiner wurde. Der Grund liegt darin, daß die in der Trennwand vom Senderraum-Luftschalleld erzeugten Biegewellen in diesem Fall fast vollständig am Rande als freie Biegewellen reflektiert werden und von neuem Schall abstrahlen, wenn deren Wellenlänge größer ist als die Wellenlänge der Luft, was jeweils erst oberhalb einer kritischen Frequenz für eintritt. Im Gegen satz dazu senkt die Abgabe der Energie am Rande den Körperschallpegel im Innern der Wand. Das gleiche läßt sich auch durch innere Verluste erreichen.
Die von GÖSELE erhaltenen Meßwerte stimmen sehr gut mit einer Formel überein, die M. HECKL zunächst am zwei-dimensionalen Modell ableitete, aber in einer späteren Arbeit /5/ auf die allgemeinere Form brachte

\[ R = (20 \log_{10} \frac{\omega m}{2Z_0} - 10 \log_{10} (0,7 f \tau T_B)) \text{ dB} \]  \( (7) \)

Darin stellt der erste Summand, in welchem \( \omega \) die Winkel- frequenz, \( m \) den Massenbelag und \( Z_0 \) den Kennwiderstand der Luft bedeutet, das sogenannte Masse-Gesetz für senkrechten Einfall dar, der zweite ergibt eine Korrektur, in die die innere Dämpfung und die Randabsorption über die leicht meßbare Nachhallzeit \( T_B \) der Biegewellen eingehen. \( f \tau \) bedeutet die erwähnte "kritische Frequenz", bei der Biegewellen- und Luftwellen-Länge gleich groß werden.

Die Neufassung von \( R \leq 140 \) empfiehlt daher, die Rand- einspannung soweit als möglich den später im Bau zu erwartenden Verhältnissen anzupassen. Das würde zugleich bedeuten, daß Prüfdecken und Wände die ganze Trennfläche einnehmen und nicht nur eine nischenartige Aussparung in einer dickeren Trennkonstruktion. Im letzten Falle gewinnt man zwar größere Meßräume und damit bessere Vor- aussetzungen für das statistisch behandelte Luftschall- feld. Die Zahl der auf einen Terzbereich \( \Delta f/f_m = 0,23 \) in einem Raum vom Volumen \( V \) einer Summe der Begrenzungsflächen \( S \) und einer Summe der Kantenlänge \( l \) entfallenden Eigenfrequenzen

\[ \Delta N_{\text{terz}} = 2,9 \frac{V}{\lambda_m^2} + 0,35 \frac{S}{\lambda_m^3} + 0,5 \frac{l}{\lambda_m} \]  \( (8) \)

beträgt beispielsweise bei einem Raum mit den Abmessungen \( 5 \times 4 \times 2,5 = 50 \text{ m}^3 \) im Terzbereich um 100 Hz nur 8. Eine Raumvergrößerung auf die Abmessungen \( 6,5 \times 5,5 \times 2,8 = 100 \text{ m}^2 \) würde diese Zahl auf wenigstens 14 erhöhen, was freilich auch noch hinter der oft geforderten Zahl 20 zurückbleibt. Aber es nützt dies wenig, wenn nicht gleichzeitig auch die Zahl der Eigenfrequenzen des Meßobjektes durch größere Fläche desselben erhöht wird. Hier gilt aber für eine am Rande aufgestützte Platte

\[ \Delta n_{\text{terz}} = 0,72 \frac{S}{\lambda_B^2} \]  \( (9) \)

d.h. bei den geforderten 10 m\(^2\) (wegen der Dispersion der Biegewellen):

\[ \Delta n_{\text{terz}} = \frac{4 f_m}{c_L h} \]  \( (10) \)
Die hier auftretende Ausbreitungsgeschwindigkeit $c_p$ für quasi-longitudinale Wellen in Platten variiert bei den Baumaterialien nur wenig. Sie beträgt für Ziegelstein etwa 2350 m/s. Dagegen variieren die Dicken $h$ der Wand- schalen beträchtlich. Sie entscheiden, ob sich das Vor- handensein von Eigenfrequenzen stark bemerkbar macht. Bei einer bimssteinstarken Vollziegelwand mit $h = 0,25$ m fällt auf den untersten Terzbereich noch nicht einmal eine Eigenfrequenz.

Die Neufassung von R 140 sieht außerdem vor, die Bau- ähnlichkeit der Randbedingungen wenigstens durch eine Körperschallnachhallmessung zu kontrollieren.

Es gibt auch die Möglichkeit, die Randverluste zu berechnen. Ähnliches gilt für die Flankenübertragung und die bei Wandverzweigungen auftretenden Energieverzwei- gungen. Bei senkrechtem Einfeld und tiefen Frequenzen er- hält man sogar relativ maßgebliche einfache Formeln /7/. Be- rücksichtigt man aber, daß die Biegewellen auch schräg auflauen, so erhält man sehr verwischte Winkeldahfähig- keiten /7,8/ und es fragt sich, ob die entsprechenden Transmissionsgrade noch im Sinne einer Gleichverteilung zu ermitteln sind.-Jedenfalls dürfte es zur Zeit keinen beratenden Akustiker geben, der mit vertretbarem Aufwand in der Lage wäre, die Schallübertragung zwischen zwei Räumen, auch wenn er alle konstruktiven Daten kennt, vor- auszuberechnen, also $D_n$ oder $R'$ anzugeben. Wieviel weniger könnte man von einem Architekten erwarten, daß er in der Lage ist, eine Normpegeldifferenz zu garantieren. Er kann nur nachweisen, daß er Bauteile verwendet hat, deren Schalldämme als ausreichend anerkannt sind. Insofern ist gerade er an Tabellen von $R$-Werten interessiert, die unter normalen Einbauverhältnissen bei möglichster Aus- schaltung der Nebenwege gemessen sind. Aber selbst wenn man $D_n$ oder $R'$ im Einzelfall vorausberechnen könnte, so bleibt dieses Ergebnis im allgemeinen doch unübertragbar auf ein anderes Haus mit zwar gleichem trennwandkonstruk- tion, aber anderen Flankenwänden oder anderen Abmessun- gen und somit uninteressant. Nur bei Fertighäusern könne- te und sollte man die ein für allemal geltenden $D_n$-Werte bestimmen und angeben. Vielleicht würde diese Forderung auch dazu führen, daß der Akustiker bei ihrer Entwick- lung etwas mehr eingeschaltet wird.

b) Das Bewertungsverfahren

Es erscheint naheliegend, aus den für die verschiedenen Terzbereiche gewonnenen $R$- (oder auch $D_n$-) Werten einen Mittelwert zu bilden und hierfür Mindestwerte festzu- setzen. Abgesehen davon, daß bereits in den gewählten Logarithmen eine Willkür liegt, da diese nicht die subjektive Empfindung repräsentieren, würde es bei solchen
Mittelwerten verhältnismäßig leicht sein, Mängel im tie-
fen und mittleren Frequenz-
bereich durch leicht errech-
bare hohe Werte bei hohen
Frequenzen zu kompensieren.
Daher wird statt der Mittel-
wertbildung ein Vergleich mit
einer Sollkurve vorgenommen,
ähnlich wie man bei Geräuschen,
deren Pegel-Oktavspektrum mit
vereinfachten Kurven gleicher
Lautstärkepegel (NR-Kurven)
vergleicht (Fig. 2). Nur be-
stimmt dort der höchste Meß-
wert die dann als maßgebend
angeschene Kurve, während es
hier auf die Unterschreitun-
gen ankommt. Außerdem läßt
man eine gewisse "mittlere
Unterschreitung" von 2 dB
zu, einmal um die Zulassung
der Konstruktion nicht von
der Unsicherheit eines Meßpunktes abhängig zu machen,
dann aber auch, um sie nicht von der genauen Wahl der, der
Einfachheit halber, aus einzelnen Geraden bestehenden Soll-
kurve zu abhängig zu machen.

Bei der Wahl der Sollkurve ging man von der als Wohn-
ungstrennwand bewährten vollsteinstarken, beiderseits
verputzten Ziegelwand aus. Diese Wahl enthielt zwei Ge-
sichtspunkte, nämlich einen subjektiven – die Wand hatte
nicht zu Klagen geführt – aber auch einen ökonomischen;
die Konstruktion stellte keine Verfeuerung gegenüber dem
am Bau üblichen dar.

(Im Streit um Schallschutzforderungen geht es immer
um ein Abwägen von subjektiv Wünschbarem und wirtschaft-
llich Zumutbarem.)

Aus subjektiven Erwägungen konnte man die Anforderun-
gen gegenüber der Ziegelwand einmal nach tiefen Frequen-
zien wegen des Nachlassens der Ohrenempfindlichkeit mildern,
dann aber auch bei hohen, wegen des Nachlassens der Ener-
gieanteile im Störspektrum, und weil in diesem Bereich
die Dämmäste schon so hoch sind, daß die von dem Trenn-
objekt durchgelassenen Anteile nur noch geringe Beiträge
liefern.

Diese Sollkurve wird auch zur Kennzeichnung nicht aus-
reichender oder bei weitem ausreichender Konstruktionen
benutzt, indem man die Verschiebung nach oben oder unten
– oder den Wert der verschobenen Kurve bei 1000 Hz – an-
gibt, die die Konstruktion als gerade zulässig
Die so gewählte Sollkurve ist auch heute noch international in Gebrauch, vielleicht weil sie noch eine andere Interpretation erlaubt. Eine ihr angepasste Konstruktion liefert im Empfangsraum eine Kurve gleicher Lautstärkeepegel, wie sie der Bewertung eines Schalldruckpegels in dBA entspricht, vorausgesetzt, im Senderraum wird in allen Terzbereichen der gleiche Pegel in dB erzeugt. Gösele u.a. /9,10/ haben darauf ein verkürztes Meßverfahren gegründet, wobei unter diesen Umständen nur noch die Gesamtpegel in dBA in beiden Räumen verglichen werden.


Trotz aller Schwierigkeiten, die sich aus den endlichen Abmessungen der Räume und Wände und aus den flankierenden Bauteilen ergeben und die nie mit einfachen Verfahren eliminiert werden können, und trotz der Problematik adäquater, subjektiver Bewertung, kann man sagen, daß man den Luftschallschutz gegen den Nachbarn zur Vermeidung von Klagen hinreichend erforscht und im Griff hat.

2. Schallschutz gegen die äußere Umwelt

a) Schallschutz durch Fenster und zugehörige Meßmethoden

Jedenfalls ist es zur Zeit viel dringender, Meßmethoden und Anforderungen für den Luftschallschutz nach außen festzulegen.

Da das Schalldämmaß einfacher Trennwände umso größer ist, je schwerer sie sind, leuchtet ein, daß bei tragenden Außenwänden nur die Fenster den Schallschutz bestimmen und daß man, sofern man einfache Scheiben verwendet, diese möglichst dick machen sollte.

Dem steht nun allerdings der Koinzidenzeffekt entgegen, d.h. die bei höheren Frequenzen mögliche Anpassung der Spurgeschwindigkeit \( c_0/sin \theta \) der einfallenden Luftschallwelle an die einer freien Biegewelle \( c_p \). Dieser Effekt ist gerade bei Fensterscheiben sehr ausgeprägt, einmal weil sie nur geringe innere Verluste und bei normaler Bettung in den hart werdenden Kitt keine Randdämpfung aufweisen, aber auch, weil vor dem Fenster bestimmte Einfallsrichtungen dominieren können.
Die Neufassung von ISO R 140 trägt dem Rechnung, indem zur Prüfung des Schallschutzes von Fenstern nicht nur der Einbau zwischen einem Sende- und einem Empfangsraum herangezogen werden kann, sondern auch zwei "Einraum"-Methoden, die sich insbesondere für Messungen im Gebäude eignen. Die eine besteht darin, daß man den vorhandenen äquivalenten Dauerschallpegel innen und außen mißt, wo bei das Mikrofon einmal 2 m vor dem Gebäude angebracht wird. Schon hierbei wird keine allseitig gleichmäßige Verteilung der einfallenden Schallwellen zu erwarten sein. Man kann aber auch, indem man einen Lautsprecher in schräger Richtung zur Fensternormalen anbringt, das Schalldämmaß unter einer Vorzugsrichtung, z.B. 45°, ermitteln. Welcher Schallpegel \( L_f \) in diesem Falle der einfallenden Schallwelle entspricht, kann man ein für allemal in einem schalltoten Raum oder im Freien bestimmen. Man erhält dann \( R(\theta^0) \) aus:

\[
R(\theta^0) = L_f - L_2(\theta^0) + (10 \log_{10} \frac{4S \cos \theta^0}{A}) \text{ dB (11)}
\]

Bild 3 zeigt zwei so von EISENBERG /12/ gewonnene Frequenzgänge von \( R(45^0) \). Man erkennt deutlich, daß bei tiefen Frequenzen die dickere Scheibe überlagert ist, dagegen bei 2000 Hz ihr Dämmaß einen tiefen Einbruch erfährt. Daß die Spuranpassung sich erst bei hohen Frequenzen bemerkbar macht, liegt daran, daß die Ausbreitungsgeschwindigkeit der freien Biegewellen

\[
c_B = \sqrt{2\pi f B/m} = \text{konst} \sqrt{f h} \text{ (12)}
\]

mit der Wurzel aus der Frequenz \( f \) wächst. Sie erreicht erst bei hohen Frequenzen die konstante Ausbreitungsgeschwindigkeit in Luft. Da \( c_B \) auch mit der Wurzel aus der Scheibendicke \( h \) wächst, geschieht dies bei der dünneren Scheibe erst oberhalb des interessierenden Bereichs. Man kann aber auch durch Wahl eines kleinen Verhältnisses von Biegesteife \( B \) zu Masse je Fläche \( m \) erreichen, daß die kritische Frequenz, oberhalb der die Spuranpassung erstmals eintreten kann, über den konventionellen Meßbereich hinausgeschoben

*Fig. 3: \( R(45^0) \) für zwei Fensterscheiben unterschiedlicher Dicke nach EISENBERG*
wird, z.B. indem man eine dicke Einfachscheibe in zwei Scheiben mit einer durchsichtigen verbindenden aber seitlich Verschleibungen nicht hindernder Schicht verwendet.

Auch und gerade bei Fensterscheiben spielen die Rand- einspannungen eine große Rolle. Hier hat eine elasti- sche Lagerung in Neoprene- streifen (siehe Fig. 5), wie sie UTLEY und FLETCHER /14/ untersucht haben, den umge- kehrten Effekt wie bei schwe- ren Trennwänden. Der Energie- abfluß in den schmalen Rah- men ist ohnehin unbedeutend. Dafür kann der Neoprene- streifen infolge seiner inne- ren Verluste den dünnen Scheiben erheblich Energie entziehen. Dies zeigt sich bei den in Fig. 5 eingetra- genen Messungen vor allem daran, daß der Energie- einbruch abgeflacht wird. Bei den Doppelscheiben kommt eine isolierende Wirkung der Schallbrücke, die der Fensterrahmen bietet, hinzu.

Der Wert schluckender La- bungen in Luftpolster ist im allgemeinen gering, da nur dünne Schichten zur Verfüg- gung stehen. Dagegen wirken poröse Schichten, die den Fensterräumen selbst seitlich in Räume untertei- len (Fig. 6) viel stärker, insbesondere auch bei tiefen Frequenzen. Macht man zudem nur einen kleinen Flügel- öffnungsbogen, so wirkt der nicht geöffnete Teil als Luftfederung oder gar als Rezonanzraum hinter dem Strömungswiderstand. Nach den in Fig. 7 wiedergegebenen Messungen von P. SCHNEIDER ist diese Wirkung auch noch bei leicht geöffnetem Fenster festzustellen.

Übrigens braucht man auch bei geschlossenem Fenster nicht auf natürliche Lüftung.
zu verziehen. Man braucht nur Lichtdurchtritt und Luft-
durchtritt zu trennen. Der letzte kann auch über geson-
derte schallschluckend ausgek
kleidete Kanäle über und un-
ter den Scheiben erfolgen, evtl. in Verbindung mit einem
Ventilator. Das natürliche
Lüften ist in lauter Umge-
bung meist ohnehin nicht zu
empfehlen. Wo "Schall" ist,
ist meist auch "Rauch" oder
eine sonstwie verschmutzte
Luft.

Daß man aber auch mit
normalen Fenstern beacht-
liche Schalldämmungen er-
zeugen kann, möge der von
EISENBERG/12/ an einem sorgfältig gearbeiteten Doppel-
fenster mit allerdings auch 24 cm Scheibenabstand ge-
messene Frequenzgang von R beweisen (Fig. 7).

Wenn ein solches Fenster nur 1/10 der Außenwandfläche
ausmacht, so kann sein Schalldämmaß auch um 10 dB unter
dem der übrigen Außenwand liegen, ohne daß durch das
Fenster mehr Schall eindringt als über den Rest der Wand.
Anders ist es bei den architektonisch so beliebten Glas-
fassaden, bei denen alles Fenster ist.

Aber auch unter den sogenannten Vorhangwänden gibt
es Konstruktionen, deren Schalldämmung weit hinter der
eines guten Kastfenstern zurückbleibt. Dabei gäbe es
heute viele Möglichkeiten, eine nicht durchsichtige
Doppelwand mit nur 50 kg/m² Masse je Fläche und Dicken
von 6 bis 8 cm herzustellen.

b) Bewertung der Meßergebnisse

Die Bemühungen um Meßverfahren über innere Wände und
Decken, die internationale Einigung über Sollkurven und
die Aufstellung von Zulassungsbedingungen, haben sicher
dazu geführt, daß die Störungen durch den Wohnungsnach-
barn heute eine geringere Rolle spielen als die durch
Lärmbelästigung von außen. Umso dringender wäre es, auch
für Außenwände Sollkurven und Mindestwerte festzulegen.

Wenn dies bisher noch nicht international geschehen
ist, so mag es in erster Linie daran liegen, daß einer-
seits die zu stellenden Anforderungen je nach Raum-
und Gebäudeart verschieden sind und mehr noch, daß die vor
einem Hause zu erwartenden Geräuschpegel je nach Wohn-
lage sehr unterschiedlich sind. BRUCKMAYER/15/ hat in
diesen Sinne für die österreichischen Normen Regeln
aufgestellt, wann Fenster mit einem mittleren Schalldämmmaß von 30 dB genügen und wann solche mit 35 dB anzunehmen wären. In einer deutschen Vornorm (DIN 18005) werden die in Tabelle I angegebenen A-bewerteten Dauerpegel als erstrebenswerte Höchstgrenzen angegeben:

<table>
<thead>
<tr>
<th></th>
<th>Tag</th>
<th>Nacht</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Reines Wohngebiet</td>
<td>50</td>
<td>35 dB(A)</td>
</tr>
<tr>
<td>2. Allgemeines Wohngebiet</td>
<td>55</td>
<td>40</td>
</tr>
<tr>
<td>3. Mischgebiet</td>
<td>60</td>
<td>45</td>
</tr>
<tr>
<td>4. Gewerbegebiet</td>
<td>65</td>
<td>50</td>
</tr>
<tr>
<td>5. Industriegebiet</td>
<td>70</td>
<td>70</td>
</tr>
</tbody>
</table>

Man beachte, daß bei 1 bis 4 zwischen Tag, das ist 6 bis 22 Uhr, und Nacht, 22 bis 6 Uhr, eine Pegeldifferenz von 15 dB(A) eingesetzt ist. Nur beim reinen Industriegebiet fehlt diese Differenz. Trotzdem gibt es, wie wir gesehen haben, durchaus Fensterkonstruktionen, die auch in diesem bei Nacht die eingangs erwähnten 30 dB(A) im Schlafzimmer garantieren könnten, sofern man auch diese hierbei als äquivalenten Dauerschallpegel auffaßt. Im "reinen Wohngebiet" würde das Fenster nur noch 5 dB(A) bei Nacht aufzubringen haben, d.h. der Bewohner könnte bei Beschränkung der Öffnung seines Fensters auf einen meist zur Lüftung ausreichenden Schlitz bereits den Wunschpegel im Schlafzimmer einhalten. Im Bereich von Großstädten scheint freilich der angestrebte Ausgangspegel

Überwautop geht die Aufteilung in unterschiedliche Geräuschzonen, je nachdem ob dort lärmerzeugende Betriebe sind, an der Tatsache vorbei, daß die meisten Geräuschklagen auf Verkehrs lärm beziehen /16/, den ja die Bewohner reiner Wohngegend auch unvermeidlich selbst erzeugen. Zumindest müßte man – wie GABLER /17/ das in 4 von ihm aufgestellten Regeln fordert – dort jeden Durchgangsverkehr vermeiden. Wie schwer übrigens schon die Aufteilung einer Stadt in Industrie-, Gewerbe- und Wohngebiete ist, zeigt das in Fig. 6 wiedergegebene Beispiel der Stadt Newbury nach THIESSEN /18/. Man müßte – ebenfalls nach THIESSEN – künftig Städte so planen, wie es Fig. 9 zeigt, eine Zumutung, die sicher bei jedem Städtebauer nur eine Gänsehaut erzeugen kann.

Übrigens nimmt auch dieses Bild nicht darauf Rücksicht, da ß kein Schallpegel an einer Grenze plötzlich um 5 dB fallen kann.


Jede feste Schirmwand aus fügenlosen dünnen Brettern ist überlegen. Ihre Wirksamkeit nach tiefen Frequenzen wird hier durch die Beugung begrenzt, die umso stärker in Erscheinung tritt, je geringer der Umweg ist, zu dem der Schall gezwungen ist.

U. KURZE /19/ der das Zusammenwirken mehrerer Geräuschquellen mit den Gesetzen der mathematischen Statistik
behandelt hat, weist darauf hin, daß die Wirkung eines Schirms endlicher Länge nicht nur auf seiner Pegelsonnung beruht, sondern auch darauf, daß aus einem stark schwankenden Pegel ein gleichmäßiges Grundgeräusch wird. Dieser Unterschied wird vom Äquivalenten Dauerschallpegel nicht erfaßt, dagegen vom eigens hierzu entwickelten Traffic Noise Index /20/, sowie dem allgemein anzuwendenden Noise Pollution Level /21/.

Wenn Schirmwände nicht so leicht und hoch gebaut werden können, wie es akustisch zweckmäßig wäre, so liegt das daran, daß sie nicht nur vertikal ihr eigenes Gewicht tragen müssen, sondern auch horizontal die Windkräfte.

Stabilisierete Schattenbilder sind daher Wälle, bzw. bei Verkehrsadern deren Verlegung in Einschnitte - hier sind auch Beplantungsmatten mit Buschwerk vorteilhaft - und Häuserzeilen. Die in Fig. 10 wiedergegebene Prinzipskizze von GABLER /17/ für eine kleine Stadt mit Umgehungsstraße zeigt beides. Auch der modernen Stadt tut eine Stadtmaur gut - gegen die Pfeile des Industrielärms und die Kanonenkugeln des Verkehrslärms. Dabei kann diese Mauer "zivil" genutzt werden durch Gewerbebetriebe, die geräuschempfindlich sind, aber nicht selbst geräuscherzeugend. Sie kann sogar bewohnt werden, wenn man zur Lärmsseite Flure, Küchen und Badezimmer legt oder den Außensteil der Kastenhäuser nicht scheut.


Auch Wohnungen eines großen Wohnhauses, das unter
EIGHTH INTERNATIONAL CONGRESS
ON ACOUSTICS, LONDON 1974

STAND DES SCHALLSCHUTZES IM WOHNUNGSBAU
(NOISE ABATEMENT IN DWELLING HOUSES)

Einschaltung einiger Park- oder Gewerbe-Etagen direkt über eine Autobahn gebaut ist, sind zumindest lufts-
Schallmäßig geringeren Belästigungen ausgesetzt als Wohnungen neben einer Autobahn.

Kurzum, es gibt vielerlei städtebauliche Maßnahmen,
die aufwendige Fensterkonstruktionen zu ersparen ge-
statten. Sie alle unterliegen aber auch den unvorherseh-
baren Umgebungen durch Schallbrechung in der Luft infolge von Temperatur- und Windunterschieden. Aber auch von diesen abgesehen, sollten die Normenausschüsse bald-
möglichst zur Frage der Bewertung von Fenster- oder ganzen Außenwandmessungen Stellung nehmen und womöglich auch hierfür Mindestforderungen stellen, die wenigstens dort beachtet sein sollten, wo Gebäude aus Mitteln Dritter errichtet werden. Vor allem aber sollte es Güte-
prüfungen für Hotels geben; für sie sollte eine Plakette am Eingang ausgestellt werden.


Das empfohlene Schallschutz-Güte-Schild an Hotels verlangt aber außer den in diesem Vortrag behandelten Problemen auch noch ausreichenden Schallschutz gegen un-
mittelbare Anregungen der Baukonstruktionen, wie das z.B. beim Begehen der Fall ist, oder wie es durch fließen-
de Wasserleitungen und schlechte Armaturen geschieht.
Die Vortragszeit erlaubte es nicht, diese Probleme ein-
zubeziehen. Es sei aber erwähnt, daß die ISO sich auch hiermit sorgfältig befaßt und es sei angenommen, daß beim nächsten ICA-Kongreß ein Übersichtsvortrag über Trittschall- und Wasserleitungsgeräusche eingeplant wird.

Literatur
(1) K. Gösele, 3th ICA, Stuttgart, Elsevier Amsterdam (1961), II. p.989; (2) O. Brandt, 4th ICA, Kopenhagen (1962); II. p.31; (3) M. Heckl, u. K. Seifert, Acustica 8 (1958) 212; (4) K.A. Mulholland and R.M. Lyon, JASA (1973) 54 867; (5) M. Heckl, Frequenz (1964) 18 299; (6) Siehe z.B. R.H. Bolt, JASA (1939) 10 228; (7) L. Cre-
EIGHTH INTERNATIONAL CONGRESS
ON ACOUSTICS, LONDON 1974

STAND DES SCHALLSCHUTZES IM WOHNUNGSBAU
(NOISE ABATEMENT IN DWELLING HOUSES)

(8) T. Kihlman, Rapport fran Byggforskningen, Stockholm
(1967) 2; (9) K. Gösele, Acustica (1965) 15 264;
(10) Quindry, T.L. u. D.R. Flynn, JASA (1973) 53 387;
1961, II, p.1041; 6th ICA Tokyo, E-3-10, 1968;
(13) Wooley, R.M. Building Res. Stat., Design pap.64;
(14) W.A. Utley, B.L. Fletcher, J.Sound a. Vibr. (1973)
26 (1) 63; (15) F. Bruckmayer, Schalltechnik im Hochbau,
Wien 1962; (16) L. Schreiber, Lärmschutz im Städtebau,
Wiesbaden 1971; (17) W. Gabler, Kampf dem Lärm, (1965)
Heft 4; (18) G.J. Thiessen, Vortr. P2 Meeting Acoust.
d. VDI-Zeitschriften (1973) Reihe 11, Nr. 15;
(20) J.D. Griffiths and F.J. Langdon, J. of Sound and
Vibr. (1968) 8 16; (21) D.W. Robinson, NPL Acoustics
ENVIRONMENTAL SOUND PROPAGATION

Zyun-iti Maekawa  Faculty of Engineering, Kobe University,
Rokko Nada Kobe 657 JAPAN

1. INTRODUCTION

It is essential to know the sound field of the source at all points, in order to solve the environmental noise problem. In "Environmental Acoustics" we have to consider the sound field with respect not to a microphone but to human ears. It is suitable to separate the consideration about sound propagation into two parts, one is wave divergence and the other is excess attenuation (1,2).

Wave divergence is used to deal with the geometrical relation between the source and the receiver, assuming that the elemental point source radiates sound energy incoherently and that the air is a homogeneous and loss-free medium. Studies of wave divergence have been presented as noise reduction by distance from various shaped sources (3,4).

The excess attenuation owing to environmental and boundary conditions is the attenuation beyond that caused by wave divergence. It includes the effects of atmospheric inhomogeneities (5-7), wind and temperature gradients, and of weather conditions (8-11), and ground effects, plants and trees (12-16), and includes the attenuation created by barriers.

The main interest of this paper is to review the studies on barrier performance and then to introduce a method to measure the environmental sound propagation, using Fourier-Transform techniques.

2. SOUND ATTENUATION BY BARRIERS

Some fifteen years ago, there was no reliable method to design a screen for the purpose of noise reduction. Theoretical works giving the foundations of this problem have been reviewed in the text by Bowman et al. (17). There remains the problem of how to simplify and apply the solution to noise control.

2.1 Half-Plane in Free Space

For the diffraction of a plane wave by a semi-infinite (half) plane, the well known exact solution was given by Sommerfeld in 1896 (18) as shown in Eqn.(1) and Fig.1 with the notation shown in Fig.2.

\[
\phi_s = \frac{e^{-ikB\cos(\theta+\phi)}}{\sqrt{\pi}} \left[ e^{-\int_{W_1} e^{2t^2} dt} + e^{-\int_{W_2} e^{2t^2} dt} \right]
\]

where \( k = \frac{\omega}{c} \), \( W_1 = -\sqrt{2kB} \cos \frac{1}{2} (\theta+\phi) \), \( \frac{\omega}{2\pi} \) is frequency, \( W_2 = +\sqrt{2kB} \cos \frac{1}{2} (\theta-\phi) \), and \( c \) is velocity of sound

(1)
Fig. 1 Sound attenuation of a plane wave by a half plane in free space calculated by the theory of Sommerfeld as shown in Eqn. (1). $\psi$ = diffraction angle.

Fig. 2 Geometry and notation of a half plane.

When a sound source is located at a finite distance from the half plane, Redfearn's approximation in 1940 (19) is well known. However, recently Jonasson (20) gave a better solution for this situation after Macdonald's exact solution in 1915 (21), as shown in Eqn. (2) and Fig. 3. Fig. 1 and Fig. 3 are reviews of Redfearn's work in a refined form.

Noise reduction by a thin screen on the ground was calculated by Fehr (22) and Rettinger (23) using the well known Kirchhoff's diffraction theory, under the condition that both the source and the receiver are on the ground. Unfortunately, sufficient experimental data was not available to verify the applicability of these theoretical results.
EIGHTH INTERNATIONAL CONGRESS ON ACOUSTICS, LONDON 1974

ENVIRONMENTAL SOUND PROPAGATION

Fig. 3 Sound attenuation of a spherical wave by a half plane in free space calculated by the theory of Jonasson (20) as shown in Eqn. (2)

\[
\phi_j = \frac{\mathrm{e}^{-i\pi}}{\sqrt{\pi}} \left[ \frac{\sqrt{2}}{k\sqrt{S(R_1+S)}} \int_0^\infty e^{i\psi u^2} \, du + \frac{\sqrt{2}}{k\sqrt{S(R_2+S)}} \int_0^\infty e^{i\psi u^2} \, du \right]
\]

where \( S = A + B \)

\[
R_1^2 = A^2 + B^2 - 2AB\cos(\theta_1),
\]

\[
R_2^2 = A^2 + B^2 - 2AB\cos(\theta_2),
\]

\[
w_1 = -\sqrt{\frac{4kAB}{S+R_1}} \cos \frac{1}{2} (\theta_1 + \theta_2),
\]

\[
w_2 = +\sqrt{\frac{4kAB}{S+R_2}} \cos \frac{1}{2} (\theta_1 - \theta_2).
\]

An experiment was performed by the author in 1960. The sound fields diffracted by a thin rigid screen were measured by the pulsed tone method in the laboratory. The condition was satisfactory for the experimental semi-infinite screen in free space. The sound pressure levels at many points in the shadow zone of the screen were measured for a wide range of frequencies (5-40 kHz) and for a wide variety of geometries (diffraction angle \( \psi = 0^\circ - 90^\circ \)). All the results are shown to lie approximately on one curve plotting against Fresnel's zone number \( N \) according to the Fresnel-Kirchhoff's diffraction theory (24) as shown in Fig. 4 (25-27). The abscissa of the figure was adjusted so that the
Experimental curve became a straight line, for convenience as a design chart. Kurze (28) expressed this result with

$$s_{\text{[Att]}}_o = 10 \log(20N) \text{ dB} \quad N > 1 \quad (3)$$

where [Att] means the attenuation in dB, and suffixes s and o refer to "of a spherical wave" and "by a half plane (zero thickness)" respectively. He showed that Eqn. (3) corresponds to the first term of the asymptotic formula derived from the geometrical theory of diffraction by Keller (29). Tatge (30) also showed an approximation of the results in a similar form,

$$s_{\text{[Att]}}_o \approx 10 \log(3 + 20N) \text{ dB} \quad N > 0 \quad (4)$$

![Diagram](image-url)

**Fig. 4.** Sound attenuation by a semi-infinite screen in free space. Horizontal scale, logarithmic scale in the region of $N > 1$, adjusted so that the experimental curve becomes a straight solid line in the region of $N < 1$. Depending on whether $N > 0$ or $N < 0$, the receiving point $P$ lies in the illuminated region or in the geometrical shadow, respectively. Attenuation is relative to propagation in free space. o, experimental value measured by pulsed line. ($\psi = 0^\circ \text{ or } 90^\circ$).

For the transition region between the shadow and bright zones, Kurze (2 or 31) showed, from Keller's theory,

$$s_{\text{[Att]}}_o \approx 20 \log \frac{\sqrt{2\pi N}}{\tanh \sqrt{2\pi N}} + 5 \text{ dB} \quad N > 0.2 \quad (5)$$

This equation is convenient for calculation by computer, though it has a small error compared to the exact values in Fig.3.
2.2 Finite-sized Screen of Arbitrary Shape

In general, even if a wall or screen has its own particular shape, the sound level in the shadow zone of the wall or screen may be integrated from all contributions from the open surface, according to the principle of Kirchhoff's integral. An approximate method for a screen shaped by right angles has been presented (26,27 & 32). Also another method, using the Fresnel-Surface-Integral (33), for a thin triangular screen which has an arbitrary vertex angle was reported (34). Fig.5 shows an example of the results, which shows that the calculated values overestimate the attenuation by about 3 dB. This fact corresponds to Fig.4 in the half-plane problem. These discrepancies must be caused by the rough approximation of the Fresnel-Kirchhoff's diffraction theory. In other words, all attenuation values calculated by the Kirchhoff's diffraction theory must be reduced by 3 dB.

![Diagram of a screen and measurements](image)

Fig.5-a Location of measurements parallel to the screen

Fig.5-b Sound level distribution in the shadow zone of the screen, -----calculated values by the Fresnel-Surface-Integral.

2.3 Effect of the Thickness of a Screen

The theoretical works mentioned above assume the screen has zero thickness. According to much experimental data, the effect of the thickness of a screen should be negligible as long as the thickness is smaller than the wavelength, for example as shown in Fig.6. However, there are many occasions when the thickness of the barrier such as a building, must be considered.

A thick plate or solid body has double edges that affect diffractional attenuation (35). Kurze (28) gave a simple but reasonable approximate method by adding the two attenuations created by the edges to obtain the whole attenuation caused by the thick barrier. Pierce (70) has reported a further theoretical study of this problem, using Keller's geometrical diffraction theory and an analytical procedure.

On the other hand, Fujiwara and the author (36) have reported a method to obtain exactly the effect of the thickness "b" of a thick
Fig. 6 Sound attenuation by a semi-infinite plate, 15 mm thickness, vs. Fresnel's zone number N. Solid curve; Kirchhoff's theory, o x □; experimental values, measured by pulsed tones.

Fig. 7 The value of \( K = n[ET]_b / \log_{10} kb \) in Eqn. (8) to calculate the effect of thickness of a noise barrier.

plate, \([ET]_b\), which must be added to the value of the attenuation by a thin half plane as shown in Eqn. (6),

\[
[\text{Att}^p_b] = [\text{Att}^o_b] + [ET]_b
\]  

(6)

The effect of thickness \([ET]_b\) is assumed to be approximated by,

\[
[ET]_b \approx [\text{Att}^p_b] - [\text{Att}^o_b]
\]  

(7)

where the suffixes \( p \) and \( o \) refer to "of a plane wave" and "by a thick plate with thickness of \( b\)" respectively. For the numerical computation, the approximate solution of \([\text{Att}^p_b]\) (37,38) and the exact solution of \([\text{Att}^o_b]\) are used. The results of the theoretical computation show the resonance effect of thickness of \( b\). This fact is not convenient for practical estimation of noise reduction by barriers. Therefore the authors present a single chart to obtain the effect of thickness for a noise having considerable band-width, after reasonable approximation as shown in Fig. 7 (39). The effect of thickness can be obtained by Eqn. (8) using the value of \( K \) shown in Fig. 7.

\[
n[ET]_b = K \cdot \log_{10} kb
\]  

(8)

It is useful to estimate the attenuation of the thick barrier for field noise.
2.4 Effect of the Angle of a Wedge

The theory of the sound diffraction by a wedge has been treated by many authors (17,21,40-44), however, the solutions are too complicated to apply to practical use. Kurze (31) showed a correction term for the asymptotic formula for a half plane, in order to obtain the attenuation by the wedge. The author (45) intended to obtain the effect of the wedge-angle $\Omega$, $[\text{FW}]_\Omega$ dB, using a graphical chart in order to get the attenuation by the wedge, in the similar way to obtaining the effect of thickness, as follows:

$$s[\text{Att}]_\Omega = s[\text{Att}]_O + [\text{FW}]_\Omega$$  \hspace{1cm} (9)
and \[ [\text{EW}]_\Omega \simeq p_{\text{[Att]}}_\Omega - p_{\text{[Att]}}_0 \] (10)

where suffix \( \Omega \) refers to "by a wedge-angle of \( \Omega \)." For the numerical computation of \( p_{\text{[Att]}}_\Omega \) and \( p_{\text{[Att]}}_0 \) the asymptotic solution by Oberhettinger (41) and the exact solution by Sommerfeld were used respectively. \([\text{EW}]_\Omega \) is a function of not only the wedge angle \( \Omega \), but also the incident angle \( \theta \), the diffraction angle \( \psi \) and \( kB \), where \( k = \omega/c, B \): distance of the receiver from the top of the wedge, with the notation as shown in Fig.8. From the results of numerical calculation for the right angle wedge \( (\Omega = 90^\circ) \) the effect of wedge angle does not depend on \( kB \) in the region of \( kB \geq 10 \). The examples of the computed values of Eqn. (10) are shown graphically in Fig.9 and Fig.10. It is clear that \([\text{EW}]_\Omega \) decreases the barrier attenuation referred to that of the half plane, but not over 6dB, and vanishes when \( \Omega \to 0 \) or \( \psi \to 0 \).

Fig.11 shows very good agreement between the estimated values of attenuation by a right angle wedge obtained by Eqn. (9) with Fig.4, 9 & 10, and the experimental values measured by the pulsed-tone method.

2.5 Effect of the End of a Wedge

Every barrier has an end. When a barrier is used for noise control, it is often necessary to know the effect of that end. Supposing the corner of a building is the barrier, for example, preliminary experiments were performed by the author (46). The situation of the measurement is shown in Fig.12a. And a portion of the results of the measurements are shown in Fig.12b, in the normalized distance scale.

In this figure, the curves seem to approach quickly to their own constant values, which agree with the estimated values for an infinitely long wedge.

![Diagram](image-url)

Fig.12-a Location of measurements. S and P are on a plane perpendicular to the edge O-Z of the right angle wedge, and moved simultaneously parallel to the edge, through the points of \((x_1, y_1)\) and \((x_2, y_2)\) respectively. The wedge has an end-plane at \( Z = 0 \), occupying the negative coordinate.

![Diagram](image-url)

Fig.12-b A portion of the results of measurements displayed on the normalized distance.

- - - - - - calculated values for the infinite long wedge
2.6 Effect of Surface Absorption of the Barrier

Until now, the surface condition of the barrier has been regarded as rigid (or hard). However walls or screens with heavily absorptive treatment have been widely used for noise barriers. The question to be answered is "What effect can be expected by the surface absorption?".

Theoretically, the word "hard", expressing $\partial \phi / \partial n = 0$, naturally means the rigid surface with no absorption, but the word "soft", expressing $\phi = 0$, does not mean an absorbing surface. When the acoustic surface impedance is introduced in the fundamental equation, the second term of Eqn. (2) must be multiplied by the reflection coefficient $Q$ of the surface. The coefficient must be the spherical reflection coefficient which is a function of the distance from the source, of the specific admittance $\beta$ of the surface and of the angle of incidence $\theta$ (47). For practical convenience, however, it is often approximated by

$$Q = (\cos \theta - \beta)/(\cos \theta + \beta)$$

(11)

The effect of surface absorption of the half plane should be obtained by the difference of the value of attenuation from Eqn. (2) with and without multiplying the second term by the coefficient $Q$. The results of this computation are complicated because $\beta$ is a complex number. If the imaginary part of $\beta$ is negligible under the condition of a far field we can get a simple chart as shown in Fig.13, which may be very useful for quick estimation of the effect of absorption in the practice of noise control (48). This fact seems to be verified by the experimental results shown in Fig.14 (25).
Table 1. Sound pressure reflection coefficients measured by the pulsed-tone method.

<table>
<thead>
<tr>
<th></th>
<th>5 kHz</th>
<th>10 kHz</th>
<th>20 kHz</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hard screen</td>
<td>0.85</td>
<td>0.92</td>
<td>0.71</td>
</tr>
<tr>
<td>Soft screen</td>
<td>0.28</td>
<td>0.19</td>
<td>0.10</td>
</tr>
</tbody>
</table>

Fig. 14 Sound attenuation in free field measured by two semi-infinite screens, whose absorption coefficients are shown in Table 1. Solid curves: the calculated values in Fig. 3. Parameter $\psi$: the diffraction angle. The data for Hard Screen were plotted in Fig. 4.

2.7 Effect of Shape and Size of the Sound Source

It is a more difficult problem to treat theoretically sound diffraction with a source of large size, because the wave front from the large source cannot be expressed exactly. There is a conventional method, however, if a large source can be replaced by one or more point sources.

When the noise is emitted incoherently from point sources, the shielding effect of the barrier for the group of sources can be expressed as follows,

$$[\text{Att}] = 10 \log \left( \frac{\sum_{\xi=1}^{n} \frac{K_{\xi}}{d_{\xi}^2}}{\sum_{\xi=1}^{n} \frac{K_{\xi}}{d_{\xi}^2}} \right) \log^{-1} \left( \frac{[\text{Att}]_{i}}{10} \right) \text{ dB} \quad (12)$$
where $K$ = the factor for each point source, $d$ = the distance from each source to the receiver and $[\text{Att}]$ = the value of attenuation due to the barrier at the receiving point for each source, which can be obtained by methods mentioned above. The experimental results measured by the scale model have verified this conventional method (35).

When a source of finite plane is considered as a rectangular array of small equally spaced incoherent sub-sources, an analytical solution by integration has been given by Tatge (30), under the condition of the receiver being very far from the barrier. This form is suitable for calculation by computer.

The street noise or highway noise is often treated as an incoherent line source. The performance of a barrier against the highway noise, therefore, should be considered with a line source parallel to the edge of the barrier. Kurze (2 or 31) showed the results calculated by integrating over point sources that make up the line source, as curve B in Fig.15. On the other hand, Yamashita et al. (49) have reported the purely experimental data measured by the scale model in an anechoic room as curve C in Fig.15. It shows that the attenuation of sound from an incoherent line source by a barrier is always a few $\text{dB}$ smaller than the attenuation for a point source which is on the closest position of the line as curve A in Fig.15. It seems to be more useful in practice to estimate for the traffic noise barrier.

![Diagram](image URL)

**Fig.15** Sound attenuation by a thin screen vs. Fresnel's number $N$ for an incoherent line source.
2.8 Field Performance of the Barrier

Theoretical and experimental performance of the barrier described above have been made under anechoic conditions, without any reflecting surface except the barrier itself. However, most of the barriers are built on the ground, and we must consider the effect of reflection from the ground as well as diffraction by the barrier. Experimental studies on the field performance of barriers have been carried out by many investigators (47,50-55). The results show the difficulty of the problem. Let us consider, therefore, a rather simple and fundamental condition.

In Fig.16, the paths of propagation from the source S to the reception point P can be obtained by drawing the images S' and P' respectively, assuming the specular reflection of the ground.

![Diagram](image)

Fig.16 Section of a plane screen on the ground.

Let $\phi_1$, $\phi_2$, $\phi_3$ and $\phi_4$ denote the values of the complex amplitudes of the sound waves via four paths S-O-P, S'-O-P, S-O-P', and S'-O-P', respectively, and let $\phi$ be the superposed value. Then the amplitude of sound wave at the receiving point P is obtained by

$$\phi = \phi_1 + Q_1\phi_2 + Q_2\phi_3 + Q_1Q_2\phi_4 \quad (13)$$

where, $Q_1$ and $Q_2$ must be the spherical reflection coefficients of the ground of each side of the barrier, as discussed in Eqn.(9).

Even though the coefficients are assumed to be unity, totally reflecting, the interference pattern calculated by Eqn.(13) shows a complicated curve. Compared with the experiments using both pure tones and 1/3 oct. band noise in near field up to 16 m from the barrier on the ground, the calculation of Eqn.(13) was not considered suitable (25). An approximation has been provided to estimate the noise reduction by the barrier, that is, the effect of ground reflection is calculated by applying the same method for obtaining $\phi_3$ as $\phi_1$ by the chart of Fig.4, assuming $Q_2=1$. In order to neglect the contributions of $\phi_2$ and $\phi_4$ for simplification, only the sound levels just above the top of the barrier for the reference level, and the inverse square law, based on the distances of the barrier and receiving point from the source, are taken into account (25)(27).

Scholes et al. (50) performed several measurements which were made over the frequency range 125 Hz - 4 kHz, over a range of receiver dis-
stances up to 120 m, receiver heights up to 12 m, and barrier heights from 1.8 to 4.9 m, and for two source distances, 25 m and 10 m, with 0.7 m above the ground. Although the presence of an absorbing ground complicates the results, the evidence of this experiment is generally that the results under zero wind conditions, show good agreement with calculated values predicted by the method mentioned above neglecting the ground reflection.

On the other hand, Jonasson (20) showed that there could be a considerable discrepancy between theory and practice when the finite acoustic impedance of the ground was not taken into account, and also that the insertion loss of the barrier is often small and even negative, though the total sound reduction is considerable by the same barrier. This is due to the fact that a barrier generally decreases the original attenuation of the ground. The ground attenuation was treated by the theories of Ingard (56) and Rudnick (57).

Naturally, the difference between the insertion loss and the total sound reduction can be quite substantial. In order to predict the effect of a screen on noise reduction, before it has been built in the field, the insertion loss of the screen should be known. There are many difficulties, however, in obtaining the complete data for a rigorous calculation, especially on the acoustic properties of the ground and its seasonal variations. In addition, when a more exact solution is treated a more detailed effect of the meteorological conditions of the atmosphere also must be taken into account. The presence of a sound velocity gradient, most likely due to wind, has a large effect on barrier performance. These velocity gradient effects are frequency dependent, as well as are the effects of the ground (50). Unfortunately, the calculation method and sufficient data to predict this effect are not yet available, though the fundamental studies have been done by several authors (5-15).

From the practical point of view, the insertion loss is not always necessary to predict the effect of a screen, but the sound level at any point in the shadow zone of the screen must be known. An approximate method to accomplish this aim, with only a simple chart without using a computer, have been proposed (25-27), and its usefulness has been ascertained (50-53). Now, we are going to find how to predict the effects of the ground attenuation and of the other factors in the following chapter.

3. MEASUREMENT OF ENVIRONMENTAL SOUND PROPAGATION

It is fundamental to know the characteristics of sound propagation from a sound source to a receiver to execute environmental sound control. However, it requires an enormous amount of work to measure the sound propagation through the atmosphere with many sound level meters, because the sound level is not stationary in space and time (8). In addition, it is very difficult to know the acoustic properties of the boundary of field (58) and the satisfactory meteorological data, in order to calculate the excess attenuation theoretically.
**Fig. 17** Geometry for diffraction

**Fig. 18** Direct pulse.

**Fig. 19** Relative sound pressure vs. frequency, by the Fourier-Transform of the direct pulse.

**Fig. 20** Amplitude of diffraction transfer function by a barrier.

Assuming that the variables affecting the sound propagation are a linear operator, all acoustical phenomena, such as diffraction, reflection, refraction, and scattering etc. can be formulated in a linear system approach to the environmental sound field. If the impulse response of a linear system is given, it is possible to find by analysis all the properties of the system, whether in the time domain or frequency domain. Our interest is to examine the applicability of this well known principle with the aid of a computer.

### 3.1. Theory

Let the input signal of a single pulse to an acoustic system be \( X_i(t) \) with a Fourier Transform of \( C_i(\omega) \), and the output signal of the pulse be \( X_o(t) \) with a Transform of \( C_o(\omega) \). The sound transfer function \( H(\omega) \), corresponding to the excess attenuation of the system, is obtained by

\[
H(\omega) = K \cdot C_o(\omega)/C_i(\omega)
\]

where \( K \) is a factor cancelling the wave divergence.

\( H(\omega) \) is a function of the position of source and receiver \((\mathbf{r}_i, \mathbf{r}_o)\), corresponding to the Green's function of the sound field. In addition, any acoustical phenomena may be included, even time varying factors such as turbulence, wind and temperature gradients, and the function is written in the form \( H(\omega; \mathbf{r}_i, \mathbf{r}_o, t) \).

In order to obtain the impulse response, the correlation technique is utilized for two fixed points, one for source input and the other for received signal pick-up, using the continuous random noise signal (59-61). However, this solution requires considerable time.

We tried to use a single pulse signal to measure the function in a very short time interval \( \Delta t \), so that it may include the time varying phenomena, assuming the system is stationary for the interval \( \Delta t \).
3.2. Single Shot Pulse used in the Field Measurements

Fig.17 shows the situation of measuring the sound attenuation by a long screen on the ground. The input signal, by a pistol shot for a starting signal in athletic games, recorded at the point R is shown in Fig.18, which is displayed by an X-Y recorder after A-D conversion with a sampling rate of 20 μsec. The linearity of this signal has been satisfied outside the region of a few meters from the source (62,63). The Fourier-Transform of the waveform, in the interval of 1.28 msec. in Fig.18, is the sound pressure in the frequency domain as shown in Fig.19. From this figure, it is useful in the frequency range of 300~3000Hz, since the curve shows many dips above 3000Hz.

3.3. Measurement of Diffraction by the Screen

The sound transfer function of diffraction by the screen, made of plastic plates with thickness of 5 mm, is measured at the point P in Fig.17. The absolute value of result D(ω) is shown in Fig.20. The small circles show the values of predicted attenuation obtained from Fig.4, correcting the transmission through the screen by the TL values of the plate.

3.4. Measurement of Reflection from the Ground and the Rice Field

The measured reflection transfer function R(ω) of the ground, which consisted of weathered granite, and also the function of the rice field are shown in Fig.21.

3.5. Measurement of the Effect of Wind

The transfer function of the atmosphere affected by a wind with (continuous) velocity of 4~5 m/sec was measured at a distance of 10 m in the directions of upwind, sideward and downwind. The result shows the difference of about 3 dB between upwind and downwind as shown in Fig.22.

These results of the experiments seem to show the applicability of this method for the measurement of sound propagation.
4. DISCUSSION

In the first half of this paper, several simple charts are given for the prediction of noise reduction of the various types of barriers. We can immediately obtain the approximate values of sound level at any point in the shadow zone of the barriers, without the aid of a computer. It is essential for environmental sound problems that anyone, not only an acoustician, can perform the calculations easily because the problems surround everyone all over the world.

On the other hand, with the aid of a computer, a method involving the measurement of the transfer function of an acoustic system has been treated in the latter half. The method, which uses an impulse signal, is also effective in time variant systems. At present, an important problem to be solved is to get a suitable impulse source, which does not exceed the linearity region of the system, covers sufficient frequency range and can be repeated with stability.

Based on the transfer function, the sound propagation from sources to the ears of a person in any environment, either in the open or closed space, may be approximately simulated by the computer (64). This simulation enables us to have systematic subjective tests of environmental acoustic fields (65,66), so that, we might be able to determine an optimum acoustic system or a boundary condition for any acoustic environment, both for noise and for room acoustics.

We cannot forget the fact, however, that the excess attenuation of noise by diffraction, absorption and scattering phenomena is rather limited. For example, any noise screen cannot shield a person from an aircraft noise flying over him. In addition, a high screen for noise reduction can break other environmental factors, such as sunshine, daylight, breeze and a fine view. Naturally no one wants this. Therefore, it is quite essential to prevent the radiation from noise sources, under the present conditions of growing noise levels day by day.

These daily noises affect not only adults but also the embryo. Recently, the different reactions of babies to noise after birth, according to their mothers' length of stay during pregnancy in a noisy area (67) has been reported, also, the birth-rate of low-birth-weight babies increased as the noise level increased (68). This fact suggests that the goal of all "Environmental Control" should give consideration to these effects on future generations.

5. ACKNOWLEDGEMENT

The author would like to thank Mr Y. Ando, Mr K. Fujiwara and our colleagues in the Laboratory of Kobe University for valuable discussions and assistance in the preparation of this paper.

REFERENCES

(1) L.L.Beranek (ed); Noise Reduction (1960 McGraw-Hill) 185-205
(2) L.L.Beranek (ed); Noise and Vibration Control (1971 McGraw-Hill) 164-193
(3) E.J.Rathe; J.Sound Vib. (1969) 10 472-479
(4) Z.Maekawa; Appl. Acoust. (1970) 3 225-238. "Noise Reduction by Distance from Sources of Various Shapes"
(7) T.M. Georges et al.; J.A.S.A. (1972) 52 1397
(8) F. Ingerslev et al.; 6th I.C.A. (1968) T-3-6
(9) A.R. Kriebel; J.A.S.A. (1972) 51 19
(10) M.J. Rudd; J.Sound Vib. (1973) 26 551
(11) D.C. Henley et al.; J.A.S.A. (1973) 54 437
(12) C.J. Moore; J.Sound Vib. (1971) 16 269
(13) M.E. Delany; J.Sound Vib. (1971) 16 315
(15) D. Aulor; J.A.S.A. (1971) 51 197
(16) W.E. Scholes et al.; Appl. Acoust. (1972) 5 145
(18) A. Sommerfeld; Math. Ann. (1896) 47 317
(19) S.W. Redfearn; Phil. Mag. 7th (1940) 30 223
(20) H.G. Jonasson; J.Sound Vib. (1972) 22 113
(21) H.M. Macdonald; Proc. London Math. Soc. (1915) 14 410
(23) M. Rettinger; Noise Control (1957) 3 No.5, 50
(24) M. Born & E. Wolf; Principle of Optics (1959 Pergamon)
(26) Z. Maekawa; 5th I.C.A. (1965) T 13
(27) Z. Maekawa; Appl. Acoust. (1968) 1 157-173
(28) U. Kurze; DAGA (1973) 291-294
(30) R.B. Tate; J.A.S.A. (1973) 53 1317-1319
(33) C. Millington et al.; IEE Mono. No.508E (1962) 430
(36) K. Fujiwara, Y. Ando and Z. Maekawa; Acustica (1973) 28 341-347, "Attenuation of a Spherical Sound Wave Diffraction by a Thick Plate"
(37) D.S. Jones; Proc. Roy. Soc. (1953) (A) 217 153
(38) B. Noble; Meth. Bas. Wiener-Hopf Tech.. (1958 Pergamon)
(39) K. Fujiwara, Y. Ando and Z. Maekawa; to be presented at this 8th I.C.A. (1974) "Effect of Thickness of Barriers on Noise Reduction"
(42) A.A. Tuzhilin; Sov. Phys. Acoust. (1963) 9 168
(44) P. Ambaud et al.; Acustica (1972) 27 291
(47) H.G. Jonasson; J.Sound Vib. (1972) 25 577
(48) K. Fujiwara et al.; (to be published)
(49) M. Yamashita et al.; Appl. Acoust. (1973) 6 233-242
(50) W. E. Scholes et al.; J. Sound Vib. (1971) 16 627-642
(51) W. E. Scholes et al.; Appl. Acoust. (1972) 5 205-222
(52) M. Rettinger; (private communication)
(54) M. Ringheim; Norway Inst. Tech. Rep. LBA 335 (1971)
(55) H. C. Jonasson; J. Sound Vib. (1973) 30 299-304
(56) U. Ingard; J. A. S. A. (1951) 23 329-335
(57) I. Rudnik; J. A. S. A. (1947) 19 348-356
(59) P. D. Schomer; J. A. S. A. (1972) 51 1127-1141
(60) R. J. Donato; J. A. S. A. (1973) 53 589-592
(61) B. Escudie et al.; Inter-Noise 73, G 24 X 6
(62) J. Igarashi; J. A. S. Japan (1956) 12 55-62
(63) A. Kosugi et al.; J. A. S. Japan (1963) 19 45-54
(64) M. R. Schroeder et al.; IEEE Intern. Conv. Record (1963) 150
(65) Y. Ando, S. Shidara and Z. Maekawa; J. A. S. Japan (1973) 29 151-159
"Some Basic Studies on the Acoustic Design of Room by Computer".
(66) Y. Ando, S. Shidara and Z. Maekawa; to be presented at this 8th
I.C.A. (1974)."Simulation of Sound Propagation with Boundary and
Subjective Test".
Intense Noise during Fetal Life upon Postnatal Adaptability (Statistical
Study of the Reactions of Babies to Aircraft Noise)".
Studies on the Effects of Intense Noise during Human Fetal
Life".
(69) K. D. Kryter; The Effect of Noise on Man (1970 Academic P.)
(70) A. D. Pierce; ASA-NBS Sympo. (Sep. 1972), submitted to the J. A. S. A.
When I accepted the invitation for this presentation back in 1972, I sincerely hoped that substantial progress in hearing research could be achieved between then and this Congress to warrant inclusion of hearing among the plenary topics. My hope has been realised beyond my expectation: While only two years ago we were still confronted with major puzzles and paradoxes of the most basic nature - even at the very periphery of the auditory channel - several large pieces of this puzzle are finally falling into place - owing to the very substantial experimental work during the last few years and recent theoretical interpretations based upon the experimental data. The beauty of this development is that very few and reassuringly simple and realistic assumptions can explain a wide range of experimental observations.

In this paper I will concentrate on 3 prominent auditory phenomena and their possible explanation:

1. The cubic difference tones (CDT's), first described by Tartini in 1740; particularly their "extraordinary" amplitude behavior.

2. The remarkable nonlinearities observed in neural discharge patterns for double-click excitations first reported by Gobelick and Pfeiffer. A possibly related nonlinearity was recently found by Hall and Lummis in psychoacoustical double-click masked-threshold experiments.

3. Some basic properties of the neural transduction process involving, cooperatively, the motions of the basilar and tectorial membranes, the stimulation of the hair cell and the firing of the attached primary acoustic nerve fiber.

NONLINEAR LOSSES AND THEIR POSSIBLE ORIGIN

It appears that the essence of the CDT-phenomena and the double-click nonlinearities can be explained by one simple addition to the Békésy-Zwislocki-Peterson-Bogert model of the basilar membrane, namely that the motion of the basilar membrane is associated with nonlinear losses such that the resistance $R$ (see Fig. 1) representing these losses is of the following form:
EIGHTH INTERNATIONAL CONGRESS
ON ACOUSTICS, LONDON 1974

\[ R = R_0 (1 + \beta I^2) \]  \hspace{1cm} (1)

where \( I \) is the current through \( R \) and represents the velocity of the basilar membrane (most likely relative to the tectorial membrane). \( \beta \) is a "constant" which may depend upon the place along the basilar membrane and the frequency of the stimulus.

\[ \begin{align*}
V_i & \quad M_i \\
L_1 & \quad R_1 \\
Q_1 & \quad C_1
\end{align*} \]

**BASE**

\[ \begin{align*}
V_i & \quad M_i \\
J_i & \quad I_i \\
L_1 & \quad R_1 \\
Q_i & \quad C_i
\end{align*} \]

**...**

\[ \begin{align*}
V_{174} & \quad M_{174} \\
L_{174} & \quad R_{174} \\
Q_{174} & \quad C_{174}
\end{align*} \]

**APEX**

**\( \Delta X \)**

Fig. 1 Electrical analog model of the basilar membrane.

Nonlinear losses of form given by eqn. (1) are also consistent with, in fact are strongly suggested – as already noted by Kim (1972) and Kim, Molnar and Pfeiffer (1973) – by the nonlinearities observed near resonance by Rhode (1971) in his Mössbauer measurements on squirrel monkey.

Even with such a simply formed nonlinearity as specified in eqn. (1), one may ask "what causes the nonlinearity?" CDT's are observed at very low sensation levels where the basilar membrane motion is of the order of 1 Ångström (10^{-8} cm). Nonlinear viscosity and a host of other nonlinear effects are simply not credible at such low amplitudes. (For comparison, perhaps not accidentally, the diameter of the hydrogen atom is also about 1 Å.)

On the other hand, observable auditory effects do occur, by definition, at amplitudes corresponding to the threshold of hearing or less than 1/10 Å. The most likely physical process that could account for the extraordinary sensitivity of our auditory receptor organ is a physico-chemical system showing cooperative behavior and existing near a phase-change boundary. Biological membranes of neurons and sense organs (including the hair cell) are believed to be such systems. Reversible phase change of model lipid membranes have been elicited by changes in the external physical parameters, including temperature, electrical fields, hydrostatic pressure and surface tension.
(Träuble, 1971; Kaatze, 1973, personal communications). Near the phase-change boundary, these membranes switch states (from "solid crystal" i.e., highly ordered, to "liquid crystal", i.e., less highly ordered and consequently permeable for certain ions) for exceedingly small perturbations in any of the controlling physical parameters. Furthermore, these phase changes, although reversible, are known to exhibit hysteresis and nonlinear losses (M. Eigen, 1973, personal communication).

The inevitable conclusion from these observations is that the moment we hear any sound, even at the threshold of hearing, nonlinear loss processes are of necessity operative at the juncture between hydro-mechanical vibrations and the electrochemical processes which cause afferent fibers to fire.

To the extent that these nonlinear losses are invariant against the direction of motion and analytic at zero velocity, eqn. (1) would be a proper first approximation. (It is conceivable that, at higher amplitudes, terms proportional to $I^2$ etc. may become significant.)

If asymmetrical loss effects occur, which is not unreasonable in view of the one-sided ("rectifying") response of the hair cell, eqn. (1) might be supplemented by a first-order term in $I$:

$$R = R_0 (1 + \alpha I + \beta I^2) \ .$$

(2)

The $\alpha I$-term would lead to inner-ear quadratic distortions which have been, in fact, observed by J. L. Hall (1972). (For two primary tones with frequencies $f_2 = 3f_1/2$, the lower CDT $2f_1-f_2$ and the quadratic difference tone (QDT) $f_1-f_1$ occur at the same frequency ($f_1/2$) but their dependence on the phase $\phi_2$ of the upper primary has opposite sign for CDT and QDT. Thus, when $\phi_2$ is changed from 0 to $\pi$ the distortion component heard at $f_1/2$ should go through a complete "beat" cycle with a pronounced amplitude "null" (and phase reversal) if CDT and QDT have comparable amplitudes - a fact observed in all detail by Hall.)

The velocity amplitude $I$, at which this complete nulling occurs (for equal-amplitude primaries) is a measure of the relative magnitude of the nonlinear coefficients $\alpha$ and $\beta$:

$$|\alpha| \leq \beta I \ ,$$

(3)

(provided $\beta I^2 < R_0$, otherwise the relation between $\alpha$ and $\beta$ might be more complicated.)
Hall's experiments also yield information on the
algebraic sign of $a$ and this in turn might be an important
clue in the unravelling of the detailed molecular process
at the transduction interface. (Thus, it may be reason-
able to assume that $a$ is negative if a "downward" velocity
($I<0$) of the basilar membrane triggers more nerve pulses
(and presumably engenders greater losses) than an upward
motion.) Finally, $a$ could be complex if the asymmetric
losses are displacement dependent. A careful analysis of
the inner-ear quadratic distortion is thus of prime
importance and may help clarify one of the most elusive
extant problems in auditory theory, namely whether basi-
lar membrane displacement or a (temporary or spatial)
derivative of displacement increases the firing probabil-
ity of afferent nerve fibers attached to the inner and
outer hair cells, respectively.

THE "EXTRAORDINARY" AMPLITUDE BEHAVIOR OF THE CDT

Although cubic nonlinearities probably occur along the
entire length of the basilar membrane, most of the dis-
tortion observed in CDT's must originate at places where
both primary amplitudes are large, i.e., near the character-
istic place of the upper primary ($f_2$). For $f_1 = f_2$,
the shunt impedance at that place is approximately equal
to $R$ and the membrane velocity, represented by $I$,
is given by:

$$I = V/R;$$  \hspace{1cm} (4)

where $V$ represents the pressure difference between the
scalae. Inserting eqn. (4) into eqn. (1) and solving for
$R$ yields

$$R(v) = \left(\frac{R_o}{3}\right) \left\{ 1 + \left[1 + v^2 + (2v^2 + v^4)^{\frac{1}{2}} \right] \right\}$$

$$+ \left[1 + v^2 - (2v^2 + v^4)^{\frac{1}{2}} \right] \right\}$$

(5)

where $v$ is a dimensionless voltage (pressure):

$$v = \left(\frac{278}{V/2 R_o}\right)^{\frac{1}{2}}$$

(5a)

For $v = 0$, $R = R_o$, and for $v \to \infty$, $R \to 0.67 R_o^{\frac{1}{2}}$

The nonlinear "input-output" (pressure-velocity)
characteristic at the characteristic place becomes:

$$I(V) = V/R(v)$$

(6)

with $R(v)$ from eqn. (5). Whereas for small amplitudes
(v << 1) I is linear in V: I = V/R_0, for large amplitudes (v >> 1) one has:

\[ I = (R_0 \beta)^{-\frac{1}{3}} V^{\frac{2}{3}}, \] (7)

i.e., in the large amplitude limit, for a 10-dB increase in the pressure (V), the velocity (I) of the membrane increases by only 3.3 dB. At finite amplitudes, correspondingly smaller compression factors obtain — in good agreement with Rhode's (1971) Mössbauer measurements. Rhode (Fig. 6) found a response increment of approximately 4 dB for a 10-dB stimulus increase from 80 to 90 dB re 0.0002 dyn/cm² near the characteristic frequency (about 7 KHz).

In fact, the entire input-output characteristic given by eqns. (5) and (6), when plotted, looks very much like Rhode's measured basilar membrane response at the characteristic frequency (Rhode, 1971, Fig. 7). The asymptotic slopes for small and large stimulus amplitudes are about 1.0 dB/dB and 0.3 dB/dB, respectively, in excellent agreement with the theoretically predicted behavior.

By matching experimental and theoretical curves a value for the nonlinearity constant \( \beta \) for squirrel monkey near 7 KHz can be found:

\[ \beta^{-\frac{1}{3}} = 0.02 \text{ cm/sec}. \]

At this velocity, corresponding to a displacement of 4.6 \( \times \) 10^{-7} cm, the nonlinear losses begin to exceed the linear (small-amplitude) losses.

Since CDT's are most likely generated in a very narrow region near the \( f_2 \)-place, their amplitude behavior should follow directly and solely from the nonlinear "input-output" characteristic specified by eqns. (5) and (6) without additional assumptions. In fact, preliminary analysis by the author indicates that, setting

\[ V = A_1 \cos \omega_1 t + A_2 \cos \omega_2 t, \] (8)

the dependence of the (2\( \omega_1 \) - \( \omega_2 \))-component of I on \( A_2 \) (the upper-harmonic amplitude), for example, is strongly nonmonotonic: first increasing with \( A_2 \) and then, for \( A_2 \) const. \( A_1 \), decreasing with increasing \( A_2 \) — precisely as first reported by Zwicker (1968).

Zwicker's "quasi-regular" amplitude behavior of the CDT for either \( A_1 \ll A_2 \) or \( A_2 \ll A_1 \) can also be deduced from eqns. (5) and (6).
Finally, when both $A_1$ and $A_2$ are varied proportionally together, the tendency of the CDT amplitude to "saturate" (Zwicker, 1955, Goldstein, 1967) at high primary levels likewise seems to be a direct consequence of the nonlinearity in the losses (eqns. 1 or 2). In order to verify this - theoretically somewhat less tractable - phenomenon, J. L. Hall (1974, to be published) has digitally simulated Schroeder's (1973) model of the basilar membrane with $R$ as in eqn. (1). In this manner Hall also found the proper phase behavior of the CDT with primary level (Goldstein, 1967) and the significant difference between CDT-level and cancellation-tone level at high primary levels (Smoorenburg, 1972).

Thus, while many important details remain to be filled in, the overall picture of the inner-ear mechanical nonlinearities finally becomes visible through the mist which has, for so long, enveloped it. As we progress along this path, we can expect to gain significant further clues to the highly elusive molecular processes which govern the mechanical-to-neural transduction in the cochlea.

THE FREQUENCY-DEPENDENCE OF THE CDT

Once the cubic distortion products are generated (near the $f_2$-place), their basic frequency behavior follows directly from linear basilar membrane mechanics. Thus, the steep decrease of the CDT-level with increasing $f_2/f_1$ is essentially a consequence of the lowering of the $f_2$-level at the $f_2$-place as $f_1$ is decreased (or $f_2$ increased). The rapid phase shift of the CDT with increasing $f_2/f_1$ (Goldstein, 1967) follows likewise in large measure from the linear response of the basilar membrane.

Somewhat puzzling in this context is the nonmonotonic amplitude of the CDT with $f_2/f_1$ found by Weber (1970) and Smoorenburg (1972) for limited amplitude ranges $A_1$, $A_2$. Weber (1972, private communication) has speculated that these "dips" in the CDT-amplitude may be due to standing waves of frequency $2f_1-f_2$ on the basilar membrane. Weber's conjecture has since been confirmed, in astonishing detail, by computer model experiments (Hall, 1974, to be published). In these simulations, a wave of frequency $2f_1-f_2$ is found to travel in the basal direction and, after reflection at the base, to act much like an external cancellation tone on the $(2f_1-f_2)$-wave travelling in the apical direction. The most pronounced minimum in the model is found for $f_2/f_1 = 1.31$ in excellent agreement with Weber's (1973) extensive data.

* * * *
DOUBLE CLICK NONLINEARITIES

A remarkable nonlinearity was observed by Gobelick and Pfeiffer (1969) in recordings from primary auditory fibers of cat. Clicks applied to the cat's ear result in "compound" post-stimulus-time (PST) histograms showing positive and negative peaks spaced approximately 1/2f₀ apart, where f₀ is the characteristic frequency of the observed neuron. By presenting a second pulse of proper amplitude with a delay of Δt = 1/2f₀ any PST-peak (but the first) can be "cancelled," i.e., neural activity in the time interval occupied by the peak can be reduced to spontaneous firing. A convenient interpretation is that the n-th peak in the mechanical response to the first click coincides with the (n-1)st peak of the mechanical response to the second click. Since two consecutive peaks have opposite sign, they can cancel each other for the proper click-amplitude ratio. For linear mechanics, the click-amplitude ratio for cancellation would in fact equal the amplitude ratio of peaks "n" and "n-1" which could thus be measured. Unfortunately, mechanical responses determined in this manner are thoroughly unbelievable. For example, calculated peak amplitudes would keep rising up to the 10-th peak - in stark contrast to overwhelming evidence from a host of experiments. Even more serious, peak-amplitude ratios calculated from different click spacings give grotesquely contradictory results (discrepancies exceeding 30 dB for a basic experimental accuracy of 0.5 dB).

Since these discrepancies should not appear in a linear system, we are safe to call them "nonlinearities." But where do they "reside"? ("hide" would be a better word.) It couldn't be any nonlinearity at the place of observation or in the mechanical-to-neural transduction process. The beauty of the Gobelick-Pfeiffer (G+P) peak-cancellation experiment is that it is independent of the nonlinearity of the mechanical-to-neural transduction process and of any (memoryless) mechanical nonlinearity at the place which stimulates the neuron. Nonlinear effects at places along the basilar membrane other than the observed place, however, cannot be excluded from responsibility. In fact, the author (1974, to be published) was able to show that all of the G+P double-click nonlinearities can be accounted for by a single, simple assumption: That in each double-click experiment the mechanical response to the second click is attenuated relative to that of the first click. This hypothesized extra attenuation depends upon the amplitude of the first click, decreasing with increasing click amplitudes for high signal levels.

But where could this hypothetical extra attenuation
come from? Not from the $f_c$-place, as already pointed out. More basally? Perhaps. For a click spacing of $\Delta t=1/2f_c$ the two mechanical click responses overlap in phase at the $2f_c$-place. Thus, any nonlinearity losses at the $2f_c$-place would certainly give different attenuations for a single click and two consecutive clicks with in-phase overlapping mechanical responses.

Specifically, the power lost into the side-branch resistance $R$ (cf. Fig. 1) would be (with eqn.(1))

$$V \cdot I = R_0 I^2 + R_0 \beta I^6,$$

where $R_0 \beta I^6$ represents the "nonlinear" losses. Obviously, the greater $I$, the greater these nonlinear losses. This acts just like the nonlinear progression in our income tax law - to the detriment of the second-click response which is completely overlapped (in phase!) by the response to the first click, and thus suffers an extra attenuation, while the front portion of the first-click response passes the 2 $f_c$-place without being overlapped and consequently without the extra attenuation. (Of course, other regions of the basilar membrane are also involved, some possibly with a reversed effect.)

Why does the double-click nonlinearity diminish for large click amplitudes (assuming my interpretation of the G+P "amplitude nonlinearity" is correct)? At high click amplitudes, when the shunt resistance according to eqn.(1) becomes larger than the characteristic impedance $Z$ of the basilar membrane "transmission line," the losses would approach $V^2/R$ and thus tend to zero as click level is further increased. By the same token, the nonlinear losses, giving rise to the extra attenuation of the second click response would vanish. Thus, we have the remarkable situation in which a perfectly regular nonlinearity gives rise to a nonlinear effect that vanishes for large amplitudes. Of course, the cause of this paradox is nothing but the familiar impedance matching: for small amplitudes the losses increase as $R$ and for large amplitudes they decrease as $1/R$.

If this interpretation is correct, as I believe it to be, then the G+P click ratios tend toward the peak-amplitude ratios of the basilar membrane impulse response for high sound levels.

The "characteristic" sound pressure level at which this decrease in double-click nonlinearity begins can be calculated from an estimate of $\beta$ in terms of the membrane velocity at threshold $I_0$. Psychoacoustic experiments tell us that the CDT begins to "saturate" (intersection of asymptotes) at about 60 dB SPL.
Thus,
\[
\beta = 10^{-6} I_0^{-2} \quad (11)
\]

Using relationships from the author's (1973) paper to express the characteristic impedance of the membrane, one obtains the characteristic sound pressure level in question as \(10 \log_{10} K + 60 \) dB SPL where \(K\) is the dimensionless constant characterizing the author's integrable membrane model. With \(K = 20\), the "characteristic" level becomes 73 dB SPL in excellent agreement with G+P (1969, Figs. 10,11) who found "characteristic" levels between 70 and 80 dB SPL.

Simulation experiments are being planned to test these admittedly somewhat rash proposals. In addition, the following two experiments (one physiological and one psychoacoustic) would constitute a strong test of my hypothesis:

1. In the psychoacoustic double-click masked-threshold experiment (Hall and Lumsds, 1973) using masking noise with a stop-band at 2f_0 (as well as around f_0) should enhance the observed click amplitude asymmetry. This experiment has already been started in collaboration with J. L. Hall.

2. In the G+P double-click peak-cancellation experiment, the addition of a masking noise around 2f_0 should decrease the observed nonlinearities.

** ** **

A MODEL OF THE NEURAL TRANSDUCTION PROCESS

Among the many outstanding properties of neural discharge patterns recorded intracellularly from primary acoustic fibers the following is particularly noteworthy and must be accounted for by any serious model of mechanical-to-neural transduction process: At amplitudes exceeding about 40 dB SPL, period-histograms are relatively independent of signal level while appearing to be reasonable replicas of the ("half-wave rectified") mechanical signal. In other words, there is a gain-control mechanism inherent in the transduction process with relatively little non-linear distortion (other than the well-known half-wave rectification). The time-constant of the gain control mechanism can be obtained from experiments with pulsed stimuli and appears to be of the order of 20 msec. Needless to say, no automatic gain control mechanism in any conventional sense has been found in physiological preparations of the hair cell and the structures surrounding it.
How then can we explain this remarkable adaptation which, in combination with the middle-ear reflex and the (amplitude-compressing) mechanical nonlinearity, enables the ear to cover a range of intensities exceeding a ratio of $10^2$ between the thresholds of hearing and feeling (pain)?

In view of the wide range and complexity of the observed phenomena (period histograms, PST-histograms, interval histograms for pure tones, composite periodic signals, tone bursts, clicks and noise) it may seem preposterous to expect a simple model of the transduction process to account for the bulk of observations in the published literature. Yet this is precisely what the "urn" model (Schroeder and Hall, 1973) specified by the following 3 rules can do:

1. "Quanta" of an (electrochemical) agent are generated in the hair cell at a fixed average rate, $r$ quanta/sec.

2. Quanta disappear and cause an attached (afferent) nerve fiber to fire with a probability per unit time proportional to their number, $n(t)$, and a "permeability" function, $p(t)$, related to the input stimulus.

3. In addition, quanta disappear independently of stimulation and without causing nerve firings with a probability per unit time equal to $g\cdot n(t)$, where $g$ is a constant with dimension sec$^{-1}$.

For the permeability function we assume, somewhat arbitrarily, the following "soft half-wave rectifier" law:

$$p(t) = \left( \frac{p_o}{2} \right) \left\{ S(t) + \left[ S^2(t) + 4 \right]^{1/2} \right\} .$$

where $S(t)$ is proportional to the mechanical stimulus.

The firing probability per unit time $f(t)$ is given by (cf. Rule 2)

$$f(t) = n(t) \cdot p(t) .$$

For sufficiently high fundamental frequency of a stimulating periodic signal, the number of quanta will remain relatively constant during the fundamental period. Thus, the firing probability per unit time, eqn. (13), can be approximated by

$$f(t) \approx n p(t)$$

Hence, the firing probability is approximately propor-
tional to \( p(t) \). For small signal amplitudes, \( p(t) - P_0 \) itself is proportional to the signal. For large signal amplitudes, \( p(t) \) is proportional to the half-wave rectified signal. For periodic signals, eqn. (14) describes the data displayed in period histograms of auditory-nerve activity.

In the steady state, as many quanta are generated on the average as disappear, i.e.,

\[
r = \bar{n}(t) \cdot p(t) + n(t)g,
\]

where \( r \) is the rate of generation of quanta (cf. Rule 1), \( n(t) \) the number of quanta, and \( g \) the probability, per quantum and unit time, of disappearance of a quantum (cf. Rule 3).

The first term on the right of eqn. (15) is the average probability, \( \bar{f} \), of nerve firing. Thus, for stationary conditions,

\[
\bar{f} = r - \bar{n}g. \tag{16}
\]

From eqn. (14), the average firing probability equals

\[
\bar{f} = \bar{n} \cdot \bar{p}. \tag{17}
\]

Thus, with eqn. (16), the average number of quanta \( \bar{n} \) is approximately equal to

\[
\bar{n} = \frac{r}{(g+\bar{p})}, \tag{18}
\]

i.e., for large signal amplitudes \( (\bar{p} \gg g) \), \( \bar{n} \) is inversely proportional to the signal amplitude. As a consequence, the firing probability becomes asymptotically independent of signal amplitude:

\[
f(t) \to rp(t)/\bar{p}. \tag{19}
\]

This "normalization" of the firing probability is one of the outstanding characteristics observed in period histograms and follows here directly from the Rules that specify the model.

Another neurophysiologically-observed property that is correctly predicted by eqn. (14) is the (for positive signal values) essentially true waveform reproduction of the firing probability—including its phase stability with signal amplitude (Rose et al., 1971, Fig. 11). This property of nerve activity is difficult to explain by threshold models (Weiss, 1966).
Finally, the initial overshoot and the final undershoot seen in PST-histograms during and after burst stimulation follow directly from the Rules: Immediately before the beginning of the burst, the number of quanta corresponds to the no-signal condition $n_o$, with an average value given by (cf. eqn.(18))

$$\bar{n}_o = r/(g+p_o).$$

(20)

The firing probabilities are accordingly high. As soon as the stimulating signal is applied, the number of quanta tends to decrease and with it the firing probabilities.

At the end of the burst, the number of quanta recovers over a few milliseconds to its no-signal condition, and with it the initially-depressed, spontaneous firing probability also grows back to the steady-stage spontaneous firing rate.

It is interesting to note the average behavior of the model can be represented by the simple RC-circuit shown in Fig. 2: a current generator with a fixed shunt conductance $g$, a shunt capacitance $C$, and a variable conductance $p(t)$. (Figures 2,3 are to be found at the end of the paper)

The source current $r$ corresponds to the rate of generation of quanta, and the current $f(t)$ through the variable conductance represents the firing probability of the nerve. The charge on the capacitor $C$ corresponds to the average number of quanta in the hair cell.

The refractory effects seen in nerve spike data can be included by making the firing probability $f(t)$ depend in part on time elapsed since the preceding neural event. Instead of eqn.(13), we have

$$f(t) = n(t) \cdot p(t) \cdot \rho(t-t_o)$$

(24)

where $\rho(t-t_o)$ is a recovery function and $t_o$ is the time at which the preceding neural event occurred. We chose $\rho(t-t_o)$ such that there was a dead-time of 1 msec followed by an exponential recovery with a time constant of 1 msec.

The response of the model (including dead-time) to a variety of mechanical stimuli has been studied by digital computer simulation. As an example, Fig. 3 shows period histograms in response to a complex input with components at 1 and 2 kHz. The phase and intensity relations are such that there are two excitatory stimulus peaks, the first three times as big as the second. The waveform is preserved in the response. At each intensity the first response peak is larger than the second, even though in the stimulus the second peak at 50 dB, for example, is
EIGHTH INTERNATIONAL CONGRESS ON ACOUSTICS, LONDON 1974

larger than the first peak was at 40 dB. Similar results have been observed for auditory-nerve fibers (Rose et al., 1971).

Among the numerous predictions of the model one is particularly straightforward: for sufficiently low-frequency sinusoidal stimulation, the maximum firing rate should be noticeably smaller than that for high frequencies. This results from quanta being lost without causing firing during the inhibitory half-cycle of the stimulus. In fact, in the low-frequency limit, the number of quanta so lost approaches one half the total number of quanta generated. Simulation showed a reduction in maximum firing rate from about 150 to 120 spikes per second at a stimulating frequency of 125 Hz. Has this low-frequency effect been observed experimentally?

** **

CONCLUSION

While the proposals made here concerning the origin of the inner-ear nonlinearities and the mechanical-to-neural transduction are subject to further confirmation and will certainly have to undergo considerable refinement (if not summary rejection) as additional experimental evidence becomes available, it is hoped that they will stimulate the sorely needed critical investigations which, in combination with detailed studies at the molecular level, are our only hope of shedding further light on one of nature's important senses: hearing.

REFERENCES


59
CURRENT = FIRING PROBABILITY

CHARGE = NUMBER OF QUANTA

Fig. 2. Equivalent electrical circuit for the average behaviour of the "urn" model, described in the text, for the neural-to-mechanical transduction process. The current source on the left simulates the creation of quanta of an electrochemical agent that causes the afferent acoustic nerve to fire. The variable conductance p(t) depends on the mechanical stimulus and may represent the variable permeability of a hair cell membrane to the electrochemical agent. (Mechanical deflection in one direction "stretches" the membrane and changes its state from nearly impermeable solid-crystalline to more permeable liquid-crystalline form thereby increasing the firing probability of the afferent nerve fiber attached to the hair cell.)
Fig. 3 Results of a digital computer simulation of the neural transduction model for a two-tone stimulus. The model combines pronounced amplitude normalization ("gain control") above 50-dB sound pressure level (the parameter in this figure) with near faithful reproduction of the excitatory half-cycle of the stimulating mechanical waveform. The model thus duplicates the remarkable "quasi-linear adaptation" behaviour seen in histograms of neural spikes in primary acoustic nerve fibers. In the model, this gain control results from the depletion property inherent in the "urn" model.
EIGHTH INTERNATIONAL CONGRESS ON ACOUSTICS, LONDON 1974

ULTRASONIC SURGERY

I.P. Goliamina  Acoustical Institute of the Academy of Science, Moscow, USSR

INTRODUCTION

Ultrasound is the most recent field of application of ultrasound in medicine. In surgical interventions irreversible changes are usually produced in the tissues and organs of the human body. According to the methods and techniques of operation two types of ultrasonic surgical application may be recognised. One is concerned with the effect of high power focused ultrasound on small-sized parts of the human body, sometimes in deeply located structures. This type is often referred to as "surgery without a knife" and the operating frequencies are usually in the MHz range. Up to the present many serious experimental investigations have been carried out using this technique and it has become routine use in the ear surgery clinic and in experimental neurosurgery. The other form of ultrasonic surgical application is based on the use of instruments vibrating at ultrasonic frequencies in the range of tens of kHz in which the devices are brought into direct contact with the affected part of the body. This work was started in the late 'sixties' and has received little mention so far in the specialised acoustical and ultrasonic literature. It should be pointed out that in this type of ultrasonic application no completely new technique is demanded of the surgeon for he is manipulating with an old form of instrument (scalpel, probe, saw, etc.) which has however acquired a new 'quality' as a result of being subjected to ultrasonic vibrations.

The use of ultrasonically vibrated instruments for the 'cutting' and 'welding' of biological tissues was first reported in 1968 when the advantage of using the vibrating scalpel (1) for cutting brain tissues was demonstrated by experiments on animals and a new method (2) was proposed for knitting broken bones (osteosynthesis) using an ultrasonic instrument. In the immediate years following rapid progress has resulted from the use of ultrasonic devices in osteosynthesis and in the cutting of bone tissues and of soft tissues in orthopaedic surgical clinics (3), (4). Similar procedures were applied successfully also in thoracic surgery (5), (6) and experiments performed on animals and in some clinics confirmed the usefulness of these techniques in neurosurgery (7), (8). Some experimental and clinical results also showed the possibilities of applying ultrasonic instruments in oto-laryngology (9)-(11). Ultrasonic phaco-emulsification, (14)-(17), and retinal detachment (18), (19), operations belong also to the field of surgical applications. A good source of references is to be found in the proceedings of the USSR conference 'Ultrasound in Physiology and Medicine' (20), where the historical development of instrumentation and experimental and clinical results in the field are reported.

63
ULTRASONIC SURGICAL INSTRUMENTS

The ultrasonic surgical instrument is essentially a rod-type resonant vibrating system, consisting of a half-wave length transducer of magnetostrictive or piezoceramic material, and a resonant wave guide in the form of a velocity transformer with the working end having a shape designed for the specific type of operation.

The operating frequency of the instruments lies in the range, often called by medical people a "low frequency ultrasound" range. Because of the small size of the working area of the instrument in comparison with the wavelength and of the small specific impedance of the acoustical load, the working parameter of the instrument is not the radiated power but the vibration amplitude of its tip. Consequently the efficiency has to be estimated by the ratio of the (tip-amplitude)$^2$ to the electrical power consumption of the transducer [21]. The value of the vibration amplitude varies in different surgical operations from 15µ to 60µ. For the calculation of the acoustical power, applied to the biological tissue, one must know the exact value of the specific mechanical resistance of the acoustical load, which is of the order $10^3$-$10^4$ dyn.sec/cm$^2$ and depends on the shape and size of the working tip, the applied force and the kind of tissue itself. As verified experimentally, the mechano-acoustical efficiency for the instruments of the best mechanical quality does not exceed 2-3% when applied to soft tissues. The instruments most widely used today are designed with compound velocity transformers, the first part of which is firmly joined to the transducer, and the second part, which includes the working tip, is screwed to the first part and thus is interchangeable, permitting the same instrument to be used for different kinds of operation. The vibrating system is put into a housing, by which the instrument is held.

Besides the common requirements for ultrasonic technological instruments - high efficiency of the transducer, low losses in the waveguide, stability of the longitudinal vibrations at high amplitudes, the vibrating system of the ultrasonic surgical instrument must satisfy some specific requirements, which are as follows: minimisation of dimensions (in particular, the cross-sectional dimensions) and of the weight; non-corrosive properties of instrument materials; possibility of sterilization; safety for the personnel; simplicity of the design; ease and reliability of monitoring; automatic maintenance of fixed parameters.

In the opinion of the author the above-mentioned requirements are best satisfied by using systems with ferrite transducers. Such transducers work at low voltages, do not corrode in chemical-active media, have low intrinsic mechanical and magnetic losses, are easily adjusted to the supply generators, with motional feedback circuits for automatic regulation of their frequency and amplitude. In the case of
heavy acoustic loading, where high power and large amplitudes are required, transducers made of metal magnetostrictive materials are advisable.

For velocity transformers titanium alloys are usually used as these alloys have low mechanical losses, do not corrode and are of relatively light mass. In some cases the first part of the velocity transformer is made of an aluminium alloy and the second is of steel. The latter provides better cutting quality of the working tip, but introduces high mechanical losses into the vibrating system. The search for materials, which are satisfactory from the acoustical point of view and at the same time meet the requirements of optimum cutting conditions, is one of the main remaining problems to be solved in the present state of development of ultrasonic surgical instruments. The front section of the velocity transformer should not as a rule have too large a transformation coefficient. The transformation coefficient of the second part, as well as its shape and the form of the working tip, are chosen according to the specific surgical operation. The design of this part as a whole should be worked out by the physicist or engineer together with the practising surgeon.

PHYSICAL BASIS OF ULTRASONIC SURGERY

One of the main applications of ultrasonic surgical instruments is the operation of cutting biological tissues. In early experiments the ultrasonic scalpel was used for cutting brain tissues but nowadays ultrasonic knives and saws are used for cutting soft tissues as well as bones. Vibrations of the knife or the saw with ultrasound frequency reduce the force of the cutting remarkably (5-10 times), which in its turn diminishes the trauma of the surgical effect. This reduction in cutting force may be explained as a consequence of the physical phenomenon, involving the reduction of friction between two surfaces sliding one upon the other when one of the surfaces is in vibration (22). (An analogy may be drawn with the effect of superposing vibrations on the motion of a cutting tool in metalwork technology, which is used practically in order to reduce power consumption and to speed up the process of cutting itself). Besides diminishing trauma, ultrasonic vibrations arouse some other useful medical effects. For example, haemostasis may be mentioned: when using an ultrasonic knife the haemorrhage is reduced in particular types of operations. This phenomenon may be explained by bactericidal action and analgesic effect, the latter being very important. The elucidation of the mechanism of this process is in the author's opinion one of the most interesting problems to be solved in this field.

The other important application of ultrasonic instruments is osteosynthesis. This process is often called "bone welding", but this name does not properly designate its mechanism. The essence of the
process is briefly as follows: for knitting bone fragments a liquid monomer, ciacrin is put into the gap; then an ultrasonic instrument is applied (its working tip having the form of a flat blade) and the monomer polymerizes in a few seconds, producing an artificial bony corn callus. In order to obtain rigid coupling a monomer filled with bone dust or bone shavings is used. In the osteosynthesis experiments optimal concentrations of bone dust, optimal sizes of bone particles, optimal processing time, etc., were established (3)–(6). The mechanism of ultrasonic "bone welding" may be explained by two effects of ultrasound. The first effect is the acceleration of the diffusion processes, which in this case facilitate and accelerate the penetration of monomer into the pores of bone tissue. Such penetration, verified by experiments with radioactive tracers (4), increase to the right degree the rigidity of the joint coupling between the callus and bone fragments. The second effect, responsible for ultrasonic osteosynthesis, is the influence of ultrasound on physico-chemical processes in the monomer, provoking its rapid polymerisation. Experiments, described in (23) proved, that the "artificial callus" does not prevent processes of natural bone regeneration and is dissolved in the passage of time. Its important function is the fixation of bone fragments in the first stage of the regeneration process, which avoids in some cases the immobilisation of the broken extremity.

The experiments showed, that when performing the ultrasonic cutting or "welding" processes at the recommended parameter rates (amplitude of vibration, time of application of the instrument, the applied force) the temperature of tissues in the immediate proximity of the instrument does not exceed 60–70°C and at a distance of about 10–20 mm the temperature changes are negligible (4), (6), (7). Numerous histological and physiological experiments performed immediately after applying ultrasonic instrument as well as after passages of time, proved that the ultrasonic methods of cutting living tissues and of the osteosynthesis are harmless for the human organism.

As to the mechanism of other types of ultrasonic surgical operations, it may be mentioned that the phacoemulsification in ophthalmology is based on the dispersive effect of ultrasound. In the operation of retinal detachment in ophthalmology and in some operations in otolaryngology (such as removing of special kinds of new growths) the thermal effect of ultrasound and partly its cavitation and mechanical displacement effects are used.

APPLICATION OF ULTRASONIC SURGICAL INSTRUMENTS IN DIFFERENT BRANCHES OF MEDICINE

Cutting of living tissues and bone knitting by means of ultrasonic instruments have already found a relatively wide application in orthopaedic surgery (3), (4). For these purposes instruments with magnetostrictive transducers made of nickel or permendur are used. Power
consumption of the instrument itself is about 100 watts and that of
the supply generator is about 200-400 watts. As a rule, generators
have no motional feedback systems, and a technician must adjust work-
ing parameters during the operation according to instructions of the
surgeon. Working frequencies in this field of application are usually
25-28 kHz. For different kinds of orthopaedic surgical operations
many modifications of the second sections of velocity transformers
were developed. Working ends of these interchangeable parts of instru-
ments may have the form of a knife, of a saw, or of a spade with
different dimensions. In some cases velocity transformers may have
curved axes. Surgeons in their turn developed special kinds of opera-
tions, performed by means of ultrasonic instruments with the purpose
to correct some congenital or incidentally acquired bone deformities,
to extract a sore part of a bone or to knit bone fragments (in some
cases by method of homotransplantation). Application of ultrasonic
instruments for cutting soft tissues in this branch of medicine is
considered most useful for the purposes of plastic surgery, for
instance — the cicatricotomy, i.e. — for extractions of scars. Some
surgeons have commented upon the peculiar ease of separating different
types of tissue by means of a sharp or of a blunt ultrasonic blade.

Active work is also being performed in the field of thoracic
surgery (5), (6). Instruments, used in this case are like those
utilised in orthopaedic surgery. Ultrasonic cutting knives and saws
are used for operations on ribs, lungs, bronchial tubes, pleura and the
thoracic wall. Of special interest is the application of ultrasonic
instruments for cutting and knitting of the breast bone: the number of
complicating diseases after such operations remarkably decreased when
ultrasonic instruments were employed. Special flexible thin wave-
guides were developed for bronchoscopy and other kinds of bronchial
operations (for example, in cases of bronchial stenosis).

In ophthalmology ultrasonic instruments may be used for different
purposes (12)-(19). In most cases the instrument must be very light-
weight, of small dimensions, of high efficiency and permitting very
fine, or even microscopic, operations to be performed. Such instru-
ments have been developed and based on ferrite transducers (12), (24).
Their working frequency is 45 or 75 kHz, power consumption does not
exceed 10 watts and a supply voltage of 10 volt. The acoustical para-
meters of the operation are maintained automatically, and the surgeon
may control the conditions by himself, without any assistance of a
technician. After experimental verification the operation of dacriocistorinostomy was put into routine use in clinics (12). In the
case of this operation chisel and hammer were substituted by the much
less traumatic fine ultrasonic saw. It is also desirable for opera-
tions on soft eye tissues to use ultrasonic instruments, in particular,
when different kinds of tissue are to be separated (12). Another
promising application for the future seems to be the operation of
phacoemulsification, i.e. the cataract extraction through a tiny perforation without making a big incision (14). For this purpose the working end of the instrument is made in the form of a small tube, which is put into the eye through a perforation of diameter about 1-2 mm. The tip of the tube, vibrating with ultrasonic frequency, produces the effect of shattering to fragments and of emulsifying the lens, the emulsion being sucked through the tube or washed out of the eye with the liquid, flowing from the tube. A number of experiments were done (15)-(17) in order to clear up some aspects of this operation to establish its harmlessness for neighbouring tissues, to find out optimal characteristics of the ultrasound and to improve the operation technique and the instrumentation. Another ultrasonic operation of a special kind in ophthalmology is the repairing of retinal detachment. The working end of the instrument in this case has a form of a thin rod with a rounded tip (12), (18), (19).

Careful experiments performed on animals over a long period of time with registrations of electrocardiogram, encephalogram, pulse-curve, arterial tension, temperature, with investigation of regeneration process, proved the possibility and the benefit of the application of ultrasonic instruments for neurosurgical operations on the head (7) and so permitted the performance of such operations in a clinic (8), (20). In these cases highly efficient instruments with ferrite transducers were used. Working frequencies were 28 kHz and 40 kHz, power consumption was about 20-25 watts, vibration amplitude — up to 45 μ. Automatic regulation of acoustical parameters have a kind of freedom to the experimenter and to the surgeon. Utilisation of the ultrasonic saw for operations on the head allowed, together with the reduction of traumatisation to diminish the time for cutting out a piece of bone to 2-3 minutes instead of 25-30 minutes when using routine methods. The ultrasonic bone-knitting with the help of ciaxcin proved very effective for traumatic complicated skull fractures. For operations on scars and for other operations on soft head tissues low power optimized instruments with resonant frequencies of 45 kHz and 75 kHz were successfully used.

Manifold possibilities for utilisation of ultrasonic instruments exist in otolaryngology (9)-(11). For this field special waveguide-concentrators were designed, from one to four half-wavelengths long. Their working end had the form of a scalpel, of a blunt blade, of a spade, of a probe, of a needle, etc. In some instruments by a special choice of their shape, flexural vibrations were excited, parallel with the longitudinal ones. The amplitude relation of longitudinal and transverse vibration components was found to have an optimum value, different for different kinds of operations. The resonant frequencies of instruments were 45 kHz and 75 kHz. They worked with ferrite transducers and motional feedback generators. As in the previous case, after carefully performed experiments with registration of the main
functions of the organism (9), instruments were put into practice (10), (11). The ultrasonic blunt blade was used for the tonsillectomy and the ultrasonic scalpel for tracheotomy. In both cases the haematoysis effect of ultrasound proved most useful and besides it reduced appreciably the time consumption of the operation. For example, the tracheotomy, performed with a specially developed technique, lasted only 1-1.5 minutes. Ultrasonic scalpel was used also for operations on scar tissues, for laryngostenosis operations. Ultrasonic probes three or four half-wavelengths long were successfully used for surgical treatment of hypertrophic rhinitis (by putting the probe into the skull) and of some kinds of new growths. In the last case the destruction of the new growth was observed after a single or repeated contact application of the probe for a few seconds, with the resulting exfoliation or resolution of the necrotic tissue.

REFERENCES

ULTRASONIC MEDICAL DIAGNOSTIC METHODS

L. Filipczynski, Institute of Fundamental Technological Research, Polish Academy of Sciences, Warsaw, Poland

INTRODUCTION

A rapid development of ultrasonic methods and their applications in medical diagnostics has been observed during the last decade. Ultrasonic methods are nowadays used in obtaining valuable medical information which would be difficult or even impossible to get otherwise, for visualising soft tissues like those of human eyes, heart, fetus, circulatory system, brain, etc.

This unique ability of ultrasound waves is due mainly to exceptionally advantageous conditions of their propagation in soft tissues. Attenuation of soft tissues in the majority of cases is directly proportional to frequency up to 100 MHz and equals, at 1 MHz from 0,1 dB/cm/in the vitreous body of the eye/ to 3,5 dB/cm/in muscles and across fibres/ [8,21,42]. Therefore, for examination of the abdomen — where the range may be as long as 30 cm — frequencies of ca 2 MHz are applied, while e.g. in the eye examination — where the penetration depths amount to ca 3 cm — higher frequencies up to 20 MHz are used.

In this frequency range the ultrasonic wavelength of 0,75 — 0,075 mm exceeds only by three or two orders of magnitude the wavelength of light. Applications of such wavelengths enables us to form parallel or focused beams within the human body, the transverse dimensions of the beams being in the mm range, and to obtain a satisfactory resolution of the order of a wavelength.

Recently, a frequency of 100 MHz was applied in an ultrasonic microscope making it possible to obtain 25 μm resolution [28]. Unfortunately, due to high attenuation, thickness of tissues examined is limited to a few mm. For frequencies higher than 100 MHz attenuation appears to increase with frequency according to a square law [25].

The next factor deciding upon the possibility of application of ultrasound in medical diagnostics consists in the fact that acoustic impedances of various soft tissues differ from each other only but little due to differing elastic properties, their densities being practically constant [8]. The acoustical mismatch is so small that only ca. 2% of the intensity is reflected when the wave direction is normal to the plane interface of two different tissues. In the case of pulses, the echo produced however is strong enough to be registered on a CR-tube thus making it possible to discern tissue boundaries. Ultrasonic waves penetrate deeper and deeper into the human body with very low intensity losses and produce successive echoes from the subsequent tissue boundaries. Thus the conditions existing in soft tissues are highly advantageous from the point of view of their examination by means of ultrasonic waves, provided the
waves do not meet air cavities /which constitute impenetrable obstacles due to the acoustic mismatch/ or bones /which have an acoustic impedance and attenuation much higher than those of soft tissues/.

The information carried by an ultrasonic wave penetrating the human body is coded in its time of passage, in amplitude, frequency and phase. It is decoded and displayed by the corresponding echo method on a CR-tube as the A-, B-, or TM-presentation /amplitude, brightness and time-motion display, respectively/, or by the methods based on the Doppler effect and ultrasonic holography.

The variety of structures of human organs and their physiological functions requires the ultrasonic methods and instrumentation to be developed from the point of view of the specific properties of the organs investigated. Taking into account the fact that we shall have to deal very soon with 3 billion brains and hearts, 6 billion eyes, etc., existing on our planet, such a development of diagnostic methods aimed at individual organ properties is fully justifiable.

An excellent review of development of these methods and their applications in various fields of medicine was presented at the Second World Congress of Medical Diagnostics held in Rotterdam in 1973 [38]. The problem of interaction of ultrasonic waves and biological tissues, and the present state of knowledge in this field, were discussed and published in the Proceedings of a Workshop held in Seattle in 1972 [36]. A historical review on the development of medical ultrasonics was published quite recently in the British Journal of Radiology [23].

Ultrasonic diagnostics is nowadays becoming a wide field of interdisciplinary science and a large field of technology owing to the close cooperation of medicine, acoustics and electronics.

The purpose of this paper is to discuss some acoustical questions of this problem and, perhaps, to suggest certain improvements in this field, at the same time eliminating from our discussion medical and electronic problems as well as system considerations, which can be found in the literature [1,7,22,24,36,38,42].

SIGNAL POWER EFFICIENCY IN THE DIAGNOSTIC PULSE METHOD

One of the fundamental problems of ultrasonic diagnostics, dealt with by numerous scientists, is the problem of the maximum admissible intensity of ultrasonic waves. The problem of influence of ultrasound on biological structures remains still unanswered; numerous mechanisms and interactions detected so far are still not explained and in the years to come the research in that domain will
ULTRASONIC MEDICAL DIAGNOSTIC METHODS

constitute an important and attractive field of activity for biologists and acousticians.

That is why the further improvement of ultrasonic methods and instrumentation, increasing their sensitivity and detectability should proceed at the expense of increased intensity of ultrasonic waves radiated into the human body. Attempts should be made to decrease the intensities and thus to increase the separation from the existing threshold values, above which the effect of ultrasound could be harmful [18,20,39].

Having this situation in view let us investigate the signal power efficiency in the ultrasonic pulse diagnostic method following the pulse from the transmitter through the transducer into the biological medium, then after reflection, back through the same transducer into the receiver /Fig.1/. The electromechanical circuit of the piezo-electric transducer is shown in Fig.2. For the sake of simplicity, the clamped capacitance $C_0$ of the transducer and all remaining capacitances at its electrical input are assumed to be compensated by a shunt inductance. From Fig.2 the electric input impedance of the transducers dynamical branch may be determined to be

$$Z = Z_0 + jZ_1$$

where

$$a = \frac{(R_A + R_B)(R_A R_B + R^2 \tan^2 \frac{\pi f}{2})}{2} \left[ \frac{R_A + R_B}{2} + 4R^2 \tan^2 \frac{\pi f}{2} \right]$$

$$b = R \left\{ \frac{k_t^2}{x} - \frac{j l}{\sin x} + \frac{\tan \frac{\pi f}{2}}{2} \left[ \frac{R_A + R_B}{2} - 2R_A R_B + 2R^2 \tan^2 \frac{\pi f}{2} \right] \right\}$$

In these formulae $R_A = \rho_A c_A A$, $R_B = \rho_B c_B A$, where $\rho_A$, $\rho_B$, $c_A$, $c_B$ denote the respective densities and wave velocities in the media on both sides of the transducer, $A$ - the contact surface area of the transducer, $k_t$ - electromechanical coupling factor for thickness vibrations, $x = \pi f/j_m$, $f$ - frequency, $j_m$ - mechanical resonance frequency.

In the case of electrical resonance, $b = 0$ and then the electric input impedance of the transducer, when $R_A + R_B \ll R$, equals to a good approximation

$$Z = \frac{(R_A + R_B)(2N)^{-2}}{2}$$

Here $N$ denotes the transformation ratio of the electro-mechanical
transformer and is equal to \( N = \kappa_c \left( \frac{2f_m C_0 A \rho}{D} \right)^{\frac{1}{2}} \), where \( \rho \) - transducer material density, \( c \) - wave velocity in transducer for electric induction \( D = 0 \) [33].

Now the system of Fig.1 may be transformed to the electric circuit of Fig.3 where the functions of the transmitting-receiving transducer are separated between two transducers, and the biological medium is represented in the form of an electric transmission line with the wave impedance \( Z_p \). From this circuit we are able to evaluate the power radiated into the biological medium, which assumes the value

\[
P_B = 16P_{\text{max}} N^2 R_B Z_p \left[ 4N^2 Z_p + \frac{(R_A + R_B)(Z_p + Z_p)}{Z_p} \right]^{-2}
\]

Consequently, the signal power obtained at the receiver input resistance \( Z_r \), is equal to

\[
P_r = 16P_B N^2 R_B Z_p (R_A + R_B + 4N^2 Z_p)^{-2}
\]

In Fig.4 are presented the calculated signal levels in a diagnostic pulse equipment intended for eye visualization as also the losses during passage of a signal from the transmitter to the receiver. The signal was assumed not to be attenuated in the biological medium.

As the intensity of ultrasonic waves radiated should be limited, it follows that we should aim at achieving the maximum value of the ratio of the power received to the power arriving at the biological medium, \( P_r/P_B \). From Eq. /6/ the condition can be determined under which the ratio attains its maximum value, namely

\[
R_A + R_B = 4N^2 Z_p
\]

Another conclusion concerns the necessity of decreasing the impedance \( R_A \) representing the backing absorption. However, in order to avoid the associated increase of the pulse time duration, \( R_B \) should be increased by applying a \( \lambda/4 \) layer matching the medium to the transducer [29].

Fulfilment of the condition /7/ sometimes requires application of an electric transformer changing the electric impedance \( Z_p \) "viewed" from the transducer. One could also proceed in a different manner and divide the transducer into two or four segments which would then be connected in series. In Fig.5 the electric admittance measurements are shown of such a transducer applied for eye visualization [18]. Owing to its division the electric impedance is increased from 4 \( \Omega \) to
Fig. 1 System of ultrasonic pulse diagnostic equipment under consideration.
T-transmitter, R-receiver, P-piezoelectric transducer, A-absorber, B-biol. medium

Fig. 2 Electromechanical circuit of piezoelectric transducer vibrating in thickness, after Mason [33]

Fig. 3 Electrical equivalent circuit of pulse equipment for eye visualization [18].
E, Z_u - electromotive force and internal resistance of transmitter, respectively, Z_A, Z_B - mechanical impedances of transducers backing absorber and of biological medium, after transformation into electrical resistances, E = 300 V, Z_u = 110Ω, Z_p = 75Ω, Z_A = 75Ω, Z_B = 25Ω, k_b = 0.33, f = 7 MHz, trans. diam. = 2 cm

Fig. 4 Signal levels calculated from Fig. 3 for pulse equipment for eye visualization. ED - electric damping of transmitter circuit with Z_p, BA - backing absorption, MTR - medium/transducer reflection, P_max - maximum power for ideal transmitter to load matching, P_B - power radiated into biological medium, P_R - power delivered on receiver input.
70 Ω.

The problem of primary importance for the axial resolution of ultrasonic pulse diagnostic methods is the frequency bandwidth which limits the pulse duration time. In our considerations, a simple compensation of the clamped capacitance \( C_0 \) was assumed by means of a shunt inductance /Fig.3/ tuned to the resonance frequency of the transducer. According to filter theory the fractional bandwidth which can be obtained without introducing loss in such a case is \( r = \frac{1}{2} \sqrt{\frac{1 - \frac{C_0^2}{L^2}}{C_0^2}} \) [41], \( r \) being the ratio of the clamped to the dynamical capacitance of the transducer equivalent circuit with lumped parameters. It is equal to

\[
r = \frac{m^2}{8} \frac{1 - \frac{C_0^2}{L^2}}{C_0^2}
\]

/8/

It follows that the widest bandwidth can be achieved with transducers having a large electromechanical coupling coefficient \( k_t \). A new transducer material which seems to be useful is lithium iodate owing not only to its high value of \( k_t = 0.51 \) but also to its low acoustic impedance \( \rho c = 18.5 \times 10^6 \) kg/m²s; ceramic transducers with the same value of \( k_t \) yield \( \rho c = 30 \times 10^6 \) kg/m²s. Thus a lithium iodate transducer may be better matched to the acoustic impedance of soft tissues, and also lower back surface loading will be needed to obtain the same bandwidth as for ceramic transducers. It reduces the value of \( R_A \) and consequently losses of the useful signal. However, the transducer made of lithium iodate would require protection against its solubility in water [34,35].

Application of more complex electric matching networks leads to a wider frequency band, though only to a limited extent. For instance, using a ladder-network composed of two coils and three capacitors, it is possible to obtain a 30% bandwidth increase compared with a single shunt coil [27].

The measurements performed by us confirmed, however, the signal level diagram of Fig.4 only qualitatively. In reality, the power of radiated signals, as also of those obtained at the receiver input, were much smaller. One of the possible reasons for this could be transducer vibrations deviating from piston-type vibrations. This doubt was removed in our earlier paper [31], by showing that the ceramic transducer when heavy loaded at its back surface performs piston-like vibrations. Results of those investigations are demonstrated in Fig.6.

Another reason for the discrepancy discussed is also possible. The electric impedance of transducers is determined on the basis of steady-state measurements. Under real conditions, however, the pulses

76
Fig. 5 Electric admittance of a divided ceramic transducer

Fig. 6 Displacement distribution along diameter of ceramic transducer vibrating in air /A/, loaded on back surface with perspex /B/ and with high impedance absorber /C/

Fig. 7 Equivalent circuit of piezoelectric transducer for transient analysis
applied in our methods consist of only three or four high frequency periods and, in such cases, the electric input impedance of the transducer varies as in a transient process. Let us estimate this effect for the case under consideration. Fig.7 shows the equivalent electric circuit of the piezoelectric transducer deduced for transients /14/ in which the electric transmission line represents the transducer itself. From this Fig. it is easy to find the input impedance for resonance in the first half of the period to be

\[
\frac{1}{(Z_A + 2)(Z_B + 2)(Z_A + Z_B + 22)^{-1}} 
\]

while in the steady state it is equal to

\[
\frac{1}{4} \frac{1}{Z_A + Z_B}
\]

In the case under consideration it means that in the first half period the input impedance of the transducer is about ten times greater than in steady state. In next half periods the input impedance becomes stepwise smaller achieving at the very end the steady-state value. This is responsible for the fact, in the pulse duration time, of the electrical and mechanical matching undergoing great changes. This effect may explain the discrepancies in the signal level evaluation quoted above. In spite of these, the general idea of reasoning presented above remains valid.

**QUANTITATIVE ULTRASONOGRAPHY**

Ultrasonic diagnostic methods, in the course of their development, have changed character from qualitative to quantitative. The only information supplied by the first echoscopes concerned the existence of tissue interfaces. Later it was possible to measure even with great accuracy distances within structures investigated. Introduction of visualization methods in which the ultrasonic beam was scanning the human body and the ultrasonic signal was applied for the CR-tube brightness modulation /B-presentation/ enabled us, in turn, to determine the forms of tissues boundaries /Fig.8 A,B/. The refinement of this technique resulted in obtaining ultrasonograms with a considerable amount of detailed information /Fig.8 C/, while the application of calibrated attenuators created the technical basis for identification of different tissues.

Ultrasonic diagnostic equipment is now required to supply quantitative rather than qualitative information [26]. This is necessary from the point of view of reproducibility and correct interpretation of results obtained by means of various equipment and different
operators working in different clinical centres. It is also necessary for evaluation and possible future standardization of diagnostic equipment being produced currently by some 50 firms all over the world.

Two parameters of particular interest for the ultrasonography should be standardized in the near future, and they will be discussed in this paper, namely the radiated wave intensity and the overall sensitivity of diagnostic equipment.

Using a number of instruments based on the radiation pressure principle it is possible to measure the ultrasonic power radiated in liquids, having acoustic properties similar to those of soft tissues. With these methods the power/average time value/ could be measured in the mW range [42] and recently the power output of 30 μW could be determined using an electrobalance for radiation pressure measurements [37].

However, in the case of pulses more sensitive at least by three orders, there are methods which make it possible to measure the peak value, since in typical diagnostic equipment it is $10^3$ times greater than the time average value. We would like to mention here two absolute methods, namely the electrodynamic and the capacitance methods [10] which allow for measuring the acoustic velocity and displacement, respectively /Fig.9/. From these values it is possible to determine the intensity of radiated pulses under the condition that the measurements were made in the entire beam cross-section of the near field or in the far field. This procedure may be shown to be, in general, inappropriate for intensity determination in the near field, since the phase and amplitude relationship show large divergencies from corresponding plane wave values [12]. However, for measuring in the entire beam cross-section the signal level in the near field is independent of distance, if one takes into account the small correction for diffraction losses [3].

Using the capacitive transducer we could measure ultrasonic intensities radiated by different diagnostic pulse equipment [11] and also by the c.w. Doppler fetal detector when pulsing its transmitter [15].

Interesting properties are shown by beams radiated with plane and weakly focusing transducers having Gaussian velocity distributions on their surface. The ratio of pressure to acoustic velocity along the beam axis is the same as in the case of a plane wave and equals $p/c$. The maxima and minima of amplitude and phase which are present when the velocity distribution on the transducer surface is uniform, do not appear in the ultrasonic beam /Fig.10 and 11/ [16]. The particular properties of acoustical beams radiated by Gaussian transducers may be useful in ultrasonic measurement problems, as also in acoustical
Fig. 8 Longitudinal ultrasonogram of a woman in 36 week pregnancy, fetus in longitudinal position /A/; ultrasonogram of eye /B/; high resolution transverse ultrasonogram of a pregnant woman in the term, visible fetal head with middle structures, placenta previa /C/; [17,19]

Fig. 9 Capacitive and electrodynamic transducers for pulse measurements in liquid

Fig. 10 Acoustic pressure calculated across beam width for plane /A/ and focused /B/ transducers. Full line denotes amplitude, dashed line - phase angle
holography for removing the artifacts, like e.g. the degrading background ring patterns [24].

Heavy loading of the transducers back surface leading to amalgamation of a large number of modes into one piston-like mode [31] may also be useful for this purpose. Independently of this the transducers with Gaussian velocity distribution were observed to vibrate more uniformly than the conventional ones [43].

In the ultrasonographs designed for visualization of internal human structures /B-presentation/, the focusing systems are used to concentrate the ultrasonic beam scanning the human body. Such a solution, being necessary to improve the lateral resolution increases the ultrasonic intensity in focus. This increase is the highest in ultrasonographs for eye visualization. The intensity focusing factor of weakly focusing transducers is expressed by the formula [18]

\[ \frac{I_F}{I_T} = \pi^2 a^4 \lambda^{-2} f^{-2} \]

where \( I_F \), \( I_T \) denote ultrasonic intensities in the focus and at the transducer's surface, respectively; \( \lambda \) - wavelength, \( a \) - transducer radius, \( f \) - focal distance. In typical ultrasonographs for abdomen and eye visualization the parameters of focusing systems are \( \alpha = 1 \) cm and respectively \( \lambda = 0.7 \) mm \( /0.2 \) mm, \( f = 20 \) cm/ 10 cm giving the focusing factors equal to 5 or 250. From these considerations it follows how important is the intensity determination in eye visualization problems. For this purpose we have developed a special capacitive transducer with spherical electrodes to measure the intensity at the surface of the radiating transducer [18]. Measuring additionally the focal distribution of the ultrasonic field by means of a small ball target we were able to determine the peak intensity at the focus. In this manner it was possible to carry out a dosage study on rabbit's eyes by means of focusing beams; it led to the determination of the threshold intensity values from the point of view of histological changes in the retina and, consequently, to reducing the output of our ultrasonograph and maintaining a wide margin of safety. We have accepted peak intensities at the focus to be equal to 0.5 or 5 W/cm², using the higher value in eyes with suspected tumors. The corresponding values averaged in time were 0.2 or 2 mW/cm², respectively [18].

The safety margin may be increased in all ultrasonic pulse and c.w. equipment in present use by varying the output of the transmitter rather than the receiver gain, thus reducing the ultrasonic dose during clinical examinations. In this way we could decrease also the power of our Doppler fetal detector by introducing three output levels of 8, 3 and 1 mW/cm².
The second parameter, closely related to the radiated intensity, is the overall sensitivity of the diagnostic equipment. This parameter describes the properties of the transmitting and receiving channels with CR-tube when connected with the ultrasonic probe. The problem consists in finding a simple definition and a simple measuring method for that magnitude, and to make its measurement possible even under clinical conditions. For this purpose we proposed an overall sensitivity scale that assumes the level of the pulse transmitted as the reference value [19]. The reference level is easily obtained by immersing a large perfect reflector e.g. of steel/ in water and placing it in the near field where the echo is practically independent of the distance [3]. The overall sensitivity used in our examinations, or its value characterizing the pulse equipment, may then easily be determined by means of the attenuator which enables us to compare the echo from the perfect reflector and the actual echo from a biological target, at the same threshold signal level obtained on the A- or B-scope.

In this way we have measured the maximal overall sensitivity of the A-echoscope DI-12 /Inco, Warsaw/ obtaining $S_{\text{max}} = 96 \text{ dB } /f=3 \text{ MHz}$, transducer diam. 1 cm /[11] and of the ultrasonograph UPO-1 /IPPT, Warsaw/ for the abdomen visualization equal to $S_{\text{max}} = 98 \text{ dB } /f=2 \text{ MHz}$, transducer diam. 2 cm/ [2].

Once the intensity of the pulse radiated and the overall sensitivity in dB are known, we are able to determine the absolute values of the sensitivity of the receiving channel. One disadvantage of the plane reflector method is a very large echo, greater by ca. 40 dB than that obtained from soft tissue interfaces [32]. The 80 dB attenuator used in the present diagnostic pulse equipment should be increased to suppress the echo from the perfect reflector which saturates the input of the receiver. In order to eliminate these difficulties we have introduced temporarily a perspex block with an oblique groove thus obtaining a 65 dB attenuation of the ultrasonic pulse due to the oblique reflection at the boundary water-perspex /Fig.12/ [13].

In the second method for determining the overall sensitivity, targets of steel balls were proposed, the balls being small enough to obtain echoes comparable with tissue echoes. The balls should be placed at uniquely defined points in the ultrasonic beams, such as the first maximum of the near field / in the case of a parallel beam/ or the focus [13] [32].

The evaluation of the overall sensitivity of Doppler equipment has to be based on a different principle since it depends on the velocity of the moving structure. The Doppler signal received is modulated both in amplitude and in phase. The phase modulation being very small,
Fig. 11 Transverse distribution of acoustic beam measured for typical /TU/ and Gaussian /TG/ focusing transducers.

Fig. 12 Attenuating reflector. 1—probe, 2—perspex, 3, 4—reflecting planes, 5—water.

Fig. 13 Overall sensitivity determination of Doppler fetal detector. T, R—transmitter and receiver, RS—reflecting system, M—modulator, AG—acoustic generator, A—attenuator, PLM—power level meter, 3—water, 4—perfect reflector.
only the amplitude demodulation is used in most instruments. Therefore, in the case of fetal detectors — the most popular ultrasonic diagnostic devices — the electroacoustic reflecting unit RS is introduced for overall sensitivity measurement /Fig.13/ [2]. The signal received by the reflecting unit RS is amplified, modulated in amplitude and radiated with the same amplitude as the signal received. In this manner we are able to replace the moving structure by the electroacoustical reflecting unit RS, and then to apply typical electrical measurement methods for the overall sensitivity determination. In order to avoid standing waves in the water container in which measurements are performed, its bottom is oblique and made of perspex, thus attenuating the reflected wave 27 dB.

In this way we have determined the maximum overall sensitivity of the Doppler fetal detector UDT-10 /Techpan, Warsaw/ to be 104 dB at the signal level on audio terminals 20 dB above the noise level [2].

We have discussed here only two parameters, in our opinion the ones which are currently important in the process of transformation of ultrasonography from a qualitative into a quantitative measurement method. Obviously, the number of actual parameters is much greater and will increase in the course of further development of ultrasonic diagnostic methods.

The most crucial problem facing the present development of these methods consists in tissue identification. In spite of a certain progress achieved in this field [4,6] the problem remains still open for discussion [9] and far from being definitely solved. It must be admitted that not all the information carried by the ultrasonic waves penetrating the human body is now utilized. In particular, measurements of absorption and, independently, scattering of ultrasonic waves in the living tissue, as well as their dependence on frequency and direction should in the future yield new and important information and widen the range of application of ultrasonic diagnostic methods [5,30]. It will create the necessity of further fundamental acoustic research, and further development in the process of transformation of the diagnostic equipment into still more sophisticated measurement devices.

REFERENCES

EIGHTH INTERNATIONAL CONGRESS
ON ACOUSTICS, LONDON 1974

ULTRASONIC MEDICAL DIAGNOSTIC METHODS

akustyki, Warszaw (1973) 8 265; [16] L.Filipczyński and J.Etienne,
Seminar,Mierzyń, Poland; (1973) 1 51; [18] L.Filipczyński et al., Proc.
[27] W.Kooij et al., Inst.Med.Ohys.TNO Utrecht, Report 24121/1 (1973);
published; [33] W.Mason, Electromechanical transducers and wave
filters. Van Nostrand 1948; [34] E.Mattiat Ed. Ultrasonic Transducer
[36] J.Reid and M.Sikov Ed. Interaction of Ultrasound and Biological
Physical Principles of Ultrasonic Diagnosis, Academic Press 1969;
MODELS OF THE UNDERWATER ACOUSTIC CHANNEL

J. Johnsen  Norwegian Defence Research Establishment,  
3190, Horten, NORWAY

INTRODUCTION

The propagation of sound in the ocean is very complex. The velocity of sound under water is a function of physical parameters such as temperature, pressure and salinity. It varies with geographical position, with depth and with season. It is time and space dependent.

The characterization of the underwater medium with a mean sound velocity may be called a macroscopic model of the medium. It is useful in sound propagation studies and in the estimation of ranges for active and passive sonar systems. Such models may also give explanations as to why it is possible that sound in some cases propagates over very long distances while in other cases under different conditions it may be impossible to measure sound generated only a short distance away.

Since the physical properties of the underwater acoustic medium are such that the sound velocity is space and time dependent it must be characterized as an inhomogeneous medium. Although it is an approximation one may say that the medium is characterized by two types of inhomogeneities, one regular and one stochastic. The regular type of inhomogeneity is the spatial variation of the mean of the medium characteristics expressed by the sound velocity. Refraction and channeling, focusing and shadowing effects in the sound propagation picture is caused by this type of inhomogeneity.

The random type of inhomogeneity is due to the deviation of the physical parameters from the mean values. This stochastic characterization may be called a microscopic medium model. The random inhomogeneities cause scattering or diffraction of the acoustic wave. The effect is observed as amplitude and phase fluctuations superimposed on a received wave and as sound energy scattered and reflected back to the transmitting position. The latter effect is called reverberation.

In an investigation of the various possible methods which can be used in the transmission of information through water the use of sound is found superior. This is due to the fact that the attenuation of transmitted energy for low frequency sound is much less than for electromagnetic waves. There are a variety of systems utilizing sound transmission capabilities in the sea. Based on one-way transmissions through the medium one may list:
communication between submersibles, divers or surface ships

telemetering as from buoys, submersibles, etc

navigation as by acoustic beacons

noise source localization by passive sonar

Based on two-way transmissions through the medium one may list:

echo sounders as depth indicators, topography mapping, etc

echo ranging as in civilian and military sonar systems

navigation as by acoustic repeaters and doppler log systems

In the general information transmission problem there are consequently two fundamentally different models. One is associated with the one-way transmission where information is transmitted through the medium between a transmitter and the receiver, fig 1 a. The other model is associated with two-ways transmission where a receiver, usually co-located with the transmitter, receives signals reflected or transmitted back by some object or reflectors in the medium, fig 1 b.

Almost all types of information transmission or acoustic detection systems may be simplified by a single block as in the middle in fig 1 a. A signal characterized by \( x(t) \) is transmitted into the underwater acoustic medium by an electro-acoustic transducer system. At the receiver a signal, \( y(t) \), picked up by the receiver transducer system, differs from the input signal. The questions to be asked are: How is the received signal related to the transmitted signal?

\[ \text{Fig. 1 - Models of the underwater acoustic information transmission situation (T-transmitter, R-receiver)} \]
Can the medium be described by a model which makes it possible to predict $y(t)$ when $x(t)$ is known? Since the medium is inhomogeneous one is confronted with a basic problem in the analysis of signal transmission systems - how to relate the statistical parameters of the output signal to those of the input signal and the transmission characteristics of the medium. There are a variety of models some to more and some to less degree of accuracy. Simple models may be sufficient for use in estimating detection ranges. Optimum signal detection methods, however, are based on the assumption that accurate statistical models of the medium are available.

THE PHYSICAL MODEL OF THE MEDIUM

Understanding of sound propagation capabilities of the underwater acoustic medium is based on knowledge of its physical model, fig 2. In general the ocean may be regarded as a fluid with two boundaries formed by the sea surface and the sea bottom. The fluid is not homogeneous with regard to density, salinity or temperature. The medium may also contain sound scatterers such as high density regions, fish, etc. Due to the spatial variations in the physical parameters, it is associated with this a spatial variation in sound velocity. The boundaries, both the upper and lower surface may be either rough or smooth and their properties may vary from one place to another. The sea surface is also extremely time variable.

The bottom may be anything from an almost perfect absorber to a perfect reflector and often it may be correct to model it as a continuous region where the physical properties gradually change from those of the ocean to those of the underlying material. In this case one can symbolize the sea floor by a bottom and a bottom interface. Its characteristics are important factors in seismic profiling and shallow water long range sound transmissions.

The properties of the sea surface may be the most complex of the model parameters. It is hard to describe in a form suitable for a working medium model. The presence of waves introducing white caps, bubbles, particle motions, etc produce scattering and shadowing effects when the surface is illuminated by a sound wave. The reflected and scattered waves have properties significantly different from those reflected from a simple boundary formed by a pure water-air interface.
MODELS OF THE UNDERWATER ACOUSTIC CHANNEL

Fig. 2 - The physical model of the underwater acoustic medium

THE CLASSICAL MODELS

There are two general, they may be called classical, methods which can be used to solve sound propagation problems connected with underwater acoustics. One approach is based on wave motion calculations and includes information on both amplitude and phase of the wave as it propagates through the medium. The other approach is based on a technique similar to ray path technique in optics. It may be derived as an approximation of the first approach or it may be considered as a technique which can be developed separately. In the ray path technique information about the phase is lost.

THE WAVE EQUATION APPROACH

If a fluid is in a state of equilibrium and the pressure, \( p \), in a certain region deviates from the equilibrium value, \( p_0 \), the fluid itself generates forces which tend towards restoring the equilibrium value. As a result vibrations are generated and propagated as waves through the fluid. The restoring forces are caused by the elasticity of the medium and the wave propagation is due to elasticity and the inertia of the displaced particles permitting a transfer of
momentum to adjoining particles. It is known that a particle system like this can be described mathematically by the wave equation. In the three-dimensional case it may be shown that the wave equation for the pressure fluctuations becomes:

\[ \nabla^2 p - \frac{1}{c^2} \frac{\partial^2 p}{\partial t^2} = 0 \]  \hspace{1cm} (1)

where the sound velocity \( c \) is defined from the density, \( \rho_0 \), and the bulk modulus, \( \kappa \), by \( c^2 = \kappa / \rho_0 \). A general solution of eqn 1 always contains arbitrary constants and functions. These can only be evaluated by considering situations with specific boundary and initial conditions. For one-dimensional motion the wave equation has a general solution of the form:

\[ p(x,t) = f_1(x-ct) + f_2(x+ct) \]  \hspace{1cm} (2)

where the \( f \)'s are arbitrary functions depending on the boundary fixed by the geometry of the medium and by the initial conditions defined by the pressure-space-time distribution of the disturbance. The waves defined by eqn 2 are called plane waves. They have common phase and amplitude at all points on any plane perpendicular to the wave propagation direction \( x \).

From a sufficiently large distance any sound source looks like a point source. A wave propagating in three dimensions and symmetrical with respect to the source cause the pressure disturbance to be a function of range \( y \). In spherical coordinates the wave equation becomes:

\[ \frac{\partial^2 p}{\partial r^2} + \frac{2}{r} \frac{\partial p}{\partial r} - \frac{1}{c^2} \frac{\partial^2 p}{\partial t^2} = 0 \]  \hspace{1cm} (3)

with a general solution of the form:

\[ p(r,t) = \frac{1}{r} f_1(r-ct) + \frac{1}{r} f_2(r+ct) \]  \hspace{1cm} (4)

The first term in eqn 4 represents a spherical wave diverging from the source. The second term represents a similar wave converging on the origin. The wavefront of many types of diverging waves in a homogeneous medium have the characteristics of a plane wave as the distance from the source becomes large.
The intensity of the acoustic wave is defined as the time average of the instantaneous rate of energy flow. For a harmonic spherical wave the following expression for the intensity can be derived:

$$I = \frac{a^2}{2 \rho_0 c r^2} \tag{5}$$

$a$ is the maximum pressure change at a distance one length unit from the source. Eqn 5 express the inverse square law of intensity loss for a spherical wave in an ideal and infinite medium. This spreading loss is observed as a decrease in sound intensity with increasing distance. Measurements indicate, however, that the intensity loss is greater than the theoretical spreading loss. This extra loss is due to the fact that the medium is far from ideal. Transmission anomalies or attenuation is a function of frequency and can be derived mathematically by treating the medium as a viscous fluid. The loss due to attenuation or absorption can then be found and expressed by the intensity variation with range:

$$I = I_o e^{-\alpha x} \tag{6}$$

$\alpha$ is a function of frequency representing the kinematic friction loss in the fluid and $I_o$ is the intensity referred to one length unit from the source. $\alpha$ is known from measurements and varies from one area of the ocean to another. The total intensity variation for a spherical wave propagation can then be written:

$$I = I_o e^{-\alpha r/r^2} \tag{7}$$

Inhomogeneities in the medium are caused by spatial changes in temperature and density. The wave equation for such an inhomogeneous medium can also be derived (1). The treatment of such an equation is, however, more complicated and several authors introduce the effect of the inhomogeneities in a different way. One method is to assume that the sound velocity is expressed by:

$$c = c_0/(1 + \mu) \tag{8}$$

where $\mu$ represents the fluctuation in the refractive index, $\eta$, defined by: $\eta = c_0/c$. Since the refractive index is
spatial dependent $\mu = \eta(x,y,z) - 1$, the following modified wave equation can be written:

$$v^2 p - \frac{(1 + \mu)^2}{c_0^2} \frac{\partial^2 p}{\partial t^2} = 0$$  \hspace{1cm} (9)$$

If the space and time characteristics of $\mu$ are specified, a solution to eqn (9) is possible to find. The methods of solution will not be treated here since the mathematics involved in interesting cases are quite formidable.

THE RAY PATH APPROACH

Assuming a point source the solution of the spherical wave equation in an infinite, homogeneous medium shows that the sound pressure fluctuations propagate as spherically expanding waves. The sound pressure decreases proportionally with range while the intensity decreases according to the inverse square law.

If equiphase surfaces are imagined these surfaces will form concentrical spherical surfaces propagating with the velocity of sound away from the source. The normals to the surface indicate the direction in which the wave propagates. If the sound velocity varies throughout the medium then the equiphase surfaces are no longer spherical. The surface normals then form curved lines instead of straight lines. These lines are called rays. Based on some physical conditions these rays give the propagation direction for the sound energy. The conditions are:

The radius of curvature of the path of the sound ray must be large compared to the acoustic wave-length, $\lambda$, i.e. the refractive index must remain approximately unchanged over a wave length.

The change in amplitude (due to loss) must be small over a wave-length.

The sound propagates along the path with a velocity equal to the local sound velocity along the path. The number of rays crossing a unit area normal to the main propagation direction gives a measure of the intensity of the wave at that point in the medium. If the form of the ray paths are calculated one can obtain an easy and convenient geometrical presentation of the direction and intensity of the sound field.
The equations for the ray paths are more easy to derive and to treat than the wave equation. Exact solutions are possible; however, information about the phase and frequency is lost and problems like interference, dispersion, etc., cannot be treated. Assuming the sound velocity is a function of depth only, Snell's law can be used to find the direction of the sound ray throughout the medium. If a linear sound velocity profile is assumed, fig 3, and if the sound velocity and the tilt angle at the source are $c_o$ and $\theta_o$ respectively, then by using Snell's law one can find the direction of the ray at any point along the path:

\[
\frac{\cos \theta}{c} = \frac{\cos \theta_o}{c_o} \quad \text{or} \quad \frac{\cos \theta(z)}{\eta(z)} = \frac{\cos \theta_o}{c_o}
\]  

(10)

The resulting path is also shown in fig 3. When the sound velocity profile is known it is possible by the aid of the mathematical tools developed for the ray acoustics to calculate ray paths for any starting angle from the source. A map of ray paths drawn for a variety of tilt angles can be used as a good approximation in estimating how the acoustic energy is distributed throughout the medium.

Fig. 3 - Sound velocity profile and acoustic ray propagation

From acoustic ray path theory it is found that the rays always are refracted away from areas with a higher sound velocity.
If acoustic ray theory is to be applied, a model of the medium and its boundaries must be constructed. The sea volume is characterized by a set of sound velocity profiles. The boundaries are characterized by its form and reflectivity. Ray path pictures are easily obtained numerically if the sound velocity profiles are approximated to linear segments and if specular reflections are assumed at the boundaries. For a quick overview of the sound propagation condition this method is very much used and if sonar coverage for a specific sonar system is to be estimated this method is very useful.

Based on the ray path method some standard models of the medium can be constructed. These models are derived from typical seasonal sound velocity profiles. In general any sound velocity profile can be approximated by a combination of one or more out of four basic types of profiles. These are:

Uniform sound velocity as a function of depth - One can obtain ideal sound propagation as indicated in fig 4 a.

Positive gradient - The sound velocity increases with depth and the rays are bent upwards as in fig 4 b. The sound energy is concentrated at the surface.

Negative gradient - The sound velocity decreases with depth - The energy is refracted down towards the bottom and shadow zones are located at the surface.

Negative gradient near surface and positive gradient at greater depths - Assuming a source located at a minimum sound velocity depth, most of the sound energy is then trapped in a sound channel as in fig 4 d.

Due to the spatial variations of the sound velocity and the often unknown characteristics of the boundaries, ray plotting is rarely very accurate for intensity distribution estimation. Scattering from the inhomogeneities in the medium also causes energy to be measured in areas predicted to be complete shadow zones by the ray path method. In deep waters these and other problems are not so severe as in shallow and coastal waters.

THE TIME DISPERSIVE MEDIUM

Ray path techniques are used to demonstrate the medium as a time dispersive medium. Experiments show that if sound of constant intensity and frequency is transmitted from one point in the ocean to another at fixed distance the received intensity varies with time. The intensity fluctu-
MODELS OF THE UNDERWATER ACOUSTIC CHANNEL

a - uniform s.v. profile  
b - positive s.v. gradient

c - negative s.v. gradient  
d - sound channeling

Fig. 4 - Basic types of sound velocity profiles with ray paths

mirror surface  
rough surface

Fig. 5 - Ideal sound propagation situation

rough surface  
Fig. 7 - A realistic sound propagation situation

Fig. 6 - Approximated realistic sound propagation situation

Intensities are due to the changing sound propagation conditions. If surface reflections are involved probably most of the variations are associated with the reflection process. The intensity variations are caused by interference between signals following different paths through the medium. Different paths contribute to differences in travel time and since the sound velocity and the ray path picture change with time, a time varying interference process is observed at the receiver. The differences in travel times associated with the various paths through the medium is said to cause time spreading.
MODELS OF THE UNDERWATER ACOUSTIC CHANNEL

All possible paths between a transmitter and a receiver define the multipath structure of the medium. The variations of the multipath propagation picture with time cause the time spread to vary with time. The medium is a time varying dispersive communication channel. If the transmitter and the receiver are mounted on moving platforms, the varying geometry also causes a changing multipath structure. A significant variation in intensity may be observed.

Based on what is said above the medium can be modelled as an ideal communication channel as in fig 5, and approximately realistic channel as in fig 6 and as a realistic channel as in fig 7.

In fig 5 the sea is assumed to be a semi infinite homogeneous medium bounded by a mirror surface. The model is valid for deep waters, relatively smooth surfaces, negligible sound velocity gradients and fixed platforms. Direct and surface reflected paths are observed. In a frozen situation no fluctuations in the intensity are observed at R if a constant signal is transmitted.

In fig 6 the surface is allowed to be rough and moderate refraction due to weak inhomogeneities in the medium may be assumed. Several surface reflected waves interfere with a bunch of almost identical direct waves. The interference picture varies as the surface and volume structure varies with time. The model describes the medium as a timespread communication channel. It may be valid for fixed underwater installations.

In a realistic picture of sound transmission between platforms, the vertical and horizontal variations in sound velocity and sea bottom configuration and reflectivity must also be considered. Factors contributing to the intensity fluctuations are: moving transducers and pitch and roll causing spatial intensity variations in transmitted signal; bottom reflections interfering with surface reflected and direct paths; the phase of the bottom reflected wave also changes with changing physical properties of the bottom and with incident angle; shadow zones, focusing effects and partly sound channeling effects caused by the spatial dependent refractive index; and finally a variable refraction picture, scattering and absorption caused by the inhomogeneities. In general numerous factors and mechanisms are involved in the propagation picture. Some are connected to the platforms and the transmitting and receiving system.
Others are connected to the medium itself including the boundaries and are very hard to predict.

THE FILTER MODEL

In the introduction it is presumed that the underwater acoustic medium can be modeled as a filter. It must characterize the channel as a dispersive, time varying filter due to the time varying multipath structure in the medium.

The simplest model can be derived from the ideal ray path picture in fig 5. For one path only the output of the medium filter is an attenuated, a, and delayed, $\tau$, replica of the input signal, $x(t)$. If all possible paths in a frozen multipath situation are combined, the output signal, $y(t)$, is a sum of $N$ such elementary signals:

$$y(t) = \sum_{i=1}^{N} a_i x(t-\tau_i) \quad (11)$$

If a fine structured multipath picture makes it impossible to resolve the paths, the summation in eqn (11) must be substituted by an integral:

$$y(t) = \int_{-\infty}^{t} h(\tau) \cdot x(t-\tau) \, d\tau \quad (12)$$

The medium characteristics are expressed by the impulse response, $h(\tau)$, and the output is obtained by convolving an arbitrary input signal and the impulse response of the medium. Sound propagation through such a dispersive time invariant medium is characterized by an impulse transmission received as a train of pulses with various strengths and arrival times or by CW transmission received as a tone at the same frequency, but with a different phase and amplitude. In all signal transmission situations the background noise must also be considered and included as an additive term in eqn 12. Here the ambient noise in the sea is omitted.

In the general case it is not sufficient to model the medium by the simple time invariant model. The propagation characteristics are both time and space variant (2). To simplify the expressions the spatial variables are omitted in the following. In a time variable dispersive medium a signal is subjected to both time and frequency spreading.
A frequency and time dispersive channel is characterized by two definitions, the time spread, L, and the frequency spread, B. The definitions are demonstrated in fig 9. In fig 9 a an impulse is transmitted. The received signal is a time delayed and smeared signal. The average length of the signal is defined as the time spread of the communication channel. In fig 9 b a tone at fixed frequency is transmitted. The received signal consists of a frequency shifted and frequency smeared signal. The average bandwidth of the received signal is defined as the frequency smear of the channel. The average frequency shift is due to relative motion between the platforms, local currents or surface effects. Due to the random fluctuations of the medium characteristics with time the form of the received signal also varies with time. These variations are observed both in the time and the frequency plane. B and L must therefore be found as statistical averages.

In a system design phase involving underwater communication, telemetering or sonar both B and L must be considered. They must be found as a function of environmental parameters such as surface wave spectrum, current conditions, etc. B and L defines the maximum time and frequency resolution which can be obtained in the medium. The use of signals with high resolution in time and frequency, range and doppler in an active sonar system, must be matched to these medium limitations.

Since the signals transmitted through the underwater acoustic medium generally are subjected to time spread and fre-
quency spread, it can be described as a linear, time-varying filter. The filter functions must be regarded as random functions due to the randomness of the medium and average values must be used to characterize the medium as a transmission channel. For acoustic channels as linear time-varying stochastic filters a set of useful system functions can be derived. The medium described as a time varying stochastic filter is a very general description and can be used to describe the medium transmission characteristics if communication or passive sonar systems are considered. The models can also be used to characterize the two-ways transmission properties for a target signal in active sonar systems and to characterize the reverberation since the reverberation signal, consisting of an infinite number of signals scattered and reflected back to the receiver, can be observed as the output of a highly dispersive filter. (In the following the signals and the systems functions are assumed represented by their complex envelopes).

First, as for time invariant systems, the time varying frequency response of the medium filter, $H(f,t)$, is defined. This system function can be interpreted as an amplitude and phase response varying with time, t. The inverse Fourier transform of $H(f,t)$ is a function of both absolute time, $t$, and delay time, $\tau$, and is called the time varying impulse response, $h(\tau,t)$. It is the output of the filter at time $t$ due to an unit impulse applied $\tau$ seconds ago. By Fourier transforming the time varying impulse response with respect to $\tau$, a new system function called the filter spreading function (3) is obtained:

$$s(\tau,\phi) = \int_{-\infty}^{\infty} h(\tau,t) \exp(-j2\pi\phi t) \, dt$$

(13)

This function gives the spectrum with $\phi$ as frequency variable, of the time variations of the impulse response. The output signal can then be written:

$$y(t) = \int_{-\infty}^{\infty} h(\tau,t) x(t-\tau) \, d\tau = \int_{-\infty}^{\infty} s(\tau,\phi) x(t-\tau) \exp(j2\pi\phi t) \, d\tau \, d\phi$$

(14)

where the inverse Fourier transform of $s(\tau,\phi)$ is used. From eqn 14 it is seen that the output signal is formed as a sum of time and frequency shifted versions of the input signal weighted with the spreading function, $s(\tau,\phi)$. Thus the spreading function determines the spread in time and frequency the signal will suffer in the medium. If rever-
berations are studied, the spreading function determines how a transmitted signal is reflected back to the receiver in a time and frequency smeared fashion.

\[ y(t) = \int h(\tau, t) \cdot x(t - \tau) \, d\tau \]

\[ = \int X(f) \cdot H(f, t) \cdot \exp(j2\pi ft) \, df \]

the underwater medium filter

Fig. 10 - Input-output relations for the underwater acoustic medium as a linear time varying filter

The Bi-frequency function is another system function obtained by Fourier transforming the time varying frequency response with respect to \( t \):

\[ B(f, \phi) = \int H(f, t) \exp(-j2\pi \phi t) \, dt \]  \hspace{1cm} (15)

The output signal can then be written:

\[ Y(f) = \int B(f-\phi, \phi) \cdot X(f-\phi) \, d\phi \]  \hspace{1cm} (16)

\( B(f, \phi) \) gives the spectrum of the variations of the frequency response. The meaning of this is seen from eqn 16. The output at frequency \( f \) is not only dependent upon the input at the same frequency, but upon components in a frequency band around \( f \), determined by \( B(f, \phi) \).

The four system functions are interconnected by Fourier transforms. Any one of the system functions may be used to completely define the time varying medium. If one of them was exactly known the medium degradation of the signal could be completely compensated for in the receiver by applying the inverse transformation to the received signal. This technique could also be used to design optimum de-reverberation filters for active sonar systems.

The time spread and the frequency spread are connected to the scattering effects associated with the time varying small-scale inhomogeneities characterizing the volume and the boundaries and it will never be possible to find the
exact system functions of the medium. They can only be
described statistically and their effect can not be com-
pletely compensated for in the receiver. The four system
functions then become random functions in two variables.
The random functions can be described by their corre-
lation functions, functions of four variables, defined by
ensemble averages or expectations as for the impulse re-
response and the spreading function:

\[ R_h(\tau, \tau', t, t') = E \left[ h(\tau, t) \ h^*(\tau', t') \right] \]

\[ R_s(\tau, \tau', \phi, \phi') = E \left[ s(\tau, \phi) \ s^*(\tau', \phi') \right] \]  (17)

Dealing with four dimensional correlation functions is
inconvenient and restrictions to the nature of the random
process are often introduced. The restrictions reduce the
functions to functions of two variables only. If it is
assumed that the random process described by the random
system functions of the medium, belongs to a class of
processes which are in a wide sense stationary in time and fre-
quency, then the correlation functions depend only on
time and frequency differences rather than absolute time
and frequency. This is expressed by writing the corre-
lation functions:

\[ R_h(\tau, \Delta t) , R_h(\Delta f, \Delta t) , R_s(\tau, \phi) , R_h(\Delta f, \phi) \]  (18)

These four functions are also Fourier transforms of each
other. The most important of the functions is the scatter-
ing function, \( R_s(\tau, \phi) \), which has received particular attention
in the analysis of sonar and radar systems. From the
assumptions above:

\[ R_s(\tau, \phi) = E \left[ |s(\tau, \phi)|^2 \right] \]  (19)

The scattering function can then be regarded as a power
density function. In underwater communication, passive
sonar, etc., it determines the average delay and frequency
spread the input signal will be subjected to in the medium.
In active sonar the scattering function can also be inter-
preted as the function which determines how the rever-
beration energy, in average, will be distributed in the
range and doppler plane. The scattering function must be
related to specific geometric and environmental conditions. It includes the influence of all coupling mechanisms between the system and the medium such as the transmitting and receiving beam-patterns. It also includes all medium parameters such as the distribution of inhomogeneities throughout the medium, boundaries, etc.

Under ideal propagation conditions the time and frequency resolution capability of a system is given by the two-dimensional correlation function of the transmitted signal:

$$ R_x(\tau, \phi) = \int_{-\infty}^{\infty} x(t) x^*(t-\tau) \exp(j2\pi\phi t) \, dt $$  \hfill (20)

The squared absolute value of this two-dimensional correlation function is called the signal ambiguity function, \( \chi(\tau, \phi) \). It may be considered as the mean squared output of a filter matched to a time shifted, \( \tau \), and frequency shifted, \( \phi \), version of \( x(t) \) when \( x(t) \) is applied to the input. The time and frequency resolution for the same system in a random inhomogeneous medium can be expressed by a new ambiguity function, \( P(\tau, \phi) \), expressing the combined effect of the signal and the medium. For narrowband signals it can be shown that the combined ambiguity function of the signal and the medium is obtained by convolving the ambiguity function of the signal with the scattering function of the medium:

$$ P(\tau, \phi) = \iint_{-\infty}^{\infty} \chi(\tau-\tau', \phi-\phi') \, R_S(\tau', \phi) \, d\tau' \, d\phi' $$  \hfill (21)

In an echo ranging system the medium can be treated as having one scattering function defining the delay-doppler distribution of the reverberation energy and another scattering function defining the two-way resolution loss of the target signal due to the propagation through the medium. The target can also be described by its scattering function which for a point target reduces to an impulse in the range-doppler plane. This situation can be modeled as a set of filters, fig 11. If a matched filter sonar receiver is assumed, the mean squared receiver signal is then distributed in delay and doppler according to the total system ambiguity function:

103
MODELS OF THE UNDERWATER ACOUSTIC CHANNEL

\[ P(\tau, \phi) = \int_{-\infty}^{\infty} R_{sr}(\tau, \phi) \chi(\tau-\tau', \phi-\phi') \, d\tau' \, d\phi' + \int_{-\infty}^{\infty} R_{sp}(\tau, \phi) R_{st}(\tau-\tau'', \phi-\phi'') \chi(\tau-\tau', \phi-\phi') \, d\tau' \, d\tau'' \, d\phi' \, d\phi'' \]

\[(22)\]

![Diagram of acoustic channel](image)

**Fig. 11 - Composite active sonar situation symbolised by a combination of system filters**

The model in Fig. 11 may be expanded to include multiple targets and scattering functions consisting of subsections equivalent to various types of scattering processes such as surface-, volume-scattering etc. The expressions for the mean square output of the matched filter sonar receiver are then composed of the same number of blocks. If all the random filters are totally incoherent, the total scattering function of the composite filter is then obtained by convolving the functions of the subfilters in series and adding the functions of the parallel channels as in eqn 22.

The assumption of a wide sense stationary process in time and frequency is rarely valid. In practice it may happen, however, that the scattering function can be approximated by a product of a delay scattering function, \( R_{ST}(\tau) \), and a frequency scattering function, \( R_{SF}(\phi) \):

\[ R_{ST}(\tau) \approx R_{ST}(\tau) \cdot R_{SF}(\phi) = \int_{-\infty}^{\infty} R_{S}(\tau, \phi) \, d\phi \int_{-\infty}^{\infty} R_{S}(\tau, \phi) \, d\tau \quad (23) \]

\( R_{S}(\tau) \) can be considered as the mean squared value of the impulse response, \( h(\tau, t) \), which no longer is a function of \( t \), and \( R_{SF}(\phi) \) can be considered as the mean squared value of the Bi-frequency function, \( B(f, \phi) \), which no longer is a function of \( f \). The true scattering function can not be
uniquely reconstructed from the knowledge of the two simplified functions, however, the assumptions do lead to some simple rules by which it is made possible to measure and obtain estimates of the scattering function, (4,7).

CONCLUSIONS

In this survey some of the modeling techniques of the underwater acoustic channel are treated. It is impossible in a survey like this to treat any of the models in detail. It is also impossible to discuss how and where the models can find their use. The various models and theories connected to the specific scattering or reverberation problems deserve a survey by themselves.

To-day much effort is concentrated on finding models with which one can predict the time and frequency characteristics of scattered waves associated with the surface and the volume. The work is concentrated on the theoretical and on the experimental side although the quantity from the latter is not impressing. Most of the experiments so far carried out concentrates on finding the macroscopic characteristics of the medium. Some newer experiments although they are few have been concentrating on finding the frequency and time spread of the acoustic channel or parts of the channels such as the surface. The hope is that these measurements lead to a confident construction of the medium scattering function.

The theoretical contributions to the field are numerous. The discrete point scattering model (5,6) not treated here is for example very attractive from a mathematical point of view. The model of the medium can in this case be expanded in considerable detail and very accurate approximations to the real system and the medium are believed to be obtainable. The degree of value of these models remain unanswered, however, as long as verifications through experiments in the real sea environment are not available.
REFERENCES

(6) V.V. Ol'shevskii, Characteristics of Sea Reverberation, Consultant Bureau, New York (1967)
THE APPLICATIONS OF SURFACE ACOUSTIC WAVE TECHNOLOGY IN ELECTRONICS

J.H. Collins  
Department of Electrical Engineering,  
University of Edinburgh

INTRODUCTION

Surface acoustic waves (SAW) in solids have their origins in the works of Lord Rayleigh. However, only in the past decade have electronic signal processing functions, notably frequency selectivity, been achieved in the VHF through microwave frequency range. Advantages claimed for SAW are small planar structures compatible with integrated circuit (IC) manufacturing techniques, having no subsequent adjustments, ruggedness and mass reproducibility. Today, certain SAW devices have reached product status. This paper surveys this situation and attempts to predict the future electronic systems applications for the technology.

THE INTERDIGITAL TRANSDUCER

The corner-stone of SAW technology is the interdigital transducer (IDT). It consists of a series of, photolithographically defined, interleaved metal electrodes on a smooth piezoelectric substrate, Fig 1. In the simplest case, when the width and spacing of these electrodes is uniform throughout, the stress output is proportional to N sin \( N \pi (f - f_0) \) where \( f_0 = \frac{V}{L} \), illustrating the frequency selective property of an \( N \) period IDT. Thus, electronic functions are performed basically by an electric generator, of impedance \( R_L \), driving an input IDT which causes a SAW to propagate to the output IDT, with subsequent reconversion into electric power in the load \( R_L \).

![Fig 1 IDT's in a SAW delay line](image)

The IDT capacitance to ensure maximum bandwidth for minimum insertion loss at \( f_0 \); and \( M_{opt} \) denotes \( W \) (Fig 1) in wavelengths which ensures then a 50 \( \Omega \) radiation resistance.

Although no ideal material yet exists, some general observations can be made. For low frequencies (<50 MHz) PZT is the cheapest material; temperature stability demands (ST, X) quartz; Bi\(_{12}\)GeO\(_{20}\) is favoured for long delay; A\(_2\)Ni/A\(_2\)O\(_3\) makes the least demand on standard photolithography; and finally (Y, Z) LiNbO\(_3\) is a good compromise choice which is enhanced by its low acoustic diffraction.
<table>
<thead>
<tr>
<th>MATERIAL</th>
<th>CUT</th>
<th>PROPAGATION DIRECTION</th>
<th>VELOCITY (km/sec)</th>
<th>K² (%)</th>
<th>N_opt</th>
<th>M_opt (in wavelengths)</th>
<th>PHOTOLITH* LIMIT (MHz)</th>
<th>ATTENUATION at 1 GHz (dB/ usec)</th>
<th>TEMP COEFF OF DELAY (ppm/°C)</th>
<th>MAIN AREAS OF APPLICATION</th>
</tr>
</thead>
<tbody>
<tr>
<td>LiNbO₃</td>
<td>Y</td>
<td>Z</td>
<td>3.485</td>
<td>4.30</td>
<td>4</td>
<td>108</td>
<td>580</td>
<td>1.6</td>
<td>85</td>
<td>WIDE BANDWIDTH DEVICES</td>
</tr>
<tr>
<td>Quartz</td>
<td>Y</td>
<td>X</td>
<td>3.173</td>
<td>0.22</td>
<td>19</td>
<td>53</td>
<td>530</td>
<td>4.1</td>
<td>-24</td>
<td>PULSE COMPRESSION</td>
</tr>
<tr>
<td>Quartz</td>
<td>ST</td>
<td>X</td>
<td>3.158</td>
<td>0.17</td>
<td>21</td>
<td>28</td>
<td>530</td>
<td>--</td>
<td>&lt;3</td>
<td>STABLE DEVICES, OSCILLATORS</td>
</tr>
<tr>
<td>Bi₁₂Ge₂₀</td>
<td>(100)</td>
<td>(011)</td>
<td>1.681</td>
<td>1.40</td>
<td>8</td>
<td>26</td>
<td>280</td>
<td>1.5</td>
<td>-122</td>
<td>LONG DELAY</td>
</tr>
<tr>
<td>PZT</td>
<td>POLED DIRECTION</td>
<td>ANY</td>
<td>2.200</td>
<td>2.40</td>
<td>6</td>
<td>--</td>
<td>370</td>
<td>6.0</td>
<td>--</td>
<td>TV FILTER</td>
</tr>
<tr>
<td>LiTaO₃</td>
<td>Z</td>
<td>Y</td>
<td>3.329</td>
<td>1.00</td>
<td>13</td>
<td>31</td>
<td>555</td>
<td>--</td>
<td>67</td>
<td>--</td>
</tr>
<tr>
<td>AaN/Al₂O₅</td>
<td>X</td>
<td>Z</td>
<td>6.120</td>
<td>0.61</td>
<td>11</td>
<td>60</td>
<td>1030</td>
<td>1.7*</td>
<td>44</td>
<td>INTEGRATED AMF'S, MICROWAVE LINES</td>
</tr>
</tbody>
</table>

* Based on finger-width limit of 1.5 μm
† at 37 MHz
* at 200 MHz

**TABLE 1**- SAW PARAMETERS OF SELECTED PIEZOELECTRIC MATERIAL CONFIGURATIONS
COMPUTER AIDED DESIGN AND FREQUENCY FILTERS

Significant computer-aided design (CAD) techniques have been developed (Hartmann, 7 and Mitchell, 9) for SAW frequency filters, spurred by their potential mass application in color TV receivers. Realisation of precision characteristics depends on the inherent IDT freedom of amplitude weighing in the W dimension and phase weighing in the L dimension (Fig 1). In the Mitchell approach each IDT electrode corresponds to a δ-function source. Fig 2 shows a typical CAD flow diagram. The first program (source synthesis) takes a frequency response and produces a truncated source distribution. This feeds a second program which designs the IDT structure. Further programs follow for acoustic reflection/piezoelectric interaction effects, and diffraction corrections. The CAD ends with a tape output which drives the pattern generator of the mask making machine. This chromium mask is used subsequently in the photolithographic definition of the IDT on the piezoelectric substrate. The degree of sophistication which SAW frequency filters have reached is well illustrated by experimental results for a British specification color TV filter shown in Fig 3.

SAW frequency filters are now being contemplated for both analog 2700 channel and PCM 32 channel communication links. For the former the following performance seems feasible:

- IF = 140 MHz; insertion loss <8 dB; passband width ±15 MHz (-1 dB), ±25 MHz (-3 dB);
- passband ripple <0.1 dB (±15 MHz); stopband attenuation >50 dB (at ±40, ±80 MHz);
- and group delay variation <0.1 nsec (±15 MHz).

Fig 2 CAD flow diagram

Fig 3 Color TV filter with British specification
SAW DEVICES AND MAJOR AREAS OF APPLICATION

Table 2 summarises the current status of basic SAW devices, their SAW device derivatives which involve either combination with other SAW devices or IC’s, and the major areas of systems applications. The reference list provides information on all aspects listed in Table 2. In this article attention is confined to a few important topics such as at the device level - the oscillator, the tapped delay line, the PSK matched filter; and at the systems level - high resolution radar, spread spectrum communications and acoustic imaging.

OSCIllATORS

Two classes of oscillators are in common use, the quartz crystal oscillator and the LC oscillator (including cavity oscillators). The former is of high stability (Q>10^6) but suffers from a number of disadvantages including mechanical fragility, low fundamental frequency operation (<50 MHz), and limited frequency modulation, FM, capability (500 ppm). In contrast, the LC is much less stable (Q<10^6) but superior FM wise. Recently the SAW oscillator has appeared on the scene (Lewis 8,11 and 12) having an intermediate stability and modulation capability and practical advantages over both. It is believed that those properties will make the SAW oscillator significant in the 20 MHz to 2 GHz frequency range as local oscillators in frequency synthesisers and telemetry.

The basic SAW oscillator comprises a SAW delay line (Fig 1) with a transistor amplifier as the feedback element. Providing there is net loop gain, oscillation will occur at a frequency, f, satisfying the condition

\[ \frac{2\pi f L}{V} + \Phi = 2n\pi \quad (n = \text{an integer}) \]

where the SAW path length, L, is the center-center spacing of the IDT’s, and \( \Phi \) is the electrical phase shift introduced by the amplifier and IDT’s.

Thus, the frequency of operation is determined by the transducer pattern and not by a dimension of the quartz crystal. Hence, the crystal is rugged and can be firmly bonded to the header. This good thermal contact enables the loop to oscillate up to 1 watt power levels. Other significant performance figures for the SAW oscillator are: frequency deviation up to 2%; short term stability of <1 in 10^8 for 1 second; medium term stability of 100 ppm for a temperature excursion of ±40°C; and single sideband FM noise of -130 dB per Hz, 10 kHz away from carrier. Limits to the long term stability still need establishing but 1 part in 10^5 over 1 year is readily achievable. Fig 4 is a photograph of an early RRE oscillator. Amplifier and delay line are housed in separate TO5 cans.

Fig 4 Photograph of RRE 720 MHz SAW oscillator
<table>
<thead>
<tr>
<th>SAW DEVICE</th>
<th>BASIC</th>
<th>DERIVATIVES</th>
<th>MAJOR AREAS OF APPLICATION</th>
</tr>
</thead>
<tbody>
<tr>
<td>INTERDIGITAL TRANSDUCER</td>
<td>FREQUENCY FILTER</td>
<td>Color TV, Integrated IF's; Satellite Multiplexers; ECM</td>
<td></td>
</tr>
<tr>
<td></td>
<td>DISCRIMINATOR</td>
<td>Analog FM Demodulator</td>
<td></td>
</tr>
<tr>
<td></td>
<td>OSCILLATOR</td>
<td>Stable Sources VHF to Microwave; Specialised Comm's</td>
<td></td>
</tr>
<tr>
<td>DELAY LINE</td>
<td>TAPPED</td>
<td>Radar Clutter Reference; ECM; IFF/SSR</td>
<td></td>
</tr>
<tr>
<td></td>
<td>BIT MATCHED FILTER</td>
<td>Digital Comm's PSK Demodulator</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ACoustIC AMPLIFIER</td>
<td>Long Delay for TV Frame Storage</td>
<td></td>
</tr>
<tr>
<td></td>
<td>PSK MATCHED FILTER</td>
<td>Spread Spectrum Comm's; ICNI</td>
<td></td>
</tr>
<tr>
<td></td>
<td>PN GENERATOR</td>
<td>Link Analysis; Spread Spectrum Comm's</td>
<td></td>
</tr>
<tr>
<td></td>
<td>DIGITAL</td>
<td>Recirculating Memory; Data and Bandwidth Compression; Sonar</td>
<td></td>
</tr>
<tr>
<td></td>
<td>TOUCH SENSITIVE</td>
<td>Glass Protection; Computer Controlled Graphics</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Signal Routing</td>
<td></td>
</tr>
<tr>
<td>MULTI-STRIP COUPLER</td>
<td>VARIABLE DELAY LINE</td>
<td>High Resolution Radar; Navigation; Group Delay Equalisers</td>
<td></td>
</tr>
<tr>
<td>DISPERSER</td>
<td>CONVOLVER</td>
<td>Range Calibration; Doppler Extraction; Electronic Timing; ECM</td>
<td></td>
</tr>
<tr>
<td></td>
<td>FFT PROCESSOR</td>
<td>Spread Spectrum Comm's; Radio Astronomy</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ACoustIC IMAGING</td>
<td>Link Diagnosis; Sonar</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ACousto-OPTIC</td>
<td>Non-Destructive Testing</td>
<td></td>
</tr>
<tr>
<td>WAVEGUIDE</td>
<td></td>
<td>Displays, Cameras.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ACOUSTIC IC's</td>
<td>Subminiature Signal Processing Sub-systems</td>
<td></td>
</tr>
</tbody>
</table>

Key: PSK, Phase Shift Keyed; ECM, Electronic Countermeasures; MTI, Moving Target Indicator; IFF/SSR, Interrogate Friend or Foe/Secondary Surveillance Radar; ICNI, Integrated Comm's Navigation and Identification

TABLE 2 - SAW DEVICES AND MAJOR AREAS OF APPLICATION
TAPPED DELAY LINES

One of the major advantages cited for SAW's over their bulk acoustic wave counterparts is the ability to display spatially a signal in real time, and permit it to be sampled arbitrarily and, if necessary, modified. Another important feature is the linear analog nature of the signal processing. The upper part of Fig 5 depicts a SAW tapped delay line (TDL) in which broadband, weakly coupled, IDT's are used for the taping function in a linear array. In the past year two significant applications for simple TDL's have been demonstrated at the Royal Radar Establishment (RRE), namely to provide pseudo-coherence in clutter reference pulse Doppler radars (Brown et al, 11) and the reduction of garbling in secondary surveillance radar (SSR) used in Air Traffic Control Systems (Moule, 11). These applications will now be discussed.

It is not always economic to produce fully coherent radars capable of extracting moving target information (MTI). This is particularly true in man-pack clutter reference radars. The requirement is to add signals from a target to those from clutter, within half a pulse length of the target. Then, movement with respect to clutter produces a Doppler frequency signal which is extracted by a box-car integrator to give the MTI. Unfortunately, at short range (several km) the clutter is often minimal thus degrading performance. Reflecting SAW TDL's at the radar IF are being used to guarantee an internal reference clutter. Design requires that the reflections are matched to the inverse of the IF amplifier swept gain law. Application to this pseudo-coherent system has particular significance for SAW, since it is a clear example of a situation in which ease of design of the SAW TDL is affecting the overall systems design. The evolution of such systems is currently incomplete but it will be interesting to note whether similar techniques are applied to medium and long range radars.

In SSR a ground-based transmitter sends out a pair of microwave interrogation pulses that are recognised by the transponder in the aircraft. Then, a coded reply for aircraft identity is emitted extending over 20.3 μsec. The reply consists of a sequence of up to 15 pulses each of 0.45 μsec duration and spaced 1.45 μsec apart. The transponder reply is electronically primitive in that amplitude modulation is used, the pulses are non-coherent and no rigid specification is imposed on oscillator stability. Unfortunately, in overloaded situations when aircraft are close together (<3km) the replies overlap in time (garble) and identification proves difficult.

Fig 5 Top shows concept of SAW TDL. Lower is shown impulse response of PSK analogue matched filter.
Pairs of suitably interconnected SAW TDL's, with electronically switched taps, because of their property of linear signal processing when added to conventional SSR can reduce this garbling problem. They act as IF (70 MHz) matched filters to the amplitude modulated waveform providing correlation at the ground station between the incoming signal and a preset code, corresponding to the required aircraft. In principle, correlation increases the ability to resolve overlapping signals from that of the pulse train length (3 km) to that of a pulse length (70 m). Basic laboratory experiments with SAW TDL's have been successful and fully operational trials are eagerly awaited. However, this enthusiasm must be tempered by the poor autocorrelation and crosscorrelation functions of the aircraft codes which will certainly add complexity to the overall electronic system design.

PSK MATCHED FILTERS

As stated earlier SAW TDL's form the basis of correlating amplitude modulated waveforms which are non-coherent. However, it is possible to include phasing effects as well and build matched filters whose correlation peak contains less ambiguity due to lower time sidelobes. One example is the phase shift keyed (FSK) matched filter.

The structure is shown in the top of Fig 5. It is now relevant to consider the interconnections of the tapping IDT's to the output bus-bar, so that the impulse response is the PSK waveform shown in the lower part of Fig 5. When an impulse is applied to the left hand IDT, a SAW packet with a corresponding number of periods, N, is produced. This packet is sampled subsequently in turn by the linear array of taps. When the taps are spaced by a distance corresponding to the duration of the wave packet, the SAW device expands the impulse into a continuous PSK waveform where the phase of the signal from an individual tap is determined simply by the polarity of the

---

Fig 6 Generation and correlation of coded waveforms: a — impulse response of a SAW PSK filter (f_p = 127 MHz, bit rate = 5 MHz); b — impulse response of a second filter coded to give the time-reverse waveform; c — output of the second filter when fed with waveform a. The waveforms correspond to the 13-bit Barker code and its time-reverse. Both devices use Si-quartz substrates.

Results due to Darby, Univ of Edinburgh

Fig 7 Programmable PSK matched filter manufactured by Rockwell International, USA. Monolithic integration on single crystal sapphire (Al2O3) of aluminium nitride (AlN) for the SAW functions and silicon for the microelectronic LSI control, switching and memory functions is used.
connections shown in the top of Fig 5. The bit rate of this waveform is given by \( f_o/N \). Fig 6 shows typical results for a SAW PSK matched filter operating at \( f_o = 127 \text{ MHz} \) and a bit rate of 5 MHz. There are 13 taps in the filter for purposes of correlating a 13 bit Barker sequence (3). To date, PSK matched filters have been built with up to 1023 taps and up to 50 MHz bit rates.

The PSK matched filter can be made programmable by using an electronic switching network between the taps and summing bus to control the phase contribution of individual taps, in accordance with a binary code stored in an associated shift register. Reprogramming to a new code is accomplished simply by reading the desired code from a read-only-memory (ROM) which stores a library of possible codes. Switching and summing of the rf signals requires a high performance switching network, in contrast to the remaining logic circuitry which simply provides DC levels except when re-programming. In certain applications there are, however, requirements for a fast reprogramming capability, necessitating consideration of both switching network and code store transient speeds. An advanced programmable PSK matched filter is shown in Fig 7. The structure has 127 taps and a bit rate of 20 MHz. It is fully monolithic and has the tremendous advantage of using silicon-on-sapphire technology for compactness and isolation of the numerous diode circuits. This filter is reprogrammable within the bit period (50 nsec).

HIGH RESOLUTION RADAR

Many modern pulsed radars spread the energy of the transmitted signal in time so that the peak power handling capacity of the transmitter is not reached before the mean power limit. This approach gives unacceptable range resolution since the ability to resolve the range of two targets depends on the bandwidth and hence the duration of a simple transmitted rf burst. This dilemma between energy and resolution is elegantly overcome by encoding the waveform on transmission, using either linear frequency modulation or a PSK waveform, and decoding it with a matched filter so that the received energy is compressed to a time interval short enough to give the required resolution.

The block diagram of a simple pulse compression radar is shown in Fig 9. The matched filter, optimizes the peak signal power to mean noise power but produces unacceptably high time-sidelobes which, for linear fm (chirp) are minimized using a spectral weighting filter before detection. Pulse length, \( T \), is typically in the range 1-100\( \mu \text{sec} \), and the bandwidth, \( B \), is 1-500 MHz. SAW dispersive devices based on quartz and lithium niobate are admirably suited to these parameters. The chirp signal can be 'actively' generated using conventional electronic circuits. Alternatively, the amplified impulse response of SAW coded delay lines allows flexible 'passive' generation of complex waveforms.

Two distinct techniques, Figs 8a and b, have been developed for SAW dispersers. In Fig 8a, the acoustic impulse generated in the interdigital transducer (IDT) A, when a sharp electrical impulse is
EIGHTH INTERNATIONAL CONGRESS ON ACOUSTICS, LONDON 1974

Fig 8a SAW dispersive delay line utilizing a dispersive transducer input and wideband output transducer.

applied, exhibits a frequency change from \( f_h \) to \( f_l \) over a time duration, \( T = T_L - T_h = \) transducer length/SAW velocity. Providing the output IDT is of sufficiently wide bandwidth the electrical impulse response follows the amplitude and phase excursion of the acoustic signal. However, IDT, B exhibits a change in insertion loss over the useful bandwidth, and the variation in path length with frequency also introduces amplitude changes through acoustic diffraction. These effects are overcome by CAD based on electrode weighting in IDT A. This technique is further extended to include the effects of matching networks and to incorporate control-led spectral weighting.

Fig 8b SAW reflecting array compressor.

A lower loss situation is achieved by producing two complementary dispersive transducers. However, CAD complications result since the acoustic impulse created under one dispersive IDT propagates under a second dispersive IDT while generating the device impulse response. The associated convolution process creates amplitude and phase errors which lead to spurious sidelobes in the compressed pulse waveform. Marked reduction in these spurious effects can be achieved by inclining the two transducers in a complementary fashion. In-line dispersive structures suffer further disadvantages when high frequencies pass under interdigital regions designed to operate at lower frequencies, since Rayleigh wave energy is converted to bulk acoustic modes which distort the device impulse response. In the inclined structure this effect is minimised.

A significant new technique (Williamson, 11) called the reflecting array compressor, RAC, is shown in Fig 8b. Here, a wideband input IDT generates a signal which passes under an array of reflective slots. The spacing and depth of these slots is chosen to give frequency selectivity \( (f_1, f_2, \text{and} \ f_3) \) and spectral weighting. The first array of slots are generated by a second group which bring the dispersed signal to the wideband output IDT. The reflection angle is determined by the anisotropic velocities of the incident and reflected acoustic waves. The SAW frequency selectivity is of particular advantage in dispersers which employ large fractional bandwidths since it eliminates spurious bulk acoustic modes. This is illustrated vividly by results on a RAC device where \( T = 10 \mu \text{sec} \) and \( B = 512 \text{ MHz} \).
Fig 9  Block diagram of a simple high resolution radar employing linear FM pulse compression

Fig 10  Spread spectrum communications system
To achieve only slight sampling of the wave at each frequency slots must have a depth in the region of one thousandth of an acoustic wavelength, be approximately one half wavelength wide, and spaced by one wavelength between centres. Although the greater periodicity compared with the dispersive IDT is advantageous from the viewpoint of pattern definition, this is negated by the small slot depths which are only 30 Å at 1 GHz. Workers at MIT Lincoln Laboratory have shown that controlled ion-beam etching gives satisfactory characteristics.

The majority of pulse compression radar systems, Fig 9, in military use deploy TB products up to a few hundred, with bandwidths less than 100 MHz. Table 3 illustrates the performance of certain engineered subsystems employing conventional printed circuit board electronics (Butler et al., 12). A double sided enclosure for the subsystem allows the maintaining of the pair of SAW dispersers back to back. Then any ambient temperature variations result in a minimum mismatch. SAW dispersers are degraded by dirt or condensation and require hermetic encapsulation. Suitable partitions in the enclosure allow isolation between circuit blocks and cover plates have NFP gaskets.

The design of the peripheral electronics is as significant as the design of the SAW dispersers with their matching networks. This is illustrated by the design of the 250:1 TB subsystem, column 4 of Table 3. In the transmitter the impulse consists of two rectangular pulses of 14 volts peak amplitude and 15 nanoseconds total duration. The CW loss in the chirp generator is 38 dB and the output amplitude is 4 mV peak to peak into 50 Ω. An overdrive of 2 dB in the limiter is needed to meet the amplitude specification of ±0.2 dB ripple. The post expansion amplifier has 49 dB gain and a noise figure of 4 dB,
giving an output of 0 dBm with 40 dB signal to noise. Detection linearity is required over a 35 dB range with a deviation not exceeding 2½ dB. This is achieved by 40 dB amplification of the compressed pulse to give 18 volts peak to peak across the 500 Ω load of the hot carrier diode detector.

Thus, SAW technology offers designers of high resolution radar significant advantages over other technologies in the controlled, yet flexible, passive generation and compression of complex coded waveforms. The deployment of thick or thin film hybrid integration will result in substantial reduction of package volume. Reproducible 40 dB sidelobe performance and TR's exceeding $10^4$ seem on the horizon.

SPREAD SPECTRUM COMMUNICATIONS

A spread spectrum (SS) system is a communication link employing a bandwidth many times larger than the digitised information bandwidth in order to achieve security, multipath rejection, low detectability, and accurate ranging whilst maintaining near optimum sensitivity and channel utilisation capability. SAW matched filters, both linear fm based on the RAC (Bush et al, 10) and fixed and programmable pseudo-noise coded PSK (Setrin et al, 10) have been developed recently for such communication applications. Interest in SAW devices has been stimulated by their asynchronous linear signal processing capability, flexible and rapid code generation and detection capability for frequency hopping, burst (as well as) continuous mode operation. The objective here is to describe the principles of SS systems, the status and limitations of the deployment of SAW devices and the likely trends in the future.

Fig 10 shows the fundamentals of a SS communications system (Hunsinger, 11). In the transmitter a wideband (10-500 MHz) signal synthesiser generates a characteristic phase and/or FM waveform - the transmitter signature. This is the role presently played by micro-electronic circuits which the SAW devices, listed above, may replace. The synthesiser output is phase modulated by the input data (typically 100 k bits/sec, upconverted, and then transmitted by the antenna. The receiver consists of an identical wideband synthesiser, a multiplier, an integrator and a synchronizer. The synthesizer in the receiver is adjusted to provide an output identical to that emitted from the transmitter synthesiser, but delayed by the propagation time. The synthesiser output is multiplied by the incoming signal and the product integrated over one bit time - the correlation process. The sign of the output voltage then determines whether a data bit 'l' or '0' was transmitted. However, it is first necessary to establish the correct phase of the receiver wideband synthesiser. This currently requires a time ranging from a fraction of a second to as high as a minute. SAW devices promise to reduce this search time by a factor proportional to TB.

Interference sources and multipath signals, due to ground or sea reflections, are discriminated against by this multiplication and integration process. This results in a signal to noise ratio (SNR) improvement equal to TB, which is proportional to the ratio of signal

118
bandwidth to data bandwidth. This SNR improvement is quantified for current SAW devices as being in the range 20-37 dB. It is important to realise that an intentional jammer, unaware of the transmitter signature, must transmit sufficient power to overcome this SNR improvement. Naturally, the closer the jammer is to the receiver, the easier is the task. This 'near-far problem' of SS systems also arises in multiple access systems employing continuous transmissions. Fortunately, systems deploying a satellite have the inherent advantage of building-in the required power systems doctrine. For terrestrial SS mobile communication systems the burst transmissions available through the asynchronous property of SAW devices allow signal formats suitable for a time division multiple access capability to assist the 'near-far problem'. In contrast, the conventional frequency division multiple access (FDMA) system succeeds through orthogonality of transmissions. However, security considerations dictate the use of baseband cryptography on each transmission. An interesting area of research in the civil field is the possibility of overlaying a SS communications link based on SAW devices over an FDMA link for emergency uses.

The latter system will depend on another SS concept which allocates specific time coded and frequency hopped pulse patterns to address each transmission (Grant et al, 7). This random access discrete address technique transfers data by further modulation of the address, for example, through quantised pulse position modulation. There complex address patterns are easily generated by impulsing SAW PSK matched filters, suitably encoded.

Another attractive feature of SS communication systems is their automatic ranging capability (Burnsweig, 7), without interruption to communications. The range is calculated by comparing the receiver synthesiser phase to the transmitter synthesiser phase by a return path link. The range resolution is determined by the system bandwidth. The target velocity can also be obtained by taking the time derivative of the range calculations. However, it is necessary to make the signature repeat time sufficiently long so that the receiver can establish the actual range without ambiguity. This concept is being pursued in a military ATC system termed ICNI, which stands for Integrated Communications Navigation and Identification.

AIR TRAFFIC CONTROL

The ATC system of the world's airways is a complex integration of airborne and ground based equipments, procedures and facilities embracing many different technologies. The prime objectives of the ATC system are to provide for safe, expeditious and economic operation of aircraft movements (Clark, 11). The four major disciplines involved in the integrated ATC system are: data acquisition; data processing; navigation, guidance and control; and communications. With the seemingly continuous growth in air traffic movements (it was predicted before the energy crisis that 2000 aircraft will simultaneously be airborne in the Los Angeles basin by 1990) the pressure to automate ATC systems has become acute. To this end, several major ATC system developments are under active consideration for deployment within the next 20 years, namely: automatic air-ground data links; aeronautical satellite communications (eg Aerosat for the North Atlantic, the
collision avoidance surveillance (CAS) ranging system), and microwave landing systems.

It is of relevance to enquire how SAW devices may impact this area. Existing ATC systems can only accommodate add-ons - one cannot stop the world! The SAW TDL matched filter, described earlier is an excellent example for SSR. Current VHF and UHF communications equipments will use SAW frequency filters and decoders providing they are cost-effective.

However, SAW devices have a real opportunity to influence the thinking on future ATC systems. Specific examples, in Aerosat are for: automatic, high integrity, position reporting from aircraft inertial navigation systems using a 4-signature SS technique (Parker, 11); and satellite ranging of aircraft which overcomes sea multipath (Burnsweig, 7). In CAS programmable PSK matched filters promise to be used for generation and identification of aircraft signatures, in a non-cooperative system centred on the powerful SS time/frequency technology. In military ICN1, as observed earlier, a clear application exists for similar filters. Also, in military SSR, called IFF (Table 2), SS techniques employing SAW matched filters have received considerable emphasis for new equipments (Setrin et al and Bush et al, 10). As a note of caution, bandwidth limitations to a few MHz will probably force the large TB's to be realised with delays in excess of 100 μsec. Thus, SAW waveguides (Mason, 11), which are still in their research infancy, may find a significant application area.

An important operational requirement which technology has not yet been able to solve is that of Clear Air Turbulence. Investigations using optical infra-red radiometry techniques have to date suffered from a high rate of false alarm. The unique integration and correlation, and other signal processing techniques available with SAW devices could be profitably translated to optical components to solve this problem. This may prove to be an application for the emerging SAW area of optical imaging techniques?

DISPLAYS

Important developments have been made recently in Displays using SAW devices. Techniques employed have been electronically scanned acoustic imaging; electronically scanned optical imaging based on acousto-optic interactions; and the touch sensitive digitiser which has become a commercial instrument. The principles of acoustic imaging are shown in 1-dimensional form in Fig 11 (Havlice et al, 14). The goal is to focus on an acoustic image object plane at a prescribed distance from the detector, without an acoustic lens and then scan the object at MHz rates. A large number of piezoelectric detectors are used to detect an acoustic image signal from an object illuminated separately, either by a low frequency acoustic wave passing through it or reflected from it. A SAW TDL is used as the basic scanning device. There is one tap for each detector. The output signal from each tap is mixed in a diode mixer with the output from the corresponding piezoelectric detector. The pulsed signal in the delay line is at frequency $f_1$, and the signal corresponding to the acoustic image is at frequency $f_2$. Individual mixers give signals at the sum frequency.
f_s + f_1 and the difference frequency f_s - f_1. One of these frequencies
is rejected by conventional electrical filtering. Unique experiments
at Stanford University (Havlice et al, 14) have been conducted with
f_s = 4 MHz, f_1 = 50 MHz, and 50 SAW taps and piezoelectric detectors. The
mixer outputs are used to intensity modulate a CRT, and hence display
a visual image corresponding to one line of the acoustic image. The
pulse in the SAW TDL thus acts like the scanning electron beam in a
vidicon.

In the 2-dimensional version, a square array of piezoelectric
detectors is used with an X-scan SAW TDL (f_1) and a Y scan SAW TDL(f_2).
A similar mixing procedure produces an output at f_1 + f_2 + f_s, when
both SAW pulses are effectively passing a particular detector element.
This corresponds to scanning a line at 45° to the X axis. By delaying
one pulse with respect to the other, a complete raster scan may be
obtained. Normally a lens would be necessary for focussing the object.
However, this device unlike camera film retains phase information
besides amplitude information. By correct choice of the signal sent
down the delay line, namely a linear FM chirp, electronic focussing
without an acoustic lens has been achieved in 1-dimension with mech-
anical scanning in the other dimension. 2-dimensional electronic
scanning seems on the horizon promising significant application for
high speed scanning in non-destructive testing. Another major
advantage is that M^2 resolvable spots can be obtained using only 2M
SAW taps.

Similar SAW techniques for converting an optical image to an
electrical signal have also been demonstrated (Mo11 et al, 15). The
principle relies on the scanning of optically induced conductivity
variations in silicon-on-sapphire films spatially adjacent to lithium
niobate, propagating a short SAW read pulse. Scanning of the conduc-
tivity is accomplished by means of the non-linear interaction
previously used in the convolver (Kino et al, 6). The scanning pulse
propagates through a relatively long contradirected reference pulse
and perturbs its amplitude by an amount depending on the film con-
ductivity in the region of pulse overlap. Thus, a replica of the
conductivity variation is imposed
on the reference pulse that is
subsequently used to intensity
modulate a CRT.
EIGHTH INTERNATIONAL CONGRESS ON ACOUSTICS, LONDON 1974

CONCLUSIONS

Certain SAW devices have made the transition out of the research laboratory to product engineering status. These SAW devices are now seriously influencing the thinking of electronic systems design groups. For example, their incorporation in military radar systems is a measure of their competitive performance, reliability, reproducibility and cost effectiveness. However, their potential breadth of application as delineated in Table 2 is much wider. In the next phase the mass production of the color TV filter, the widespread communications utilisation of the SAW oscillator, the deployment of the SAW TDL as on SSR decoder in Air-Traffic Control Systems, and acoustic imaging for Non-Destructive Testing can be predicted with confidence.

The intrinsic features of SAW that guarantee its future may be summarised as follows:

- spatial display and controlled sampling of real time waveforms
- passive linear signal processing at high dynamic range (>50 dB)
- IF bandwidth covering VHF, UHF and low microwave frequencies
- planar, simple, IC compatible manufacturing techniques
- underlying physics well understood so that CAD allows demanding specifications to be met
- radiation hardness coupled with graceful degradation of performance

Finally, it is noteworthy that the early research in SAW was conducted by less than 100 people. Although not a stand-alone technology its leverage in the marketplace promises to be large. One forecaster has stated that by 1980 over £50 million per annum of electronics business will be SAW dependent!

ACKNOWLEDGEMENTS

The author thanks IPC Science & Technology Press Ltd, and Microwave Systems News, and M Lewis for certain graphics included in this article.

REFERENCES

These references serve as both overviews of the evolution and the detailed implementation of SAW technology.

EIGHTH INTERNATIONAL CONGRESS
ON ACOUSTICS, LONDON 1974

L'INTELLIGIBILITE ET CERTAINES DONNEES PSYCHOPHYSIOLOGIQUES
SOUS-ESTIMEES

R.G. Busnel * Laboratoire de Physiologie Acoustique,
INRA, Jouy-en-Josas, France

INTRODUCTION

L'intelligibilité de la parole est le thème de nombreuses recherches
sur lequel oeuvrent surtout des ingénieurs et des phonéticiens.

Ce problème passionnant, complexe en soi, nous paraît encombré d'in-
terprétations différentes au sein d'intérêts disciplinaires parfois
très divergents, et qui se sous-estiment souvent au travers de leurs
méthodes de mesure. Nous formons le vœu que cet article soit con-
sidéré comme une approche de conciliation.

Récemment, en France, le Ministère de l'Environnement (1973) a confié
à un ingénieur le soin de faire une synthèse des données actuelles,
et ce choix montre bien que la tendance est orientée vers la techni-
que, même si, en définitive, les épreuves se terminent par un sujet
humain qui fournirait la réponse, et que délibérément on prenne celui-ci
pour un individu standard ayant des caractéristiques statistiques, ce
qui avec une certaine logique est relativement justifiable. Or, les
résultats sont si fréquemment divergents, que le Président du Groupe-
ment des Acousticiens de Langue Française a pu dire, avec quelques
raisons que nous accepterons provisoirement, que l'idéal consisterait
à remplacer le sujet humain par une machine objective. Car, en effet,
jusqu'à présent, malgré toutes les précautions techniques dont s'en-
tourent ceux qui se préoccupent de ce problème, ils sont, en bout de
chaîne, obligés d'utiliser des mesures subjectives ; c'est à ce ni-
veau que nous pensons qu'il est nécessaire de faire part de données
qui relèvent du domaine de la biologie ** et qui ne semblent pas être
intégrées dans ce genre de problème par ceux qui l'étudient fonction-
nellement.

Toutefois, il nous paraît nécessaire ici de bien rendre à César ce qui
est à César.

Les ingénieurs des télécommunications ont les premiers, en acoustique,
attiré l'attention et fait des recherches sur ce problème de l'intelli-
gibilité (Campbell, 1910 - Chavasse, 1962). Leurs nombreux travaux sur
les caractéristiques du message parlé, l'adaptation des moyens de
transmission aux possibilités de la perception périphérique ne sont
pas délibérément minimisés ou négligés dans notre analyse. Nous n'ou-
blions pas, en effet, qu'ils ont créé la théorie metrologique dont

* Laboratoire de Physiologie Acoustique-INRA-ERA n°332 du CNRS et
EPHE.Jouy-en-Josas, 78350 France.

** on voudra bien entendre par ce terme toutes les sciences biologi-
ques dont la physiologie, la neurologie, la psychophysiologie etc...
L'INTELLIGIBILITE ET CERTAINES DONNEES PSYCHOPHYSIOLOGIQUES
SOUS-ESTIMEES

l'usage, entre autres, pour la mesure de la surdité, est d'un considé-
vable intérêt. Ils ont bien perçu, et Chavasse parmi eux, la notion
différentielle qui doit théoriquement séparer netteté, d'intelligibi-
ité ; l"intelligibilité" étant la compréhension des idées qui se
mesure avec des phrases et des mots significatifs et diffère de la
" netteté" qui se réfère à des associations conventionnelles de sons
ou de mots dépourvus de sens (logatomes).

Pour l'ingénieur (L.Pimonow, communication personnelle), l'intelli-
gibilité est exclusivement relative au courant d'information et à ses
variations au travers des systèmes de transformation et de transmis-
sion, par exemple, en téléphonie par le microphone, la ligne, l'écou-
teur) dont la mesure (ou le contrôle) doit être considérée comme
constante, bien qu'elle se fasse au travers d'un récepteur humain,
qui ne devrait, théoriquement, " que mesurer " le courant d'informa-
tion, sans pouvoir l'influencer.

Si en technique, et nous devons admettre ce point de vue de l'ingé-
nieur, le fait que les propriétés psychophysiologiques du récepteur
sont à la base de la mesure et qu'il est très difficile qu'il les
domine pour qu'elles restent constantes, nous devons donc tenir compte
de ces propriétés pour, peut être, éliminer leurs variations dans les
études technologiques. Nous admettrions très volontiers que la notion
d'intelligibilité au travers du récepteur soit donc, par essence,
différente de celle qui ne se rapporte qu'au courant d'information,
et ce sera cet aspect de la question qui sera plus spécialement trai-
té ici. Il y a donc une nette distinction à établir entre le système
de mesure (dans notre cas, le récepteur) et le phénomène qu'on mesure,
le courant d'information.

Pour bien poser le problème, il semble utile de le définir dans le
schéma simplifié n°1 qui représente un canal d'information entre une
source et un récepteur biologique.

L'intelligibilité du domaine de l'ingénieur concerne des mesures de ce
qui se passe dans le canal ; même si elles sont actuellement faites
par un récepteur biologique, supposé à réponse stable, pris comme ins-
strument métrologique, on doit tendre à définir un modèle mathématique
et du canal et du récepteur et, à la limite, remplacer celui-ci par
un calculateur. Nous pensons que ceci sera réalisé dans un avenir pro-
che.

Dans notre schéma, le récepteur biologique est formé d'un ensemble.
Celui-ci comprend un premier étage de transducteurs (oreilles externe,
EIGHTH INTERNATIONAL CONGRESS
ON ACOUSTICS, LONDON 1974

L'INTELLIGIBILITE ET CERTAINES DONNEES PSYCHOPHYSIOLOGIQUES
SOUS-ESTIMEES

moyenne, et interne) déjà très complexe, mais néanmoins à partir du-
quel on peut faire des mesures relativement reproductibles. Cet étage
pourrait être appelé le canal de la physiologie acoustique. Il est re-
lie à l'étage central que l'on peut assimiler à un ordinateur, mais
dont les propriétés fonctionnelles et les programmes, qui sont du do-
maine de la neurologie, ne sont que très imparfaitement connus. C'est
au travers de l'ensemble des parties II et III que sera définie l'in-
telligibilité qui relève des intérêts du biologiste. Elle est d'autant
plus difficile à mesurer que l'étage central n'est pas directement
accessible, mais seulement via le premier étage ; ses propriétés ne
sont pas complètement définies, elles sont éminemment variables, dans
des limites actuellement encore imprécises, et il sera ardu d'établir
un modèle mathématique de cet ensemble, bien que ce soit la tendance
souhaitable.

AUDITION ET INTELLIGIBILITE DE LA PAROLE

Si la perception de la parole passe obligatoirement par le système
auditif, il nous semble nécessaire de mettre en évidence que l'intelli-
gibilité est un phénomène dans lequel la réception auditive périphéri-
que et la physique du signal n'ont pas l'importance que leur attri-
buent les spécialistes de l'audition ou les phonéticiens, et que les
mécanismes de base sont essentiellement d'origine centrale.

Sur quoi peut-on se baser pour appuyer ce point de vue ?. D'abord sur
des phénomènes banaux, classiques, bien qu'un peu trop oubliés ou sous-
estimés : - Le langage étant la résultante d'un apprentissage et donc
phénomène entièrement acquis, si la compréhension, c'est-à-dire l intel-
ligibilité passe par l'audition, elle n'en est pas complètement dépend-
dante.*

.../...

* La mesure de la fidélité ("speech fidelity" de Nakatani, par exemple),
de l'articulation, référent du même problème, bien qu'en étant plus
eloignés. Ils traduisent néanmoins, nous le pensons, des mesures com-
paratives avec un stock de données antérieurement mémorisées.

** Si on se reporte aux références de la langue, au travers des diction-
naires, on s'aperçoit d'ailleurs de ce genre de dualités et donc d'amb-
biguité. Ainsi, on trouve dans le grand Larousse : intelligibilité :
terme caractérisant la compréhension d'un message. Mais dans le Quillet
Philos : qui ne peut être connu que par l'intelligence, l'entendement
et non par les sens. On y trouve aussi, à la suite des phrases classi-
ques " parler de façon peu intelligible, ou parler à haute et intelli-
gible voix ":" qui peut être perçu directement par l'ouie ". Sans
compter que le mot "entendement" n'a rien à faire avec entendre :
"faculté par laquelle on comprend". Ne pas oublier non plus :"ils en-
tendent mais n'écoutent pas".
L'INTELLIGIBILITE ET CERTAINES DONNEES PSYCHOPHISIOLOGIQUES 
SOUS-ESTIMEES

On rappellera, parmi les travaux français, la thèse de R. Lehmann (1962) qui montre bien que l'intelligibilité des chiffres et des phrases est plus grande que celle des mots, et très supérieure à celle des logatomes (10 à 20 dB quand il s'agit des niveaux d'émission par exemple).

Les résultats donnés par les logatomes ne sont certainement pas intéressants et peuvent probablement être acceptés comme valables pour des besoins techniques en mesurant la netteté en télécommunication, mais pour nous, ils ne font pas autre chose que de traduire la capacité centrale de raccrocher les sons des logatomes à des sons du langage connu ; ils sont donc fonction de nombreuses variables dont, par exemple, la somme des connaissances linguistiques des individus (la culture).

A ce sujet, il paraît nécessaire de rappeler ici les très intéressantes expériences des Warren (1970) sur les illusions et confusions auditives. Si on écoute en boucle répétitive sans silence un très bon enregistrement d'un mot ou d'une phrase, sans autre contexte, il apparaît assez rapidement une illusion acoustique qui fait comprendre d'autres mots que ceux enregistrés. Tout mot ou toute phrase, dans ces conditions, sont sujets à produire ces illusions, avec des distorsions phonétiques considérables et fréquemment avec des associations sémantiques. Ces mots illusoire sont perçus auditivement et centralement ; ils sont intelligibles et les sujets trouvent difficile de croire qu'ils ont entendu un mot ou un signal isolé répété sur la boucle enregistrée. Par exemple, un sujet écoutant le mot "tress" répété sans silence, entend et comprend distinctement après quelques minutes, des mots illusoire tels que : dress, stress, joyce, floris, florist, purse. Cette illusion a été appelée par ces auteurs "effet de transformation verbale" ; elle met bien en relief que l'intelligibilité n'a que peu de relation avec la physique du message. Ils ont, de plus, montré que l'âge du sujet avait des incidences importantes sur les résultats quantitatifs obtenus dans la mesure de ces "illusions acoustiques" ; ce-ci reflète que les processus d'analyses centraux changent ; les enfants de cinq ans n'ont que peu ou pas de transformation verbale. À six ans.../...

* En fonction du niveau d'émission dans le silence, l'intelligibilité des chiffres est légèrement plus grande que celle des phrases, ainsi que Lehmann l'a montré (1962, Fig.8). On remarquera que pour les chiffres, il y a peu de choix possible, donc une prévisibilité relativement forte, ce qui peut expliquer les meilleurs résultats.
la moitié des enfants entendent les illusions, et les plus âgés les obtiennent le plus vite. À huit ans, tous les enfants ont ces illusions. Le taux de ces illusions reste à peu près constant jusqu'aux vingt ans et décroît ensuite progressivement, et pour les sujets de plus de 65 ans, le taux d'illusion est seulement du 1/5 ème de celui des jeunes adultes, soit approximativement égal à celui des enfants de cinq ans. Il n'y a cependant aucune corrélation avec une perte auditive. Mais plutôt que de compter le nombre d'illusions, si on examine les mots perçus pour déterminer les unités d'organisation perceptuelle, les enfants répondent par des sons de langue anglaise, mais peuvent les grouper dans des associations qui n'existent pas dans cette langue, par exemple, avec le mot "tress", un enfant comprendra "sreb" bien que la séquence initiale "sr" n'existe pas en anglais. Le groupe de jeunes adultes assemble les sons seulement selon les règles de l'anglais, mais ils prononcent des syllabes sans signification. Avec "tress", ils signalent, par exemple, "tresh", les sujets plus âgés, au contraire, ne donnant que des vrais mots de la langue. Le "tress" est entendu correctement et presque continuellement et si l'illusion apparaît, ils perçoivent des formes proches, telles que "dress". Il est très intéressant de noter que, à ces mêmes sujets âgés, on donne la répétition d'une syllabe sans signification (flime, par exemple), ils distordent le mot en le rapprochant d'un mot de la langue anglaise courante comme "slime" et tendront toujours à rester à des mots illusories ayant un sens.

L'absence du phénomène d'illusion à cinq ans suggère que les jeunes enfants n'ont pas atteint le niveau de l'apprentissage du langage dans lequel la mémorisation est associée avec les possibilités de combinaison. Les sujets âgés qui resistent à l'illusion auditive montrent qu'ils n'ont plus les capacités fonctionnelles de ce mécanisme qui doit être associé avec la mémoire instantanée qui, on le sait, est moins efficace chez les sujets âgés quand une activité (ici la perception de l'intelligibilité ou son équivalent) doit intervenir entre la réception du message (ou du stimulus) et sa traduction.

Il y a presqu'un siècle qu'il a été observé que les meilleurs télégraphistes receveurs de morse ne transcrivaient les signaux auditifs qui constituaient les mots qu'après en avoir entendu 6 à 12, s'aidant ainsi du contexte, sans aucun doute ; ce temps de latence correspond en fait à la perception de l'intelligibilité et à son intégration dans une réponse motrice. Regardez votre secrétaire taper à la machine, elle porte son attention visuelle quatre ou cinq mots en avance de celle qu'elle consacre à ses doigts. (cité par Warren et Warren).
L'intelligibilité et certaines données psychophysiologiques sous-estimées

Ces faits expérimentaux montrent bien que cette intelligibilité met en oeuvre des processus uniquement centraux qui n'ont pas grand chose à faire avec la physique du signal, ni avec la physiologie auditive périphérique, bases cependant sur lesquelles s'appuient la plupart de ceux qui se préoccupent de ces problèmes.

Venons en à d'autres exemples, également résultants d'expériences qui montrent bien ces dualités.

EXPERIENCES SUR L'INTELLIGIBILITE ET LA PHYSIQUE DU SIGNAL

Ce sont les téléphonistes qui ont montré les premiers la disjonction considérable entre la physique du signal et l'intelligibilité dont, au travers de quelques exemples personnels, nous voudrions souligner certains traits.

Le spectre de la voix humaine peut présenter une étendue considérable, puisque s'étendant jusqu'au domaine ultrasonore, certains ayant même prétendu qu'elle atteignait jusqu'à 60 kHz (Husson). Mais quel est donc le rapport entre ce spectre et l'intelligibilité ? Nous pensons qu'il est presque nul.

Il faut tout d'abord bien préciser que l'intelligibilité est la perception intégrée (compréhension) par le récepteur biologique de la somme de toutes les informations contenues dans un signal de communication d'un émetteur ; cette définition est, dans notre esprit, valable pour tous les supports physico-chimiques des messages, et pour toutes les espèces animales, homme compris (qu'il s'agisse donc d'un signal chimique, optique, électrique ou kinesthésique).

Or, quelles sont les informations transmises : certaines sont intrinsèques à l'émetteur, d'autres extrinsèques, et on peut en compter dix dans le cas des signaux acoustiques de langue humaine (1), et de l'ordre de 6 à 10 selon les autres espèces animales ; elles concernent, déchiffrées (intelligibles) par le récepteur :
1. l'espèce.
2. l'origine géographique (dans la parole, deux niveaux : par exemple c'est du français et c'est d'un méridional).
3. le sexe.
4. l'âge relatif (trois classes).

* on a omis volontairement, à cause de sa subjectivité, l'information esthétique.
Planche n°1 : Tracés sonographiques "saucisson". 1 : Témoin 2 : A la sortie d’un tuyau de plastique de 25 m, utilisé en téléphone. 3 : A la sortie d’un téléphone, type 1930. 4 : Après passage dans deux filtres P.B. et P.H. réglés chacun sur 500 Hz (intelligibilité toujours présente pour un récepteur informé). 5 : Sur 1 kHz, avec la même valeur d’intelligibilité.
5. l'état physiopathologique (enrhumé, nerveux, parkinsonien, etc...)
6. l'humeur (affectivité).
7. la culture.
8. l'identité (signature).
9. la localisation spatiale (distance, orientation).
10. la sémantique du message.

Certaines de ces notions sur la nature des diverses informations sont présentes à l'esprit des chercheurs préoccupés de l'intelligibilité, par exemple, l'identité, et on sait tout ce qui a été tenté par certains pour définir la physique de l'identité du locuteur, par exemple au travers du système Voice-Print (Kersta) et les excès de certaines des interprétations qui en ont été faites, soulignées par Ladefoged et Vanderslice. (voir aussi les travaux de Atal, des Bell Labs, et en France de R.Carré).

Mais étudions de plus près ce qui se passe, disons dans un téléphone. Il y a 50 ans, chez mon père, nous utilisions, en guise de téléphone intérieur, un système de tuyau dans lequel on parlait, le receveur à 20 ou 25 m. comprenant correctement (avec une intelligibilité parfaite) les messages. Nous avons fait la manipulation avec un tube de tuyau d'arrosage de 25 m. ; on analyse le signal de parole à l'entrée et à la sortie, après avoir mesuré, qu'en effet, l'intelligibilité est bien de 100 %. Nous avons utilisé le mot "saucisson" pour y mettre les ultrasons (?) des "s" bien sifflants. Les sonagrammes 1 et 2 de la planche 1 (bande 50 Hz - 8 kHz) sont ceux des deux signaux : 1, le témoin - 2 le même mot à la sortie. Le n°3 est le signal recueilli à la sortie d'un écouteur téléphonique. Le n°4 est le même mot, filtré avec un passe-bas et un passe-haut, centrés chacun sur 500 Hz et le n°5, une autre opération de filtrage sur 1 kHz. Or, que se passe-t-il dans ces cas expérimentaux ? Bien que la physique du signal soit très altérée par rapport à l'original, la sémantique est parfaitement conservée dans tous les cas, et sauf la localisation spatiale qui n'a pas à intervenir dans ce contexte, la plupart des autres types d'informations restent perceptibles à l'auditeur.

Des dégradations encore plus poussées feront disparaître les informations concernant l'âge, l'individu, le sexe, l'affectivité, la sémantique restant néanmoins présente à l'extrême.

.../...

* Ce qui veut dire que le squelette physique de la sémantique, dont la structure est d'ailleurs à déterminer, reste suffisamment intact puisqu'il est encore intelligible.
Avec un système de synthèse comme ordinateur ou un vocoder, à partir duquel on peut assez aisément modifier les divers aspects physiques d'un signal de parole », nous avons obtenu des signaux sans pitch, avec fondamental modifié (x2 ou supprimé), avec des bandes de fréquences éliminées, etc...(Examplifiés dans la planche II). On a également fait varier par 2 (x ou :) la vitesse d'un message sans changer ses autres constances. Ces altérations physiques, considérables, montrent qu'avec cette méthode on pourrait, par des études exhaustives faire ressortir les éléments physiques du signal qui supportent les 9 types d'informations, et il semble regrettable que ce genre de recherche n'ait pas été entrepris systématiquement. On s'aperçoit d'ailleurs que les diverses informations sont très difficilement dégradables, avec par ordre successif de disparition : l'identité, le sexe, l'origine géographique du 2ème niveau, la sémantique étant extrêmement résistante. Mêmes les changements de vitesse dans les proportions indiquées n'altèrent aucune des informations, et ce même au niveau d'un phonème isolé. D'autres que nous ont déjà signalé de tels faits (Leipp par exemple).

Dans toutes ces expériences, et c'est le point qui nous paraît capital, l'auditeur qui connaît le message est celui qui le reconnaît le mieux, et il existe des écarts considérables entre un écouteur naïf et l'écouteur averti. Autrement dit, on ne reconnaît que ce que l'on connaît préalablement, et la reconnaissance (intelligibilité) est donc à plusieurs niveaux ; il y a, au niveau central, un processus de comparaison qui est fonction du taux de mémorisation, et qui permet la reconstitution des informations, même si les signaux sont très altérés. Si avec l'expérience des filtres on fait entendre le signal le plus dégradé, encore complètement intelligible pour l'écouteur averti, à un écouteur naïf, il ne perçoit rien. Si on lui fait entendre le signal témoin et qu'on le dégrade progressivement, il obtiendra les mêmes résultats que l'écouteur averti. C'est cet aspect mémoriel, cette reconnaissance que nous avons voulu souligner, qui se révélait déjà par les différences des tests sur les logatomes et sur les mots ou les chiffres ; ceci traduit essentiellement un phénomène central, et non un phénomène purement acoustique.

Travail réalisé avec l'aide du laboratoire d'Acoustique du C.N.E.T. à Lannion, dont nous remercions vivement le directeur P. Lorand et son collaborateur M. Cartier.

Vocoder à 12 canaux, bandes de fréquences de 250 à 3500 Hz.
Planche n°II : Signal de parole utilisé à la sortie d'un vocoder.

1 : Témoin. 2. : Fréquence du fondamental x 2.
3 : Pitch modifié (de 80 à 180 fois 64 microseconds, par pas de 2 en diminuant, puis en augmentant) 4 : Fondamental généralisé.

L'intelligibilité globale reste pour tous ces signaux. Seuls changent le sexe (3 en 2 et 5), qui devient féminin et change l'individualité, mais l'origine géographique de 2ème niveau est toujours présente. Phrase : un message du ...
Nous citerons un autre exemple, avec un signal physiquement plus simple, celui d'une langue sifflée, qui est principalement basée sur la modulation de fréquence et la durée. Le signal est comparé à sa source et à sa réception, dans la nature, à 2 km. Physiquement encore, on note des altérations considérables, et pourtant le signal reste parfaitement intelligible pour des individus utilisant ces langues, dont la portée à l'île canarienne de la Goméra, par exemple, atteint près de 10 km. Mais pour un écouteur non éduqué, aucune intelligibilité n'est perceptible, et il n'entend (perception même pas le signal.)

LA MOTIVATION DU RECEPTEUR

Colin Cherry, en 1950, a décrit un phénomène extrêmement intéressant, qui ne semble pas avoir mérité, de la part des acousticiens s'occupant d'intelligibilité toute l'attention qu'il mérite. Il s'agit de l'effet de Cocktail-Party. Seule la motivation personnelle d'un récepteur déterminé, lui permet d'entendre et de comprendre les paroles d'un autre membre de l'assemblée, dont le niveau est au moins inférieur de 3 dB au bruit de fond moyen. Cette motivation personnelle est déclenchée par une vigilance accrue, soit par la présence physique d'une personne à laquelle le récepteur s'associe, soit par un mot qu'il perçoit au hasard, et qui fait partie d'un thème auquel il est intéressé. Le voisin immédiat du récepteur a beau recevoir à sa périphérie auditive le même bruit de fond formé de tous les mêmes signaux de parole, s'il n'a pas la même motivation pour le même locuteur, ou pour le même thème, il n'aura aucune perception ni aucune intelligibilité, et pour lui, dans ce bruit de fond rien ne se détachera. Dans un orchestre, le profane n'entend qu'un ensemble, alors que le récepteur entraîné et motivé (le chef d'orchestre par exemple ou un mélomane) reconnaît chaque instrument bien que les signaux de l'instrument soient noyés dans le bruit de fond de l'orchestre.

L'intelligibilité est donc dirigée dans ces cas par un processus centripète du système nerveux central, qui met en œuvre d'une manière préférentielle un traitement du signal à un niveau (présumé actuellement) supérieur à la périphérie qui permet, probablement par sommation ou corrélation, l'extraction préférentielle de signaux dans le bruit. Mais là encore, l'intelligibilité apparaît dirigée par divers facteurs difficilement cernables qui sont la motivation, la vigilance et la mémoire, c'est-à-dire l'intérêt et le choix par rapport à des références connues, et dans lesquels la physique du signal seule, n'intervient que d'une manière minime.

.../...
L’INTELLIGIBILITE ET CERTAINES DONNEES PSYCHOPHYSIOLOGIQUES
SOUS-ESTIMEES
C'est pourquoi, en acoustique, les tests d'intelligibilité, dans leur ensemble, présenteront de grands écarts, si on prend des locuteurs et des écouteurs qui se connaissent bien, en comparaison avec des individus qui ne se connaissent pas ; ou qu'on utilise, pour un écouteur donné, des mots ou des textes qui font partie de ses préoccupations intellectuelles et de son vocabulaire courant, plutôt que des mots qui lui sont peu familiers (logatomes) ; ceci s'accorde également avec les résultats décrits par les Warren sur les illusions.

LE CERVEAU ET L'INTELLIGIBILITE

Il y a de nombreuses données anatomo-pathologiques qui depuis près d'un siècle ont mis en évidence que les aspects centraux de l'intelligibilité sont évidents et qu'ils n'ont avec la fonction auditive périphérique que des rapports à minima. Notre propos n'est pas de faire ici une revue exhaustive de la question, mais de rappeler seulement quelques exemples pertinents à notre problème.

Wernicke (1874) a décrit une lésion spécifique de l'hémisphère gauche, la parole (d'après A.R. Luria). située à l'opposé de l'aire de Broca qui provoque la perte de l'intelligibilité de la parole.

C'est le cas de "surdité verbale" ainsi dénommée par Kussmaul, en 1884, et étudiée plus en détail par Arnaud de la Jasse en 1887. L'individu qui en est atteint a une audition parfaitement intacte ; il entend les sons, mais il ne comprend plus le sens des mots, ils ont perdu leur intelligibilité. C'est également le cas de certains polyglottes présentant cette lésion à un degré moindre pour qui une seule langue reste intelligible.
Les études systématiques de lésions cérébrales par des méthodes neuro-psychologiques, développées par Luria (1970) pour l'analyse de l'organisation fonctionnelle du cerveau ont bien confirmé l'importance de cette zone de Wernicke pour l'intelligibilité (Fig.5).

On trouvera, dans un récent article de Geschwind (1972) une étude de synthèse sur les centres associés au langage et à la compréhension de la parole et de l'écriture (intelligibilité), révélés par l'examen de pathologies diverses, qui sont directement en charge de notre problème, et qui confirme ce point de vue.

D'ailleurs, la même suggestion existe pour le langage écrit au travers de la vision et, Dejerine en 1891, a décrit l'alexie associée à l'agraphie dans laquelle, malgré une vision parfaite, une lésion du gyrus angularis de l'hémisphère gauche détruit l'intelligibilité de l'écriture et la possibilité d'écrire.

C'est donc bien au niveau central que se situe le mécanisme de l'intelligibilité que l'analyse neurophysiologique n'apprécie encore qu'imparfaitement, et nous ignorons presque complètement les processus mis en jeu.

Des travaux récents, basés sur des épreuves psychologiques, ont confirmé ces notions. Nous rappellerons ici, par exemple, les importants travaux des écoles américaines de Studdert-Kennedy et Shankweiler, de Mattingly et Liberman (1970), qui ont montré que les deux hémisphères, dans le domaine de l'intelligibilité de la parole, étaient dissymétriques.

L'intelligibilité des sons du langage dépend de processus unilatéraux localisés dans l'hémisphère cérébral dominant pour le langage. Cette spécialisation de l'hémisphère dominant est sous la dépendance d'une organisation anatomo-physiologique qui reste à découvrir ; elle possède cette fonction spéciale qui est dépendante des capacités de l'analyse auditive. Studdert-Kennedy et Shankweiler ont conclu de leurs travaux que si le système auditif, dans son ensemble, est commun aux deux hémisphères et est équipé pour extraire les paramètres auditifs du signal de parole, l'hémisphère dominant pouvait être spécialisé pour le traitement de leurs aspects linguistiques. .../...

* ce type de lésion ne permet plus l'accès aux mémoires des gestalts de parole, on les a détruites.
Les travaux anatomiques de Levitsky et Geschwind (1972) confirment d'ailleurs que les deux hémisphères sont dissymétriques au niveau de la surface supérieure du lobe temporal, et qu'il s'agit vraisemblablement d'un caractère génétique. Quelques observations sur des patients comissurectomisés semblent confirmer que le cerveau droit perçoit mal l'intelligibilité verbale (Sparks et Geschwind, 1968 ; Milner et al., 1968) et, en particulier, que ces opérés n'ont plus d'analyse syntaxique ou grammaticale (Glass et al., 1973).

EXTENSION A TOUTES LES FONCTIONS SENSORIELLES DE LA NOTION D'INTELLIGIBILITE D'UN SIGNAL :

Par quelques exemples pris dans d'autres domaines sensoriels, nous voudrions montrer que les lois, encore à préciser, qui sont sous-jacentes à l'intelligibilité, doivent avoir une valeur générale propre à tous les systèmes de bio-communication.

Qu'il s'agisse de l'effet de cocktail-party qui se retrouve immédiatement en vision, vraisemblablement par un procédé de traitement d'image que l'on peut simuler par méthodes holographiques, mettant en évidence le rôle de la motivation (voir avec quelle vitesse et quelle précision on trouve son nom dans la lecture d'une bibliographie); qu'il s'agisse de la reconstitution de l'intelligibilité d'une image à partir d'une réduction considérable des données physiques de celles-ci quand on connaît le sujet (voir planche V), par exemple un visage; qu'il s'agisse de l'intelligibilité de ces graphismes pour enfants dans lesquels on cherche et on trouve, noyés dans le bruit de fond, telle ou telle figure, et ce d'autant plus vite qu'on la connaît. Qu'il s'agisse de l'intelligibilité de certaines informations telle que l'identité, l'humeur, la culture, l'origine géographique dans l'écriture, déterminables au travers de la graphologie, ce sont toujours des phénomènes de même nature.

Dans les autres domaines sensoriels, l'intelligibilité est également importante dans les technologies qui y sont rattachées. On citera à l'exemple de la téléphonie pour l'acoustique, celle de la qualité organoleptique pour le goût et, plus spécialement, en oenologie, et les qualités odorantes des essences pour la parfumerie (les "nez"). Le Magren (communication personnelle) nous a indiqué qu'un "nez" expérimenté peut analyser (intelligibilité) plusieurs milliers d'essences odorantes, et que certains des meilleurs spécialistes en oenologie peuvent connaître plusieurs centaines de crus (en estimation approximative 4 à 600) et les déchiffrer chacun sur 5 années, ce qui nous donne un vocabulaire de 2 à 3000 signaux; on rejoint ici les sommes des caractères mémorisés des vocabulaires de langue. .../...
L'INTELLIGIBILITE ET CERTAINES DONNEES PSYCHOPHYSIOLOGIQUES
SOU ESTIMEES
D'ailleurs toutes ces propriétés attachées au système nerveux central du récepteur, à partir de la somme de ses connaissances préalables à la reconnaissance, existent dans le règne animal. Par exemple, le chien policier motivé qui fait du cocktail-party effect, en pistage olfactif, le test d'intelligibilité des expériences de Tinbergen sur les jeunes goélands qui, malgré la déformation physique d'un signal optique, montre que les oiseaux "compriment" cependant celui-ci (planchè VI). Il s'agit ici d'un modèle de tête, dont on fait varier la forme globale ou celle du bec, ainsi que la couleur d'une tache basale.

Nous pensons nécessaire de souligner ainsi l'ubiquité de l'intelligibilité qui ne doit pas être limitée à celle des seuls signaux acoustiques, et dont les caractères fondamentaux sont généraux à tous les signaux de communication.

Le problème de l'intelligibilité paraît donc concerner essentiellement la reconnaissance de formes, d'images, acoustiques ou autres. Celle-ci peut résulter que d'opérations de traitements comparatifs de données par rapport à celles qui ont été enregistrées en mémoires (et on peut accepter sans difficulté une mémoire génétique à côté d'une mémoire d'apprentissage).

Les opérations sont obligatoirement centrales, et la tendance actuelle nous oriente vers un traitement par corrélation, selon les principes mathématiques qu'on utilise en théorie du signal.

Le fait qu'un récepteur homme ou animal soit capable d'avoir encore un certain taux de reconnaissance d'une forme altérée ou partiellement détruite dépend d'un grand nombre de variables individuelles qui concernent outre sa motivation (qui déclenche une certaine programmation), les capacités de son ordinateur central, c'est-à-dire le contenu de la ou des mémoires, et l'ordre (rangement) dans les mémoires, enfin la prévisibilité en fonction du contexte bio-écologique et culturel (social) du récepteur.

La méthodologie qui est sous-jacente au problème, et que nous venons de rappeler très sommairement, peut s'identifier à celle déjà utilisée dans les gros ordinateurs ; ceux-ci peuvent être programmés pour reconnaître des formes optiques et acoustiques, et intégrer également des notions qui sont celles des comportements biologiques que nous avons soulignés. *

Planche n° VI : Experiences de Tinbergen sur la reaction de picorage du bec d'un parent, chez le Geelband, par la jeune poussin. Le signal auquel il reagit est constitué par la tete, prolongée du bec, marquée à sa partie inferieure d'une tache rouge. Avec des signaux artificiels (leurre) découps dans du papier, en réalisant des tetes en faisant varier la fréquence relative des réactions est indiquée ; elle traduit l'intelligibilité du signal.

I - Bec jaune, tete blanche, tache rouge dans les trois images superieures. En bas, (variations des couleurs de la tete et du bec) couleur grise uniforme.

II - Variations des couleurs de la tete du bec. Bec jaune, tete blanche, pas de tache.
CONCLUSIONS

Dans son exposé aux journées d'Aix (1972), Cartier a souligné que la notion d'intelligibilité échappait à la mesure, et ce serait notre tendance dans l'état actuel de nos connaissances et de nos moyens d'approche métrologique usuels.

Il semble bien que la plupart de ceux qui se préoccupent de ces problèmes soient d'accord pour estimer qu'il est absolument inutile de postuler un décodage qui passerait par l'analyse physique fine du signal de parole, même si elle existe, si perfectionnée soit-elle, au niveau de la périphérie, ce qui n'est d'ailleurs pas le cas. Toutes les expériences, et nombreux sont les chercheurs qui les ont rapportées, par exemple, avec la méthode des indices d'articulation (Kryter) confirmant l'extraction du signal dans le bruit, et l'intelligibilité à partir d'éléments partiels du signal de parole.

Il nous semble d'ailleurs que les divergences de vues qui apparaissent, selon la formation d'origine des chercheurs, viennent partiellement de leurs sémantiques. C'est ce que Leipp a déjà signalé et avec raison. Toutefois, il est admissible que chaque groupe professionnel ait son vocabulaire spécialisé, et on admettra qu'il est même parfois difficile de le clarifier et de l'harmoniser (voir, par exemple, les vocabulaires terminologiques individuels des "nez" ou des goûteurs de vins ou d'alcool).

Nous sommes également d'accord avec le point de vue très judicieux de Leipp sur son interprétation de l'intelligibilité au travers de la théorie de la reconnaissance des formes (pattern recognition) ; il nous paraît valable pour toutes les formes, qu'elles soient acoustiques, optiques et chimiques, et ce, non pas qu'elles soient perçues comme telles par la périphérie, mais au travers des codages qu'en font les périphéries à partir du signal reçu, et leur interprétation centrale.

Ceci conduit à bien préciser les abus que l'on est tenté de faire, à défaut d'autre chose, de vouloir définir la structure physique d'un mot à l'aide de fréquences absolues et d'espace temporel, puisqu'on peut modifier ces diverses composantes informatives d'une manière considérable, sans changer l'intelligibilité. Il nous paraît en découler que le phénomène auditif en soi est donc distinct du phénomène de l'intelligibilité de la parole, alors qu'il est associé à la reconnaissance des notes de la musique par exemple, qui est pourtant également apprise, mais qui n'accepte que de faibles distorsions au niveau des fréquences ou des rapports harmoniques (à l'inverse de la phrase musicale qui est reconnue même chantée faux ou abrégée). .../...
De nos contacts avec nos collègues, nous avons senti le besoin et l'intérêt de confrontations, voire de coopération. Le point de vue des ingénieurs, compte-tenu de leurs intérêts pragmatiques, est parfois éloigné de celui des biologistes et ceux-ci, avec leurs sommes de méconnaissances devant la difficulté d'expérimenter le cerveau et l'homme, ont assez tendance à éviter ces approches communautaires. Nous pensons que d'avoir rappelé quelques faits expérimentaux, qui montrent l'importance que nous attachons à l'apprentissage, à la motivation, aux illusions acoustiques, pouvait être utile. L'impact des études des neuropsychologues qui précisent ces aires centrales, que Leipp appelle les "anamorphoseurs", nous paraît également plus important, car ces localisations concrétisent bien cette idée que c'est le phénomène central qui est capital dans les analyses des informations qui constituent l'intelligibilité ; elles appartiennent alors davantage, pour la parole, au domaine de la biologie que de l'acoustique, et ce dans un sens d'autant plus large que les principes même de ces analyses (apprentissage, mémorisation, motivation), se retrouvent pour tous les systèmes informatifs sensoriels en provenance des périphéries, quelque soit le véhicule physico-chimique des messages au travers des canaux.

Si les physiologistes ont également montré, sans toutefois que tous les mécanismes soient encore très clairs ou même convaincants, à quel point les récepteurs périphériques étaient perfectionnés et d'une exquise sensibilité, en fait, il s'agit d'une grande redondance du système à ce niveau, par rapport aux non-moins énormes capacités du système central d'analyser à partir de caricatures. Et, dans notre domaine, on ne voit plus très bien la nécessité (en phonétique, par exemple) de définir l'image spectrographique ambiguë et trop riche de pseudo-invariants physiques dont le centre d'analyse n'a qu'un usage statistique global ; d'ailleurs le décodage ne se fait même pas au niveau des unités discrètes des linguistes, qui paraissent trop obnubilés par les éléments écrits du langage en le décomposant au niveau de la lettre, qui n'a pas de rapport avec le phonème (en chinois, par exemple, le mot est un idéogramme et donc un tout).

Nous ne pensons pas que ce rapport contribue à augmenter beaucoup l'intelligibilité du problème. Dans notre esprit, il n'a voulu qu'attirer l'attention sur des points de vue, certainement discutables, d'un biologiste conscient de ses méconnaissances, et souligner que la machine vivante est encore une boîte "très" noire, par rapport à celles que les ingénieurs construisent avec tant d'ingéniosité pour la supplanter ; l'ère nouvelle des ordinateurs qu'ils imaginent...
et réalisent avec sagacité nous donnera probablement, avant la biologie, une réponse à nos interrogations.

REMERCIEMENTS

Je remercie mes collègues : M. CARTIER, H. ETZOLD, R. LAMORAL, E. LEIPP J. LE MAGNEN, P. LORAND, L. PIMONOW, J. PUJOLLE, qui ont bien voulu apporter leurs constructives remarques à l'élaboration de ce texte, ainsi que F. CANTE pour sa collaboration technique.

RESUME

La notion d'intelligibilité, définie par les ingénieurs des télécommunications, mérite d'être à la fois reconsidérée et étendue. S'il est nécessaire, au stade de la mesure du courant d'information, que le récepteur humain soit statistiquement standardisé, en fait, la pratique montre qu'il y a des difficultés qui sont inhérentes aux caractéristiques biologiques. Il semble que celles-ci doivent, dans une certaine mesure, être prises en compte, car le cerveau doué de mémoire intervient dans l'analyse des images sonores en reconstruisant ou rattachant aux formes apprises et stockées les informations reçues, même si celles-ci proviennent de signaux déformés, altérés ou réduits. On devra donc considérer une intelligibilité, du domaine de la biologie, complémentaire de celle du domaine des ingénieurs.

Des facteurs tels que la motivation, la culture et l'âge interviennent dans la compréhension des messages, et les expériences du type cocktail party effect ou celles sur les illusions et les confusions auditives les mettent bien en évidence. Il apparaît alors de plus en plus que l'intelligibilité de la parole et des autres types de signaux de nature physico-chimique diverses, qui procèdent du même modèle, est avant tout un phénomène central, dans lequel les systèmes périphériques bien qu'essentiels n'ont qu'une faible part. L'expérience montre également que les reconstructions faites au niveau central rendent partiellement vaine la finesse des détails d'analyse résultant de la physique. Les centres supérieurs, révélés par les études pathologiques doivent traiter les informations par des procédés de reconnaissance de forme, qui paraissent analogues à ceux qui sont programmés dans les ordinateurs, à partir des mémoires.

Ceci permet d'espérer des technologies nouvelles basées sur les capacités des ordinateurs qui conduiront pour l'ingénieur, à une mesure plus objective de l'intelligibilité sensu stricto. L'approche biologique de l'intelligibilité, au sens large, met en évidence que le...
L'INTELLIGIBILITE ET CERTAINES DONNEES PSYCHOPHYSIOLOGIQUES
SOUSS-ESTIMEES

champ des analyses physico-chimiques des signaux (dans notre domaine,
en acoustique physiologique et en phonétique) est à reconsidérer,
complètement, et doit tendre à restreindre la profusion des données
métrologiques qui éloignent inutilement du problème de la reconnais-
sance des formes qui est à la base du phénomène.

Tenté par une conclusion paradoxale, on a presque envie d'écrire que,
finalement, l'intelligibilité de la parole, sur le plan biologique,
est presque en dehors du domaine de l'acoustique.
L'INTELLIGIBILITE ET CERTAINES DONNEES PSYCHOPHYSIOLOGIQUES 
SOUS-ESTIMEES

BIBLIOGRAPHIE

ATAL (B.S.) : Automatic speaker Recognition Based on Pitch contours 

CARRE (R.) : Identification des locuteurs : exploitation des données 

CARRE (R.) : Rapport sur les journées d'Etudes sur la Parole 
Aix-en-Provence, 1971, in Revue d'Acoastique, 1972,n°19


CHAVASSE (P.): De la notion d'intelligibilité 
Audioprothesiste francais, 1962, n°3, Colloque: "Intelligibilité et Prothèse auditive 

CHOCHOLLE(R.) - in Leipp, 1970.

GESCHWIND (N.): Language and the brain. 

GLASS (A.V.), GAZZANICA (M.S.), PREMACK (D.): Artificial Language 
training in global aphasics. 

HABER (R.N.): How we remember what we see. 

KRYTER (K.D.): Methods for the calculators and use of the articula-
tion index, 

KRYTER (K.D.) and WHITMAN (E.C.): Some comparisons between rhyme and 
P.B. word intelligibility tests. 

LADEFOGED (P.) VANDERSLICE (R.): The Voice-Print mystique. 

LEHMANN (R.): Contribution à l'étude du masquage de la parole. 
Electroacoustique, Bull. Lab. Electroac., Université 
de Liège, 1962, n°7.

.../...


RISSET (J.C.): Methodes d'évaluation de l'intelligibilité et de la parole. in Carré, 1972.


L'INTELLIGIBILITE ET CERTAINES DONNEES PSYCHOPHYSIOLOGIQUES
SOUS-ESTIMEES

WAJSKOP (M.) : La perception de la parole : orientation et perspecti-
ves. 
Revue d'Acoustique, 1971, n°13, 4-5.

WARREN (R.M.), WARREN (R.P.) : Auditory illusions and confusions. 

WISNER (A.) et Coll. : Influence de divers types de distorsion sur la 
compréhension des messages verbaux en milieu bruyant. 
Contrat C.E.C.A. n°6242 - 12.3.043, rapport 36, mars 