

# NOBEL LECTURES

# PHYSICS



1971 – 1980

World Scientific

NOBEL LECTURES  
INCLUDING PRESENTATION SPEECHES  
AND LAUREATES' BIOGRAPHIES

**PHYSICS**

**1971-1980**

EDITOR

**STIG LUNDQVIST**



**World Scientific**

Singapore • New Jersey • London • Hong Kong

*Published for the Nobel Foundation in 1992 by*

World Scientific Publishing Co. Pte. Ltd.

P O Box 128, Farrer Road, Singapore 9128

USA Office: Suite 1B, 1060 Main Street, River Edge, NJ 07661

UK office: 73 Lynton Mead, Totteridge, London N20 8DH

First published 1992

First reprint 1994

**NOBEL LECTURES IN PHYSICS (1971-1980)**

All rights reserved

ISBN 981-02-0726-3  
981-02-0727-1 (pbk)

Printed in Singapore.

NOBEL LECTURES  
INCLUDING PRESENTATION SPEECHES  
AND LAUREATES BIOGRAPHIES

PHYSICS  
CHEMISTRY  
PHYSIOLOGY OR MEDICINE  
LITERATURE  
PEACE  
ECONOMIC SCIENCES

NOBEL LECTURES IN PHYSICS  
1971-1980

## **Foreword**

Since 1901 the Nobel Foundation has published annually "Les Prix Nobel" with reports from the Nobel Award Ceremonies in Stockholm and Oslo as well as the biographies and Nobel lectures of the laureates. In order to make the lectures available to people with special interests in the different prize fields the Foundation gave Elsevier Publishing Company the right to publish in English the lectures for 1910-1970, which were published in 1964-1972 through the following volumes:

Physics 1901-1970	4 vols.
Chemistry 1901-1970	4 vols.
Physiology or Medicine 1901-1970	4 vols.
Literature 1910-1967	1 vol.
Peace 1910-1970	3 vols.

Elsevier decided later not to continue the Nobel project. It is therefore with great satisfaction that the Nobel Foundation has given World Scientific Publishing Company the right to bring the series up to date.

The Nobel Foundation is very pleased that the intellectual and spiritual message to the world laid down in the laureates' lectures will, thanks to the efforts of World Scientific, reach new readers all over the world.

Lars Gyllensten  
Chairman of the Board

Stig Ramel  
Executive Director

*Stockholm, June 1991*

## PREFACE

The Nobel Foundation publishes annually the proceedings of the year's Prize ceremonies in a volume called *Les Prix Nobel*.

It contains the speeches given at the prize ceremony, the autobiography of each laureate as well as the Nobel Lecture. These books contain much material of great interest to the scientific community. However, they are printed only in a small number of copies, very few scientists even know that they exist and there is no advertisement.

From 1992 the material is now becoming available through a deal between the Nobel Foundation and World Scientific to publish the material from the 70's and the 80's in a series of volumes. This volume contains the materials in physics for the period 1970-1979.

The contents in this volume are not identical with the original in *Les Prix Nobel*. We have written to all the physics laureates and let them modify and update the material.

The reader may be surprised about the very short speeches of presentation. The reason is that only a few minutes are allotted to these speeches which did not permit any description of the discovery.

Stig Lundqvist

## CONTENTS

Foreword	v
Preface	vii
1971 DENIS GABOR	1
Presentation by Erik Ingelstam	3
Biography of Denis Gabor	7
<i>Holography, 1948 -1971</i>	11
1972 JOHN BARDEEN, LEON N COOPER and	
J ROBERT SCHRIEFFER	45
Presentation by Stig Lundqvist	47
Biography of John Bardeen	51
<i>Electron-Phonon Interactions and Superconductivity</i>	54
Biography of Leon N Cooper	71
<i>Microscopic Quantum Interference Effects in the Theory</i>	
<i>of Superconductivity</i>	73
Biography of J Robert Schrieffer	95
<i>Macroscopic Quantum Phenomena from Pairing in</i>	
<i>Superconductors</i>	97
1973 LEO ESAKI, IVAR GIAEVER and BRIAN D JOSEPHSON	109
Presentation by Stig Lundqvist	111
Biography of Leo Esaki	115
<i>Long Journey into Tunneling</i>	116
Biography of Ivar Giaever	135
<i>Electron Tunneling and Superconductivity</i>	137
Biography of Brian D Josephson	155
<i>The Discovery of Tunnelling Supercurrents</i>	157

1974 MARTIN RYLE and ANTONY HEWISH	165
Presentation by Hans Wilhelmsson	167
Biography of Antony Hewish	171
<i>Pulsars and High Density Physics</i>	174
Biography of Martin Ryle	185
<i>Radio Telescopes of Large Resolving Power</i>	187
 1975 AAGE BOHR, BEN R MOTTELSON and	
JAMES RAINWATER	205
Presentation by Sven Johansson	207
Biography of Aage Bohr	211
<i>Rotational Motion in Nuclei</i>	213
Biography of Ben R Mottelson	235
<i>Elementary Modes of Excitation in the Nucleus</i>	236
Biography of James Rainwater	257
<i>Background for the Spheroidal Nuclear Model Proposal</i>	259
 1976 BURTON RICHTER and SAMUEL C C TING	
Presentation by Gösta Ekspong	273
Biography of Burton Richter	277
<i>From the Psi to Charm -The Experiments of 1975 and 1976</i>	281
Biography of Samuel Chao Chung Ting	313
<i>The Discovery of the J Particle: A Personal Collection</i>	316
 1977 PHILIP W ANDERSON, NEVILL F MOTT and	
JOHN H VAN VLECK	345
Presentation by Per-Olov Löwdin	347
Biography of John Hasbrouck Van Vleck	351
<i>Quantum Mechanics : The Key to Understanding Magnetism</i>	353
Biography of Philip W Anderson	371
<i>Local Moments and Localized States</i>	376
Biography of Nevill Francis Mott	401
<i>Electrons in Glass</i>	403
 1978 PETER LEONIDOVITCH KAPITZA, ARNO A PENZIAS	
and ROBERT W WILSON	415
Presentation by Lamek Hulthén	417

	Biography of Peter Leonidovitch Kapitza	421
	<i>Plasma and the Controlled Thermonuclear Reaction</i>	424
	Biography of Arno A Penzias	439
	<i>The Origin of Elements</i>	444
	Biography of Robert W Wilson	459
	<i>The Cosmic Microwave Background Radiation</i>	463
1979	SHELDON L GLASHOW, ABDUS SALAM and STEVEN WEINBERG	485
	Presentation by Bengt Nagel	487
	Biography of Sheldon Lee Glashow	491
	<i>Towards a Unified Theory - Threads in a Tapestry</i>	494
	Biography of Abdus Salam	507
	<i>Gauge Unification of Fundamental Forces</i>	513
	Biography of Steven Weinberg	541
	<i>Conceptual Foundations of the Unified Theory of Weak and Electromagnetic Interactions</i>	543
1980	JAMES W CRONIN and VAL L FITCH	561
	Presentation by Costa Ekspong	563
	Biography of James W Cronin	567
	<i>CP Symmetry Violation - The Search for Its Origin</i>	570
	Biography of Val Logsdon Fitch	591
	<i>The Discovery of Charge-Conjugation Parity Symmetry</i>	594

Physics 1971

**DENNIS GABOR**

*for his invention and development of the holographic method*

# THE NOBEL PRIZE FOR PHYSICS

Speech by professor **ERIK INGELSTAM** of the Royal Academy of Sciences  
Translation

Your Majesty, Your Royal Highnesses, Ladies and Gentlemen,

Our five senses give us knowledge of our surroundings, and nature herself has many available resources. The most obvious is light which gives us the possibility to see and to be pleased by colour and shape. Sound conveys the speech with which we communicate with each other and it also allows us to experience the tone-world of music.

Light and sound are wave motions which give us information not only on the sources from which they originate, but also on the bodies through which they pass, and against which they are reflected or deflected. But light and sound are only two examples of waves which carry information, and they cover only very small parts of the electromagnetic and acoustic spectra to which our eyes and ears are sensitive.

Physicists and technologists are working continuously to improve and broaden the methods and instruments which give us knowledge about waves which lie outside our direct perception capacity. The electron microscope resolves structures which are a thousand times smaller than the wavelength of visible light. The photographic plate preserves for us a picture of a fleeting moment, which perhaps we may make use of over a long time period for measurements, or it transforms a wave-field of heat rays, X rays, or electron rays to a visible image.

And yet, important information about the object is missing in a photographic image. This is a problem which has been a key one for Dennis Gabor during his work on information theory. Because the image reproduces only the effect of the intensity of the incident wave-field, not its nature. The other characteristic quantity of the waves, phase, is lost and thereby the three dimensional geometry. The phase depends upon from which direction the wave is coming and how far it has travelled from the object to be imaged.

Gabor found the solution to the problem of how one can retain a wave-field with its phase on a photographic plate. A part of the wave-field, upon which the object has not had an effect, namely a reference wave, is allowed to fall on the plate together with the wave-field from the object. These two fields are superimposed upon one another, they interfere, and give the strongest illumination where they have the same phase, the weakest where they extinguish each other by having the opposite phase. Gabor called this plate a hologram, from the Greek *holos*, which means whole or complete, since the plate contains the whole information. This information is stored in the plate in a coded form. When the hologram is irradiated only with the reference wave, this wave is deflected in the hologram structure, and the original ob-

ject's field is reconstructed. The result is a three dimensional image.

Gabor originally thought of using the principle to make an electron microscope image in two steps: first to register an object's field with electron rays in a hologram, and then to reconstruct this with visible light to make a three dimensional image with high resolution. But suitable electron sources for this were not available, and also for other technical reasons the idea could not be tested. However, through successful experiments with light Gabor could show that the principle was correct. In three papers from 1948 to 1951 he attained an exact analysis of the method, and his equations, even today, contain all the necessary information.

Holography, as this area of science is called, made its break-through when the tool, which had so far been missing, became available, namely the laser as a light source. The first laser was successfully constructed in 1960, and the basic ideas were rewarded by the 1964 Nobel Prize in physics. The laser generates continuous, coherent wave-trains of such lengths that one can reconstruct the depth in the holographic image. At about the same time a solution was discovered to the problem of getting rid of disturbing double images from the field of view. A research group at Michigan University in the United States, led by Emmett Leith, initiated this development.

The fascinated observer's admiration when he experiences the three dimensional space effect in a holographic image is, however, an inefficient acknowledgement for the inventor. More important are the scientific and technical uses to which his idea has led. The position of each object's point in space is determined to a fraction of a light wave-length, a few tenths of a millimetre, thanks to the phase in the wave-field. With this, the hologram has, in an unexpected way, enriched optical measurement techniques, and particularly interferometric measurements have been made possible on many objects. The shape of the object at different times can be stored in one and the same hologram, through illumination of it several times. When they are reconstructed simultaneously, the different wave-fields interfere with each other, and the image of the object is covered with interference lines, which directly, in light wavelengths, correspond to changes of shape between the exposures. These changes can also be, for example, vibrations in a membrane or a musical instrument.

Also, very rapid sequences of events, even in plasma physics, are amenable to analysis through hologram exposures at certain times with short light flashes from modern impulse lasers.

Gabor's original thought to use different waves for both steps within holography, has been taken up in many connections. It is especially attractive to use ultra sound waves for exposures, so that, in the second step, a sound field is reconstructed in the shape of an optical image. Despite significant difficulties there is work, with a certain amount of progress, being done in this area. Such a method should be of value for medical diagnosis, since the deflected sound field gives different information from that in X ray radiography.

Professor Gabor,

You have the honour and pleasure to have founded the basic ideas of the

holographic method. Through your work and assiduous contributions of ideas you continue to add to the development of this field, and this applies especially now that you have the freedom of a professor emeritus. Your activity as a writer on culture shows that you belong to the group of physicists and technologists who are concerned about the use or damage to which technical development can lead for mankind.

The Royal Swedish Academy of Sciences wishes to give you hearty congratulations, and I now ask you to receive the Nobel Prize in physics from the hand of His Majesty the King.



Dennis Gabor

## DENNIS GABOR

I was born in Budapest, Hungary, on June 5, 1900, the oldest son of Bertalan Gabor, director of a mining company, and his wife Adrienne. My life-long love of physics started suddenly at the age of 15. I could not wait until I got to the university, I learned the calculus and worked through the textbook of Chwolson, the largest at that time, in the next two years. I remember how fascinated I was by Abbe's theory of the microscope and by Gabriel Lippmann's method of colour photography, which played such a great part in my work, 30 years later. Also, with my late brother George, we built up a little laboratory in our home, where we could repeat most experiments which were modern at that time, such as wireless X-rays and radioactivity. Yet, when I reached university age, I opted for engineering instead of physics. Physics was not yet a profession in Hungary, with a total of half-a-dozen university chairs-and who could have been presumptuous enough to aspire to one of these?

So I acquired my degrees, (Diploma at the Technische Hochschule Berlin, 1924, Dr-Ing. in 1927), in electrical engineering, though I sneaked over from the TH as often as possible to the University of Berlin, were physics at that time was at its apogee, with Einstein, Planck, Nernst and v. Laue. Though electrical engineering remained my profession, my work was almost always in applied physics. My doctorate work was the development of one of the first high speed cathode ray oscilloscopes and in the course of this I made the first iron-shrouded magnetic electron lens. In 1927 I joined the Siemens & Halske AG where I made my first of my successful inventions; the high pressure quartz mercury lamp with superheated vapour and the molybdenum tape seal, since used in millions of street lamps. This was also my first exercise in serendipity, (the art of looking for something and finding something else), because I was not after a mercury lamp but after a cadmium lamp, and that was not a success.

In 1933, when Hitler came to power, I left Germany and after a short period in Hungary went to England. At that time, in 1934, England was still in the depths of the depression, and jobs for foreigners were very difficult. I obtained employment with the British Thomson-Houston Co., Rugby, on an inventor's agreement. The invention was a gas discharge tube

with a positive characteristic, which could be operated on the mains. Unfortunately, most of its light emission was in the short ultraviolet, so that it failed to give good efficiency with the available fluorescent powders, but at least it gave me a foothold in the BTH Research Laboratory, where I remained until the end of 1948. The years after the war were the most fruitful. I wrote, among many others, my first papers on communication theory, I developed a system of stereoscopic cinematography, and in the last year, 1948 I carried out the basic experiments in holography, at that time called "wavefront reconstruction". This again was an exercise in serendipity. The original objective was an improved electron microscope, capable of resolving atomic lattices and seeing single atoms. Three year's work, 1950-53, carried out in collaboration with the AEI Research Laboratory in Aldermaston, led to some respectable results, but still far from the goal. We had started 20 years too early. Only in recent years have certain auxiliary techniques developed to the point when electron holography could become a success. On the other hand, optical holography has become a world success after the invention and introduction of the laser, and acoustical holography has now also made a promising start.

On January 1, 1949 I joined the Imperial College of Science & Technology in London, first as a Reader in Electronics, later as Professor of Applied Electron Physics, until my retirement in 1967. This was a happy time. With my young doctorands as collaborators I attacked many problems, almost always difficult ones. The first was the elucidation of Langmuirs Paradox, the inexplicably intense apparent electron interaction. in low pressure mercury arcs. The explanation was that the electrons exchanged energy not with one another, by collisions, but by interaction with an oscillating boundary layer at the wall of the discharge vessel. We made also a Wilson cloud chamber, in which the velocity of particles became measurable by impressing on them a high frequency, critical field, which produced time marks on the paths, at the points of maximum ionisation. Other developments were: a holographic microscope, a new electron-velocity spectroscope an analogue computer which was a universal, non-linear "learning" predictor, recognizer and simulator of time series, a flat thin colour television tube, and a new type of thermionic converter. Theoretical work included communication theory, plasma theory, magnetron theory and I spent several years on a scheme of fusion, in which a critical high-temperature plasma would have been established by a 1000 ampere space charge-compensated ion beam, fast enough to run over the many unstable modes which arise during its formation. Fortunately the theory showed that

at least one unstable mode always remained, so that no money had to be spent on its development.

After my retirement in 1967 I remained connected with the Imperial College as a Senior Research Fellow and I became Staff Scientist of CBS Laboratories, Stamford, Conn. where I have collaborated with the President, my life-long friend, Dr. Peter C. Goldmark in many new schemes of communication and display. This kept me happily occupied as an inventor, but meanwhile, ever since 1958, I have spent much time on a new interest; the future of our industrial civilisation. I became more and more convinced that a serious mismatch has developed between technology and our social institutions, and that inventive minds ought to consider social inventions as their first priority. This conviction has found expression in three books, *Inventing the Future*, 1963, *Innovations*, 1970, and *The Mature Society*, 1972. Though I still have much unfinished technological work on my hands, I consider this as my first priority in my remaining years.

#### *Honours*

Fellow of the Royal Society, 1956.

Hon. Member of the Hungarian Academy of Sciences, 1964.

D.Sc. Univ. of London, 1964, Hon. D.Sc. Univ. of Southampton, 1970,  
and Technological University Delft, 1971.

Thomas Young Medal of Physical Society London, 1967.

Cristoforo Colombo Prize of Int. Inst. Communications, Genoa, 1967.

Albert Michelson Medal of The Franklin Institute, Philadelphia, 1968.

Rumford Medal of the Royal Society, 1968.

Medal of Honor of the Institution of Electrical and Electronic Engineers,  
1970.

Prix Holweck of the French Physical Society, 1971.

Commander of the Order of the British Empire, 1970.

Married since 1936 to Marjorie Louise, daughter of Joseph Kennard  
Butler and Louise Butler of Rugby.

The following details about his last years were obtained from Ms Anne Barrett, College Archivist at the Imperial College:

Professor Denis Gabor was awarded the Nobel Prize for Physics in 1971 and gave his Nobel Lecture on Holography. In the years following his Nobel award he received honours from universities and institutions internationally. He travelled widely giving lectures, many on holography or the subject of his book *The Mature Society*.

Between 1973 and 1976 Gabor and Umberto Columbo jointly chaired a working party on the possible contribution of Science to the regeneration of natural resources. The results were published in 1978 as *Beyond the Age of Waste*.

In 1974 Gabor suffered a cerebral haemorrhage so he was unable to personally present lectures he had prepared for that year, but his large group of eminent friends rallied to present them for him. Gabor lost the power to read and write himself and his speech deteriorated but his intellect and hearing were acute, so he remained involved in the scientific world. He was able to visit the new Museum of Holography in New York and the Royal Academy holography exhibition in 1977. He became Honorary Chairman of the Board of Trustees of the New York Museum of Holography in 1978 and also sat for his holographic portrait by Hart Perry.

His health deteriorated during the latter part of 1978 and he died in 1979.

# HOLOGRAPHY, 1948-1971

Nobel Lecture, December 11, 1971

by

DENNIS GABOR

Imperial Colleges of Science and Technology, London

I have the advantage in this lecture, over many of my predecessors, that I need not write down a single equation or show an abstract graph. One can of course introduce almost any amount of mathematics into holography, but the essentials can be explained and understood from physical arguments.

Holography is based on the wave nature of light, and this was demonstrated convincingly for the first time in 1801 by Thomas Young, by a wonderfully simple experiment. He let a ray of sunlight into a dark room, placed a dark screen in front of it, pierced with two small pinholes, and beyond this, at some distance, a white screen. He then saw two darkish lines at both sides of a bright line, which gave him sufficient encouragement to repeat the experiment, this time with a spirit flame as light source, with a little salt in it, to produce the bright yellow sodium light. This time he saw a number of dark lines, regularly spaced; the first clear proof that light added to light can produce darkness. This phenomenon is called interference. Thomas Young had expected it because he believed in the wave theory of light. His great contribution to Christian Huygens's original idea was the intuition that mono-

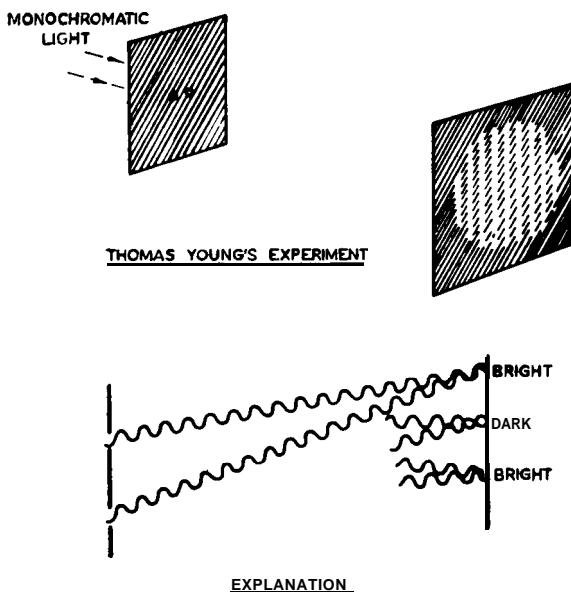


Fig. 1.  
Thomas Young's Interference Experiments, 1801

chromatic light represents regular, sinusoidal oscillations, in a medium which at that time was called "the ether". If this is so, it must be possible to produce more light by adding wavecrest to wavecrest, and darkness by adding wavecrest to wavethrough.

Light which is capable of interferences is called "coherent", and it is evident that in order to yield many interference fringes, it must be very monochromatic. Coherence is conveniently measured by the path difference between two rays of the same source, by which they can differ while still giving observable interference contrast. This is called the coherence length, an important quantity in the theory and practice of holography. Lord Rayleigh and Albert Michelson were the first to understand that it is a reciprocal measure of the spectroscopic line width. Michelson used it for ingenious methods of spectral analysis and for the measurement of the diameter of stars.

Let us now jump a century and a half, to 1947. At that time I was very interested in electron microscopy. This wonderful instrument had at that time produced a hundredfold improvement on the resolving power of the best light microscopes, and yet it was disappointing, because it had stopped short of resolving atomic lattices. The de Broglie wavelength of fast electrons, about  $1/20$  Ångström, was short enough, but the optics was imperfect. The best electron objective which one can make can be compared in optical perfection to a raindrop than to a microscope objective, and through the theoretical work of O. Scherzer it was known that it could never be perfected. The theoretical limit at that time was estimated at 4 Å, just about twice what was

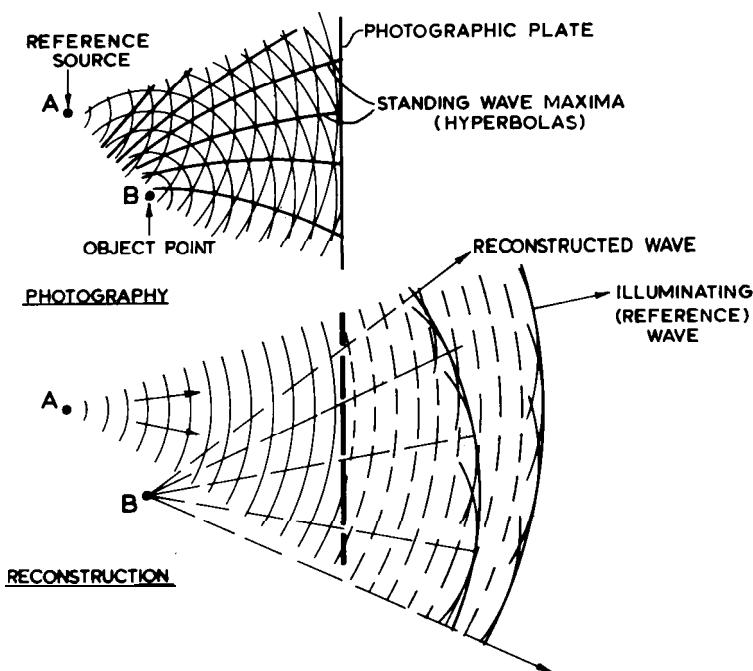


Fig. 2.

The Basic Idea of Holography, 1947.

needed to resolve atomic lattices, while the practical limit stood at about 12 Å. These limits were given by the necessity of restricting the aperture of the electron lenses to about 5/1000 radian, at which angle the spherical aberration error is about equal to the diffraction error. If one doubles this aperture so that the diffraction error is halved, the spherical aberration error is increased 8 times, and the image is hopelessly blurred.

After pondering this problem for a long time, a solution suddenly dawned on me, one fine day at Easter 1947, more or less as shown in Figure 2. Why not take a bad electron picture, but one which contains the *whole* information, and correct it by optical means? It was clear to me for some time that this could be done, if at all, only with coherent electron beams, with electron waves which have a definite phase. But an ordinary photograph loses the phase completely, it records only the intensities. No wonder we lose the phase, if there is nothing to compare it with! Let us see what happens if we add a standard to it, a "coherent background". My argument is illustrated in Figure 2, for the simple case when there is only one object point. The interference of the object wave and of the coherent background or "reference wave" will then produce interference fringes. There will be maxima wherever the phases of the two waves were identical. Let us make a hard positive record, so that it transmits only at the maxima, and illuminate it with the reference source alone. Now the phases are of course right for the reference source A, but as at the slits the phases are identical, they must be right also for B; therefore the wave of B must also appear, *reconstructed*.

A little mathematics soon showed that the principle was right, also for more than one object point, for any complicated object. Later on it turned out that in holography Nature is on the inventor's side; there is no need to take a hard positive record; one can take almost any negative. This encouraged me to complete my scheme of electron microscopy by reconstructed wavefronts, as I then called it and to propose the two-stage process shown in Figure 3. The electron microscope was to produce the interference figure between the object beam and the coherent background, that is to say the non-diffracted part of the illuminating beam. This interference pattern I called a "hologram", from the Greek word "holos"-the whole, because it contained the whole information. The hologram was then reconstructed with light, in an optical system which corrected the aberrations of the electron optics (1).

In doing this, I stood on the shoulders of two great physicists, W. L. Bragg and Frits Zernike. Bragg had shown me, a few years earlier, his "X-ray microscope" an optical Fourier-transformer device. One puts into it a small photograph of the reciprocal lattice, and obtains a projection of the electron densities, but only in certain exceptional cases, when the phases are all real, and have the same sign. I did not know at that time, and neither did Bragg, that Mieczislav Wolfke had proposed this method in 1920, but without realising it experimentally.<sup>1</sup> So the idea of a two-stage method was inspired by Bragg. The coherent background, on the other hand, was used with great success by Frits

<sup>1</sup>M. Wolfke, Phys. Zeits. 21, 495-7, Sept. 15, 1920.

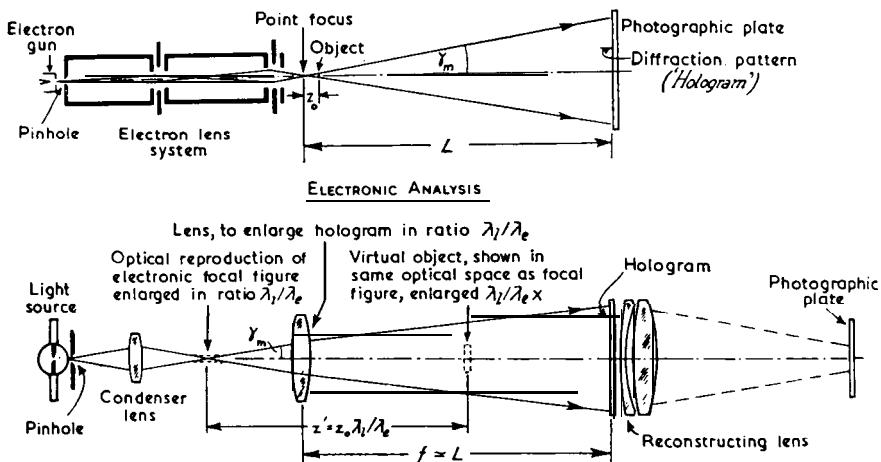


Fig. 3.

The Principle of Electron Microscopy by Reconstructed Wavefronts (Gabor, Proc. Royal Society, A, 197, 454, 1949).

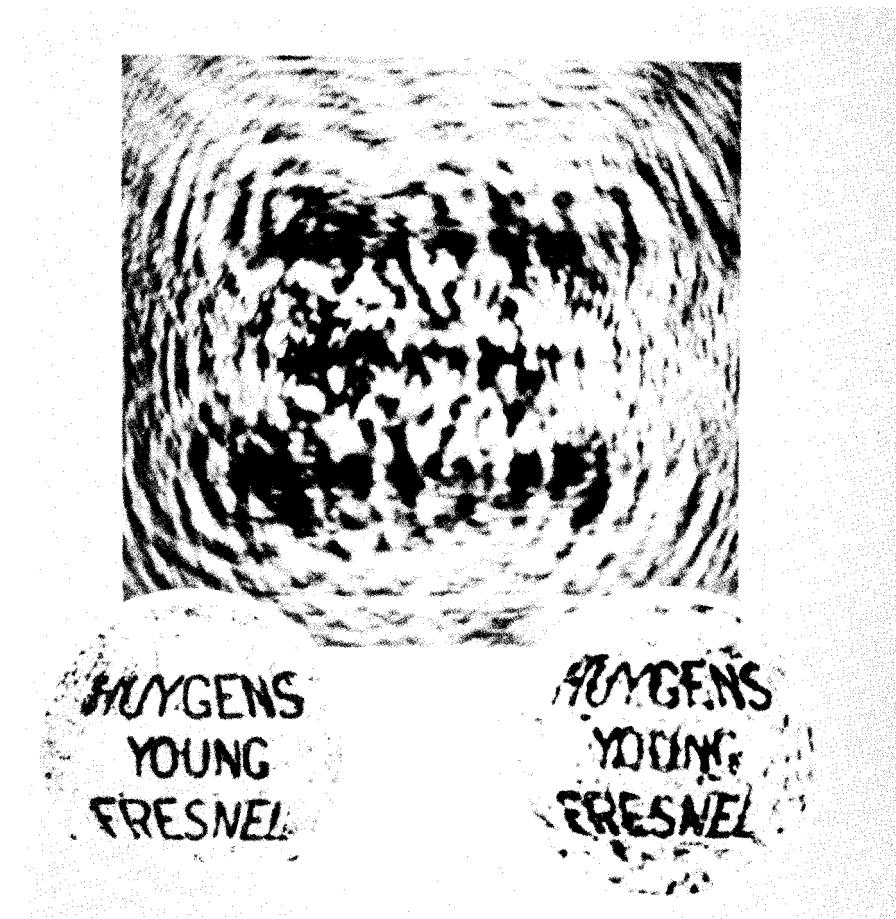


Fig. 4.

First Holographic Reconstruction, 1948

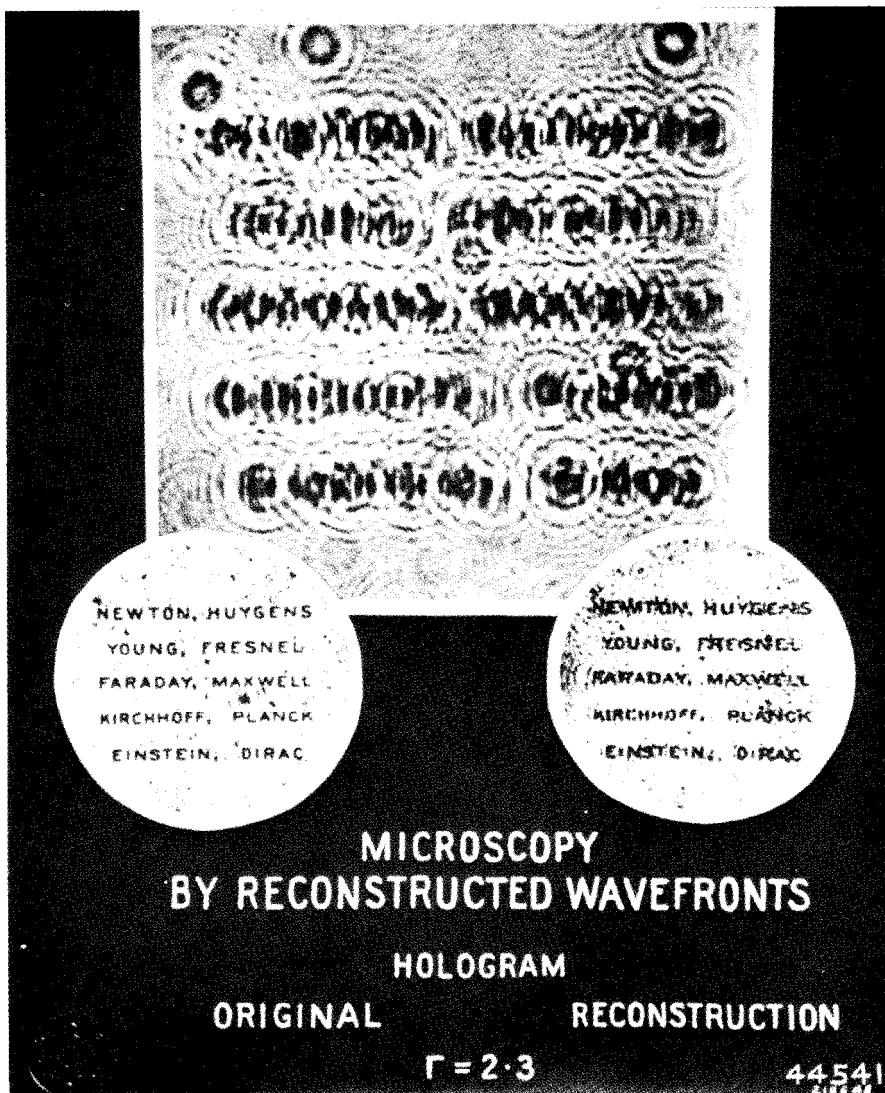


Fig. 5.

Another Example of Early Holography, 1948 (Gabor, Proc. Roy. Soc. A, 197, 454, 1949).

Zernike in his beautiful investigations on lens aberrations, showing up their phase, and not just their intensity. It was only the reconstruction principle which had escaped them.

In 1947 I was working in the Research Laboratory of the British Thomson-Houston Company in Rugby, England. It was a lucky thing that the idea of holography came to me via electron microscopy, because if I had thought of optical holography only, the Director of Research, L. J. Davies, could have objected that the BTH company was an electrical engineering firm, and not in the optical field. But as our sister company, Metropolitan Vickers were makers of electron microscopes, I obtained the permission to carry out some

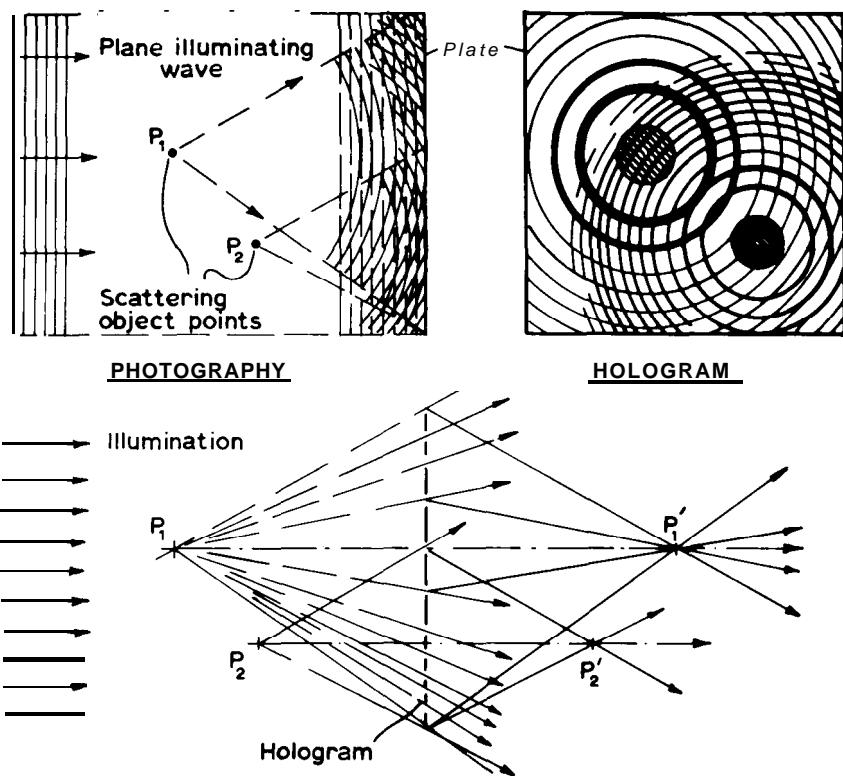


Fig. 6.

The Second Image. Explanation in Terms of Soret-lenses as Holograms of Single Object Points.

optical experiments. Figure 4 shows one of our first holographic reconstructions. The experiments were not easy. The best compromise between coherence and intensity was offered by the high pressure mercury lamp, which had a coherence length of only 0.1 mm, enough for about 200 fringes. But in order to achieve spatial coherence, we (my assistant Ivor Williams and I) had to illuminate, with one mercury line, a pinhole of 3 microns diameter. This left us with enough light to make holograms of about 1 cm diameter of objects, which were microphotographs of about 1 mm diameter, with exposures of a few minutes, on the most sensitive emulsions then available. The small coherence length forced us to arrange everything in one axis. This is now called "in line" holography, and it was the only one possible at that time. Figure 5 shows a somewhat improved experiment, the best of our series. It was far from perfect. Apart from the *schlieren*, which cause random disturbances, there was a systematic defect in the pictures, as may be seen by the distortion of the letters. The explanation is given in Figure 6. The disturbance arises from the fact that there is not one image but two. Each point of the object emits a spherical secondary wave, which interferes with the background and produces a system of circular Fresnel zones. Such a system is known after the optician who first produced it, a Soret lens. This is, at the same time, a positive and a

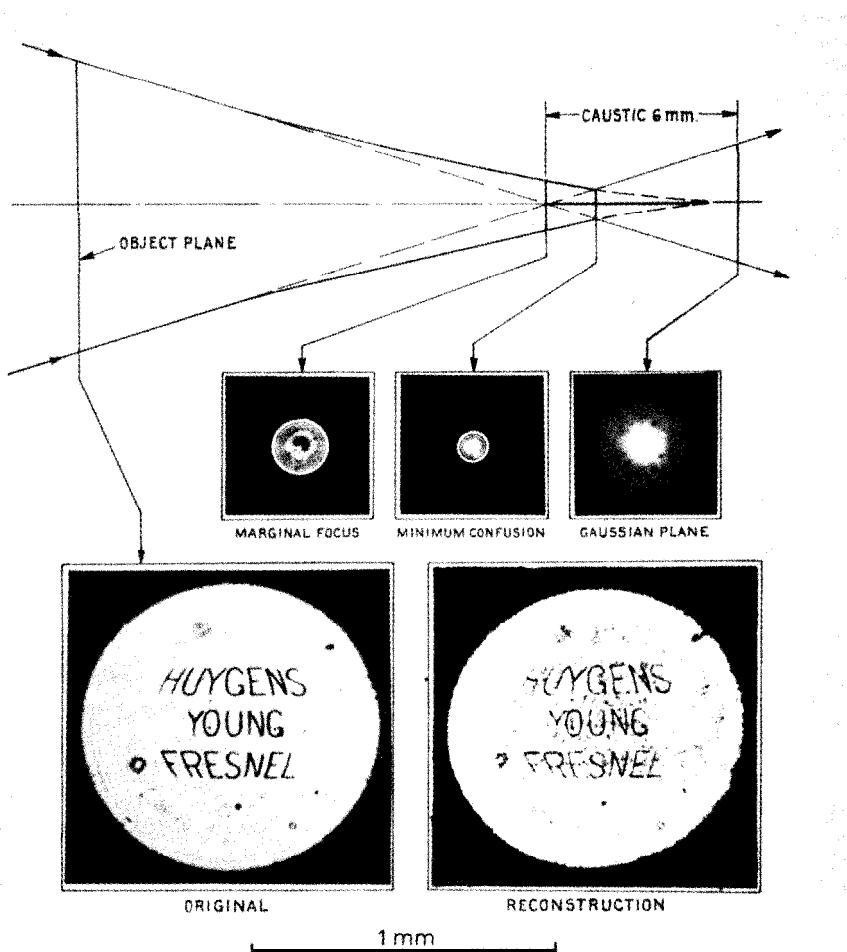


Fig. 7.  
Elimination of the Second Image, by Compensation of the Spherical Aberration in the Reconstruction (Gabor, 1948; published 1951).

negative lens. One of its foci is in the original position of the object point, the other in a position conjugate to it, with respect to the illuminating wavefront. If one uses "in-line holography" both images are in line, and can be separated only by focusing. But the separation is never quite perfect, because in regular, coherent illumination every point leaves a "wake" behind it, which reaches to long distances.

I will tell later with what ease modern laser holography has got rid of this disturbance, by making use of the superior coherence of laser light which was not at my disposal in 1948. However, I was confident that I could eliminate the second image in the application which alone interested me at that time: seeing atoms with the electron microscope. This method, illustrated in Figure 7, utilized the very defect of electron lenses, the spherical aberration, in order to defeat the second image. If an electron hologram is taken with a lens with spherical aberration, one can afterwards correct one of the two images by

suitable optics, and the other has then twice the aberration, which washes it out almost completely. Figure 7 shows that a perfectly sharp reconstruction, in which as good as nothing remains of the disturbance caused by the second image, can be obtained with a lens so bad that its definition is at least 10 times worse than the resolution which one wants to obtain. Such a very bad lens was obtained using a microscope objective the wrong way round, and using it again in the reconstruction.

So it was with some confidence that two years later, in 1950 we started a programme of holographic electron microscopy in the Research Laboratory of the Associated Electrical Industries, in Aldermaston, under the direction of Dr T. E. Allibone, with my friends and collaborators M. W. Haine, J. Dyson and T. Mulvey.<sup>2</sup> By that time I had joined Imperial College, and took part in the work as a consultant. In the course of three years we succeeded in considerably improving the electron microscope, but in the end we had to give up, because we had started too early. It turned out that the electron microscope was still far from the limit imposed by optical aberrations. It suffered from vibrations, stray magnetic fields, creep of the stage, contamination of the object, all made worse by the long exposures required in the weak coherent electron beam. Now, 20 years later, would be the right time to start on such a programme, because in the meantime the patient work of electron microscopists has overcome all these defects. The electron microscope resolution is now right up to the limit set by the spherical aberration, about 3.5 Å, and only an improvement by a factor of 2 is needed to resolve atomic lattices. Moreover, there is no need now for such very long exposures as we had to contemplate in 1951, because by the development of the field emission cathode the coherent current has increased by a factor of 3-4 orders of magnitude. So perhaps I may yet live to see the realisation of my old ideas.

My first papers on wavefront reconstruction evoked some immediate responses. G. L. Rogers (2) in Britain made important contributions to the technique, by producing among other things the first phase holograms, and also by elucidating the theory. In California Alberto Baez (3) Hussein El-Sum and P. Kirckpatrick (4) made interesting forays into X-ray holography. For my part, which my collaborator W. P. Goss, I constructed a holographic interference microscope, in which the second image was annulled in a rather complicated way by the superimposition of two holograms, "in quadrature" with one another. The response of the optical industry to this was so disappointing that we did not publish a paper on it until 11 years later, in 1966 (5). Around 1955 holography went into a long hibernation.

The revival came suddenly and explosively in 1963, with the publication of the first successful laser<sup>3</sup> holograms by Emmett N. Leith and Juris Upatnieks

<sup>2</sup>Supported by a grant of the D.S.I.R. (Direction of Scientific and Industrial Research) the first research grant ever given by that body to an industrial laboratory.

<sup>3</sup>I have been asked more than once why I did not invent the laser. In fact, I have thought of it. In 1950, thinking of the desirability of a strong source of coherent light, I remembered that in 1921, as a young student, in Berlin, I had heard from Einstein's own lips his wonderful derivation of Planck's law which postulated the existence of

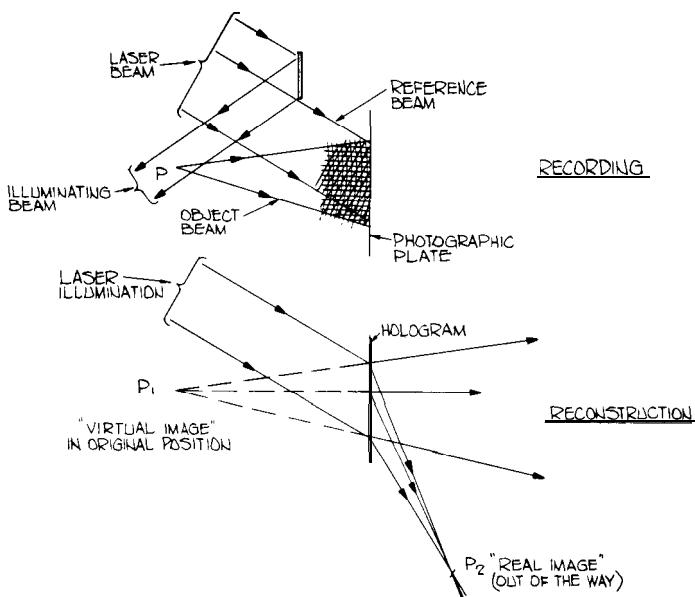


Fig. 8.  
Holography with Skew Reference Beam. E. N. Leith and J. Upatnieks, 1963.

of the University of Michigan, Ann Arbor. Their success was due not only to the laser, but to the long theoretical preparation of Emmett Leith, which started in 1955. This was unknown to me and to the world, because Leith, with his collaborators Cutrona, Palermo, Porcello and Vivian applied his ideas first to the problem of the "side-looking radar" which at that time was classified (6). This was in fact two-dimensional holography with electromagnetic waves, a counterpart of electron holography. The electromagnetic waves used in radar are about 100,000 times longer than light waves, while electron waves are about 100,000 times shorter. Their results were brilliant, but to my regret I cannot discuss them for lack of time.

When the laser became available, in 1962, Leith and Upatnieks could at once produce results far superior to mine, by a new, simple and very effective method of eliminating the second image (7). This is the method of the "skew reference wave", illustrated in Figure 8. It was made possible by the great coherence length of the helium-neon laser, which even in 1962 exceeded that of the mercury lamp by a factor of about 3000. This made it possible to separate the reference beam from the illuminating beam; instead of going through the object, it could now go around it. The result was that the two reconstructed images were now separated not only in depth, but also angularly, by twice the incidence angle of the reference beam. Moreover, the intensity of

stimulated emission. I then had the idea of the pulsed laser: Take a suitable crystal, make a resonator of it by a highly reflecting coating, fill up the upper level by illuminating it through a small hole, and discharge it explosively by a ray of its own light. I offered the idea as a Ph.D. problem to my best student, but he declined it, as too risky, and I could not gainsay it, as I could not be sure that we would find a suitable crystal.

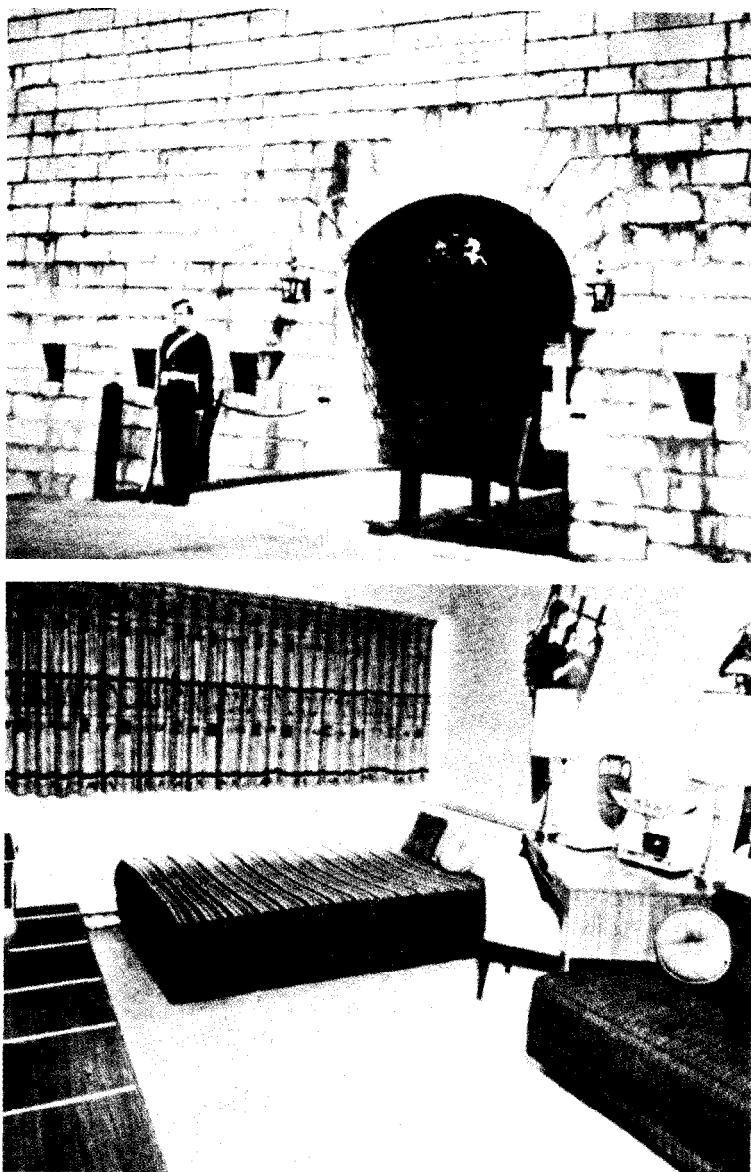


Fig. 9.

First Example of Multiple Image Storage in One Hologram. E. N. Leith and J. Upatnieks, Journal Optical Society of America, November 1964.

the coherent laser light exceeded that of mercury many millionfold. This made it possible to use very fine-grain, low speed photographic emulsions and to produce large holograms, with reasonable exposure times.

Figure 9 shows two of the earliest reconstructions made by Leith and Upatnieks, in 1963, which were already greatly superior to anything that I could produce in 1948. The special interest of these two images is, that they are reconstructions from one hologram, taken with different positions of the reference beam. This was the first proof of the superior storage capacity of

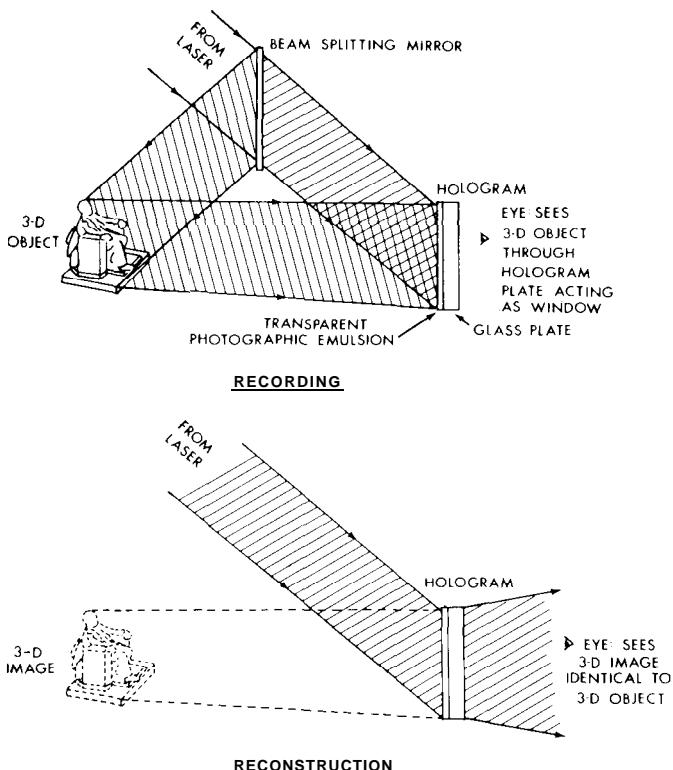


Fig. 10.  
3-D Holography of a Diffusing Object with Laser Light.

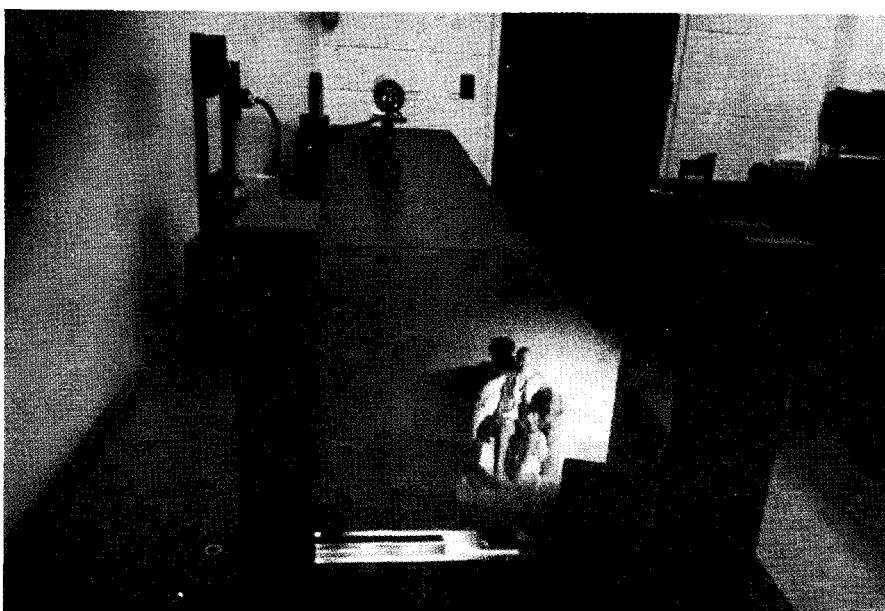


Fig. 11.  
Three dimensional Reconstruction of a Small Statue of Abraham Lincoln. (Courtesy of Professor G. W. Stroke, State University of New York, Stony Brook).



Fig. 12.

Strongly Magnified Image of a Hologram taken with Diffused Illumination. The Information is conveyed in a noiselike code. E. N. Leith and J. Upatnieks, 1964.

holograms. Leith and Upatnieks could soon store 12 different pictures in one emulsion. Nowadays one can store 100 or even 300 pages of printed matter in an area which by ordinary photography would be sufficient for one.

From then on progress became very rapid. The most spectacular result of the first year was the holography of three dimensional objects, which could be seen with two eyes. Holography was of course three dimensional from the start, but in my early, small holograms one could see this only by focusing through the field with a microscope or short-focus eyepiece. But it was not enough to make the hologram large, it was also necessary that every point of the photographic plate should see every point of the object. In the early holograms, taken with regular illumination, the information was contained in a small area, in the diffraction pattern.

In the case of rough, diffusing objects no special precautions are necessary. The small dimples and projections of the surface diffuse the light over a large cone. Figure 10 shows an example of the setup in the case of a rough object, such as a statuette of Abraham Lincoln. The reconstruction is shown in Figure 11. With a bleached hologram ("phase hologram") one has the impression of looking through a clear window at the statuette itself.

If the object is non-diffusing, for instance if it is a transparency, the information is spread over the whole hologram area by illuminating the object through a diffuser, such as a frosted glass plate. The appearance of such a "diffused" hologram is extraordinary; it looks like noise. One can call it "ideal Shannon coding", because Claude E. Shannon has shown in his Communication Theory that the most efficient coding is such that all regularities seem to have disappeared in the signal: it must be "noise-like". But where is the information in this chaos? It can be shown that it is not as irregular as it appears. It is not as if grains of sand had been scattered over the plate at random. It is rather a complicated figure, the diffraction pattern of the object, which is repeated at random intervals, but always in the same size and same orientation.

A very interesting and important property of such diffused holograms is that any small part of it, large enough to contain the diffraction pattern, contains information on the whole object, and this can be reconstructed from the fragment, only with more noise. A diffuse hologram is therefore a *distributed memory*, and this was evoked much speculation whether human memory is not perhaps, as it were, holographic, because it is well known that a good part of the brain can be destroyed without wiping out every trace of a memory. There is no time here to discuss this very exciting question. I want only to say that in my opinion the similarity with the human memory is functional only, but certainly not structural.

It is seen that in the development of holography the holograms has become always more unlike the object, but the reconstruction always more perfect. Figure 13 shows an excellent reconstruction by Leith and Upatnieks of a



Fig. 13.

Reconstruction of a Plane Transparency, Showing a Restaurant, from a Hologram taken with Diffused Illuminations (E. N. Leith and J. Upatnieks, 1964).

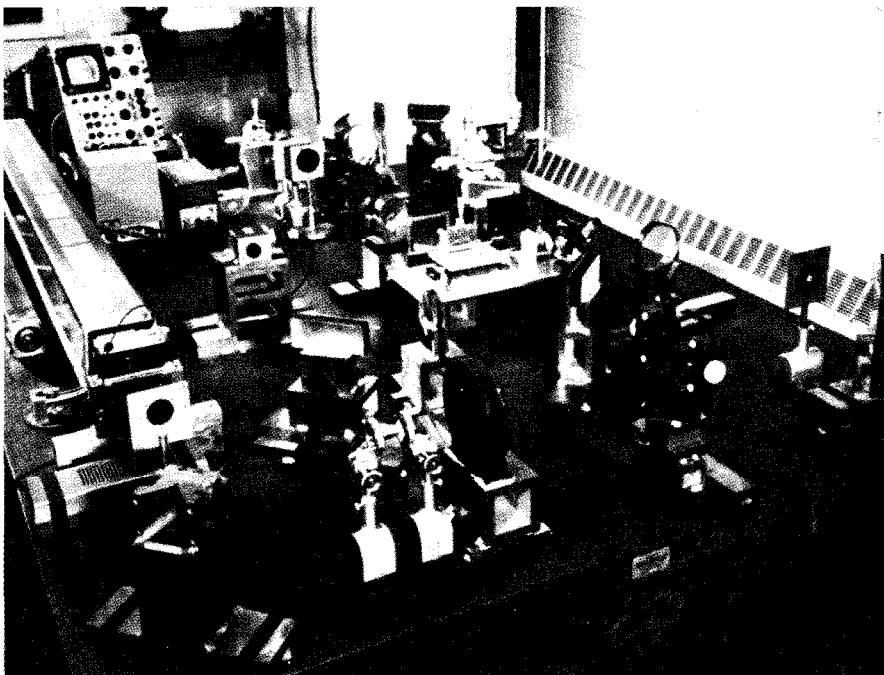


Fig. 14.  
Modern Holographic Equipment.

photograph, from a diffuse hologram like the one in the previous figure.

The pioneer work carried out in the University of Michigan, Ann Arbor, led also to the stabilization of holographic techniques. Today hundreds if not thousands of laboratories possess the equipment of which an example is shown in Figure 14; the very stable granite slab or steel table, and the various optical devices for dealing with coherent light, which are now manufactured by the optical industry. The great stability is absolutely essential in all work carried out with steady-state lasers, because a movement of the order of a quarter wavelength during the exposure can completely spoil a hologram.

However, from 1965 onwards there has developed an important branch of holography where high stability is not required, because the holograms are taken in a small fraction of a microsecond, with a pulsed laser.

Imagine that you had given a physicist the problem: "Determine the size of the droplets which issue from a jet nozzle, with a velocity of 2 Mach. The sizes are probably from a few microns upwards." Certainly he would have thrown up his hands in despair! But all it takes now, is to record a simple in-line hologram of the jet, with the plate at a safe distance, with a ruby laser pulse of 20-30 nanoseconds. One then looks at the "real" image (or one reverses the illuminating beam and makes a real image of the virtual one), one dives with a microscope into the three-dimensional image of the jet and focuses the particles, one after the other. Because of the large distance, the disturbance by the second image is entirely negligible. Figure 15 shows a fine example.

As the research workers of the TRW laboratories have shown, it is possible

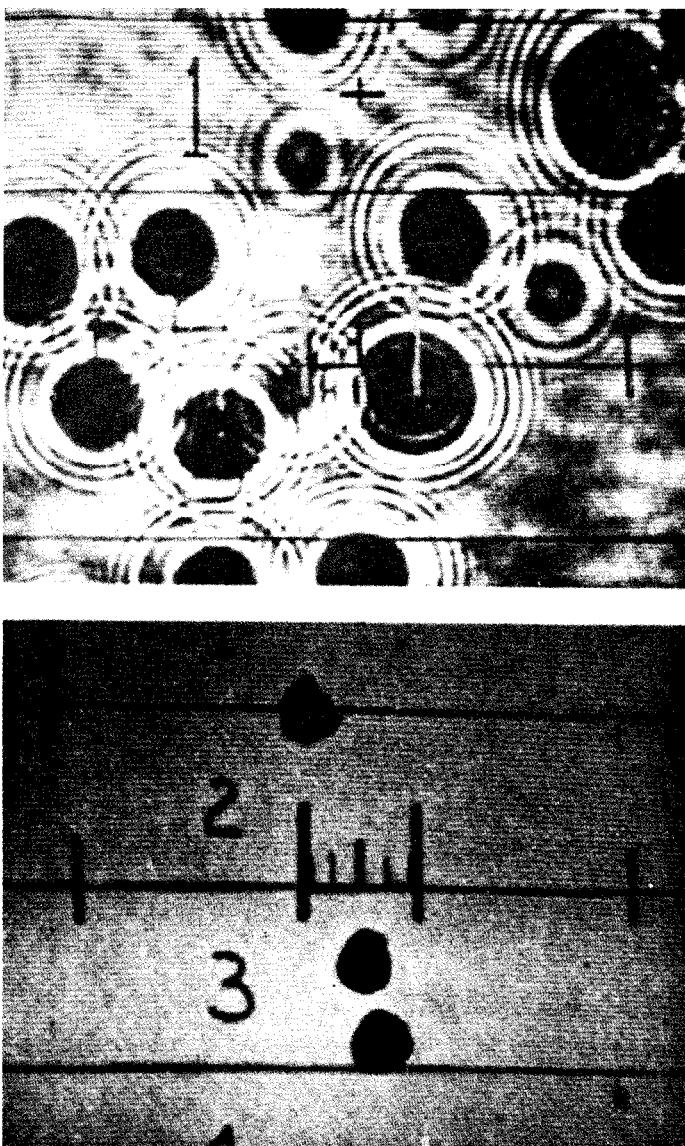


Fig. 15.  
Holography of Jets. (Courtesy of Laser Holography Inc., Santa Barbara, California.)

to record in one hologram the infusoriae in several feet of dirty water, or insects in a meter of air space. Figure 16 shows two reconstructions of insects from one hologram, focusing on one after the other. The authors, C. Knox and R. E. Brooks, have also made a cinematographic record of a holographic film, in which the flight of one mosquito is followed through a considerable depth, by refocusing in every frame (9).

Another achievement of the TRW group, Ralph Wuerker and his colleagues, leads us into another branch of holography, to holographic interferometry. Figure 17 shows a reconstruction of a bullet, with its train of shockwaves, as

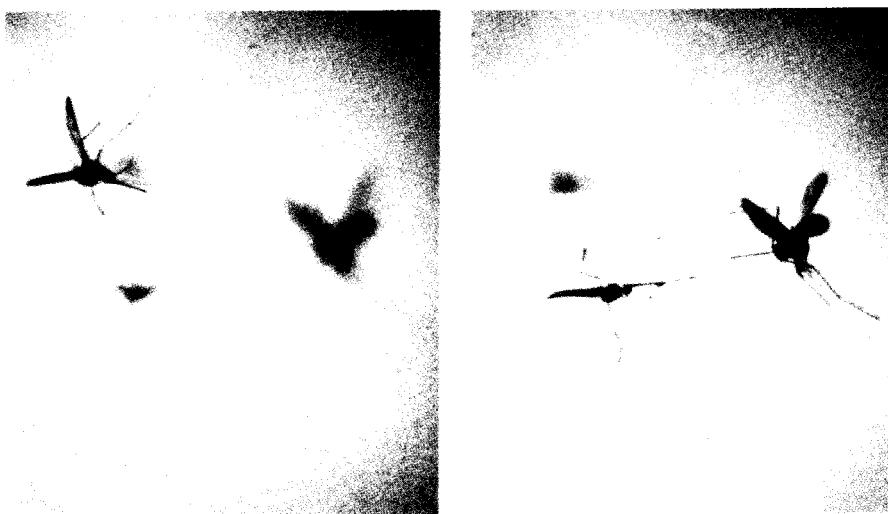


Fig. 16.

Observation of Mosquitos in Flight. Both Pictures are extracted from one Hologram. (Courtesy of C. Knox and R. E. Brooks, TRW, Redondo Beach, California.

it meets another shockwave. But it is not just an image, it is an interferometric image. The fringes show the *loci* at which the retardation of light is by integer wavelengths, relative to the quiet air, before the event. This comparison standard is obtained by a previous exposure. This is therefore a double-exposure hologram, such as will be discussed in more detail later (10).

Figure 18 shows another high achievement of pulse holography: a holographic, three-dimensional portrait, obtained by L. Siebert in the Conductron Corporation (now merged into McDonnel-Douglas Electronics Company, St Charles, Missouri). It is the result of outstanding work in the development of lasers. The ruby laser, as first realised by T. H. Maiman, was capable of short pulses, but its coherence length was of the order of a few cm only. This is no obstacle in the case of in-line holography, where the reference wave proceeds almost in step with the diffracted wavelets, but in order to take a scene of, say, one meter depth with reflecting objects one must have a coherence length of at least one meter. Nowadays single-mode pulses of 30 nanosecond duration with 10 joule in the beam and coherence lengths of 5-8 meters are available, and have been used recently for taking my holographic portrait shown in the exhibition attached to this lecture.

In 1965 R. L. Powell and K. A. Stetson in the University of Michigan, Ann Arbor, made an interesting discovery. Holographic images taken of moving objects are washed out. But if double exposure is used, first with the object at rest, then in vibration, fringes will appear, indicating the lines where the displacement amounted to multiples of a half wavelength. Figure 19 shows vibrational modes of a loudspeaker membrane, recorded in 1965 by Powell and Stetson (11), Figure 20 the same for a guitar, taken by H. A. Stetson in the laboratory of Professor Erik Ingelstam (12).

Curiously, both the interferograms of the TRW group and the vibrational



Fig. 17.

Dynamic Holographic Interferometry. This Reconstruction of a Holographic Interferogram shows the interaction of two air shock fronts and their associated flows. (Courtesy of Dr R. F. Wuerker and his associates, TRW Physical Electronics Laboratory, Redondo Beach, Calif.)

records of Powell and Stetson preceded what is really a simpler application of the interferometrical principle, and which historically ought to have come first—if the course of science would always follow the shortest line. This is the observation of small deformations of solid bodies, by double exposure holograms. A simple explanation is as follows: We take a hologram of a body in State A. This means that we freeze in the wave A by means of a reference beam. Now let us deform the body so that it assumes the State B and take a second hologram in the same emulsion with the same reference beam. We develop the hologram, and illuminate it with the reference beam. Now the two waves A and B, frozen in at different times, and which have never seen one another, will be revived simultaneously, and they interfere with one an-

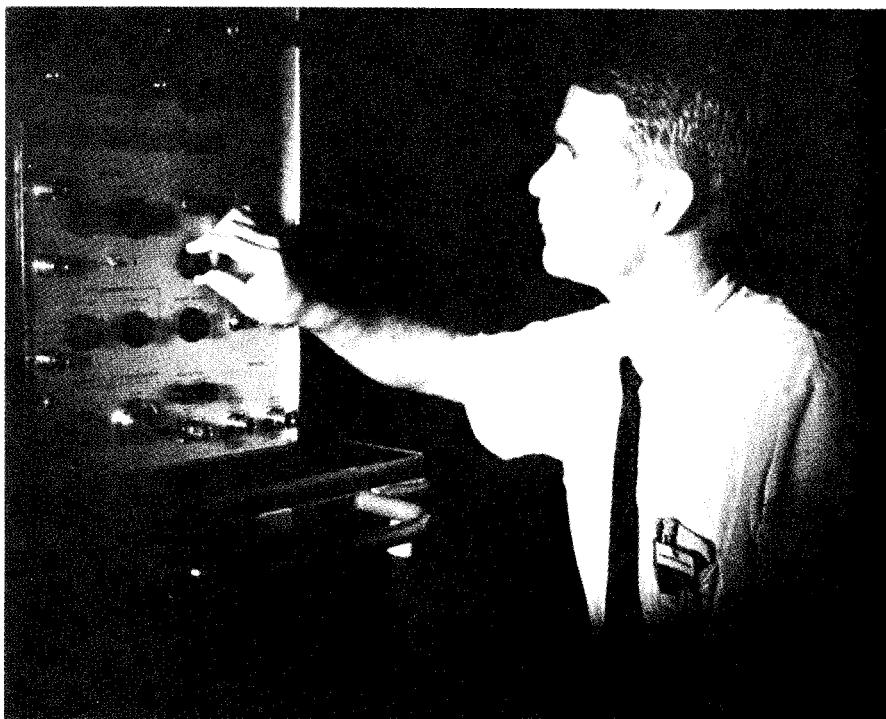


Fig. 18.

Holographic Portrait. (I. Siebert, Conductron Corporation, now merged into McDonnell-Douglas Electronics Company, St Charles, Missouri.)

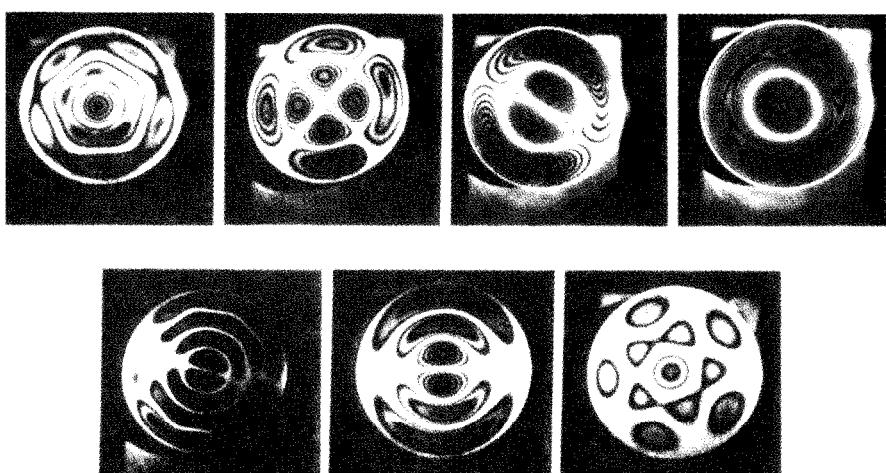


Fig. 19.

Vibrational Modes of a Loudspeaker membrane, obtained by Holographic Interferometry. (R. L. Powell and K. A. Stetson, University of Michigan, Ann Arbor, 1965.)

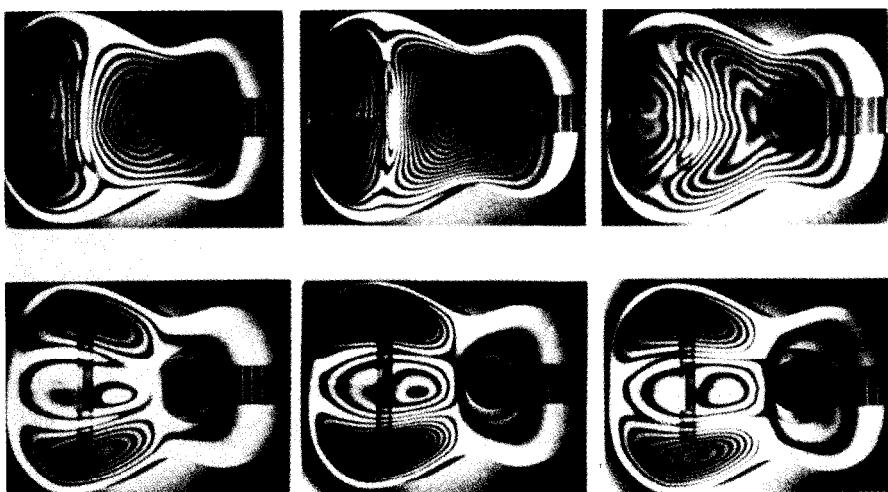


Fig. 20.

Vibrational Modes of a Guitar, Recorded by Holographic Interferometry. (Courtesy of Dr K. A. Stetson and Professor E. Ingelstarn.)

other. The result is that Newton fringes will appear on the object, each fringe corresponding to a deformation of a half wavelength. Figure 21 shows a fine example of such a holographic interferogram, made in 1965 by Haines and Hildebrand. The principle was discovered simultaneously and independently also by J. M. Burch in England, and by G. W. Stroke and A. Labeyrie in Ann

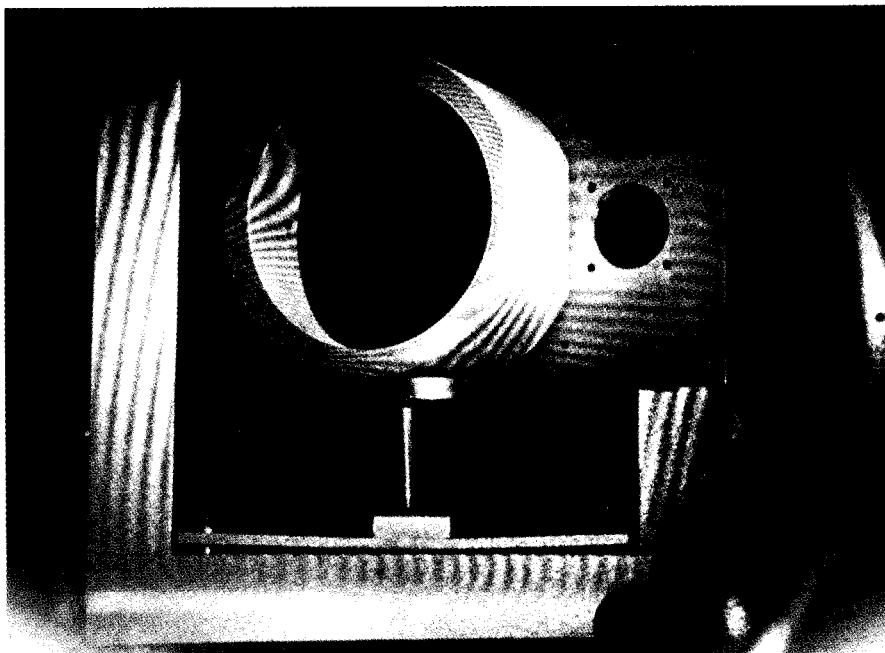


Fig. 21.

An early Example of Holographic Interferometry by Double Exposure. (Haines and Hildebrand, University of Michigan, Ann Arbor, 1965.)



Fig. 22.

Non-destructive Testing by Holography. Double Exposure Hologram, revealing two flaws in a tyre (Courtesy of Dr Ralph Grant and GCO, Ann Arbor, Michigan).

Arbor, Michigan.

Non-destructive testing by holographic interferometry is now by far the most important industrial application of holography. It gave rise to the first industrial firm based on holography, GCO (formerly G. C. Optronics), in Ann Arbor, Michigan, and the following examples are reproduced by courtesy of GCO. Figure 22 shows the testing of a motor car tyre. The front of the tyre is photographed directly, the sides are seen in two mirrors, right and left. First a little time is needed for the tyre to settle down and a first hologram is taken. Then a little hot air is blown against it, and a second exposure is made, on the same plate. If the tyre is perfect, only a few, widely spaced fringes will appear, indicating almost uniform expansion. But where the cementing of the rubber sheets was imperfect, a little blister appears, as seen near the centre and near the top left corner, only a few thousandths of a millimeter high, but indicating a defect which could become serious. Alternatively, the first hologram is developed, replaced exactly in the original position, and the expansion of the tyre is observed "live".

Other examples of non-destructive testing are shown in Figure 23; all defects which are impossible or almost impossible to detect by other means, but which reveal themselves unmistakably to the eye. A particularly impressive piece of equipment manufactured by GCO is shown in Figure 24. It is a holographic analyser for honeycomb sandwich structures (such as shown in the middle of

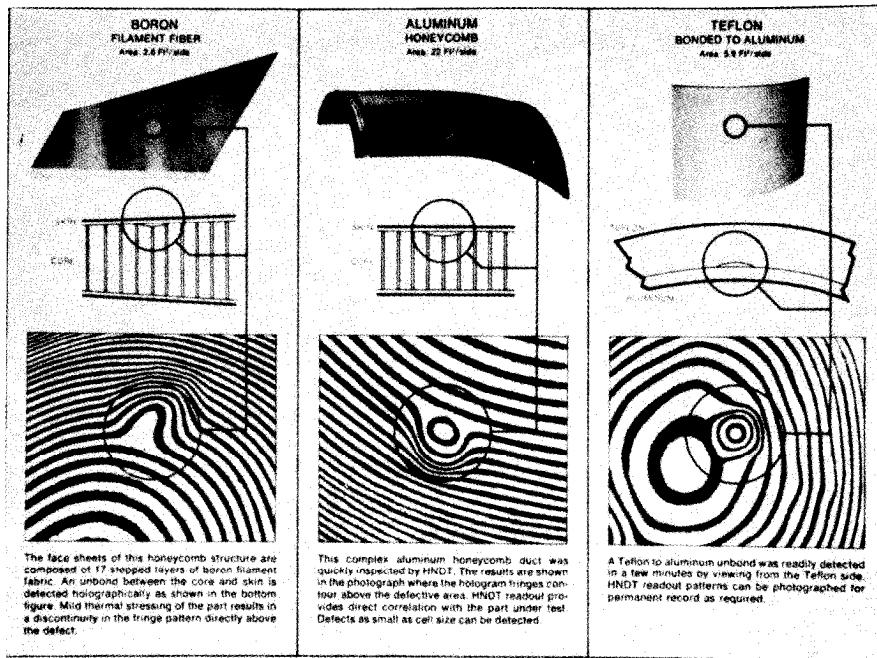


Fig. 23.  
Examples of Holographic Non-destructive Testing. (Courtesy of GCO, Ann Arbor, Michigan.)

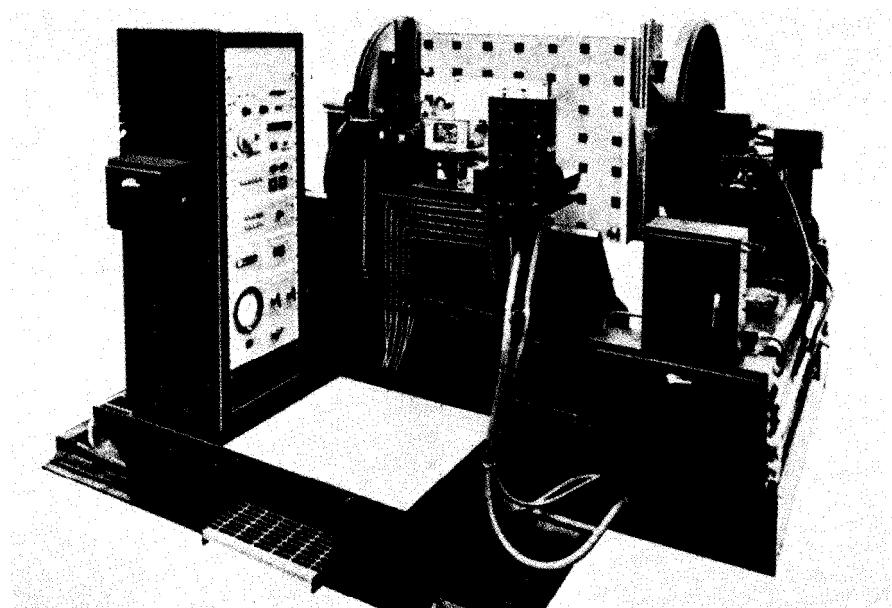


Fig. 24.  
Holographic Analyzer Mark II for Sandwich Structures, GCO, Ann Arbor, Michigan.



Fig. 25.

Holographic Contour Map, made by a method initiated by B. P. Hildebrand and K. A. Haines (Journal, Optical Society of America, 57, 155, 1967). Improved by J. Varner, University of Michigan, Ann Arbor, 1969.

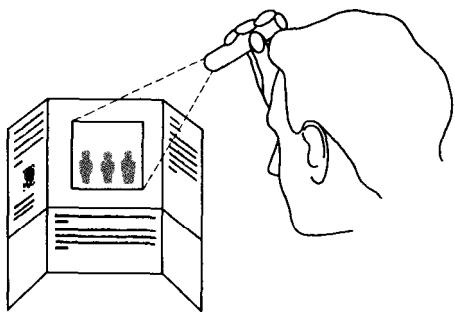
Figure 23) which are used in aeroplane wings. The smallest welding defect between the aluminum sheets and the honeycomb is safely detected at one glance.

While holographic interferometry is perfectly suited for the detection of very small deformations, with its fringe unit of  $1/4000$  mm, it is a little too fine for the checking of the accuracy of workpieces. Here another holographic technique called "contouring" is appropriate. It was first introduced by Haines and Hildebrand, in 1965, and has been recently much improved by J. Varner, also in Ann Arbor, Michigan. Two holograms are taken of the same object, but with two wavelengths which differ by e.g. one percent. This produces beats between the two-fringe system, with fringe spacings corresponding to about  $1/40$  mm, which is just what the workshop requires (Figure 25).

From industrial applications I am now turning to another important development in holography. In 1962, just before the "holography explosion" the Soviet physicist Yu. N. Denisyuk published an important paper (13) in which he combined holography with the ingenious method of photography in natural colours, for which Gabriel Lippmann received the Nobel Prize in 1908. Figure 26 a illustrates Lippmann's method and Denisyuk's idea. Lippmann produced a very fine-grain emulsion, with colloidal silver bromide, and backed the emulsion with mercury, serving as a mirror. Light falling on the emulsion was

## How to view the Lippmann type reflection hologram

---



For maximum brightness (due to fulfillment of the Bragg criterion) the hologram shall be illuminated diagonally from the upper righthand corner. An ordinary penlight at a distance of about 25 cm is recommended, see figure. Other approximately point source lighting can be used, such as spotlight, slide projector, or even direct unclouded sunlight.

NB: The hologram ought to be viewed in subdued lighting, and direct overhead light be avoided. The side screens (partly book pages), as indicated in the figure, are good for screening off room light.

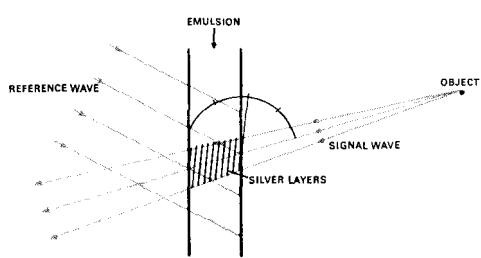
---

THE HOLOGRAM IS NOT REPRODUCED HERE DUE TO COMMERCIAL UNAVAILABILITY.

---

## How the Lippman type reflection hologram has been constructed

The figure shows how the reference wave comes from one side of the emulsion, the signal wave from the object from the other



side. The dotted line indicates how, at the reconstruction, a wave reflected from the silver layers in the emulsion is obtained, and you see in its extension backwards the object as it was at the registration. (Stroke-Labeyrie, see References.)

In fact, at the practical registration of a reflection hologram, the signal wave comes from the different points of the illuminated object. In order to have the reconstructed image of the object close to the hologram included, an image of the object has been transported there by means of a special lens. This gives localization of the image closely in front of and behind the hologram.

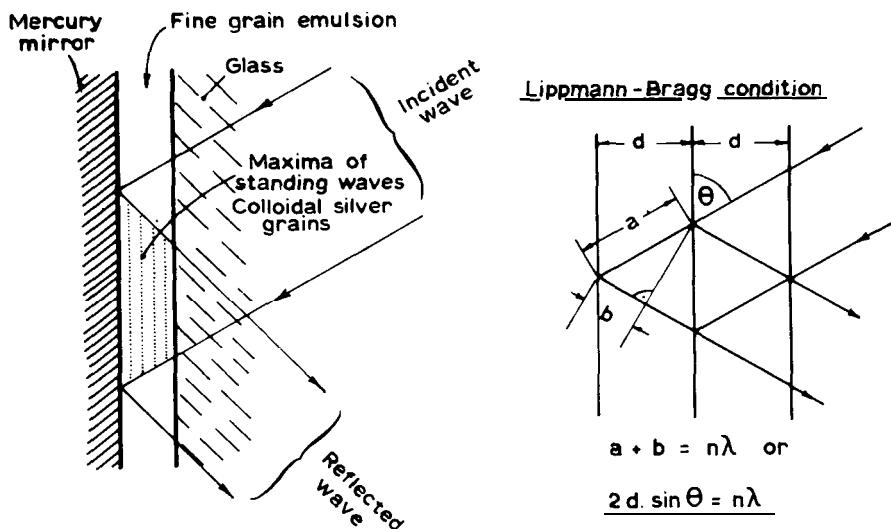


Fig. 26 a. Gabriel Lippmann's method of photography in natural colours.

reflected at the mirror, and produced a set of standing waves. Colloidal silver grains were precipitated in the maxima of the electric vector, in layers spaced by very nearly half a wavelength. After development, the complex of layers, illuminated with white light, reflected only a narrow waveband around the original colour, because only for this colour did the wavelets scattered at the Lippmann layers add up in phase.

Denisyuk's suggestion is shown in the second diagram. The object wave and the reference wave fall in from opposite sides of the emulsion. Again standing waves are produced, and Lippman layers, but these are no longer parallel to the emulsion surface, they bisect the angle between the two wavefronts. If now, and this is Denisyuk's principle, the developed emulsion is illuminated by the

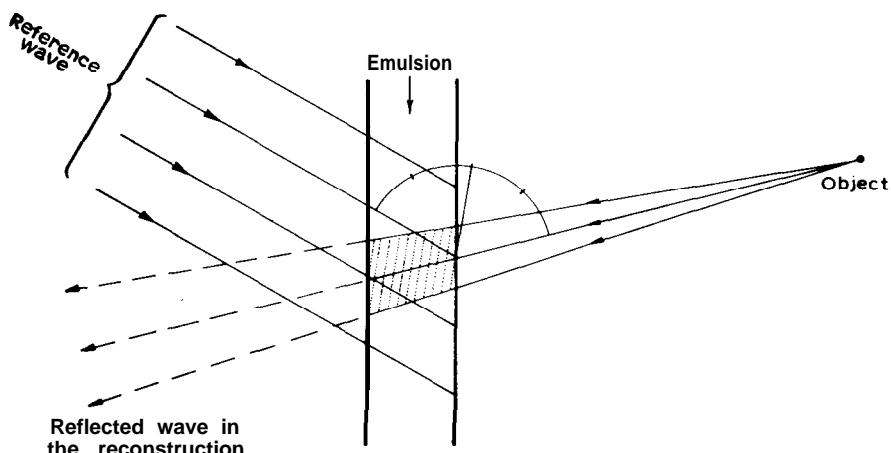


Fig. 26 b. Lippmann-Denisyuk-Stroke Reflection Hologram.

reference wave, the object will appear, in the original position and (unless the emulsion has shrunk) in the original colour.

Though Denisyuk showed considerable experimental skill, lacking a laser in 1962 he could produce only an "existence proof". A two-colour reflecting hologram which could be illuminated with white light was first produced in 1965 by G. W. Stroke and A. Labeyrie (14) and is shown in Figure 27.

Since that time single-colour reflecting holograms have been developed to high perfection by new photographic processes, by K. S. Pennington (15) and others, with reflectances approaching 100 percent, but two; and even more, three-colour holograms are still far from being satisfactory. It is one of my chief preoccupations at the present to improve this situation, but it would take too long, and it would be also rather early to enlarge on this.

An application of holography which is certain to gain high importance in the next years is information storage. I have mentioned before that holography allows storing 100-300 times more printed pages in a given emulsion than ordinary microphotography. Even without utilizing the depth dimension, the factor is better than 50. The reason is that a diffused hologram represents almost ideal coding, with full utilization of the area and of the gradation of the emulsion, while printed matter uses only about 5-10% of the area, and the gradation not at all. A further factor arises from the utilization of the third dimension, the depth of the emulsion. This possibility was first pointed out in an ingenious paper by P. J. van Heerden (16) in 1963. Theoretically it appears possible to store one bit of information in about one wavelength cube. This is far from being practical, but the figure of 300, previously mentioned, is entirely realistic.

However, even without this enormous factor, holographic storage offers important advantages. A binary store, in the form of a checkerboard pattern on microfilm can be spoiled by a single grain of dust, by a hair or by a scratch, while a diffused hologram is almost insensitive to such defects. The holographic store, illustrated in Figure 28, is according to its author L. K. Anderson (17) (1968) only a modest beginning, yet it is capable of accessing for instance any one of  $64 \times 64$  printed pages in about a microsecond. Each hologram, with a diameter of 1.2 mm can contain about  $10^4$  bits. Reading out this information

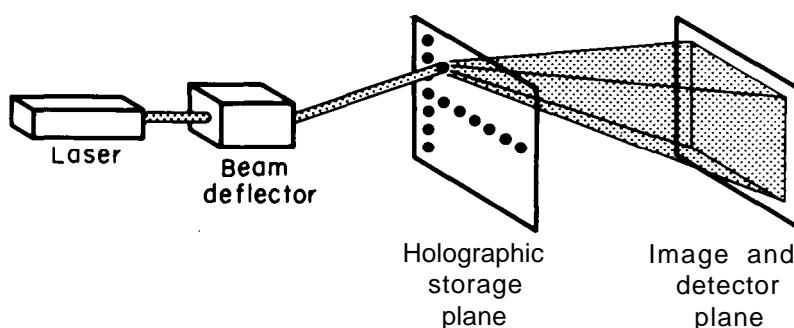


Fig. 28.

Holographic Flying Spot Store. L. K. Anderson and R. J. Collier, Bell Telephone Laboratories, 1968.



Fig. 27.

First two-colour Reflecting Hologram, Reconstructed in White Light. G. W. Stroke and A. Labeyrie, 1965.

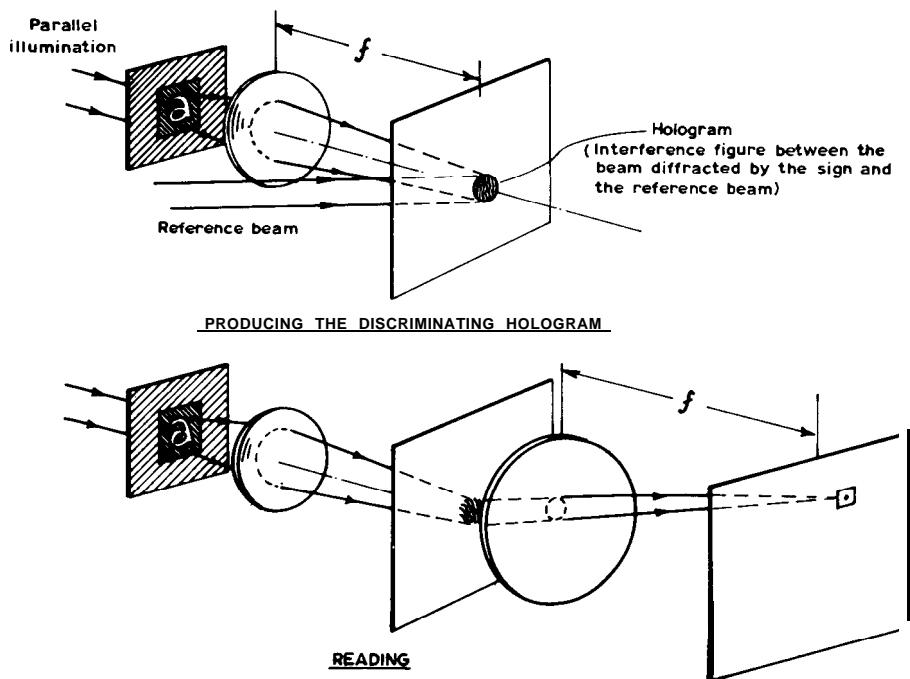


Fig. 29.

The Principle of Character Recognition by Holography.

sequentially in a microsecond would of course require an impossible waveband, but powerful parallel reading means can be provided. One can confidently expect enormous extensions of these "modest beginnings" once the project of data banks will be tackled seriously.

Another application of holography, which is probably only in an early stage, is pattern and character recognition. I can only briefly refer to the basic work which A. Vander Lugt (18) has done in the field of pattern recognition. It will be sufficient to explain the basic principle of character recognition with the aid of Figure 29.

Let us generalize a little the basic principle of holography. In all previous examples a complicated object beam was brought to interference with a simple or spherical reference beam, and the object beam was reconstructed by illuminating the hologram with the reference beam. But a little mathematics shows that this can be extended to any reference beam *which correlates sharply with itself*. The autocorrelation function is an invariant of a beam; it can be computed in any cross section. One can see at once that a spherical wave correlates sharply with itself, because it issues from a "point". But there are other beams which correlate sharply with themselves, for instance those which issue from a fingerprint, or from a Chinese ideogram, in an extreme case also those which issue from a piece of frosted glass. Hence it is quite possible for instance to translate, by means of a hologram, a Chinese ideogram into its corresponding English sentence and *vice versa*. J. N. Butters and M. Wall in Loughborough University have recently created holograms which

from a portrait produce the signature of the owner, and vice versa.<sup>4</sup> In other words, a hologram can be a fairly universal translator. It can for instance translate a sign which we can read to another which a machine can read.

Figure 29 shows a fairly modest realisation of this principle. A hologram is made of a letter "a" by means of a plane reference beam. When this hologram is illuminated with the letter "a" the reference beam is reconstructed, and can activate for instance a small photocell in a certain position. This, I believe, gives an idea of the basic principle. There are of course many ways of printing letters, but it would take me too long to explain how to deal with this and other difficulties.

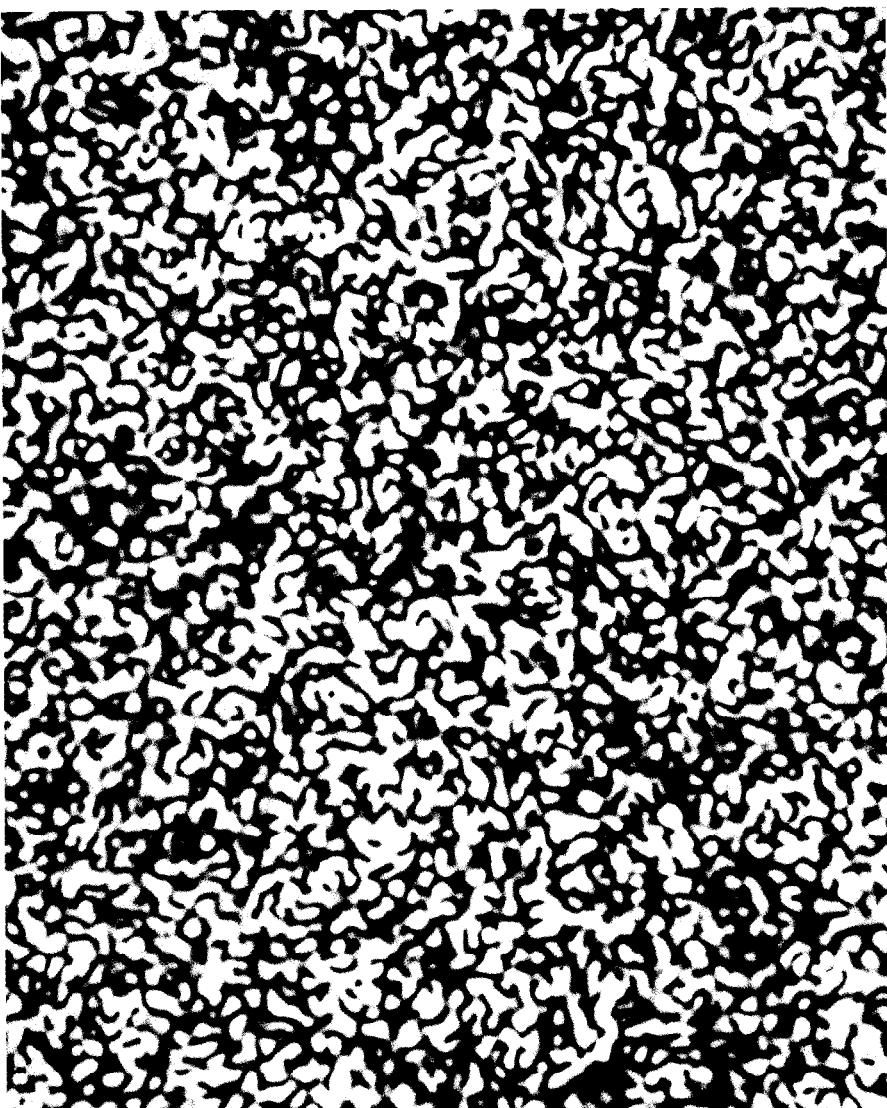


Fig. 30.

Laser Speckle. The appearance of e.g. a white sheet of paper, uniformly illuminated by laser light.

With character recognition devices we have already taken half a step into the future, because these are likely to become important only in the next generation of computers or robots, to whom we must transfer a little more of human intelligence. I now want to mention briefly some other problems which are half or more than half in the future.

One, which is already very actual, is the overcoming of laser speckle. Everybody who sees laser light for the first time is surprised by the rough appearance of objects which we consider as smooth. A white sheet of paper appears as if it were crawling with ants. The crawling is put into it by the restless eye, but the roughness is real. It is called "laser speckle" and Figure 30 shows a characteristic example of it. This is the appearance of a white sheet of paper in laser light, when viewed with a low-power optical system. It is not really noise; it is information which we do not want, information on the microscopic unevenness of the paper in which we are not interested. What can we do against it?

In the case of rough objects the answer is, regrettably, that all we can do is to average over larger areas, thus smoothing the deviations. This means that we must throw a great part of the information away, the wanted with the unwanted. This is regrettable but we can do nothing else, and in many cases we have enough information to throw away, as can be seen by the fully satisfactory appearance of some of the reconstructions from diffuse holograms which I have shown. However, there are important areas in which we can do much more, and where an improvement is badly needed. This is the area of microholograms, for storing and for display. They are made as diffused holograms, in order to ensure freedom from dust and scratches, but by making them diffused, we introduce speckle, and to avoid this such holograms are made nowadays much larger than would be ideally necessary. I have shown recently (19), that advantages of diffuse holograms can be almost completely retained, while the speckle can be completely eliminated by using, instead of a frosted glass, a special illuminating system. This, I hope will produce a further improvement in the information density of holographic stores.

Now let us take a more radical step into the future. I want to mention briefly two of my favourite holographic brainchildren. The first of this is Panoramic Holography, or one could also call it Holographic Art.

All the three-dimensional holograms made so far extend to a depth of a few meters only. Would it not be possible to extend them to infinity? Could one not put a hologram on the wall, which is like a window through which one looks at a landscape, real or imaginary? I think it can be done, only it will not be a photograph but a work of art. Figure 31 illustrates the process. The artist makes a model, distorted in such a way that it appears perspective, and extending to any distance when viewed through a large lens, as large as the hologram. The artist can use a smaller lens, just large enough to cover both his eyes when making the model. A reflecting hologram is made of it, and illuminated with a strong, small light source. The viewer will see what the plate has seen through the lens; that is to say a scene extending to any distance, in natural colours. This scheme is under development, but considerable work

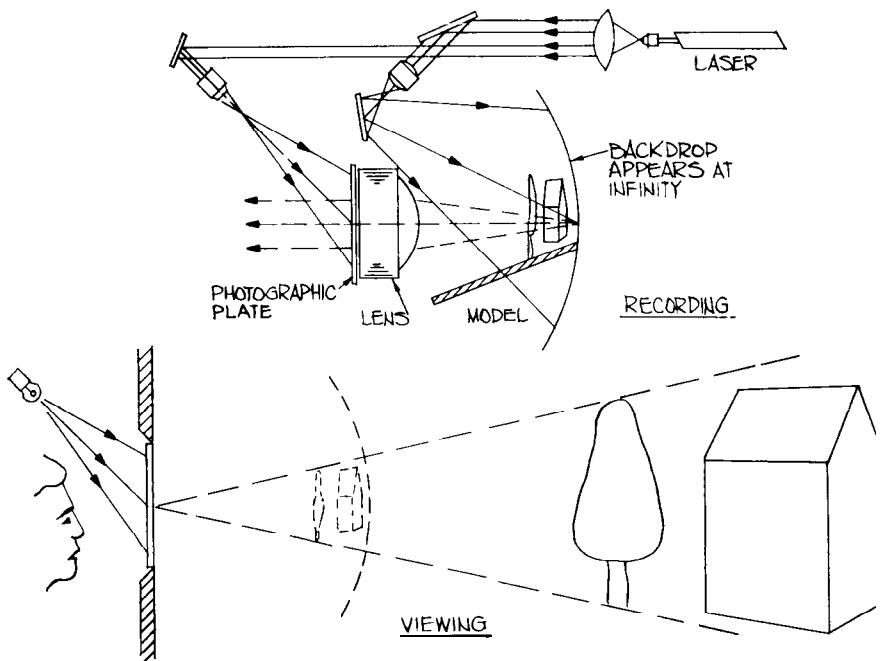


Fig. 31.  
Panoramic Holography.

will be needed to make it satisfactory, because we must first greatly improve the reflectance of three-colour holograms.

An even more ambitious scheme, probably even farther in the future, is three-dimensional cinematography, without viewing aids such as Polaroids. The problem is sketched out in Figure 32. The audience (in one plane or two)

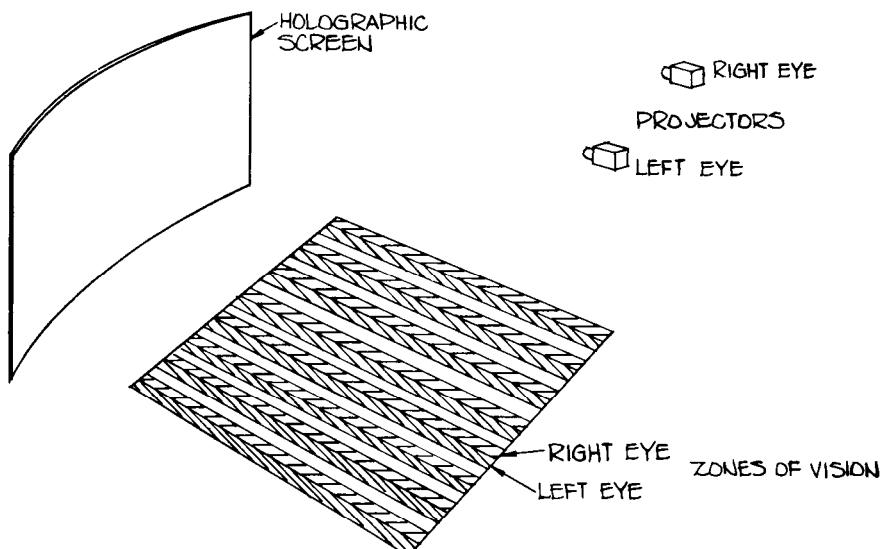


Fig. 32.  
3-D Cinematography with Holographic Screen.

is covered by zones of vision, with the width of the normal eye spacing, one for the right eye, one for the left, with a blank space between two pairs. The two eyes must see two different pictures; a stereoscopic pair. The viewer can move his head somewhat to the right or left. Even when he moves one eye into the blank zone, the picture will appear dimmer but not flat, because one eye gives the impression of "stereoscopy by default".

I have spent some years of work on this problem, just before holography, until I had to realise that it is strictly unsolvable with the orthodox means of optics, lenticules, mirrors, prisms. One can make satisfactorily small screens for small theatres, but with large screens and large theatres one falls into a dilemma. If the lenticules, or the like, are large, they will be seen from the front seats; if they are small, they will not have enough definition for the back seats.

Some years ago I realised to my surprise, that holography can solve this problem too. Use a projector as the reference source, and for instance the system of left viewing zones as the object. The screen, covered with a Lippmann emulsion, will then make itself automatically into a very complicated optical system such that when a picture is projected from the projector, it will be seen only from the left viewing zones. One then repeats the process with the right projector, and the right viewing zones. Volume, (Lippmann-Denisuk) holograms display the phenomenon of directional selectivity. If one displaces the illuminator from the original position by a certain angle, there will be no reflection. We put the two projectors at this angle (or a little more) from one another, and the effect is that the right picture will not be seen by the left eye and vice versa.

There remains of course one difficulty, and this is that one cannot practise holography on the scale of a theatre, and with a plate as large as a screen. But this too can be solved, by making up the screen from small pieces, not with the theatre but with a *model* of the theatre, seen through a lens, quite similar to the one used in panoramic holography.

I hope I have conveyed the feasibility of the scheme, but I feel sure that I have conveyed also its difficulties. I am not sure whether they will be overcome in this century, or in the next.

Ambitious schemes, for which I have a congenital inclination, take a long time for their realisation. As I said at the beginning, I shall be lucky if I shall be able to see in my lifetime the realisation of holographic electron microscopy, on which I have started 24 years ago. But I have good hope, because I have been greatly encouraged by a remarkable achievement of G. W. Stroke (20), which is illustrated in Figure 33. Professor Stroke has recently succeeded in deblurring micrographs taken by Professor Albert Crewe, Chicago, with his scanning transmission electron microscope, by a holographic filtering process, improving the resolution from 5 Angstrom to an estimated 2.5 Angstrom. This is not exactly holographic electron microscopy, because the original was not taken with coherent electrons, but the techniques used by both sides, by A. Crewe and by G. W. Stroke are so powerful, that I trust them to succeed also in the next, much greater and more important step.

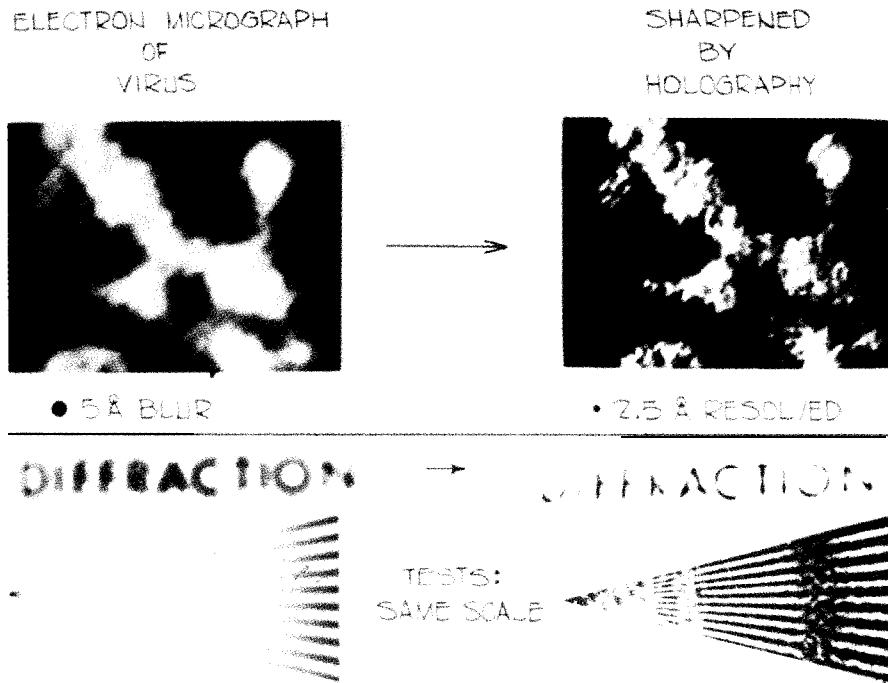
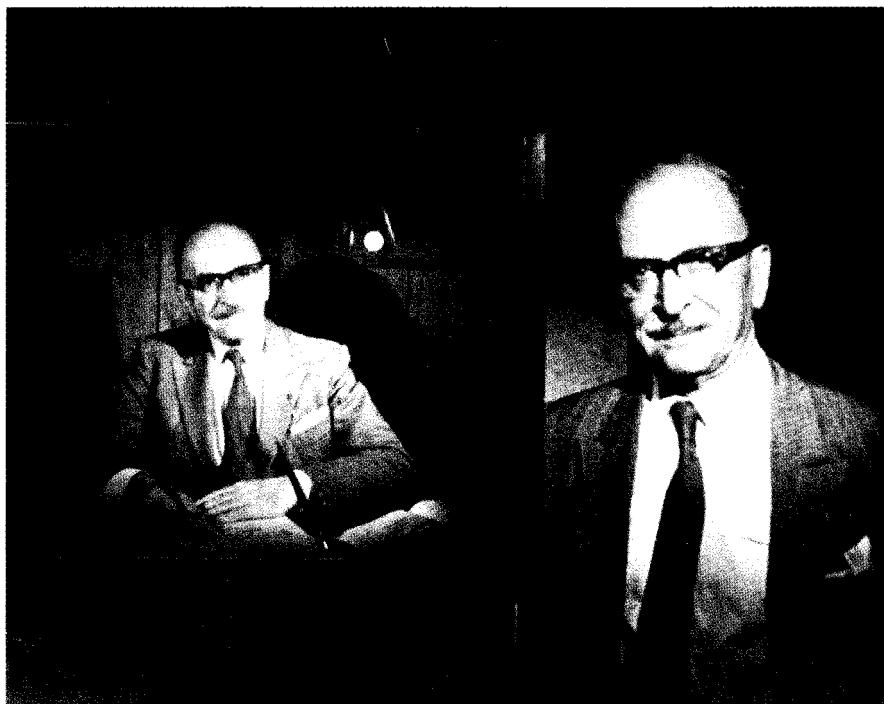


Fig. 33.

Scanning Transmission Electron Micrograph. Professor Albert Crewe, University of Chicago, holographically deblurred by Professor G. W. Stroke, 1971. The bottom photographs prove that the effect could not be obtained by hard printing, because some spatial frequencies which appear in the original with reversed phase had to be phase-corrected.

Summing up, I am one of the few lucky physicists who could see an idea of theirs grow into a sizeable chapter of physics. I am deeply aware that this has been achieved by an army of young, talented and enthusiastic researchers, of whom I could mention only a few by name. I want to express my heartfelt thanks to them, for having helped me by their work to this greatest of scientific honours.



#### REFERENCES

It is impossible to do justice to the hundreds of authors who have significantly contributed to the development of holography. The number of articles exceeds 2 000, and there are more than a dozen books in several languages.

An extensive bibliography may be found for instance in: T. Kallard, Editor, *Holography* (Optosonic Press, New York, 1969 and 1970)

#### BOOKS

- Barrekette, E. S., Kock, W. E., Ose, T., Tsujiuchi, J. and Stroke, G. W., Eds. *Applications of Holography* (Plenum Press, New York, 1971 )
- Caulfield, H. J. and Sun Lu, *The Applications of Holography* (Wiley Interscience, New York, 1970).
- Collier, R. J., Burckhardt, C. B. and Lin, L. H., *Optical Holography* (Academic Press, New York, 1971).
- DeVelis, J. B. and Reynolds, G. O., *Theory and Applications of Holography* (Addison Wesley, Reading, Massachusetts, 1967).
- Francon, M., *Holographie* (Masson et Cie, Paris, 1969).
- Kiemle, H. und Röss, D., *Einführung in die Technik der Holographie* (Akademische Verlagsgesellschaft, Frankfurt am Main, 1969). *Ibid.* *Introduction to Holographic Techniques* (Plenum Press, 1972, in print).
- Kock, W. E., *Lasers and Holography (An Introduction to Coherent Optics)* (Doubleday & Co., Garden City, NY, 1969).
- Ostrovsky, Yu. I., *Holography* (in Russian) Nauka, Leningrad, 1970).
- Robertson E. R. and Harvey, J. M., Eds., *The Engineering Uses of Holography* (Cambridge University Press, 1970).

- Stroke, G. W., An Introduction to Coherent Optics and Holography (Academic Press, New York: 1st Edition, 1966, Second Edition, 1969).
- Viénot, J. Ch., Smigelski, P. et Royer, H., Holographie Optique (Developments, Applications) (Dunod, Paris, 1971).
1. Gabor, D., "A New Microscopic Principles", *Nature* 161, No. 4098, 777-778 (1948).
  2. Gabor, D., "Microscopy by Reconstructed Wavefronts", *Proc. Roy. Soc. (London)* **A 197**, 454-487 (1949).
  3. Gabor, D., "Microscopy by Reconstructed Wavefronts: II", *Proc. Phys. Soc. (London)* 64 (Pt. 6) No. 378 B, 449-469 (1951).
  4. Rogers, G. L., "Experiments in Diffraction Microscopy", *Proc. Roy. Soc. (Edinburgh)* 63 **A**, 193 (1952).
  5. Baez, A., "Resolving Power in Diffraction Microscopy", *Nature*, **169**, 963-964 (1952).
  6. El-Sum, H. M. A. and Kirkpatrick, P., "Microscopy by Reconstructed Wavefronts" *Phys. Rev.*, 85, 763 (1952).
  7. Gabor, D. and Goss, W. P., "Interference Microscope with Total Wavefront Reconstruction", *J. Opt. Soc. Am.* 56, 849-858 (1966).
  8. Cutrona, L. J., Leith, E. N., Porcello, L. J. and Vivian, W. E., "On the Application of Coherent Optical Processing Techniques to Synthetic Aperture Radar", *Proc. IEEE*, 54, 1026-1032 (1966).
  9. Leith, E. N. and Upatnieks, J., "Wavefront Reconstruction with Continuous Tone Transparencies", *J. Opt. Soc. Am.* 53, 522 (1963) (Abstract).
  10. Leith, E. N. and Upatnieks, J., "Wavefront Reconstruction with Continuous-Tone Objects", *J. Opt. Soc. Am.* 53, 1377-1381 (1963).
  11. Leith, E. N., and Upatnieks, J., "Wavefront Reconstruction with Diffused Illumination and Three-Dimensional Objects" *J. Opt. Soc. Amer.*, 54, 1295 (1964).
  12. An early reference is:  
Stroke, G. W., "Theoretical and Experimental Foundations of High-Resolution Optical Holography" (Presented in Rome on 14 September 1964) in Vol. II, pp. 53-63 of *Pubblicazioni IV Centenario della Nascita di Galileo Galilei* (G. Barbera, Firenze, 1966).
  13. Knox, C. and Brooks, R. E., "Holographic Motion Picture Microscopy" *Proc. Roy. Soc. (London)*, **B 174**, 115-121 (1969).
  14. Heflinger, L. O., Wuerker, R. F. and Brooks, R. E., *J. Appl. Physics*, 37, No. 2, 642-649 (February 1966).
  15. Burch, J. M., "The Application of Lasers in Production Engineering", *Production Engineer*, 44, 431-442 (1965).
  16. Powell, R. L. and Stetson, K. A., "Interferometric Vibration Analysis by Wavefront Reconstruction", *J. Opt. Soc. Am.* 55, 1593-1598 (1965).
  17. Stroke, G. W. and Labeyrie, A. E., "Two-Beam Interferometry by Successive Recording of Intensities in a Single Hologram", *Appl. Physics Letters*, 8, No. 2, 42-44 (15 January 1966).
  18. Stetson, K. A., Thesis (under direction of E. Ingelstam), Royal Institute of Technology, Stockholm (1969).
  19. Denisyuk, Yu. N., "Photographic Reconstruction of the Optical Properties of an Object in its Own Scattered Radiation", *Dokl. Akad. Nauk. SSR*, 144, 1275-1278 (1962).
  20. Stroke, G. W. and Labeyrie, A. E., "White-Light Reconstruction of Holographic Images using the Lippmann-Bragg Diffraction Effect", *Physics Letters*, 20, 368-370 (1966).
  21. Pennington, K. S. and Harper, J. S., "Techniques for Producing Low-Noise, Improved-Efficiency Holograms", *Appl. Optics*, 9, 1643-1650 (1970).
  22. Van Heerden, P. J., "A New Method of Storing and Retrieving Information", *Appl. Optics*, 2, 387-392 (1963).

17. Anderson, L. K., "Holographic Optical Memory for Bulk Data Storage", Bell Lab Rep. 46, 318 (1968).
18. Vander Lugt, A., "Signal Detection by Complex Spatial Filtering", IEEE Trans. on Information Theory, IT-10, 139-145 (1964).
19. Gabor, D., "Laser Speckle and its Elimination", IBM J. of Res. and Dev. 14, 509-514 (Sept. 1970).
20. Stroke, G. W., "Image Deblurring and Aperture Synthesis Using 'A Posteriori' Processing by Fourier-Transform Holography", Optica Acta, 16, 401-422 (1969).  
Stroke, G. W. and Halioua, M., "Attainment of Diffraction-Limited Imaging in High-Resolution Electron Microscopy by 'A Posteriori' Holographic Image Sharpening", Optik (1972) (in print).

Physics 1972

**JOHN BARDEEN, LEON N COOPER and  
J ROBERT SCHRIEFFER**

*for their jointly developed theory of superconductivity, usually called the  
BCS-theory*

## THE NOBEL PRIZE FOR PHYSICS

Speech by professor **STIG LUNDQVIST**, Chalmers University of Technology  
Translation from the Swedish text

Your Royal Highnesses, Ladies and Gentlemen,

The 1972 Nobel Prize for physics has been awarded to Drs John Bardeen, Leon N. Cooper and J. Robert Schrieffer for their theory of superconductivity, usually referred to as the BCS-theory.

Superconductivity is a peculiar phenomenon occurring in many metallic materials. Metals in their normal state have a certain electrical resistance, the magnitude of which varies with temperature. When a metal is cooled its resistance is reduced. In many metallic materials it happens that the electrical resistance not only decreases but also suddenly disappears when a certain critical temperature is passed which is a characteristic property of the material.

This phenomenon was discovered as early as 1911 by the Dutch physicist Kamerlingh Onnes, who was awarded the Nobel Prize for physics in 1913 for his discoveries.

The term superconductivity refers to the complete disappearance of the electrical resistance, which was later verified with an enormous accuracy. A lead ring carrying a current of several hundred ampères was kept cooled for a period of 2 1/2 years with no measurable change in the current.

An important discovery was made in the thirties, when it was shown that an external magnetic field cannot penetrate a superconductor. If you place a permanent magnet in a bowl of superconducting material, the magnet will hover in the air above the bowl, literally floating on a cushion of its own magnetic field lines. This effect may be used as an example for the construction of friction-free bearings.

Many of the properties of a metal change when it becomes superconducting and new effects appear which have no equivalent in the former's normal state. Numerous experiments have clearly shown that a fundamentally new state of the metal is involved.

The transition to the superconductive state occurs at extremely low temperatures, characteristically only a few degrees above absolute zero. For this reason practical applications of the phenomenon have been rare in the past and superconductivity has been widely considered as a scientifically interesting but exclusive curiosity confined to the low temperature physics laboratories. This state of affairs is rapidly changing and the use of superconducting devices is rapidly increasing. Superconducting magnets are often used for example in particle accelerators. Superconductivity research has in recent years resulted in substantial advances in measuring techniques and an extensive use in the computer field is also highly probable. Advanced plans for the use of superconductivity in heavy engineering are also in existence. By way of an example, it may be mentioned that the transport of electric energy to the major cities of the world with the use of superconductive lines is being planned. Looking

further ahead one can see, for example, the possibility of building ultrarapid trains that run on superconducting tracks.

Superconductivity has been studied experimentally for more than sixty years. However, the central problem, the question of the physical mechanism responsible for the phenomenon remained a mystery until the late fifties. Many famous physicists tackled the problem with little success. The difficulties were related to the very special nature of the mechanism sought. In a normal metal the electrons move around individually at random, somewhat similar to the atoms in a gas, and the theory is, in principle, fairly simple. In superconductive metals the experiments suggested the existence of a collective state of the conduction electrons-a state in which the electrons are strongly coupled and their motion correlated so that there is a gigantic coherent state of macroscopic dimension containing an enormous number of electrons. The physical mechanism responsible for such a coupling remained unknown for a long time. An important step towards the solution was taken in 1950 when it was discovered simultaneously on theoretical and experimental grounds that superconductivity must be connected with the coupling of the electrons to the vibrations of the atoms in the crystal lattice. The conduction electrons are coupled to each other via these vibrations. Starting from this fundamental coupling of the electrons Bardeen, Cooper and Schrieffer developed their theory of superconductivity, published in 1957, which gave a complete theoretical explanation of the phenomenon of superconductivity.

According to their theory the coupling of the electrons to the lattice oscillations leads to the formation of bound pairs of electrons. These pairs play a fundamental role in the theory. The complete picture of the mechanism of superconductivity appeared when Bardeen, Cooper and Schrieffer showed that the motion of the different pairs is very strongly correlated and that this leads to the formation of a gigantic coherent state in which a large number of electrons participate. It is this ordered motion of the electrons in the superconductive state in contrast to the random individual motion in a normal crystal that gives superconductivity its special properties.

The theory developed by Bardeen, Cooper and Schrieffer together with extensions and refinements of the theory, which followed in the years after 1957, succeeded in explaining in considerable detail the properties of superconductors. The theory also predicted new effects and it stimulated intense activity in theoretical and experimental research which opened up new areas. These latter developments have led to new important discoveries which are being used in a number of interesting ways especially in the sphere of measuring techniques.

Developments in the field of superconductivity during the last fifteen years have been greatly inspired by the fundamental theory of superconductivity and have strikingly verified the validity and great range of the concepts and ideas developed by Bardeens, Cooper and Schrieffer.

Drs. Bardeen, Cooper and Schrieffer,

You have in your fundamental work given a complete theoretical explanation of the phenomenon of superconductivity. Your theory has also predicted

new effects and stimulated an intensive activity in theoretical and experimental research. The further developments in the field of superconductivity have in a striking way confirmed the great range and validity of the concepts and ideas in your fundamental paper from 1957.

On behalf of the Royal Academy of Sciences I wish to convey to you the warmest congratulations and I now ask you to receive your prizes from the Hands of His Royal Highness the Crown Prince.



John Bardeen

## JOHN BARDEEN

John Bardeen was born in Madison, Wisconsin, May 23, 1908.

He attended the University High School in Madison for several years, and graduated from Madison Central High School in 1923. This was followed by a course in electrical engineering at the University of Wisconsin, where he took extra work in mathematics and physics. After being out for a term while working in the engineering department of the Western Electric Company at Chicago, he graduated with a B.S. in electrical engineering in 1928. He continued on at Wisconsin as a graduate research assistant in electrical engineering for two years, working on mathematical problems in applied geophysics and on radiation from antennas. It was during this period that he was first introduced to quantum theory by Professor J. H. Van Vleck.

Professor Leo J. Peters, under whom his research in geophysics was done, took a position at the Gulf Research Laboratories in Pittsburgh, Pennsylvania. Dr. Bardeen followed him there and worked during the next three years (1930-33) on the development of methods for the interpretation of magnetic and gravitational surveys. This was a stimulating period in which geophysical methods were first being applied to prospecting for oil.

Because he felt his interests were in theoretical science, Dr. Bardeen resigned his position at Gulf in 1933 to take graduate work in mathematical physics at Princeton University. It was here, under the leadership of Professor E. P. Wigner, that he first became interested in solid state physics. Before completing his thesis (on the theory of the work function of metals) he was offered a position as Junior Fellow of the Society of Fellows at Harvard University. He spent the next three years there working with Professors Van Vleck and Bridgman on problems in cohesion and electrical conduction in metals and also did some work on the level density of nuclei. The Ph.D. degree at Princeton was awarded in 1936.

From 1938-41 Dr. Bardeen was an assistant professor of physics at the University of Minnesota and from 1941-45 a civilian physicist at the Naval Ordnance Laboratory in Washington, D. C. His war years were spent working on the influence fields of ships for application to underwater ordnance and minesweeping. After the war, he joined the solid-state research group at the Bell Telephone Laboratories, and remained there until 1951, when he was appointed Professor of Electrical Engineering and of Physics at the University of Illinois. Since 1959 he has also been a member of the Center for Advanced Study of the University.

Dr. Bardeen's main fields of research since 1945 have been electrical conduction in semiconductors and metals, surface properties of semiconductors, theory of superconductivity, and diffusion of atoms in solids. The Nobel Prize in Physics was awarded in 1956 to John Bardeen, Walter H. Brattain, and William Shockley for "investigations on semiconductors and the discovery of the transistor effect," carried on at the Bell Telephone Laboratories. In 1957, Bardeen and two colleagues, L. N. Cooper and J. R. Schrieffer, proposed the first successful explanation of superconductivity, which has been a puzzle since its discovery in 1908. Much of his research effort since that time has been devoted to further extensions and applications of the theory.

Dr. Bardeen died in 1991.

Born Madison, Wisconsin, May 23, 1908. Son of Dr. Charles R. and Althea Harmer Bardeen, both deceased. Dr. Bardeen was Professor of Anatomy and Dean of the Medical School at the University of Wisconsin. Stepmother, Mrs. Kenelm McCauley, Milwaukee, Wisconsin. Married Jane Maxwell, 1938. Children: James M., William A., Elizabeth A. Attended public schools and university in Madison, Washington School, Madison, 1914-17, University High School, Madison, 1917-21, Madison Central High, 1921-23.

B.S. and M.S. in E. E., University of Wisconsin, 1928 and 1929.

Geophysicist, Gulf Research and Development Corp., Pittsburgh, PA, 1930-33.

Attended Graduate College, Princeton University, 1933-35, received Ph.D. in Math. Phys., 1936.

Junior Fellow, Society of Fellows, Harvard University, 1935-38.

Assistant Professor of Physics, University of Minnesota, 1938-41.

Physicist, Naval Ordnance Laboratory, 1941-45.

Research Physicist, Bell Telephone Laboratories, 1945-51.

Professor of Electrical Engineering and of Physics, University of Illinois, and a member of the Center for Advanced Study of the University, 1951-1975. Emeritus, 1975-1991.

Served on U.S. President's Science Advisory Committee, 1959-62.

Member of the Board of Directors, Xerox Corporation, Rochester, New York, 1961-1974. Consultant, 1952-82.

Lorentz Professor, University of Leiden, Netherlands, 1975.

Visiting Professor, Karlsruhe, 1978; Grenoble, 1981; University of California, Santa Barbara, 1981, 1984; Nihon University, Tokyo, 1982.

White House Science Council, 1982-83.

Honorary degrees Union College (1955) Wisconsin (1960), Rose Polytechnic Inst. (1966), Western Reserve (1966), Univ. of Glasgow (1967), Princeton (1968), Rensselaer Polytechnic Inst. (1969), Notre Dame (1970), Harvard (1973) Minnesota (1973) Illinois (1974) Michigan (1974) Pennsylvania

(1976), Delhi, India (1977), Indian Inst. of Tech., Madras, India (1977), Cambridge (U.K.) (1977) Georgetown (1980), St. Andrews (1980) Clarkson (1981).

**Awards:** Stuart Ballantine Medal, Franklin Inst. (1952); Buckley Prize, Am. Physical Soc. (1954); John Scott Medal, Philadelphia (1955); Nobel Prize (Physics) shared with W. H. Brattain and W. Shockley (1956); Vincent Bendix Award, Amer. Soc. Eng. Educ. (1964); National Medal of Science (1965); Michelson-Morley Award, Case-Western Reserve (1968); Medal of Honor, Inst. of Electrical and Electronics Eng. (1971); Nobel Prize (Physics) shared with L. N. Cooper and J. R. Schrieffer (1972); James Madison Medal, Princeton (1973); National Inventors Hall of Fame (1974); Franklin Medal, Franklin Inst. (1975); Presidential Medal of Freedom (1977); Washington Award, Western Soc. Eng. (1983); Founders Awards, Nat. Acad. Eng. (1984); Lomonosov Prize, USSR Acad. of Sci. (1988); The Harold Pender Award, University of Pennsylvania ( 1988).

**Academic and Professional Societies:** American Physical Society (President, 1968-69); IEEE (hon. mem.); National Academy of Sciences; National Academy of Engineering; American Academy of Sciences; American Philosophical Society; Foreign Member, Royal Society of London; Foreign Member, Indian Nat. Sci. Acad.; Honorary Fellow, The Institute of Physics (London); Foreign Member, Inst. of Electronics and Telecommunications (India); Hon. Member, Japan Academy; Hon. Doctor, Venezuelan Academy; Foreign Member, USSR Academy of Sciences; Pakistan Academy of Sciences, Corr. Mem., Hungarian Acad. Sci. and Austrian Acad. Sci.

# ELECTRON-PHONON INTERACTIONS AND SUPERCONDUCTIVITY

Nobel Lecture, December 11, 1972

By JOHN BARDEEN

Departments of Physics and of Electrical Engineering  
University of Illinois  
Urbana, Illinois

## INTRODUCTION

Our present understanding of superconductivity has arisen from a close interplay of theory and experiment. It would have been very difficult to have arrived at the theory by purely deductive reasoning from the basic equations of quantum mechanics. Even if someone had done so, no one would have believed that such remarkable properties would really occur in nature. But, as you well know, that is not the way it happened, a great deal had been learned about the experimental properties of superconductors and phenomenological equations had been given to describe many aspects before the microscopic theory was developed. Some of these have been discussed by Schrieffer and by Cooper in their talks.

My first introduction to superconductivity came in the 1930's and I greatly profited from reading David Shoenberg's little book on superconductivity, [1] which gave an excellent summary of the experimental findings and of the phenomenological theories that had been developed. At that time it was known that superconductivity results from a phase change of the electronic structure and the Meissner effect showed that thermodynamics could be applied successfully to the superconductive equilibrium state. The two fluid Gorter-Casimir model was used to describe the thermal properties and the London brothers had given their famous phenomenological theory of the electrodynamic properties. Most impressive were Fritz London's speculations, given in 1935 at a meeting of the Royal Society in London, [2] that superconductivity is a quantum phenomenon on a macroscopic scale. He also gave what may be the first indication of an energy gap when he stated that "the electrons be coupled by some form of interaction in such a way that the lowest state may be separated by a finite interval from the excited ones." He strongly urged that, based on the Meissner effect, the diamagnetic aspects of superconductivity are the really basic property.

My first abortive attempt to construct a theory, [3] in 1940, was strongly influenced by London's ideas and the key idea was small energy gaps at the Fermi surface arising from small lattice displacements. However, this work was interrupted by several years of wartime research, and then after the war I joined the group at the Bell Telephone Laboratories where my work turned to semiconductors. It was not until 1950, as a result of the discovery of the

isotope effect, that I again began to become interested in superconductivity, and shortly after moved to the University of Illinois.

The year 1950 was notable in several respects for superconductivity theory. The experimental discovery of the isotope effect [4, 5] and the independent prediction of H. Fröhlich [6] that superconductivity arises from interaction between the electrons and phonons (the quanta of the lattice vibrations) gave the first clear indication of the directions along which a microscopic theory might be sought. Also in the same year appeared the phenomenological Ginzburg-Landau equations which give an excellent description of superconductivity near  $T_c$ , in terms of a complex order parameter, as mentioned by Schrieffer in his talk. Finally, it was in 1950 that Fritz London's book [7] on superconductivity appeared. This book included very perceptive comments about the nature of the microscopic theory that have turned out to be remarkably accurate. He suggested that superconductivity requires "a kind of solidification or condensation of the average momentum distribution." He also predicted the phenomenon of flux quantization, which was not observed for another dozen years.

The field of superconductivity is a vast one with many ramifications. Even in a series of three talks, it is possible to touch on only a few highlights. In this talk, I thought that it might be interesting to trace the development of the role of electron-phonon interactions in superconductivity from its beginnings in 1950 up to the present day, both before and after the development of the microscopic theory in 1957. By concentrating on this one area, I hope to give some impression of the great progress that has been made in depth of understanding of the phenomena of superconductivity. Through developments by many people, [8] electron-phonon interactions have grown from a qualitative concept to such an extent that measurements on superconductors are now used to derive detailed quantitative information about the interaction and its energy dependence. Further, for many of the simpler metals and alloys, it is possible to derive the interaction from first principles and calculate the transition temperature and other superconducting properties.

The theoretical methods used make use of the methods of quantum field theory as adopted to the many-body problem, including Green's functions, Feynman diagrams, Dyson equations and renormalization concepts. Following Matsubara, temperature plays the role of an imaginary time. Even if you are not familiar with diagrammatic methods, I hope that you will be able to follow the physical arguments involved.

In 1950, diagrammatic methods were just being introduced into quantum field theory to account for the interaction of electrons with the field of photons. It was several years before they were developed with full power for application to the quantum statistical mechanics of many interacting particles. Following Matsubara, those prominent in the development of the theoretical methods include Kubo, Martin and Schwinger, and particularly the Soviet physicists, Migdal, Galitski, Abrikosov, Dzyaloshinski, and Gor'kov. The methods were first introduced to superconductivity theory by Gor'kov [9] and a little later in a somewhat different form by Kadanoff and Martin. [10] Problems of

superconductivity have provided many applications for the powerful Green's function methods of many-body theory and these applications have helped to further develop the theory.

Diagrammatic methods were first applied to discuss electron-phonon interactions in normal metals by Migdal [11] and his method was extended to superconductors by Eliashberg. [12] A similar approach was given by Nambu. [13] The theories are accurate to terms of order  $(m/M)^{1/2}$ , where  $m$  is the mass of the electron and  $M$  the mass of the ion, and so give quite accurate quantitative accounts of the properties of both normal metals and superconductors.

We will first give a brief discussion of the electron-phonon interactions as applied to superconductivity theory from 1950 to 1957, when the pairing theory was introduced, then discuss the Migdal theory as applied to normal metals, and finally discuss Eliashberg's extension to superconductors and subsequent developments. We will close by saying a few words about applications of the pairing theory to systems other than those involving electron-phonon interactions in metals.

#### DEVELOPMENTS FROM 1950 - 1957

The isotope effect was discovered in the spring of 1950 by Reynolds, Serin, et al, [4] at Rutgers University and by E. Maxwell [5] at the U. S. National Bureau of Standards. Both groups measured the transition temperatures of separated mercury isotopes and found a positive result that could be interpreted as  $T_c M^{1/2} \simeq \text{constant}$ , where  $M$  is the isotopic mass. If the mass of the ions is important, their motion and thus the lattice vibrations must be involved.

Independently, Fröhlich, [6] who was then spending the spring term at Purdue University, attempted to develop a theory of superconductivity based on the self-energy of the electrons in the field of phonons. He heard about the isotope effect in mid-May, shortly before he submitted his paper for publication and was delighted to find very strong experimental confirmation of his ideas. He used a Hamiltonian, now called the Fröhlich Hamiltonian, in which interactions between electrons and phonons are included but Coulomb interactions are omitted except as they can be included in the energies of the individual electrons and phonons. Fröhlich used a perturbation theory approach and found an instability of the Fermi surface if the electron-phonon interaction were sufficiently strong.

When I heard about the isotope effect in early May in a telephone call from Serin, I attempted to revive my earlier theory of energy gaps at the Fermi surface, with the gaps now arising from dynamic interactions with the phonons rather than from small static lattice displacements. [14] I used a variational method rather than a perturbation approach but the theory was also based on the electron self-energy in the field of phonons. While we were very hopeful at the time, it soon was found that both theories had grave difficulties, not easy to overcome. [15] It became evident that nearly all of the self-energy is included in the normal state and is little changed in the transition. A theory

involving a true many-body interaction between the electrons seemed to be required to account for superconductivity. Schafroth [16] showed that starting with the Fröhlich Hamiltonian, one cannot derive the Meissner effect in any order of perturbation theory. Migdal's theory, [I II] supposedly correct to terms of order  $(m/M)^{1/2}$ , gave no gap or instability at the Fermi surface and no indication of superconductivity.

Of course Coulomb interactions really are present. The effective direct Coulomb interaction between electrons is shielded by the other electrons and the electrons also shield the ions involved in the vibrational motion. Pines and I derived an effective electron-electron interaction starting from a Hamiltonian in which phonon and Coulomb terms are included from the start. [17] As is the case for the Fröhlich Hamiltonian, the matrix element for scattering of a pair of electrons near the Fermi surface from exchange of virtual phonons is negative (attractive) if the energy difference between the electron states involved is less than the phonon energy. As discussed by Schrieffer, the attractive nature of the interaction was a key factor in the development of the microscopic theory. In addition to the phonon induced interaction, there is the repulsive screened Coulomb interaction, and the criterion for superconductivity is that the attractive phonon interaction dominate the Coulomb interaction for states near the Fermi surface. [18]

During the early 1950's there was increasing evidence for an energy gap at the Fermi surface. [19] Also very important was Pippard's proposed non-local modification [20] of the London electrodynamics which introduced a new length the coherence distance,  $\xi_0$ , into the theory. In 1955 I wrote a review article [17] on the theory of superconductivity for the *Handbuch der Physik*, which was published in 1956. The central theme of the article was the energy gap, and it was shown that Pippard's version of the electrodynamics would likely follow from an energy gap model. Also included was a review of electron-phonon interactions. It was pointed out that the evidence suggested that all phonons are involved in the transition, not just the long wave length phonons, and that their frequencies are changed very little in the normal-superconducting transition. Thus one should be able to use the effective interaction between electrons as a basis for a true many-body theory of the superconducting state. Schrieffer and Cooper described in their talks how we were eventually able to accomplish this goal.

## 3

## GREEN'S FUNCTION METHOD FOR NORMAL METALS

By use of Green's function methods, Migdal [11] derived a solution of Fröhlich's Hamiltonian,  $H = H_{el} + H_{ph} + H_{el-ph}$ , for normal metals valid for arbitrarily strong coupling and which involves errors only of order  $(m/M)^{1/2}$ . The Green's functions are defined by thermal average of time ordered operators for the electrons and phonons, respectively

$$G = -i \langle T\psi(1)\psi^+(2) \rangle \quad (1a)$$

$$D = -i \langle T\theta(1)\theta^+(2) \rangle \quad (1b)$$

Here  $y(r,t)$  is the wave field operator for electron quasi-particles and  $O(r,t)$  for the phonons, the symbols 1 and 2 represent the space-time points  $(r_1, t_1)$  and  $(r_2, t_2)$  and the brackets represent thermal averages over an ensemble.

Fourier transforms of the Green's functions for  $H_{\text{el}} = H_{\text{el}} + H_{\text{ph}}$  for non-interacting electrons and phonons are

$$G_0(P) = \frac{1}{\omega_n - \epsilon_0(k) + i\delta_k}$$

$$D_0(Q) = \left\{ \frac{1}{v_n - \omega_0(q) + i\delta} - \frac{1}{v_n + \omega_0(q) - i\delta} \right\}, \quad (2b)$$

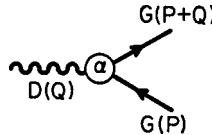
where  $P = (k, \omega_n)$  and  $Q = (q, v_n)$  are four vectors,  $\epsilon_0(k)$  is the bare electron quasiparticle energy referred to the Fermi surface,  $\omega_0(q)$  the bare phonon frequency and  $\omega_n$  and  $v_n$  the Matsubara frequencies

$$\omega_n = (2n+1)\pi k_B T; \quad v_n = 2n\pi k_B T \quad (3)$$

for Fermi and Bose particles, respectively.

As a result of the electron-phonon interaction,  $H_{\text{el-ph}}$ , both electron and phonon energies are renormalized. The renormalized propagators,  $G$  and  $D$ , can be given by a sum over Feynman diagrams, each of which represents a term in the perturbation expansion. We shall use light lines to represent the bare propagators,  $G_0$  and  $D_0$ , heavy lines for the renormalized propagators,  $G$  and  $D$ , straight lines for the electrons and curly lines for the phonons.

The electron-phonon interaction is described by the vertex



which represents scattering of an electron or hole by emission or absorption of a phonon or creation of an electron and hole by absorption of a phonon by an electron in the Fermi sea. Migdal showed that renormalization of the vertex represents only a small correction, of order  $(m/M)IP$ , a result in accord with the Born-Oppenheimer adiabatic approximation. If terms of this order are neglected, the electron and phonon self-energy corrections are given by the lowest order diagrams provided that fully renormalized propagators are used in these diagrams.

The electron self-energy  $L'(P)$  in the Dyson equation:

$$\overline{\text{---}} = \overline{\text{---}} + \overline{\text{---}} \circlearrowleft \overline{\text{---}}$$

$$G(P) = G_0(P) + G_0(P)\Sigma(P)G(P) \quad (4)$$

is given by the diagram

$$\Sigma = \text{---} \circlearrowleft \text{---} \quad (5)$$

The phonon self-energy,  $n(Q)$ , defined by

$$\text{---} \circlearrowleft \text{---} + \text{---} = \text{---} \quad (6)$$

is given by

$$\pi = \begin{array}{c} G(P+Q) \\ \text{---} \\ \text{---} \end{array} \quad \text{---} \quad \begin{array}{c} \text{---} \\ \text{---} \\ \text{---} \end{array} \quad \begin{array}{c} \text{---} \\ \text{---} \\ \text{---} \end{array}$$

Since to order  $(m/M)^{1/2}$  one can use an unrenormalized vertex function  $a = a_{\mu\mu}$ , the Dyson equations form a closed system such that both  $\Sigma(P)$  and  $n(Q)$  can be determined. The phonon self-energy,  $n(Q)$ , gives only a small renormalization of the phonon frequencies. As to the electrons, Migdal noted that we are interested in states  $k$  very close to  $k_F$ , so that to a close approximation  $\Sigma(k, \omega)$  depends only on the frequency. For an isotropic system,

$$\Sigma(k, \omega) \simeq \Sigma(k_F, \omega) \equiv \Sigma(\omega) \quad (7)$$

The renormalized electron quasi-particle energy,  $\omega_k$ , is then given by a root of

$$\varepsilon(k) = \omega_k = \varepsilon_0(k) + \Sigma(\omega_k) \quad (8)$$

In the thermal Green's function formalism, one may make an analytic continuation from the imaginary frequencies,  $\omega_n$ , to the real  $\omega$  axis to determine  $\Sigma(\omega)$ .

Although  $\Sigma(\omega)$  is small compared with the Fermi energy,  $E_F$ , it changes rapidly with energy and so can affect the density of states at the Fermi surface and thus the low temperature electronic specific heat. The mass renormalization factor  $m^*/m$ , at the Fermi surface may be expressed in terms of a parameter  $\lambda$ :

$$m^*/m = \zeta(k_F) = 1 + \lambda = (d\varepsilon_0/dk)_F / (d\varepsilon/dk)_F \quad (9)$$

In modern notation, the expression for  $\lambda$  is

$$\lambda = 2 \int_0^\infty d\omega \frac{\alpha^2(\omega) F(\omega)}{\omega} \quad (10)$$

where  $F(\omega)$  is the density of phonon states in energy and  $\alpha^2(\omega)$  is the square of the electron-phonon coupling constant averaged over polarization directions of the phonons. Note that  $\lambda$  is always positive so that the Fermi surface is stable if the lattice is stable. Values of  $\lambda$  for various metals range from about 0.5 to 1.5. The parameter  $\lambda$  corresponds roughly to the  $N(0)V_{\text{phonon}}$  of the BCS theory.

#### 4 NAMBU-ELIASHBERG THEORY FOR SUPERCONDUCTORS

Migdal's theory has important consequences that have been verified experimentally for normal metals, but gave no clue as to the origin of superconductivity. Following the introduction of the BCS theory, Gor'kov showed that pairing could be introduced through the anomalous Green's function

$$F(P) = i < T\psi_{\uparrow}\psi_{\downarrow} >, \quad (11)$$

Nambu showed that both types of Green's functions can be conveniently included with use of a spinor notation

$$\psi = \begin{pmatrix} \psi_{\uparrow}(r,t) \\ \psi_{\downarrow}^{+}(r,t) \end{pmatrix} \quad (12)$$

where  $\psi_{\uparrow}$  and  $\psi_{\downarrow}$  are wave field operators for up and down spin electrons and a matrix Green's function with components

$$\tilde{G}_{\alpha\beta} = -i < T\psi_{\alpha}\psi_{\beta}^{+} > \quad (13)$$

Thus  $G_{11}$  and  $G_{22}$  are the single particle Green's functions for up and down spin particles and  $G_{12} = G_{21}^* = F(P)$  is the anomalous Green's function of Gor'kov.

There are two self-energies,  $\Sigma_1$  and  $\Sigma_2$ , defined by the matrix

$$\tilde{\Sigma} = \begin{pmatrix} \Sigma_1 & \Sigma_2 \\ \Sigma_2 & \Sigma_1 \end{pmatrix} \quad (14)$$

Eliashberg noted that one can describe superconductors to the same accuracy as normal metals if one calculates the self-energies with the same diagrams that Migdal used, but with Nambu matrix propagators in place of the usual normal state Green's functions. The matrix equation for  $\tilde{G}$  is

$$\tilde{G} = \tilde{G}_0 + \tilde{G}_0 \tilde{\Sigma} \tilde{G} \quad (15)$$

The matrix equation for  $\tilde{\Sigma}$  yields a pair of coupled integral equations for  $\Sigma_1$  and  $\Sigma_2$ . Again  $\Sigma_1$  and  $\Sigma_2$  depend mainly on the frequency and are essentially independent of the momentum variables. Following Nambu, [13] one may define a renormalization factor  $\zeta_s(\omega)$  and a pair potential,  $\Delta(\omega)$ , for isotropic systems through the equations :

$$\omega \zeta_s(\omega) = \omega + \Sigma_1(\omega) \quad (16)$$

$$\Delta(\omega) = \Sigma_2(\omega)/\zeta(\omega). \quad (17)$$

Both  $\zeta_s$  and  $\Delta$  can be complex and include quasi-particle life-time effects. Eliashberg derived coupled non-linear integral equations for  $\zeta_s(\omega)$  and  $\Delta(\omega)$  which involve the electron-phonon interaction in the function  $a^2(\omega)F(\omega)$ .

The Eliashberg equations have been used with great success to calculate the properties of strongly coupled superconductors for which the frequency dependence of  $\zeta$  and  $\Delta$  is important. They reduce to the BCS theory and to the nearly equivalent theory of Bogoliubov [21] based on the principle of "compensation of dangerous diagrams" when the coupling is weak. By weak coupling is meant that the significant phonon frequencies are very large compared with  $k_B T_c$  so that  $\Delta(\omega)$  can be regarded as a constant independent of frequency in the important range of energies extending to at most a few  $k_B T_c$ . In weak coupling one may also neglect the difference in quasi-particle energy renormalization and assume that  $\zeta_s = \zeta_n$ .

The first solutions of the Eliashberg equations were obtained by Morel and Anderson [22] for an Einstein frequency spectrum. Coulomb interactions were included, following Bogoliubov, by introducing a parameter  $\mu^*$  which renormalizes the screened Coulomb interaction to the same energy range as the phonon interaction. In weak coupling,  $N(0)V = \lambda - \mu^*$ . They estimated  $\lambda$  from electronic specific heat data and  $\mu^*$  from the electron density and thus the transition temperatures,  $T_c$ , for a number of metals. Order-of-magnitude

agreement with experiment was found. Later work, based in large part on tunneling data, has yielded precise information on the electron-phonon interaction for both weak and strongly-coupled superconductors.

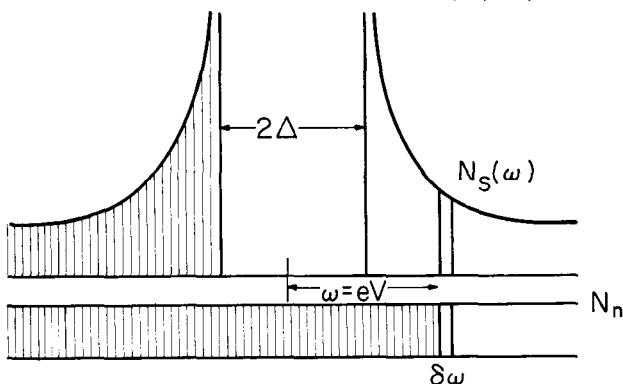
## 4

## ANALYSIS OF TUNNELING DATA

From the voltage dependence of the tunneling current between a normal metal and a superconductor one can derive  $\Delta(\omega)$  and thus get direct information about the Green's function for electrons in the superconductor. It is possible to go further and derive empirically from tunneling data the electron-phonon coupling,  $a^2(\omega)F(\omega)$ , as a function of energy. That electron tunneling should provide a powerful method for investigating the energy gap in superconductors was suggested by I. Giaever, [23] and he first observed the effect in the spring of 1960.

The principle of the method is illustrated in Fig. 1. At very low temperatures, the derivative of the tunneling current with respect to voltage is proportional to the density of states in energy in the superconductor. Thus the ratio of the density of states in the metal in the superconducting phase,  $N_s$ , to that of the same metal in the normal phase,  $N_n$ , at an energy eV above the Fermi surface is given by

$$\frac{N_s(eV)}{N_n} = \frac{(dI/dV)_s}{(dI/dV)_n} \quad (18)$$



$$\left(\frac{dI}{dV}\right)_s \sim N_s(\omega) \sim \frac{\omega}{\sqrt{\omega^2 - \Delta^2}}$$

## Tunneling from a normal metal into a superconductor

Fig. 1.

Schematic diagram illustrating tunneling from a normal metal into a superconductor near  $T = 0^\circ\text{K}$ . Shown in the lower part of the diagram is the uniform density of states in energy of electrons in the normal metal, with the occupied states shifted by an energy  $eV$  from an applied voltage  $V$  across the junction. The upper part of the diagram shows the density of states in energy in the superconductor, with an energy gap  $2\Delta$ . The effect of an increment of voltage  $\delta V$  giving an energy change  $\delta\omega$  is to allow tunneling from states in the range  $\delta\omega$ . Since the tunneling probability is proportional to density of states  $N_s(\omega)$ , the increment in current  $\delta I$  is proportional to  $N_s(\omega)\delta V$ .

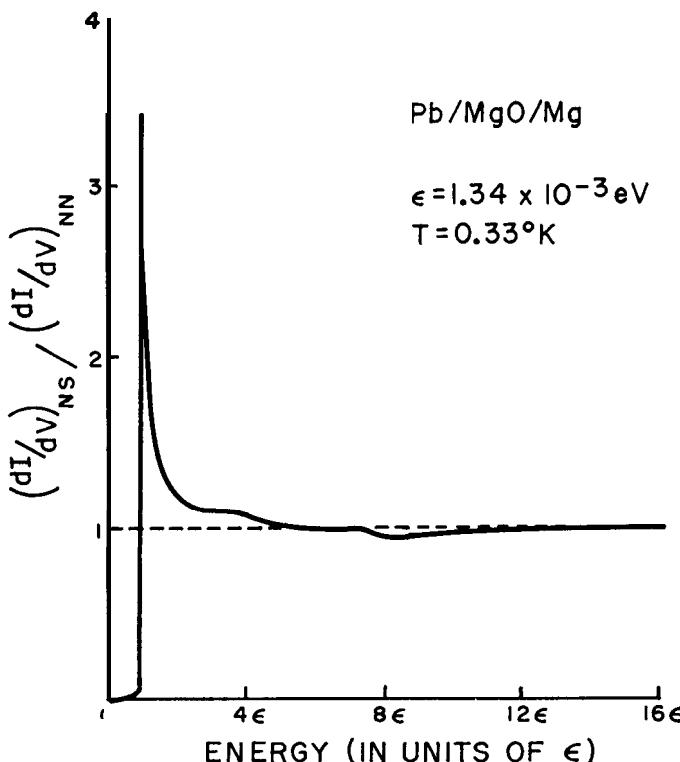


Fig. 2.

Conductance of a Pb-Mg junction as a function of applied voltage (from reference 24).

The normal density is essentially independent of energy in the range involved (a few meV). In weak coupling superconductors, for a voltage  $V$  and energy  $\omega = eV$ ,

$$\frac{N_s(\omega)}{N_n} = \frac{\omega}{\sqrt{\omega^2 - \Delta^2}}. \quad (19)$$

As  $T \rightarrow 0$  K, no current flows between the normal metal and the superconductor until the applied voltage reaches  $\Delta/e$ , when there is a sharp rise in  $dI/dV$  followed by a drop. This is illustrated in Fig. 2 for the case of Pb.

The first experiments of Giaever were on aluminum, which is a weak coupling superconductor. Good agreement was found between theory and experiment. In later measurements on tunneling into Pb, a strongly coupled superconductor, Giaever, Hart and Megerle [24] observed anomalies in the density of states that appeared to be associated with phonons, as shown in Fig. 2. These results were confirmed by more complete and accurate tunneling data on Pb by J. M. Rowell et al. [25].

In the meantime, in the summer of 1961, Schrieffer had derived numerical solutions of the Eliashberg equations working with a group engaged in developing methods for computer control using graphical display methods. [26] He and co-workers calculated the complex  $A(w)$  for a Debye frequency

spectrum. Later, at the University of Pennsylvania, he together with J. W. Wilkins and D. J. Scalapino [27] continued work on the problem with a view to explaining the observed anomalies on Pb. They showed that for the general case of a complex  $d(w)$

$$\frac{(dI/dV)_{ns}}{(dI/dV)_{nn}} = \frac{N_s(\omega)}{N_n} = \text{Re} \left\{ \frac{\omega}{\sqrt{\omega^2 - \Delta^2(\omega)}} \right\} \quad (20)$$

where  $\text{Re}$  represents the real part. From measurements of the ratio over the complete range of voltages, one can use Kramers-Kronig relations to obtain both the real and imaginary parts of  $d(w) = d(a) + id(cu)$ . From analysis of the data, one can obtain the Green's functions which in turn can be used to calculate the various thermal and transport properties of superconductors. This has been done with great success, even for such strongly-coupled superconductors as lead and mercury.

For lead, Schrieffer et al, used a phonon spectrum consisting of two Lorentzian peaks, one for transverse waves and one for longitudinal and obtained a good fit to the experimental data for  $T < < T_c$ . The calculations were extended up to  $T_c$  for Pb, Hg, and Al by Swihart, Wada and Scalapino, [28] again finding good agreement with experiment.

In analysis of tunneling data, one would like to find a phonon interaction spectrum,  $a''(w)F(w)$ , and a Coulomb interaction parameter,  $p^*$ , which when inserted into the Eliashberg equations will yield a solution consistent with the tunneling data. W. L. McMillan devised a computer program such that one could work backwards and derive  $a''(w)F(w)$  and  $p^*$  directly from the tunneling data. His program has been widely used since then and has been applied to a number of superconducting metals and alloys, including, Al, Pb, Sn, the transition elements Ta and Nb, a rare earth, La, and the compound Nb<sub>3</sub>Sn. In all cases it has been found that the phonon mechanism is dominant with reasonable values of  $p^*$ . Peaks in the phonon spectrum agree with peaks in the phonon density of states as found from neutron scattering data, as shown in Fig. 3 for the case of Pb. In Fig. 4 is shown the real and imaginary parts of  $\Delta(\omega)$  for Pb as derived from tunneling data.

One can go further and calculate the various thermodynamic and other properties. Good agreement with experiment is found for strongly coupled superconductors even when there are significant deviations from the weak coupling limits. For example, the weak-coupling BCS expression for the condensation energy at  $T = 0$  K is

$$E_{BCS} = \frac{1}{2} N(O) \zeta_n A_0^2 \quad (21)$$

where  $N(O)$  is the phonon enhanced density of states and  $A_0$  is the gap parameter at  $T = 0$  K. The theoretical expression with  $\langle s(u) \rangle$  and  $A(w)$  derived from tunneling data, again for the case of Pb, gives [29, 30, 31]

$$\bar{E}_{\text{theor}} = 0.78 E_{BCS} \quad (22)$$

in excellent agreement with the experimental value

$$E_{\text{exp}} = (0.76 + 0.02) E_{BCS}. \quad (23)$$

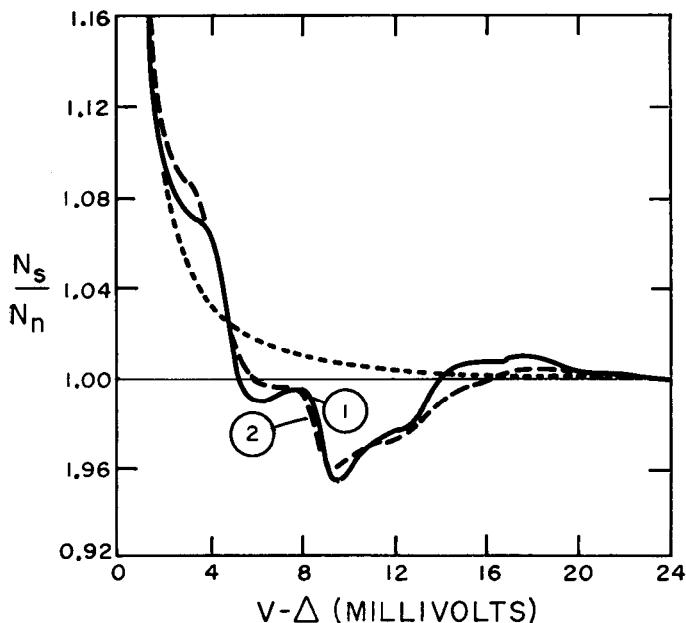
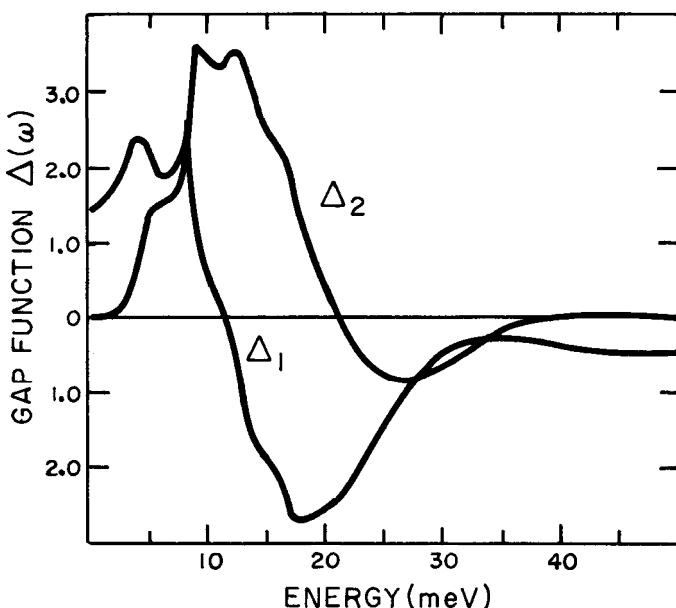


Fig. 3.

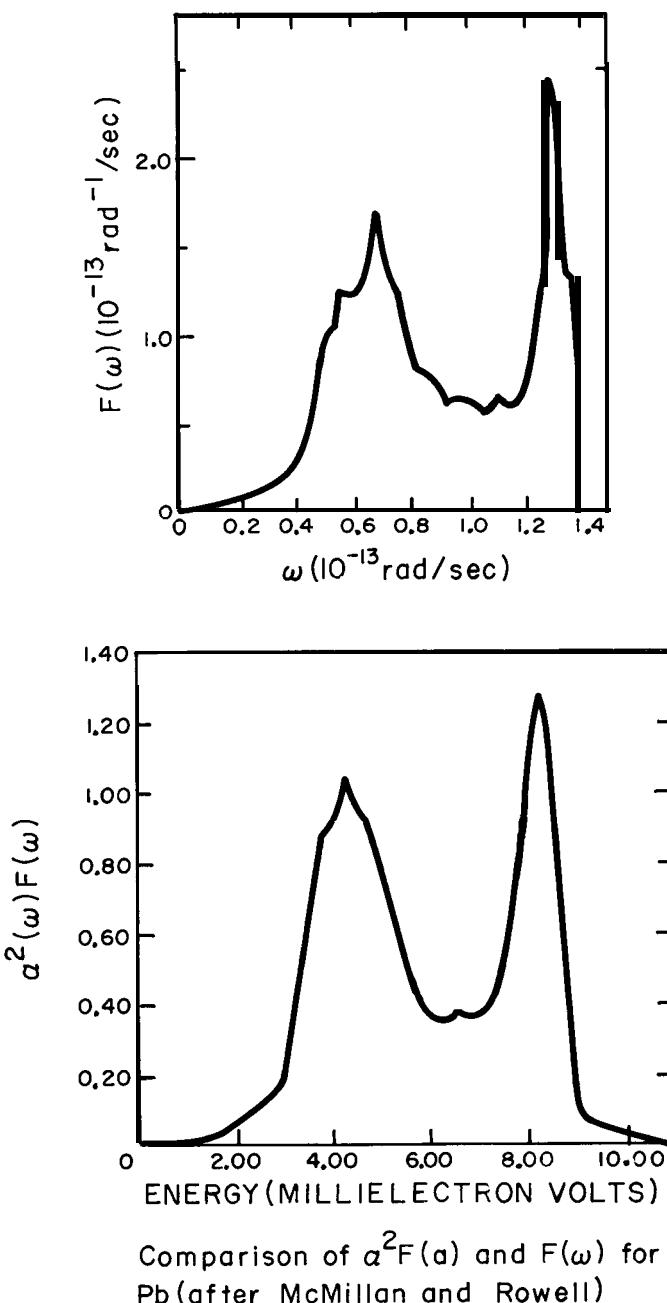
Density of states versus energy for Pb. Solid line, calculated by Schrieffer et al; long dashed line, observed from tunneling; short dashed line, BCS weak coupling theory.



Real and imaginary parts of  $A$  versus  $\text{co-}A_0$  for Pb.

Fig. 4.

Real and imaginary parts of  $d(\omega) = d_r(\omega) + i d_i(\omega)$  versus energy for Pb. (After McMillan & Rowell).

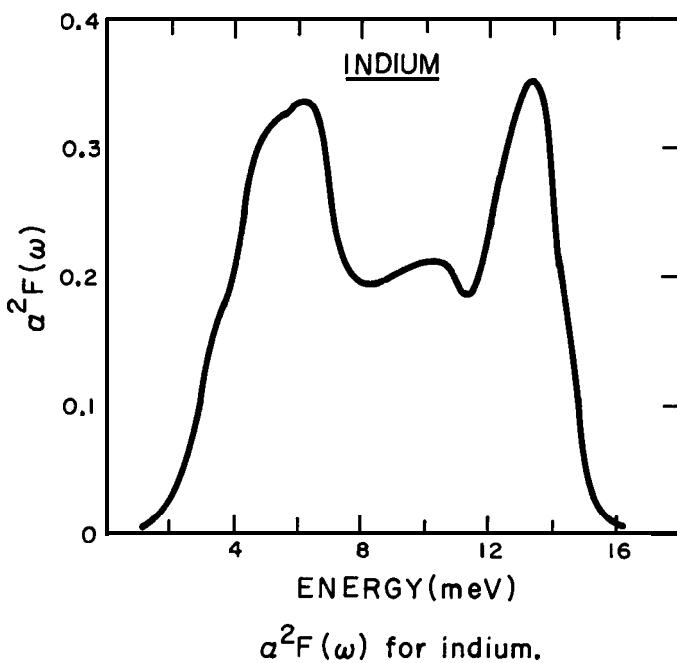


Comparison of  $\alpha^2 F(\omega)$  and  $F(\omega)$  for  
Pb (after McMillan and Rowell)

Fig. 5.

Comparison of  $\alpha^2 F$  for Pb derived from tunneling data with phonon density of states from neutron scattering data of Stedman et al. [8]

In Figs. 5, 6, 7, and 8 are shown other examples of  $\alpha^2(\omega)F(\omega)$  derived from tunneling data for Pb, In, [31] La, [32] and Nb<sub>3</sub>Sn. [33] In all cases the results are completely consistent with the phonon mechanism. Coulomb interactions play only a minor role, with  $\mu^*$  varying only slowly from one metal to another, and generally in the range 0.1-02.



$\alpha^2 F(\omega)$  for indium.

Fig. 6.

$\alpha^2 F$  for In (after McMillan and Rowell).

As a further check, it is possible to derive the phonon density of states,  $F(o)$  from neutron scattering data and use pseudo-potential theory to calculate the electron-phonon interaction parameter  $\alpha_q(\omega)$ . From these values, one can use the Eliashberg equations to calculate  $Z_s(\omega)$  and  $O(\omega)$  and the various superconducting properties, including the transition temperature,  $T_c$ . Extensive calculations of this sort have been made by J. P. Carbotte and co-workers [34] for several of the simpler metals and alloys. For example, for the gap edge,  $A_{\text{c}}$ , in Al at  $T = 0$  K they find 0.19 meV as compared with an experimental value of 0.17. The corresponding values for Pb are 1.49 meV from theory as compared with 1.35 meV from experiment. These are essentially first principles calculations and give convincing evidence that the theory as formulated is essentially correct. Calculations made for a number of other metals and alloys give similar good agreement.

#### CONCLUSIONS

In this talk we have traced how our understanding of the role of electron-phonon interactions in superconductivity has developed from a concept to a precise quantitative theory. The self-energy and pair potential, and thus the Green's functions, can be derived either empirically from tunneling data or directly from microscopic theory with use of the Eliashberg equations. Physicists, both experimental and theoretical, from different parts of the world have contributed importantly to these developments.

All evidence indicates that the electron-phonon interaction is the dominant mechanism in the cases studied so far, which include many simple metals,

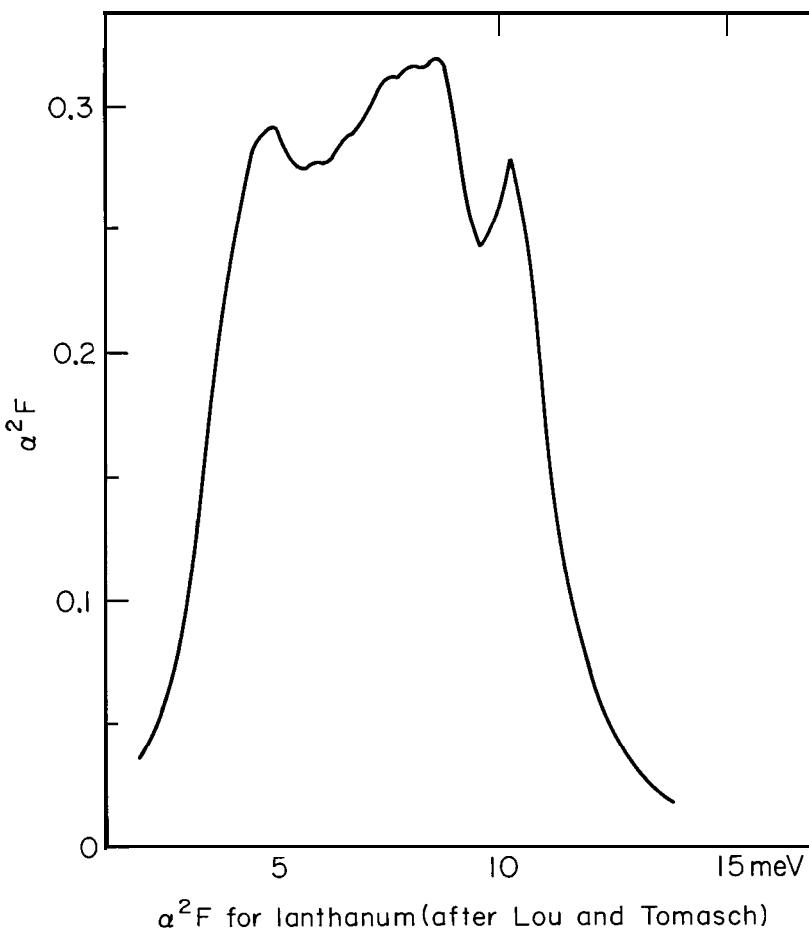


Fig. 7.  
 $\alpha^2 F$  for La (after Lou and Tomasch).

transition metals, a rare earth, and various alloys and compounds. Except possibly for the metallic form of hydrogen, [35] which is presumed to exist at very high pressures, it is unlikely that the phonon mechanism will yield substantially higher transition temperatures than the present maximum of about 21 K for a compound of Nb, Al and Ge.

Other mechanisms have been suggested for obtaining higher transition temperatures. One of these is to get an effective attractive interaction between electrons from exchange of virtual excitons, or electron-hole pairs. This requires a semiconductor in close proximity to the metal in a layer or sandwich structure. At present, one can not say whether or not such structures are feasible and in no case has the exciton mechanism been shown to exist. As Ginzburg has emphasized, this problem (as well as other proposed mechanisms) deserves study until a definite answer can be found. [36]

The pairing theory has had wide application to Fermi systems other than electrons in metals. For example, the theory has been used to account for

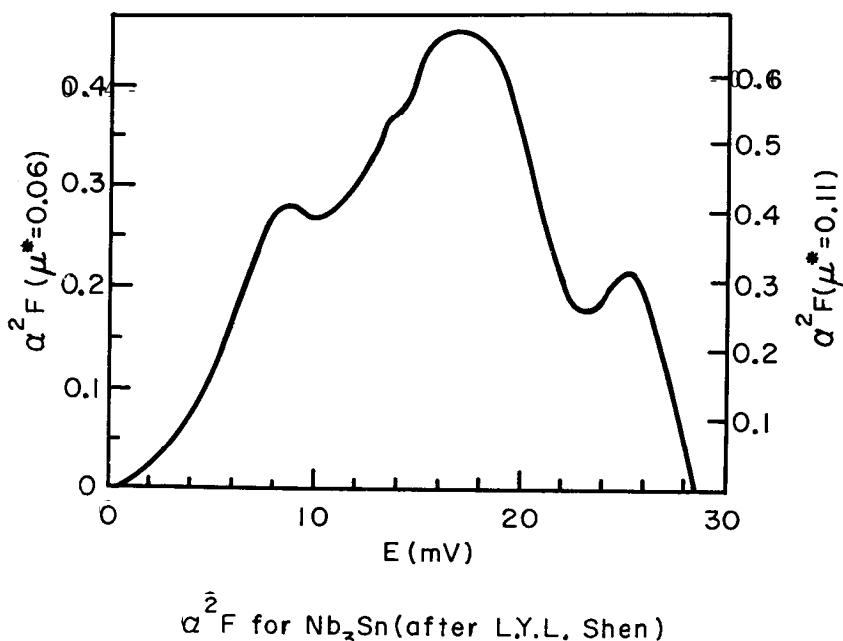


Fig. 8.  
 $\bar{\alpha}^2 F$  for  $Nb_3Sn$  (after Y. L. Y. Shen).

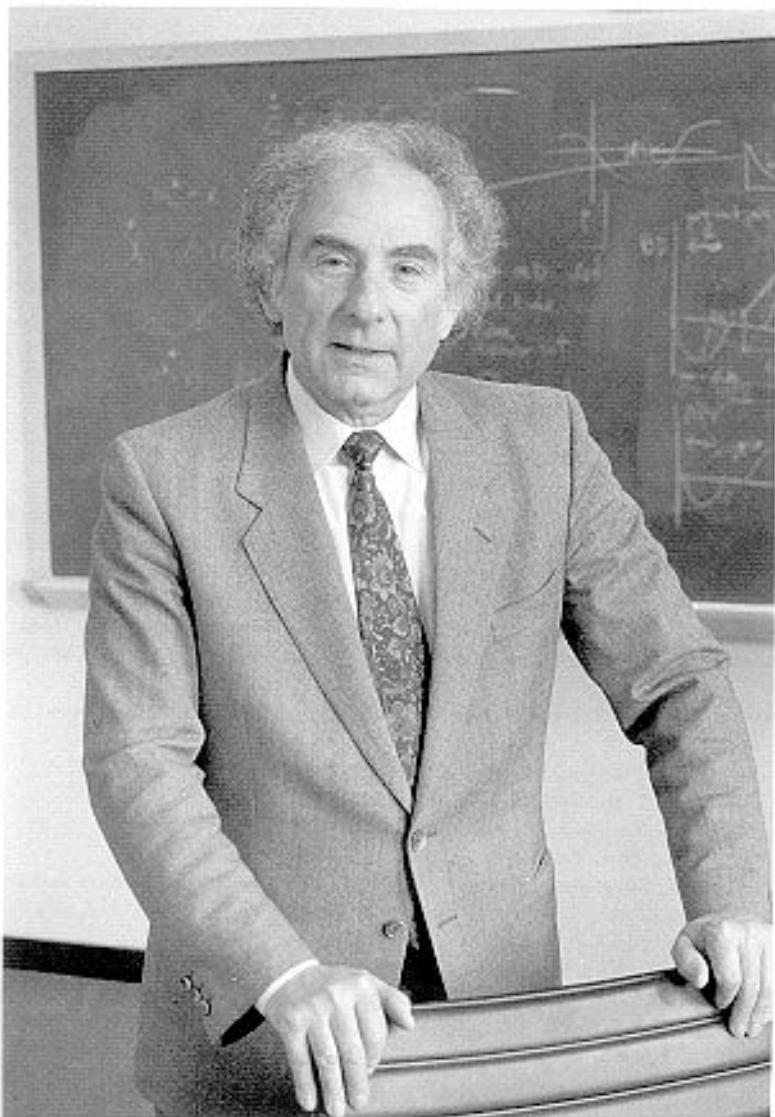
many aspects of nuclear structure. It is thought the nuclear matter in neutron stars is superfluid. Very recently, evidence has been found for a possible pairing transition in liquid  $He^3$  at very low temperatures [37]. Some of the concepts, such as that of a degenerate vacuum, have been used in the theory of elementary particles. Thus pairing seems to be a general phenomenon in Fermi systems.

The field of superconductivity is still a very active one in both basic science and applications. I hope that these lectures have given you some feeling for the accomplishments and the methods used.

#### REFERENCES

1. Shoenberg, D. Superconductivity, Cambridge Univ. Press, Cambridge (1938). Second edition, 1951.
2. London, F. Proc. Roy. Soc. (London) 152A, 24 (1935).
3. Bardeen, J. Phys. Rev. 59, 928A (1941).
4. Reynolds, C. A., Serin, B., Wright W. H. and Nesbitt, L. B. Phys. Rev. 78, 487 (1950).
5. Maxwell, E., Phys. Rev. 78, 477 (1950).
6. Frohlich, H., Phys. Rev. 79, 845 (1950); Proc. Roy. Soc. (London) Ser. A 213, 291 (1952).
7. London, F., Superfluids, New York, John Wiley and Sons, 1950.
8. For recent review articles with references, see the chapters by D. J. Scalapino and by W. L. McMillan and J. M. Rowe in Superconductivity, R. D. Parks, ed., New York, Marcel Dekker, Inc., 1969, Vol. 1. An excellent reference for the theory and earlier experimental work is J. R. Schrieffer, Superconductivity, New York, W. A. Benjamin, Inc., 1964. The present lecture is based in part on a chapter by the author in Cooperative Phenomena, H. Haken and M. Wagner, eds. to be published by Springer.

9. Gor'kov, L. P., Zh. Eksper i. teor. Fiz. 34, 735 (1958). (English transl. Soviet Phys. - JETP 7, 505 (1958)).
10. Kadanoff L. P. and Martin, P. C. Phys. Rev. 124, 670 (1961).
11. Migdal, A. B., Zh. Eksper i. teor. Fiz. 34, 1438 (1958). (English transl. Soviet Phys. - JETP 7, 996 (1958)).
12. Eliashberg, G. M., Zh. Eksper i. teor. Fiz. 38, 966 (1960). Soviet Phys. - JETP **11**, 696 (1960).
13. Nambu, Y., Phys. Rev. 117, 648 (1960).
14. Bardeen, J., Phys. Rev. 79, 167 (1950); 80, 567 (1950); 81 829 (1951).
15. Bardeen, J., Rev. Mod. Phys. 23, 261 (1951).
16. Schafroth, M. R., Helv. Phys. Acta 24, 645 (1951); Nuovo Cimento, 9, 291 (1952).
17. For a review see Bardeen, J., Encyclopedia of Physics, S. Flugge, ed., Berlin, Springer-Verlag, (1956) Vol. XV, p. 274.
18. Bardeen, J., L. N. Cooper and J. R. Schrieffer, Phys. Rev. 108, 1175 (1957).
19. For references, see the review article of M. A. Biondi, A. T. Forrester, M. B. Garfunkel and C. B. Satterthwaite, Rev. Mod. Phys. 30, 1109 (1958).
20. Pippard, A. B., Proc. Roy. Soc. (London) **A216**, 547 (1954).
21. See N. N. Bogoliubov, V. V. Tolmachev and D. V. Shirkov, A New Method in the Theory of Superconductivity, New York, Consultants Bureau, Inc., 1959.
22. Morel P. and Anderson, P. W., Phys. Rev. **125**, 1263 (1962).
23. Giaever, I., Phys. Rev. Letters, 5, 147; 5, 464 (1960).
24. Giaever, I., Hart H. R., and Megerle K., Phys. Rev. 126, 941 (1962).
25. Rowell, J.M., Anderson P. W. and Thomas D. E., Phys. Rev. Letters, **10**, 334 (1963).
26. Culler, G. J., Fried, B. D., Huff, R. W. and Schrieffer, J. R., Phys. Rev. Letters 8,339 (1962).
27. Schrieffer, J. R., Scalapino, D. J. and Wilkins, J. W., Phys. Rev. Letters **10**, 336 (1963); D. J. Scalapino, J. R. Schrieffer, and J. W. Wilkins, Phys. Rev. 148, 263 (1966).
28. Scalapino, D. J., Wada,Y. and Swihart, J. C., Phys. Rev. Letters, 14, 102 (**1965**);**14,106** (1965).
29. Eliashberg, G. M., Zh. Eksper i. teor. Fiz. 43, 1005 (1962). English transl. Soviet Phys. - JETP **16**, 780 (1963).
30. Bardeen, J. and Stephen, M., Phys. Rev. 136, A1485 (1964).
31. McMillan, W. L. and Rowell, J. M. in Reference 8.
32. Lou, L. F. and Tomasch, W. J., Phys. Rev. Lett. 29, 858 (1972).
33. Shen, L. Y. L., Phys. Rev. Lett. 29, 1082 (1972).
34. Carbotte, J. P., Superconductivity, P. R. Wallace, ed., New York, Gordon and Breach, 1969, Vol. 1, p. 491; J. P. Carbotte and R. C. Dynes, Phys. Rev. **172**, 476 (1968); C. R. Leavens and J. P. Carbotte, Can. Journ. Phys. 49, 724 (1971).
35. Ashcroft N. W., Phys. Rev. Letters, 21, 1748 (1968).
36. See V. L. Ginzburg, "The Problem of High Temperature Superconductivity," **Annual Review of Materials Science**, Vol. 2, p. 663 (1972).
37. Osheroff D. D., Gully W. J., Richardson R. C. and Lee, D. M., Phys. Rev. Lett. 29, 1621 (1972).



James Clapp

## LEON N COOPER

Leon Cooper was born in 1930 in New York where he attended Columbia University (A.B. 1951; A.M. 1953; Ph.D. 1954). He became a member of the Institute for Advanced Study (1954-55) after which he was a research associate of Illinois (1955-57) and later an assistant professor at the Ohio State University (1957-58). Professor Cooper joined Brown University in 1958 where he became Henry Ledyard Goddard University Professor (1966-74) and where he is presently the Thomas J. Watson, Sr. Professor of Science (1974-).

Professor Cooper is Director of Brown University's Center for Neural Science. This Center was founded in 1973 to study animal nervous systems and the human brain. Professor Cooper served as the first director with an interdisciplinary staff drawn from the Departments of Applied Mathematics, Biomedical Sciences, Linguistics and Physics. Today, Cooper, with members of the Brown Faculty, postdoctoral fellows and graduate students with interests in the neural and cognitive sciences, is working towards an understanding of memory and other brain functions, and thus formulating a scientific model of how the human mind works.

Professor Cooper has received many forms of recognition for his work in 1972, he received the Nobel Prize in Physics (with J. Bardeen and J. R. Schrieffer) for his studies on the theory of superconductivity completed while still in his 20s. In 1968, he was awarded the Comstock Prize (with J. R. Schrieffer) of the National Academy of Sciences. The Award of Excellence, Graduate Faculties Alumni of Columbia University and Descartes Medal, Academic de Paris, Universite Rene Descartes were conferred on Professor Cooper in the mid 1970s. In 1985, Professor Cooper received the John Jay Award of Columbia College. He holds seven honorary doctorates.

Professor Cooper has been an NSF Postdoctoral Fellow, 1954-55, Alfred P. Sloan Foundation Research Fellow, 1959-66 and John Simon Guggenheim Memorial Foundation Fellow, 1965-66. He is a fellow of the American Physical Society and American Academy of Arts and Sciences; Sponsor, Federation of American Scientists; member of American Philosophical Society, National Academy of Sciences, Society of Neuroscience, American Association for the Advancement of Science, Phi Beta Kappa, and Sigma Xi. Professor Cooper is also on the Governing Board and Executive Committee of the International Neural Network Society and a member of the Defense Science Board.

Professor Cooper is Co-founder and Co-chairman of Nestor, Inc., an industry leader in applying neural-network systems to commercial and military applications. Nestor's adaptive pattern-recognition and risk-assessment systems simulated in small conventional computers *learn by example* to accurately classify complex patterns such as targets in sonar, radar or imaging systems, to emulate human decisions in such applications as mortgage origination and to assess risks.

Born: February 28, 1930, New York City, married, two children.

A.B. 1951; A.M. 1953; Ph.D. 1954, all at Columbia University.

Member, Institute for Advanced Study, 1954-55.

Research Associate, University of Illinois, 1955-57.

Assistant Professor, Ohio State University, 1957-58.

Associate Professor, 1958-62; Professor, 1962-66; Henry Ledyard Goddard University Professor, 1966-74; Thomas J. Watson, Sr., Professor of Science, 1974-, Brown University.

Visiting Professor, various universities and summer schools.

Consultant, various governmental agencies, industrial and educational organizations, various public lectures, international conferences and symposia.

Executive Committee of the International Neural Network Society.

Member, Defense Science Board.

NSF Postdoctoral Fellow, 1954-55.

Alfred P. Sloan Foundation Research Fellow, 1959-66.

Doctor of Sciences (*honoris causa*), Columbia University, 1973; University of Sussex, 1973; University of Illinois, 1974; Brown University, 1974; Gustavus Adolphus College, 1975; Ohio State University, 1976; Universite Pierre et Marie Curie, Paris, 1977.

Comstock Prize (with J. R. Schrieffer), National Academy of Sciences, 1968.

Nobel Prize (with J. Bardeen and J. R. Schrieffer), 1972.

Award of Excellence, Graduate Faculties Alumni of Columbia University, 1974.

Descartes Medal, Academic de Paris, Université Rene Descartes, 1977.

John Jay Award, Columbia College, 1985.

Award for Distinguished Achievement, Columbia University, 1990.

Who's Who, Who's Who in America, Who's Who in the World, various other listings.

Fellow, American Academy of Arts and Sciences, American Physical Society; Member, American Philosophical Society, National Academy of Sciences, Society of Neuroscience, American Association for the Advancement of Science, Phi Beta Kappa, Sigma Xi; Sponsor, Federation of American Scientists.

# MICROSCOPIC QUANTUM INTERFERENCE EFFECTS IN THE THEORY OF SUPERCONDUCTIVITY

Nobel Lecture, December 11, 1972

by

LEON N COOPER

Physics Department, Brown University, Providence, Rhode Island

It is an honor and a pleasure to speak to you today about the theory of superconductivity. In a short lecture one can no more than touch on the long history of experimental and theoretical work on this subject before 1957. Nor can one hope to give an adequate account of how our understanding of superconductivity has evolved since that time. The theory (1) we presented in 1957, applied to uniform materials in the weak coupling limit so defining an ideal superconductor, has been extended in almost every imaginable direction. To these developments so many authors have contributed (2) that we can make no pretense of doing them justice. I will confine myself here to an outline of some of the main features of our 1957 theory, an indication of directions taken since and a discussion of quantum interference effects due to the singlet-spin pairing in superconductors which might be considered the microscopic analogue of the effects discussed by Professor Schrieffer.

## NORMAL METAL

Although attempts to construct an electron theory of electrical conductivity date from the time of Drude and Lorentz, an understanding of normal metal conduction electrons in modern terms awaited the development of the quantum theory. Soon thereafter Sommerfeld and Bloch introduced what has evolved into the present description of the electron fluid. (3) There the conduction electrons of the normal metal are described by single particle wave functions. In the periodic potential produced by the fixed lattice and the conduction electrons themselves, according to Bloch's theorem, these are modulated plane waves :

$$\phi_{\mathbf{k}}(\mathbf{r}) = u_{\mathbf{k}}(\mathbf{r}) e^{i\mathbf{k}\cdot\mathbf{r}},$$

where  $u_{\mathbf{k}}(\mathbf{r})$  is a two component spinor with the lattice periodicity. We use  $K$  to designate simultaneously the wave vector  $\mathbf{k}$ , and the spin state  $\sigma : K \equiv k, \uparrow ; -K \equiv -k, \downarrow$ . The single particle Bloch functions satisfy a Schrödinger equation

$$\left[ -\frac{\hbar^2}{2m} \nabla^2 + V_0(\mathbf{r}) \right] \phi_{\mathbf{k}} = \varepsilon_{\mathbf{k}} \phi_{\mathbf{k}}$$

where  $V_0(\mathbf{r})$  is the periodic potential and in general might be a linear operator to include exchange terms.

The Pauli exclusion principle requires that the many electron wave function be antisymmetric in all of its coordinates. As a result no two electrons can be

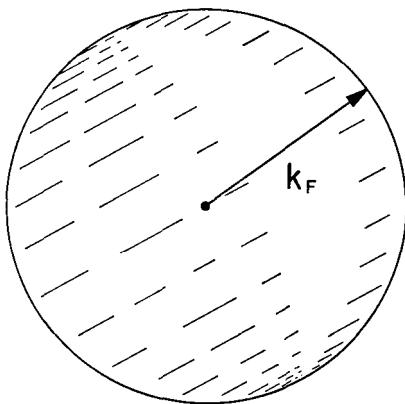


Fig. 1.

The normal ground state wavefunction,  $\Phi_0$ , is a filled Fermi sphere for both spin directions.

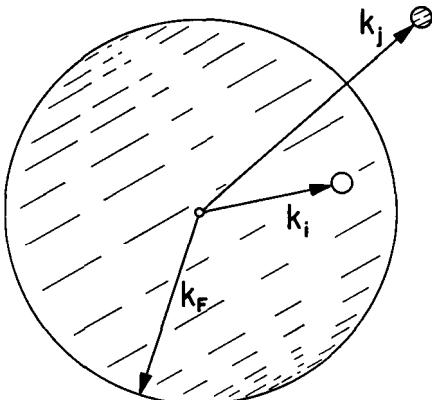


Fig. 2.

An excitation of the normal system.

in the same single particle Bloch state. The energy of the entire system is

$$W = \sum_{i=1}^{2N} \varepsilon_i$$

where  $\varepsilon_i$  is the Bloch energy of the  $i^{\text{th}}$  single electron state. The ground state of the system is obtained when the lowest  $N$  Bloch states of each spin are occupied by single electrons; this can be pictured in momentum space as the filling in of a Fermi sphere, Fig. 1. In the ground-state wave function there is no correlation between electrons of opposite spin and only a statistical correlation (through the general anti-symmetry requirement on the total wave function) of electrons of the same spin.

Single particle excitations are given by wave functions identical to the ground state except that one electron states  $k_i < k_F$  are replaced by others  $k_j < k_F$ . This may be pictured in momentum space as opening vacancies below the Fermi surface and placing excited electrons above, Fig. 2. The energy difference between the ground state and the excited state with the particle excitation  $k_j$  and the hole excitation  $k_i$  is

$$\varepsilon_j - \varepsilon_i = \varepsilon_j - \varepsilon_F - (\varepsilon_i - \varepsilon_F) = \varepsilon_j - \varepsilon_i = |\varepsilon_j| + |\varepsilon_i|$$

where we define  $\varepsilon$  as the energy measured relative to the Fermi energy

$$\varepsilon_i = \varepsilon_i - \varepsilon_F.$$

When Coulomb, lattice-electron and other interactions, which have been omitted in constructing the independent particle Bloch model are taken into account, various modifications which have been discussed by Professor Schrieffer are introduced into both the ground state wave function and the excitations. These may be summarized as follows: The normal metal is described by a ground state  $\Phi_0$  and by an excitation spectrum which, in addition to the various collective excitations, consists of quasi-fermions which satisfy the usual anticommutation relations. It is defined by the sharpness of the Fermi surface, the finite density of excitations, and the continuous decline of the single particle excitation energy to zero as the Fermi surface is approached.

**ELECTRON CORRELATIONSTHAT PRODUCE SUPERCONDUCTIVITY**

For a description of the superconducting phase we expect to include correlations that are not present in the normal metal. Professor Schrieffer has discussed the correlations introduced by an attractive electron-electron interaction and Professor Bardeen will discuss the role of the electron-phonon interaction in producing the electron-electron interaction which is responsible for superconductivity. It seems to be the case that any attractive interaction between the fermions in a many-fermion system can produce a superconducting-like state. This is believed at present to be the case in nuclei, in the interior of neutron stars and has possibly been observed (4) very recently in He<sup>3</sup>. We will therefore develop the consequences of an attractive two-body interaction in a degenerate many-fermion system without enquiring further about its source.

The fundamental qualitative difference between the superconducting and normal ground state wave function is produced when the large degeneracy of the single particle electron levels in the normal state is removed. If we visualize the Hamiltonian matrix which results from an attractive two-body interaction in the basis of normal metal configurations, we find in this enormous matrix, sub-matrices in which all single-particle states except for one pair of electrons remain unchanged. These two electrons can scatter via the electron-electron interaction to all states of the same total momentum. We may envisage the pair wending its way (so to speak) over all states unoccupied by other electrons. [The electron-electron interaction in which we are interested is both weak and slowly varying over the Fermi surface. This and the fact that the energy involved in the transition into the superconducting state is small leads us to guess that only single particle excitations in a small shell near the Fermi surface play a role. It turns out, further, that due to exchange terms in the electron-electron matrix element, the effective interaction in metals between electrons of singlet spin is much stronger than that between electrons of triplet spin—thus our preoccupation with singlet spin correlations near the Fermi surface.] Since every such state is connected to every other, if the interaction is attractive and does not vary rapidly, we are presented with submatrices of the entire Hamiltonian of the form shown in Fig. 3. For purposes of illustration we have set all off diagonal matrix elements equal to the constant-V and the diagonal terms equal to zero (the single particle excitation energy at the Fermi surface) as though all the initial electron levels were completely degenerate. Needless to say, these simplifications are not essential to the qualitative result.

Diagonalizing this matrix results in an energy level structure with M-1 levels raised in energy to  $E = + V$  while one level (which is a superposition of all of the original levels and quite different in character) is lowered in energy to

$$E = -(M-1)V.$$

Since M, the number of unoccupied levels, is proportional to the volume of the container while V, the scattering matrix element, is proportional to 1/volume, the product is independent of the volume. Thus the removal of

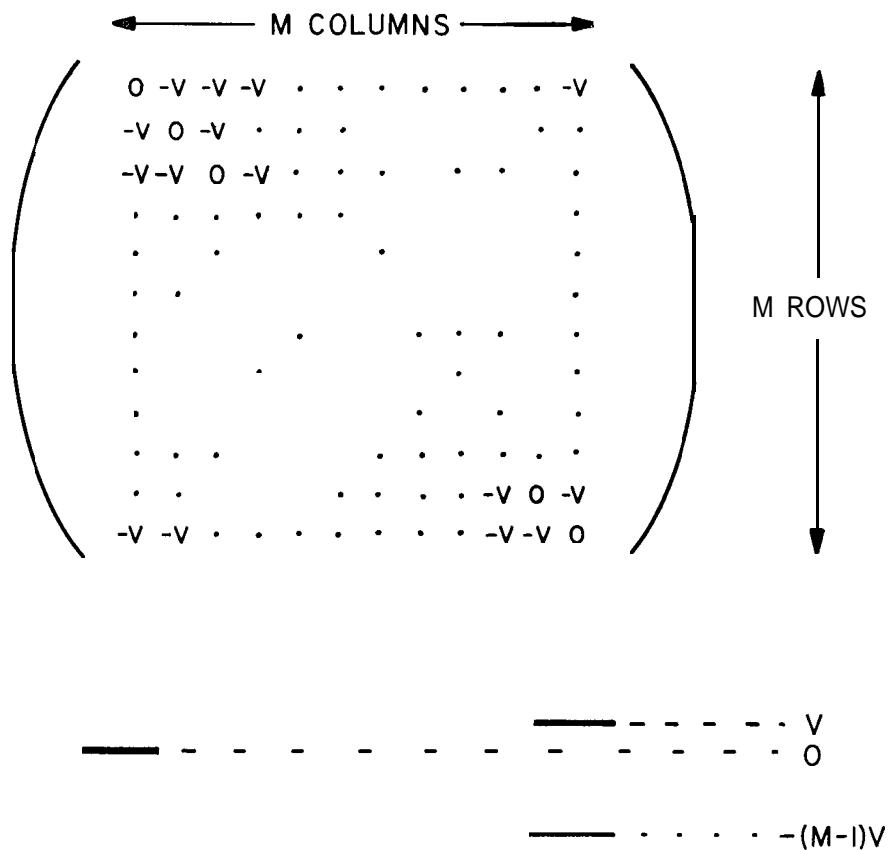


Fig. 3

the degeneracy produces a single level separated from the others by a volume independent energy gap.

To incorporate this into a solution of the full Hamiltonian, one must devise a technique by which all of the electron pairs can scatter while obeying the exclusion principle. The wave function which accomplishes this has been discussed by Professor Schrieffer. Each pair gains an energy due to the removal of the degeneracy as above and one obtains the maximum correlation of the entire wave function if the pairs all have the same total momentum. This gives a coherence to the wave function in which for a combination of dynamical and statistical reasons there is a strong preference for momentum zero, singlet spin correlations, while for statistical reasons alone there is an equally strong preference that all of the correlations have the same total momentum.

In what follows I shall present an outline of our 1957 theory modified by introducing the quasi-particles of Bogoliubov and Valatin. (5) This leads to a formulation which is generally applicable to a wide range of calculations

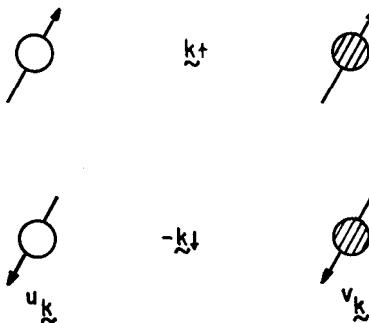


Fig. 4.

The ground state of the superconductor is a linear superposition of states in which pairs ( $k\uparrow - k\downarrow$ ) are occupied or unoccupied.

in a manner analogous to similar calculations in the theory of normal metals.

We limit the interactions to terms which scatter (and thus correlate) singlet zero-momentum pairs. To do this, it is convenient to introduce the pair operators :

$$b_k = c_{-k} c_k$$

$$b_k^* = c_k^* c_{-k}^*$$

and using these we extract from the full Hamiltonian the so-called reduced Hamiltonian

$$H_{\text{reduced}} = \sum_{k < k_f} 2|\epsilon| b_k b_k^* + \sum_{k > k_f} 2\epsilon b_k^* b_k + \sum_{kk'} V_{kk'} b_{k'}^* b_k$$

where  $V_{..}$  is the scattering matrix element between the pair states  $k$  and  $k'$ .

#### GROUND STATE

As Professor Schrieffer has explained, the ground state of the superconductor is a linear superposition of pair states in which the pairs ( $k\uparrow - k\downarrow$ ) are occupied or unoccupied as indicated in Fig. 4. It can be decomposed into two disjoint vectors-one in which the pair state  $k$  is occupied, 0, and one in which it is unoccupied,  $0_{\text{ex}}$ :

$$\psi_0 = u_k \mathcal{O}_{(k)} + v_k \mathcal{O}_{(k)}$$

The probability amplitude that the pair state  $k$  is (is not) occupied in the ground state is then  $v_k(u_k)$ . Normalization requires that  $\int u_k^2 + (-1)^k v_k^2 = 1$ . The phase of the ground state wave function may be chosen so that with no loss of generality  $u_k$  is real. We can then write

$$u = (1-h)^{1/2}$$

$$v = h^{1/2} e^{i\varphi}$$

where

$$0 \leq h \leq 1.$$

A further decomposition of the ground state wave function of the superconductor in which the pair states  $k$  and  $k'$  are either occupied or unoccupied Fig. 5 is:

$$\psi_0 = u_k u_{k'} \mathcal{O}_{(k), (k')} + u_k v_{k'} \mathcal{O}_{(k), k'} + v_k u_{k'} \mathcal{O}_{k, (k')} + v_k v_{k'} \mathcal{O}_{k, k'} .$$

This is a Hartree-like approximation in the probability amplitudes for the occupation of pair states. It can be shown that for a fermion system the wave

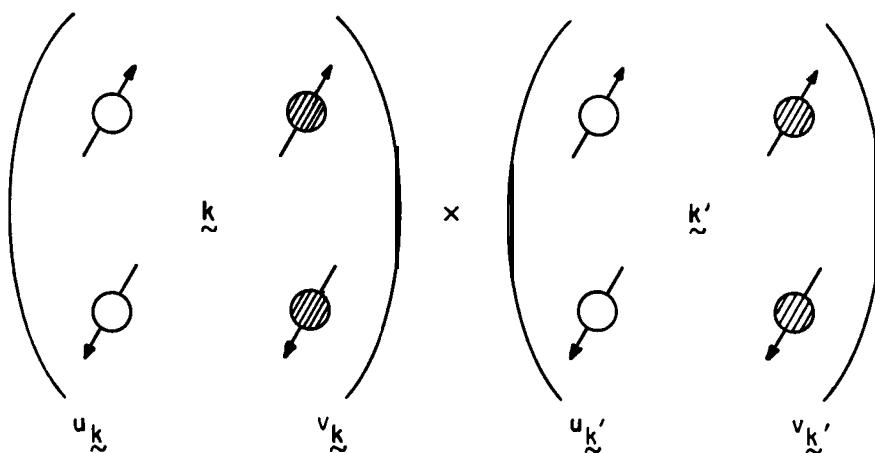


Fig. 5.

A decomposition of the ground state of the superconductor into states in which the pair states  $k$  and  $k'$  are either occupied or unoccupied.

function cannot have this property unless there are a variable number of particles. To terms of order  $1/N$ , however, this decomposition is possible for a fixed number of particles; the errors introduced go to zero as the number of particles become infinite. (6)

The correlation energy,  $W_c$ , is the expectation value of  $H_{\text{red}}$  for the state  $\psi_0$

$$W_c = (\psi_0, H_{\text{red}}\psi_0) = W_c [h, \varphi].$$

Setting the variation of  $W_c$  with respect to  $h$  and  $\varphi$  equal to zero in order to minimize the energy gives

$$h = 1/2 (l - \epsilon/E)$$

$$E = (\epsilon^2 + |\Delta|^2)^{1/2}$$

where

$$\Delta = |\Delta| e^{i\varphi}$$

satisfies the integral equation

$$\Delta(\mathbf{k}) = -1/2 \sum_{\mathbf{k}'} V_{\mathbf{k}\mathbf{k}'} \frac{\Delta(\mathbf{k}')}{E(\mathbf{k}')},$$

If a non-zero solution of this integral equation exists,  $W_c < 0$  and the "normal" Fermi sea is unstable under the formation of correlated pairs.

In the wave function that results there are strong correlations between pairs of electrons with opposite spin and zero total momentum. These correlations are built from normal excitations near the Fermi surface and extend over spatial distances typically of the order of  $10^4$  cm. They can be constructed due to the large wave numbers available because of the exclusion principle. Thus with a small additional expenditure of kinetic energy there can be a greater gain in the potential energy term. Professor Schrieffer has discussed some of the properties of this state and the condensation energy associated with it.

#### SINGLE-PARTICLE EXCITATIONS

In considering the excited states of the superconductor it is useful, as for the

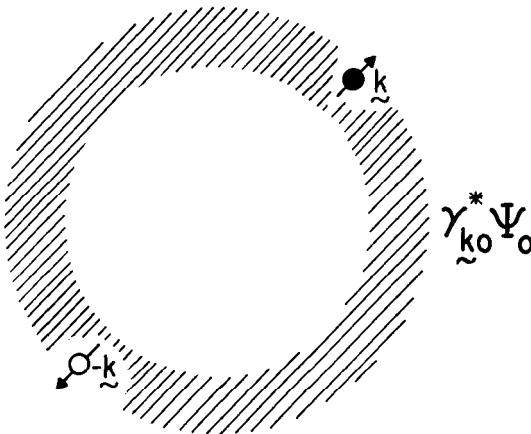


Fig. 6.

A single paricle excitation of the superconduc or in one-to-one correspondenct with an excitation of the normal fermion system.

normal metal, to make a distinction between single-particle and collective excitations; it is the single-particle excitation spectrum whose alteration is responsible for superfluid properties. For the superconductor excited (quasi-particle) states can be defined in one-to-one correspondence with the excitations of the normal metal. One finds, for example, that the expectation value of  $H_{\text{red}}$  for the excitation Fig. 6 is given by

$$E_k = \sqrt{\varepsilon_k^2 + |\mathcal{A}|^2}.$$

In contrast to the normal system, for the superconductor even as  $\varepsilon$  goes to zero  $E$  remains larger than zero, its lowest possible value being  $E = |\mathcal{A}|$ . One can therefore produce single particle excitations from the superconducting ground state only with the expenditure of a small but finite amount of energy. This is called the energy gap; its existence severely inhibits single particle processes and is in general responsible for the superfluid behavior of the electron gas. [In a gapless superconductor it is the finite value of  $A(r)$ , the order parameter, rather than the energy gap as such that becomes responsible for the superfluid properties.] In the ideal superconductor, the energy gap appears because not a single pair can be broken nor can a single element of phase space be removed without a finite expenditure of energy. If a single pair is broken, one loses its correlation energy; if one removes an element of phase space from the system, the number of possible transitions of all the pairs is reduced resulting in both cases in an increase in the energy which does not go to zero as the volume of the system increases.

The ground state of the superconductor and the excitation spectrum described above can conveniently be treated by introducing a linear combination of  $c^*$  and  $c$ , the creation and annihilation operators of normal fermions. This is the transformation of Bogoliubov and Valatin (5) :

$$\begin{aligned}\gamma_{k0}^* &= u_k c_{\mathbf{k}}^* - v_k c_{-\mathbf{k}}^* \\ \gamma_{kl}^* &= v_k c_{\mathbf{k}} + u_k c_{-\mathbf{k}}^*\end{aligned}$$

It follows that

$$\gamma_{ki} \psi_0 = 0$$

so that the  $\gamma_{ki}$  play the role of annihilation operators, while the  $\gamma_{ki}^*$  create excitations

$$\gamma_{k_1}^* \cdots \gamma_{m_j}^* \psi_0 = \psi_{k_1} \cdots m_j.$$

The  $\gamma$  operators satisfy Fermi anti-commutation relations so that with them we obtain a complete orthonormal set of excitations in one-to-one correspondence with the excitations of the normal metal.

We can sketch the following picture. In the ground state of the superconductor all the electrons are in singlet-pair correlated states of zero total momentum. In an  $m$  electron excited state the excited electrons are in "quasi-particle" states, very similar to the normal excitations and not strongly correlated with any of the other electrons. In the background, so to speak, the other electrons are still correlated much as they were in the ground state. The excited electrons behave in a manner similar to normal electrons; they can be easily scattered or excited further. But the background electrons-those which remain correlated-retain their special behavior; they are difficult to scatter or to excite.

Thus, one can identify two almost independent fluids. The correlated portion of the wave function shows the resistance to change and the very small specific heat characteristic of the superfluid, while the excitations behave very much like normal electrons, displaying an almost normal specific heat and resistance. When a steady electric field is applied to the metal, the superfluid electrons short out the normal ones, but with higher frequency fields the resistive properties of the excited electrons can be observed. [7]

#### Thermodynamic Properties, THE IDEAL SUPERCONDUCTOR

We can obtain the thermodynamic properties of the superconductor using the ground state and excitation spectrum just described. The free energy of the system is given by

$$F[h, \varphi, f] = W_c(T) - TS,$$

where  $T$  is the absolute temperature and  $S$  is the entropy;  $f$  is the superconducting Fermi function which gives the probability of single-particle excitations. The entropy of the system comes entirely from the excitations as the correlated portion of the wave function is non-degenerate. The free energy becomes a function of  $f(k)$  and  $h(k)$ , where  $f(k)$  is the probability that the state  $k$  is occupied by an excitation or a quasi-particle, and  $h(k)$  is the relative probability that the state  $k$  is occupied by a pair given that it is not occupied by a quasi-particle. Thus some states are occupied by quasi-particles and the unoccupied phase space is available for the formation of the coherent background of the remaining electrons. Since a portion of phase space is occupied by excitations at finite temperatures, making it unavailable for the transitions of bound pairs, the correlation energy is a function of the temperature,  $W_c(T)$ . As  $T$  increases,  $W_c(T)$  and at the same time  $\Delta$  decrease until the critical temperature is reached and the system reverts to the normal phase.

Since the excitations of the superconductor are independent and in a one-to-one correspondence with those of the normal metal, the entropy of an

excited configuration is given by an expression identical with that for the normal metal except that the Fermi function,  $f(k)$ , refers to quasi-particle excitations. The correlation energy at finite temperature is given by an expression similar to that at  $T = 0$  with the available phase space modified by the occupation functions  $f(k)$ . Setting the variation of  $F$  with respect to  $h$ ,  $\varphi$ , and  $f$  equal to zero gives:

$$h = 1/2 (1 - \varepsilon/E)$$

$$E = \sqrt{\varepsilon^2 + |\Delta|^2}$$

and

$$f = \frac{1}{1 + \exp(E/k_B T)}$$

where

$$\Delta = |\Delta| e^{i\varphi}$$

is now temperature-dependent and satisfies the fundamental integral equation of the theory

$$A_{\mathbf{k}}(T) = -1/2 \sum_{\mathbf{k}'\mathbf{k}''} V_{\mathbf{k}'\mathbf{k}''} \frac{A_{\mathbf{k}'}(T)}{E_{\mathbf{k}'}(T)} \tanh\left(\frac{E_{\mathbf{k}'}(T)}{2k_B T}\right).$$

The form of these equations is the same as that at  $T = 0$  except that the energy gap varies with the temperature. The equation for the energy gap can be satisfied with non-zero values of  $A$  only in a restricted temperature range. The upper bound of this temperature range is defined as  $T_c$ , the critical temperature. For  $T < T_c$ , singlet spin zero momentum electrons are strongly correlated, there is an energy gap associated with exciting electrons from the correlated part of the wave function and  $E(k)$  is bounded below by  $|\Delta|$ . In this region the system has properties qualitatively different from the normal metal.

In the region  $T > T_c$ ,  $A = 0$  and we have in every respect the normal solution. In particular  $f$ , the distribution function for excitations, becomes just the Fermi function for excited electrons  $k > k_F$ , and for holes  $k < k_F$

$$f = \frac{1}{1 + \exp(|\varepsilon|/k_B T)}.$$

If we make our simplifications of 1957, (defining in this way an ‘ideal’ superconductor)

$$V_{\mathbf{k}'\mathbf{k}} = -V \quad |\varepsilon| < \hbar\omega_{av}$$

$$= 0 \quad \text{otherwise}$$

and replace the energy dependent density of states by its value at the Fermi surface,  $N(0)$ , the integral equation for  $A$  becomes

$$1 = N(0) V \int_0^{\hbar\omega_{av}} \frac{d\varepsilon}{\sqrt{\varepsilon^2 + |\Delta|^2}} \tanh\left(\frac{\sqrt{\varepsilon^2 + |\Delta|^2}}{2k_B T}\right).$$

The solution of this equation, Fig. 7, gives  $A(T)$  and with this  $\mathbf{f}$  and  $h$ . We can then calculate the free energy of the superconducting state and obtain the thermodynamic properties of the system.

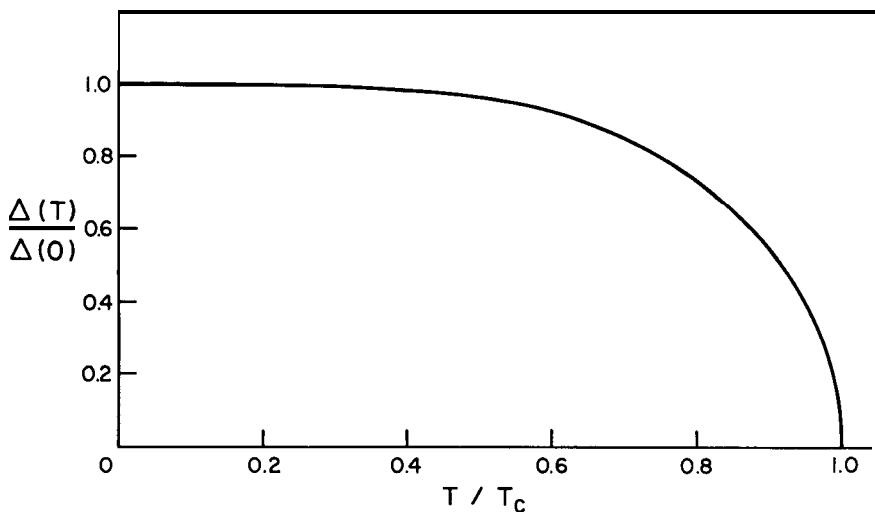


Fig. 7.

Variation of the energy gap with temperature for the ideal superconductor.

In particular one finds that at  $T_c$  (in the absence of a magnetic field) there is a second-order transition (no latent heat :  $W_c = 0$  at  $T$ ) and a discontinuity in the specific heat. At very low temperatures the specific heat goes to zero exponentially. For this ideal superconductor one also obtains a law of corresponding states in which the ratio

$$\frac{\gamma T_c^2}{H_0^2} = 0.170,$$

where

$$\gamma = 2/3\pi^2 N(0) k_B^2.$$

The experimental data scatter about the number 0.170. The ratio of  $A$  to  $k_B T_c$  is given as a universal constant

$$A/k_B T_c = 1.75.$$

There are no arbitrary parameters in the idealized theory. In the region of empirical interest all thermodynamic properties are determined by the quantities  $\gamma$  and  $\hbar\omega_{av} e^{-1/N(0)V}$ . The first,  $\gamma$ , is found by observation of the normal specific heat, while the second is found from the critical temperature, given by

$$k_B T_c = 1.14 \hbar\omega_{av} e^{-1/N(0)V}.$$

At the absolute zero

$$A = \hbar\omega_{av} / \sinh \left( \frac{1}{N(0)V} \right).$$

Further, defining a weak coupling limit [ $N(0) V \ll 1$ ] which is one region of interest empirically, we obtain

$$A \simeq 2\hbar\omega_{av} e^{-1/N(0)V}.$$

The energy difference between the normal and superconducting states becomes (again in the weak coupling limit)

$$W_s - W_n = W_c = -2N(0)(\hbar\omega_{av})^2 e^{-2/N(0)V}.$$

The dependence of the correlation energy on  $(\hbar\omega_{\text{av}})^2$  gives the isotope effect, while the exponential factor reduces the correlation energy from the dimensionally expected  $N(0)(\hbar\omega_{\text{av}})^2$  to the much smaller observed value. This, however, is more a demonstration that the isotope effect is consistent with our model rather than a consequence of it, as will be discussed further by Professor Bardeen.

The thermodynamic properties calculated for the ideal superconductor are in qualitative agreement with experiment for weakly coupled superconductors. Very detailed comparison between experiment and theory has been made by many authors. A summary of the recent status may be found in reference (2). When one considers that in the theory of the ideal superconductor the existence of an actual metal is no more than hinted at (We have in fact done all the calculations considering weakly interacting fermions in a container.) so that in principle (with appropriate modifications) the calculations apply to neutron stars as well as metals, we must regard detailed quantitative agreement as a gift from above. We should be content if there is a single metal for which such agreement exists. [Pure single crystals of tin or vanadium are possible candidates.]

To make comparison between theory and experiments on actual metals, a plethora of detailed considerations must be made. Professor Bardeen will discuss developments in the theory of the electron-phonon interaction and the resulting dependence of the electron-electron interaction and superconducting properties on the phonon spectrum and the range of the Coulomb repulsion. Crystal symmetry, Brillouin zone structure and the actual wave function (*S*, *P* or *D* states) of the conduction electrons all play a role in determining real metal behavior. There is a fundamental distinction between superconductors which always show a Meissner effect and those (type II) which allow magnetic field penetration in units of the flux quantum.

When one considers, in addition, specimens with impurities (magnetic and otherwise) superimposed films, small samples, and so on, one obtains a variety of situations, developed in the years since 1957 by many authors, whose richness and detail takes volumes to discuss. The theory of the ideal superconductor has so far allowed the addition of those extensions and modifications necessary to describe, in what must be considered remarkable detail, all of the experience actually encountered.

#### MICROSCOPIC INTERFERENCE EFFECTS

In its interaction with external perturbations the superconductor displays remarkable interference effects which result from the paired nature of the wave function and are not at all present in similar normal metal interactions. Neither would they be present in any ordinary two-fluid model. These "coherence effects" are in a sense manifestations of interference in spin and momentum space on a microscopic scale, analogous to the macroscopic quantum effects due to interference in ordinary space which Professor Schrieffer discussed. They depend on the behavior under time reversal of the perturbing fields. (8) It is intriguing to speculate that if one could somehow amplify them

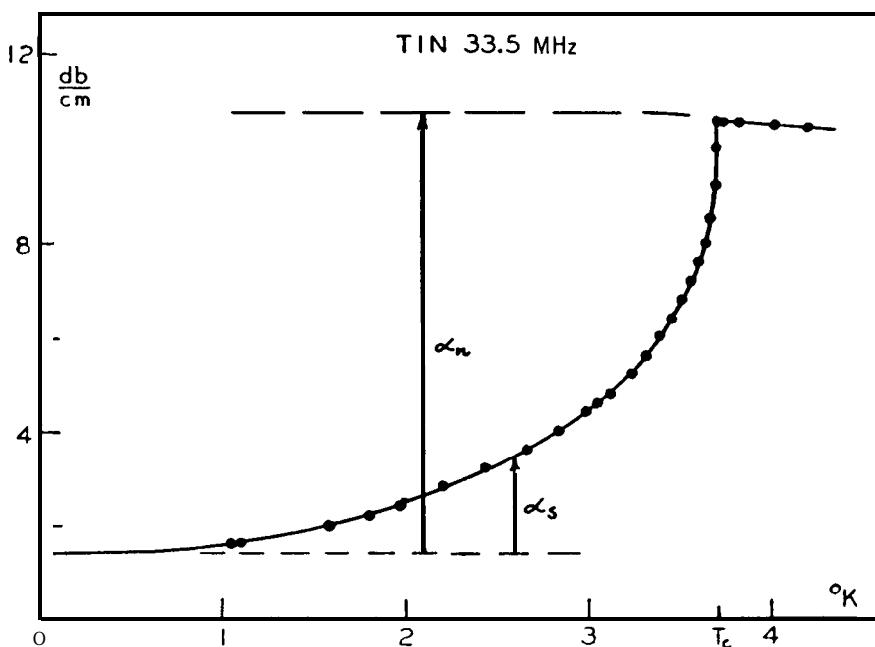


Fig. 8.

Ultrasonic attenuation as a function of temperature across the superconducting transition as measured by Morse and Bohm.

properly, the time reversal symmetry of a fundamental interaction might be tested. Further, if helium 3 does in fact display a phase transition analogous to the superconducting transition in metals as may be indicated by recent experiments (4) and this is a spin triplet state, the coherence effects would be greatly altered.

Near the transition temperature these coherence effects produce quite dramatic contrasts in the behavior of coefficients which measure interactions with the conduction electrons. Historically, the comparison with theory of the behavior of the relaxation rate of nuclear spins (9) and the attenuation of longitudinal ultrasonic waves in clean samples (10) as the temperature is decreased through  $T_c$  provided an early test of the detailed structure of the theory.

The attenuation of longitudinal acoustic waves due to their interaction with the conduction electrons in a metal undergoes a very rapid drop (10a) as the temperature drops below  $T_c$ . Since the scattering of phonons from "normal" electrons is responsible for most of the acoustic attenuation, a drop was to be expected; but the rapidity of the decrease measured by Morse and Bohm (10b) Fig. 8 was difficult to reconcile with estimates of the decrease in the normal electron component of a two-fluid model.

The rate of relaxation of nuclear spins was measured by Hebel and Slichter (9a) in zero magnetic field in superconducting aluminum from 0.94 K to 4.2 K just at the time of the development of our 1957 theory. Redfield and Anderson (9b) confirmed and extended their results. The dominant relaxation mechanism is provided by interaction with the conduction electrons so that one would expect, on the basis of a two-fluid model, that this rate should

decrease below the transition temperature due to the diminishing density of "normal" electrons. The experimental results however show just the reverse. The relaxation rate does not drop but increases by a factor of more than two just below the transition temperature. Fig 13. This observed increase in the nuclear spin relaxation rate and the very sharp drop in the acoustic attenuation coefficient as the temperature is decreased through  $T_c$  impose contradictory requirements on a conventional two-fluid model.

To illustrate how such effects come about in our theory, we consider the transition probability per unit time of a process involving electronic transitions from the excited state  $k$  to the state  $k'$  with the emission to or absorption of energy from the interacting field. What is to be calculated is the rate of transition between an initial state  $|i\rangle$  and a final state  $|f\rangle$  with the absorption or emission of the energy  $\hbar\omega_{|k'-k|}$  (a phonon for example in the interaction of sound waves with the superconductor). All of this properly summed over final states and averaged with statistical factors over initial states may be written:

$$\omega = \frac{2\pi}{\hbar} \frac{\sum_{i,f} \exp(-W_i/k_B T) |< f | H_{\text{int}} | i >|^2 \delta(W_f - W_i)}{\sum_i \exp(-W_i/k_B T)}$$

We focus our attention on the matrix element  $< f | H_{\text{int}} | i >$ . This typically contains as one of its factors matrix elements between excited states of the superconductor of the operator

$$B = \sum_{K' K} B_{K' K} c_{K'}^* c_K$$

where  $c_{K'}^*$  and  $c_K$  are the creation and annihilation operators for electrons in the states  $K'$  and  $K$ , and  $B_{K' K}$  is the matrix element between the states  $K'$  and  $K$  of the configuration space operator  $B(r)$

$$B_{K' K} = < K' | B(r) | K >.$$

The operator  $B$  is the electronic part of the matrix element between the full final and initial state

$$< f | H_{\text{int}} | i > = m_{fi} < f | B | i >.$$

In the normal system scattering from single-particle electron states  $K$  to  $K'$  is independent of scattering from  $-K'$  to  $-K$ . But the superconducting states are linear superpositions of  $(K, -K)$  occupied and unoccupied. Because of this states with excitations  $k\uparrow$  and  $k'\uparrow$  are connected not only by  $c_{k'\uparrow}^* c_{k\uparrow}$  but also by  $c_{-k\downarrow}^* c_{-k'\downarrow}$ ; if the state  $|f\rangle$  contains the single-particle excitation  $k'\uparrow$  while the state  $|i\rangle$  contains  $k\uparrow$ , as a result of the superposition of occupied and unoccupied pair states in the coherent part of the wave function, these are connected not only by  $B_{K' K} c_{K'}^* c_K$  but also by  $B_{-K-K'} c_{-K}^* c_{-K'}$ .

For operators which do not flip spins we therefore write:

$$B = \sum_{K' K} (B_{K' K} c_{K'}^* c_K + B_{-K-K'} c_{-K}^* c_{-K'}).$$

Many of the operators,  $B$ , we encounter (e.g., the electric current, or the charge density operator) have a well-defined behavior under the operation of time reversal so that

$$B_{K' K} = \pm B_{-K-K'} \equiv B_{K' K}$$

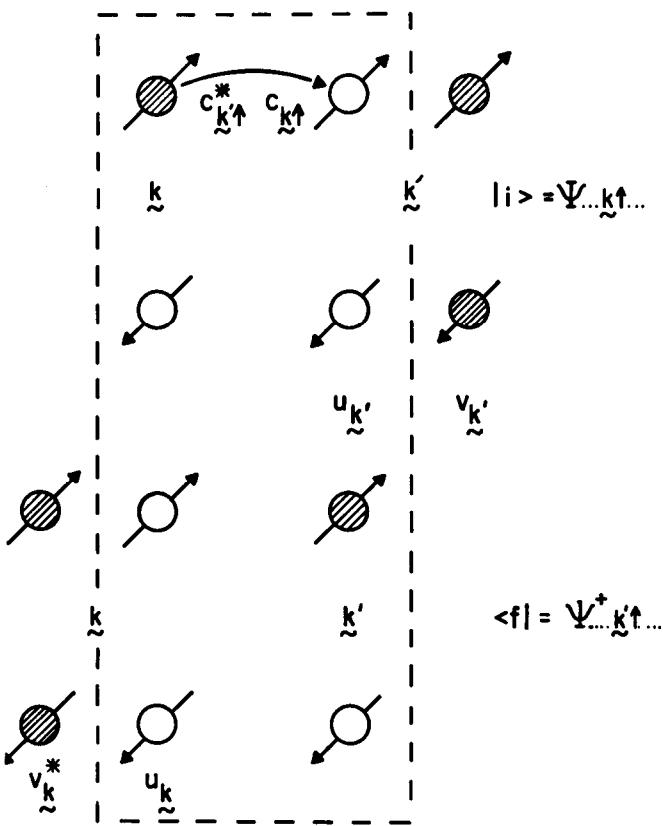


Fig. 9.

The two states  $|i\rangle$  and  $\langle f|$  shown are connected by  $C_i^* C_{ik}$  with the amplitude  $U_{ik'}$ .

Then  $B$  becomes

$$B = \sum_{\mathbf{k}\mathbf{k}'} B_{\mathbf{k}\mathbf{k}'} (c_{\mathbf{k}'\uparrow}^* c_{\mathbf{k}\uparrow} \pm c_{-\mathbf{k}\downarrow}^* c_{-\mathbf{k}'\downarrow})$$

where the upper (lower) sign results for operators even (odd) under time reversal.

The matrix element of  $B$  between the initial state,  $y \dots k \dots$ , and the final state  $y \dots k \dots$  contains contributions from  $c_{ik}^* c_{ik}$  Fig. 9 and unexpectedly from  $c_{ik}^* c_{ik}$  Fig. 10. As a result the matrix element squared  $1 \langle f | B | i \rangle$  contains terms of the form

$$|B_{\mathbf{k}\mathbf{k}'}|^2 |(u_{\mathbf{k}'} u_{\mathbf{k}} \mp v_{\mathbf{k}'} v_{\mathbf{k}}^*)|^2,$$

where the sign is determined by the behavior of  $B$  under time reversal:

upper sign       $B$  even under time reversal

lower sign       $B$  odd under time reversal.

Applied to processes involving the emission or absorption of boson quanta such as phonons or photons, the squared matrix element above is averaged with the appropriate statistical factors over initial and summed over final states; subtracting emission from absorption probability per unit time, we obtain typically

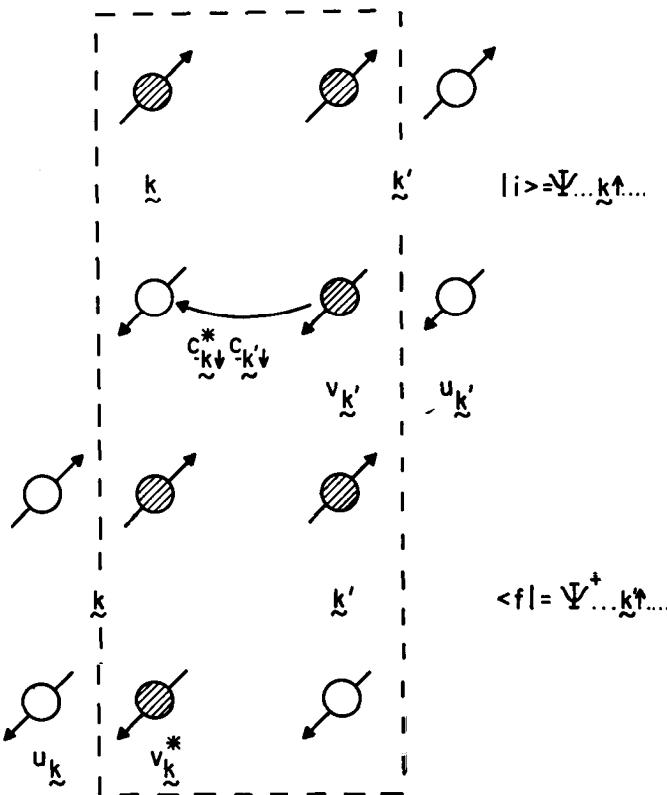


Fig. 10.  
The two states  $|i\rangle$  and  $\langle f|$  are also connected by  $C^* k J c - k' J$  with the amplitude  $-UWU^-$ .

$$\alpha = \frac{4\pi}{\hbar} |m|^2 \sum_{kk'} |(u_k u_k \mp v_k v_k^*)|^2 (f_{k'} - f_k) \delta(E_{k'} - E_k - \hbar\omega_{(k-k')})$$

where  $f_i$  is the occupation probability in the superconductor for the excitation  $kT$  or  $kJ$ . [In the expression above we have considered only quasiparticle or quasi-hole scattering processes (not including processes in which a pair of excitations is created or annihilated from the coherent part of the wave function) since  $Au, \sim, \sim, < A$ , is the usual region of interest for the ultrasonic attenuation and nuclear spin relaxation we shall contrast.]

For the ideal superconductor, there is isotropy around the Fermi surface and symmetry between particles and holes; therefore sums of the form  $\sum_k Z$  can be converted to integrals over the superconducting excitation energy,  $E$ :

$$\sum_k \rightarrow 2N(0) \int_A^\infty \frac{E}{\sqrt{E^2 - \Delta^2}} dE$$

where  $N(\theta) = \frac{E}{\pi} \text{erfc}(\frac{\Delta}{E})$  is the density of excitations in the superconductor, Fig. 11.

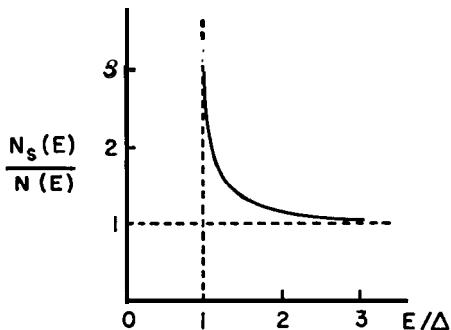


Fig. 11.

Ratio of superconducting to normal density of excitations as a function of  $E/\Delta$ .

The appearance of this density of excitations is a surprise. Contrary to our intuitive expectations, the onset of superconductivity seems initially to enhance rather than diminish electronic transitions, as might be anticipated in a reasonable two-fluid model.

But the coherence factors  $|(\langle u' u \mp v' v^* \rangle)|^2$  are even more surprising; they behave in such a way as to sometimes completely negate the effect of the increased density of states. This can be seen using the expressions obtained above for  $u$  and  $v$  for the ideal superconductor to obtain

$$(u' u \mp v' v)^2 = \frac{1}{2} \left( 1 + \frac{\varepsilon \varepsilon' \mp \Delta^2}{E E'} \right).$$

In the integration over  $k$  and  $k'$  the  $\varepsilon \varepsilon'$  term vanishes. We thus define  $(\langle u' u \mp v' v \rangle_s)^2$ ; in usual limit where  $\hbar \omega_{|k-k'|} \ll \Delta$ ,  $\varepsilon \simeq \varepsilon'$  and  $E \simeq E'$ , this becomes

$$(\langle u^2 - v^2 \rangle_s)^2 \rightarrow \frac{1}{2} \left( \frac{\varepsilon^2}{E^2} \right) \quad \text{operators even under time reversal}$$

$$(\langle u^2 + v^2 \rangle_s)^2 \rightarrow \frac{1}{2} \left( 1 + \frac{E^2}{\Delta^2} \right) \quad \text{operators odd under time reversal.}$$

For operators even under time reversal, therefore, the decrease of the coherence factors near  $\varepsilon = 0$  just cancels the increase due to the density of states. For the operators odd under time reversal the effect of the increase of the density of states is not cancelled and should be observed as an increase in the rate of the corresponding process.

In general the interaction Hamiltonian for a field interacting with the superconductor (being basically an electromagnetic interaction) is invariant under the operation of time reversal. However, the operator  $B$  might be the electric current  $j(r)$  (for electromagnetic interactions) the electric charge density  $\varrho(r)$  (for the electron-phonon interaction) or the  $z$  component of the electron spin operator,  $\sigma_z$  (for the nuclear spin relaxation interaction). Since under time-reversal

$$j(r, t) \rightarrow -j(r, -t) \quad (\text{electromagnetic interaction})$$

$$\varrho(r, t) \rightarrow +\varrho(r, -t) \quad (\text{electron-phonon interaction})$$

$$\sigma_z(t) \rightarrow -\sigma_z(-t) \quad (\text{nuclear spin relaxation interaction})$$

these show strikingly different interference effects.

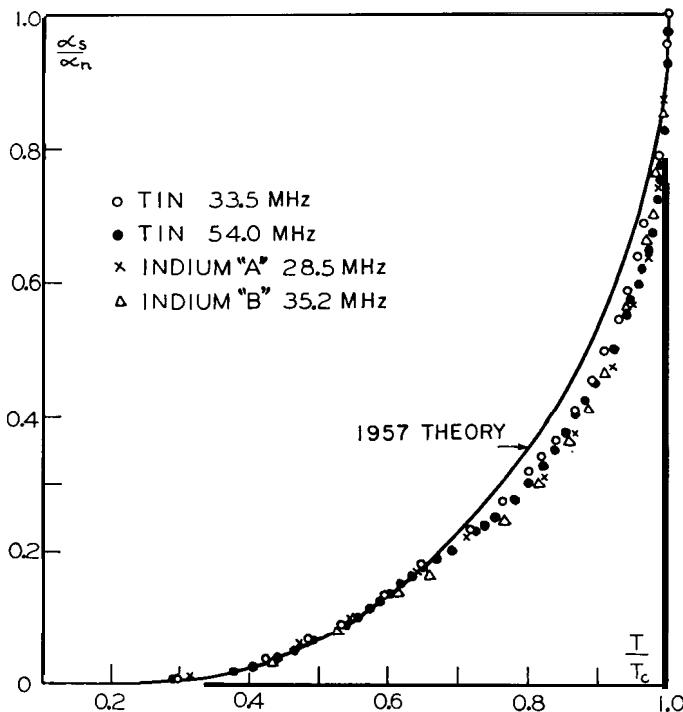


Fig. 12.  
Comparison of observed ultrasonic attenuation with the ideal theory. The data are due to Morse and Bohm.

Ultrasonic attenuation in the ideal pure superconductor for  $qI \gg 1$  (the product of the phonon wave number and the electron mean free path) depends in a fundamental way on the absorption and emission of phonons. Since the matrix elements have a very weak dependence on changes near the Fermi surface in occupation of states other than  $k$  or  $k'$  that occur in the normal to superconducting transition, calculations within the quasi-particle model can be compared in a very direct manner with similar calculations for the normal metal, as  $B_{kk'}$  is the same in both states. The ratio of the attenuation in the normal and superconducting states becomes :

$$\frac{a_s}{a_n} = -4 \int_A^\infty dE (u^2 - v^2)_s^2 \left( \frac{E}{\varepsilon} \right)^2 \frac{df(E)}{dE}.$$

Since  $(u^2 - v^2)_s^2 \rightarrow \frac{1}{2} \left( \frac{\varepsilon}{E} \right)^2$ , the coherence factors cancel the density of states giving

$$\frac{a_s}{a_n} = 2f(\Delta(T)) = \frac{2}{1 + \exp\left(\frac{\Delta(T)}{k_B T}\right)}.$$

Morse and Bohm (10b) used this result to obtain a direct experimental determination of the variation of  $A$  with  $T$ . Comparison of their attenuation data with the theoretical curve is shown in Figure 12.

In contrast the relaxation of nuclear spins which have been aligned in a magnetic field proceeds through their interaction with the magnetic moment of the conduction electrons. In an isotropic superconductor this can be shown to depend upon the z component of the electron spin operator

$$B_{\mathbf{K}'\mathbf{K}} = B(c_{\mathbf{k}'\uparrow}^* c_{\mathbf{k}\uparrow} - c_{-\mathbf{k}\downarrow}^* c_{-\mathbf{k}'\downarrow})$$

so that

$$B_{\mathbf{K}'\mathbf{K}} = -B_{-\mathbf{K}-\mathbf{K}''}$$

This follows in general from the property of the spin operator under time reversal

$$\sigma_z(t) = -\sigma_z(-t).$$

The calculation of the nuclear spin relaxation rate proceeds in a manner not too different from that for ultrasonic attenuation resulting finally in a ratio of nuclear spin relaxation rates in superconducting and normal states in the same sample:

$$\frac{R_s}{R_n} = -4 \int_A^\infty dE (u^2 + v^2)_s^2 \left(\frac{E}{\varepsilon}\right)^2 \frac{df(E)}{dE}.$$

But  $(u^2 + v^2)_s^2$  does not go to zero at the lower limit so that the full effect of the increase in density of states at  $E = A$  is felt. Taken literally, in fact, this expression diverges logarithmically at the lower limit due to the infinite density of states. When the Zeeman energy difference between the spin up and spin down states is included, the integral is no longer divergent but the integrand is much too large. Hebel and Slichter, by putting in a broadening of levels phenomenologically, could produce agreement between theory and experiment. More recently Fibich (11) by including the effect of thermal phonons has obtained the agreement between theory and experiment shown in Fig. 13.

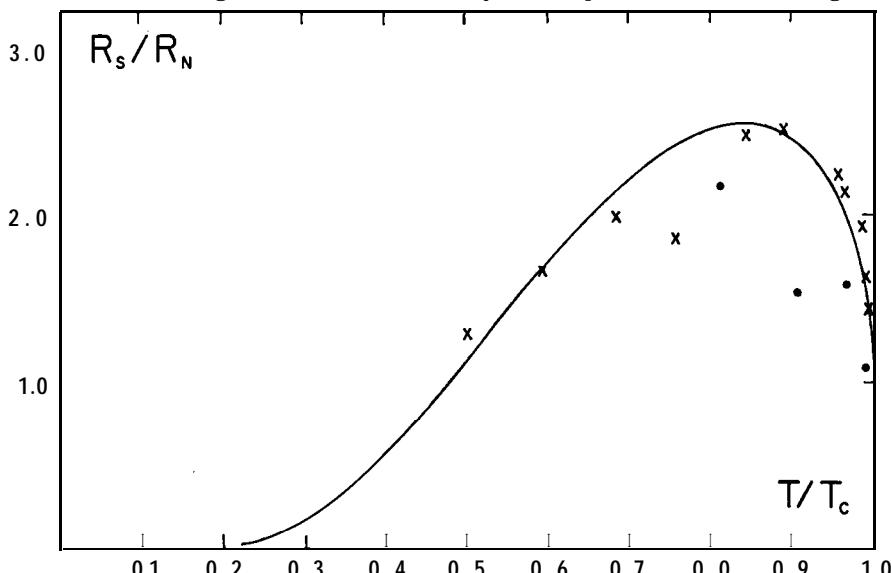


Fig. 13.

Comparison of observed nuclear spin relaxation rate with theory. The circles represent experimental data of Hebel and Slichter, the crosses data by Redfield and Anderson.

Interference effects manifest themselves in a similar manner in the interaction of electromagnetic radiation with the superconductor. Near  $T_c$  the absorption is dominated by quasi-particle scattering matrix elements of the type we have described. Near  $T = 0$ , the number of quasi-particle excitations goes to zero and the matrix elements that contribute are those in which quasi-particle pairs are created from  $\psi_0$ . For absorption these latter occur only when  $\hbar\omega > 2\Delta$ . For the linear response of the superconductor to a static magnetic field, the interference occurs in such a manner that the paramagnetic contribution goes to zero leaving the diamagnetic part which gives the Meissner effect.

The theory developed in 1957 and applied to the equilibrium properties of uniform materials in the weak coupling region has been extended in numerous directions by many authors. Professor Schrieffer has spoken of Josephson junctions and macroscopic quantum interference effects; Professor Bardeen will discuss the modifications of the theory when the electron-phonon interactions are strong. The treatment of ultrasonic attenuation, generalized to include situations in uniform superconductors in which  $qL < 1$ , gives a surprisingly similar result to that above. (12) There have been extensive developments using Green's function methods (13) appropriate for type II superconductors, materials with magnetic impurities and non-uniform materials or boundary regions where the order parameter is a function of the spatial coordinates. (14) With these methods formal problems of gauge invariance and/or current conservation have been resolved in a very elegant manner. (15) In addition, many calculations (16) of great complexity and detail for type II superconductors have treated ultrasonic attenuation, nuclear spin relaxation and other phenomena in the clean and dirty limits (few or large numbers of impurities). The results cited above are modified in various ways. For example, the average density of excitation levels is less sharply peaked at  $T_c$  in a type II superconductor; the coherence effects also change somewhat in these altered circumstances but nevertheless play an important role. Overall one can say that the theory has been amenable to these generalizations and that agreement with experiment is good.

It is now believed that the finite many-nucleon system that is the atomic nucleus enters a correlated state analogous to that of a superconductor. (17) Similar considerations have been applied to many-fermion systems as diverse as neutron stars, (18) liquid He<sup>3</sup>, (19) and to elementary fermions. (20) In addition the idea of spontaneously broken symmetry of a degenerate vacuum has been applied widely in elementary particle theory and recently in the theory of weak interactions. (21) What the electron-phonon interaction has produced between electrons in metals may be produced by the van der Waals interaction between atoms in He<sup>3</sup>, the nuclear interaction in nuclei and neutron stars, and the fundamental interactions in elementary fermions. Whatever the success of these attempts, for the theoretician the possible existence of this correlated paired state must in the future be considered for any degenerate many-fermion system where there is some kind of effective attraction between fermions for transitions near the Fermi surface.

In the past few weeks my colleagues and I have been asked many times: "What are the practical uses of your theory?" Although even a summary inspection of the proceedings of conferences on superconductivity and its applications would give an immediate sense of the experimental, theoretical and developmental work in this field as well as expectations, hopes and anticipations -from applications in heavy electrical machinery to measuring devices of extraordinary sensitivity and new elements with very rapid switching speeds for computers - I, personally, feel somewhat uneasy responding. The discovery of the phenomena and the development of the theory is a vast work to which many scientists have contributed. In addition there are numerous practical uses of the phenomena for which theory rightly should not take credit. A theory (though it may guide us in reaching them) does not produce the treasures the world holds. And the treasures themselves occasionally dazzle our attention; for we are not so wealthy that we may regard them as irrelevant.

But a theory is more. It is an ordering of experience that both makes experience meaningful and is a pleasure to regard in its own right. Henri Poincaré wrote (22) :

Le savant doit ordonner; on fait la science  
avec des faits comme une maison avec des  
pierres; mais une accumulation de faits  
n'est pas plus une science qu'un tas de  
pierres n'est une maison.

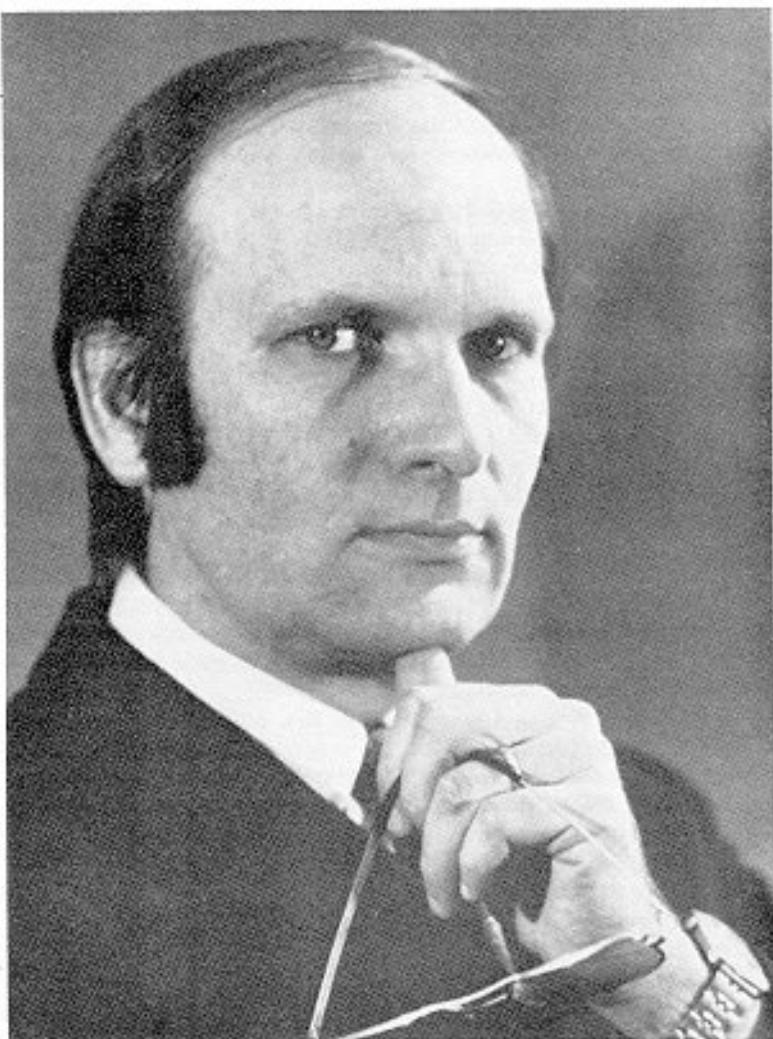
One can build from ordinary stone a humble house or the finest chateau. Either is constructed to enclose a space, to keep out the rain and the cold. They differ in the ambition and resources of their builder and the art by which he has achieved his end. A theory, built of ordinary materials, also may serve many a humble function. But when we enter and regard the relations in the space of ideas, we see columns of remarkable height and arches of daring breadth. They vault the fine structure constant, from the magnetic moment of the electron to the behavior of metallic junctions near the absolute zero; they span the distance from materials at the lowest temperatures to those in the interior of stars, from the properties of operators under time reversal to the behavior of attenuation coefficients just beyond the transition temperature.

I believe that I speak for my colleagues in theoretical science as well as myself when I say that our ultimate, our warmest pleasure in the midst of one of these incredible structures comes with the realization that what we have made is not only useful but is indeed a beautiful way to enclose a space.

#### REFERENCES AND NOTES

1. Bardeen, J., Cooper, L. N. and Schrieffer, J. R., Phys. Rev. **108**, 1175 (1957).
2. An account of the situation as of 1969 may be found in the two volumes: Superconductivity, edited by R. D. Parks, Marcel Dekker, Inc., New York City (1969).
3. Sommerfeld, A., Z. Physik 47, 1 (1928). Bloch, F., Z. Physik 52 555 (1928).
4. Osheroff, D. D., Gully, W. J., Richardson R. C. and Lee, D. M., Phys. Rev. Letters 29, 920 (1972); Leggett, A. J., Phys. Rev. Letters 29, 1227 (1972).

5. Bogoliubov, N. N., Nuovo Cimento 7, 794 (1958) ; Usp. Fiz. Nauk 67, 549 (1959) ; Valatin, J. G., Nuovo Cimento 7, 843 (1958).
6. Bardeen, J. and Rickayzen, G., Phys. Rev. 118, 936 (1960); Mattis, D. C. and Lieb, E., J. Math. Phys. 2, 602 (1961); Bogoliubov, N. N., Zubarev D. N. and Tserkovnikov, Yu. A., Zh. Eksperim. i Teor. Fiz. 39, 120 (1960) translated: Soviet Phys. JETP 12, 88 (1961).
7. For example, Glover, R. E. III and Tinkham, M., Phys. Rev. 108, 243 (1957) ; Biondi, M. A. and Garfunkel, M. P., Phys. Rev. 116, 853 (1959).
8. The importance of the coupling of time reversed states in constructing electron pairs was emphasized by P. W. Anderson; for example, Anderson, P. W., J. Phys. Chem. Solids II, 26 (1959)
9. a) Hebel, L. C. and Slichter, C. P., Phys. Rev. **113**, 1504 (1959).  
b) Redfield, A. G. and Anderson, A. G., Phys. Rev. **116**, **583** (1959).
10. a) Bommel, H. E., Phys. Rev. 96, 220 (1954).  
b) Morse, R. W. and Bohm, H. V., Phys. Rev. 108 1094 (1957).
11. Fibich, M., Phys. Rev. Letters **14**, 561 (1965).
12. For example, Tsuneto T., Phys. Rev. **121**, **402** (1961).
13. Gor'kov, L. P., Zh. Eksperim. i Teor. Fiz. 34, 735 (1958) translated: Soviet Physics JETP 7, 505 (1958) ; also Martin P. C. and Schwinger J., Phys. Rev. **115**, 1342 (1959) ; Kadanoff L. P. and Martin P. C., Phys. Rev. 124, 670 (1961).
14. Abrikosov, A. A. and Gor'kov, L. P., Zh. Eksperim. i Teor. Fix. 39, 1781 (1960) translated: Soviet Physics JETP 12, 1243 (1961) ; de Gennes P. G., Superconductivity of Metals and Alloys, Benjamin, New York (1966).
15. For example, Ambegaokar V. and Kadanoff L. P., Nuovo Cimento 22, 914 (1961).
16. For example, Caroli C. and Matricon J., Physik Kondensierten Materie 3, 380 (1965) ; Maki K., Phys. Rev. 141, 331 (1966); 1.56, 437 (1967); Groupe de Supraconductivite d'Orsay, Physik Kondensierten Materie 5, 141 (1966); Eppel D., Pesch W. and Teuwoldt L., Z. Physik 197; 46 (1966) ; McLean F. B. and Houghton A., Annals of Physics **48**, **43** (1968).
17. Bohr, A., Mottelson, B. R. and Pines, D., Phys. Rev. 110, 936 (1958) ; Migdal, A. B., Nuclear Phys. **13**, **655** (1959).
18. Ginzburg, V. L. and Kirzhnits, D. A., Zh. Eksperim. i Theor. Fiz. 47, 2006 (1964) translated: Soviet Physics JETP 20, 1346 (1965) ; Pines D., Baym G. and Pethick C., Nature 224, 673 (1969).
19. Many authors have explored the possibility of a superconducting-like transition in  $\text{He}^3$ . Among the most recent contributions see reference 4.
20. For example. Nambu Y. and Jona-Lasinio G., Phys. Rev. 122, 345 (1961).
21. Goldstone, J., Nuovo Cimento 19, 154 (1961) ; Weinberg, S., Phys. Rev. Letters 19, 1264 (1967).
22. Poincaré, H., La Science et l'Hypothèse, Flammarion, Paris, pg. 168 (1902). "The scientist must order; science is made with facts as a house with stones; but an accumulation of facts is no more a science than a heap of stones is a house."



J. Robert Schieffer

## J. ROBERT SCHRIEFFER

J. Robert Schrieffer was born in Oak Park, Illinois on May 31, 1931, son of John H. Schrieffer and his wife Louis (nee Anderson). In 1940, the family moved to Manhasset, New York and in 1947 to Eustis, Florida where they became active in the citrus industry.

Following his graduation from Eustis High School in 1949, Schrieffer was admitted to Massachusetts Institute of Technology, where for two years he majored in electrical engineering, then changed to physics in his junior year. He completed a bachelor's thesis on the multiple structure in heavy atoms under the direction of Professor John C. Slater. Following up on an interest in solid state physics developed while at MIT, he began graduate studies at the University of Illinois, where he immediately began research with Professor John Bardeen. After working out a problem dealing with electrical conduction on semiconductor surfaces, Schrieffer spent a year in the laboratory, applying the theory to several surface problems. In the third year of graduate studies, he joined Bardeen and Cooper in developing the theory of superconductivity, which constituted his doctoral dissertation.

He spent the academic year 1957-58 as a National Science Foundation fellow at the University of Birmingham and the Niels Bohr Institute in Copenhagen, where he continued research in superconductivity. Following a year as assistant professor at the University of Chicago, he returned to the University of Illinois in 1959 as a faculty member. In 1960 he returned to the Bohr Institute for a summer visit, during which he became engaged to Anne Crete Thomsen whom he married at Christmas of that year.

In 1962 Schrieffer joined the faculty of the University of Pennsylvania in Philadelphia, where in 1964 he was appointed Mary Amanda Wood Professor in Physics. In 1980 he was appointed Professor at the University of California, Santa Barbara and to the position of Chancellor Professor in 1984. He served as Director of the Institute for Theoretical Physics in Santa Barbara from 1984-89. In 1992 he was appointed University Professor at Florida State University and Chief Scientist of the National High Magnetic Field Laboratory.

He holds honorary degrees from the Technische Hochschule, Munich and the Universities of Geneva, Pennsylvania, Illinois, Cincinnati, Tel-Aviv, Alabama. In 1969 he was appointed by Cornell to a six-year term as a Andrew D. White Professor-at-Large.

He is a member of the American Academy of Arts and Sciences, the National Academy of Sciences of which he is a member of their council, the

American Philosophical Society, the Royal Danish Academy of Sciences and Letters and the Academy of Sciences of the USSR.

His awards include the Guggenheim Fellowship, Oliver E. Buckley Solid State Physics Prize, Comstock Prize, National Academy of Science, the Nobel Prize in Physics shared with John Bardeen and Leon N. Cooper in 1972, John Ericsson Medal, American Society of Swedish Engineers, University of Illinois Alumni Achievement Award, and in 1984 the National Medal of Science. The main thrust of his recent work has been in the area of high-temperature superconductivity, strongly correlated electrons, and the dynamics of electrons in strong magnetic fields.

The Schrieffers have three children, Bolette, Paul, and Regina.

# MACROSCOPIC QUANTUM PHENOMENA FROM PAIRING IN SUPERCONDUCTORS

Nobel Lecture, December 11, 1972

by

J. R. SCHRIEFFER

University of Pennsylvania, Philadelphia, Pa.

## I. INTRODUCTION

It gives me great pleasure to have the opportunity to join my colleagues John Bardeen and Leon Cooper in discussing with you the theory of superconductivity. Since the discovery of superconductivity by H. Kamerlingh Onnes in 1911, an enormous effort has been devoted by a spectrum of outstanding scientists to understanding this phenomenon. As in most developments in our branch of science, the accomplishments honored by this Nobel prize were made possible by a large number of developments preceding them. A general understanding of these developments is important as a backdrop for our own contribution.

On December 11, 1913, Kamerlingh Onnes discussed in his Nobel lecture (1) his striking discovery that on cooling mercury to near the absolute zero of temperature, the electrical resistance became vanishingly small, but this disappearance "did not take place gradually but abruptly." His Fig. 17 is reproduced as Fig. 1. He said, "Thus, mercury at 4.2 K has entered a new state

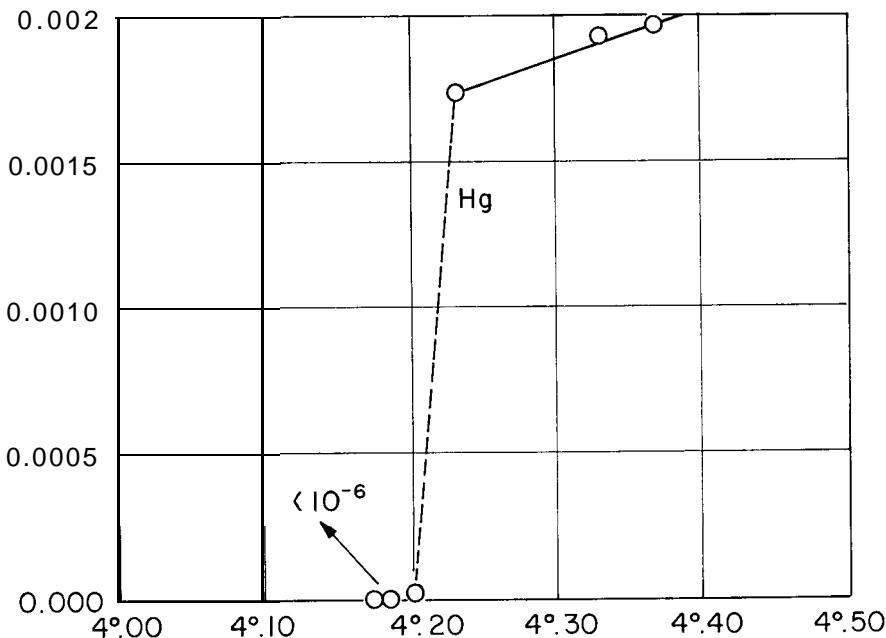


Fig. 1

which owing to its particular electrical properties can be called the state of superconductivity." He found this state could be destroyed by applying a sufficiently strong magnetic field, now called the critical field  $H_c$ . In April-June, 1914, Onnes discovered that a current, once induced in a closed loop of superconducting wire, persists for long periods without decay, as he later graphically demonstrated by carrying a loop of superconducting wire containing a persistent current from Leiden to Cambridge.

In 1933, W. Meissner and R. Ochsenfeld (2) discovered that a superconductor is a perfect diamagnet as well as a perfect conductor. The magnetic field vanishes in the interior of a bulk specimen, even when cooled down below the transition temperature in the presence of a magnetic field. The diamagnetic currents which flow in a thin penetration layer near the surface of a simply connected body to shield the interior from an externally applied field are stable rather than metastable. On the other hand, persistent currents flowing in a multiply connected body, e.g., a loop, are metastable.

An important advance in the understanding of superconductivity occurred in 1934, when C. J. Gorter and H. B. G. Casimir (3) advanced a two fluid model to account for the observed second order phase transition at  $T_c$  and other thermodynamic properties. They proposed that the total density of electrons  $\varrho$  could be divided into two components

$$\varrho = \varrho_s + \varrho_n \quad (1)$$

where a fraction  $\varrho_s/\varrho_n$  of the electrons can be viewed as being condensed into a "superfluid," which is primarily responsible for the remarkable properties of superconductors, while the remaining electrons form an interpenetrating fluid of "normal" electrons. The fraction  $\varrho_s/\varrho_n$  grows steadily from zero at  $T_c$  to unity at  $T = 0$ , where "all of the electrons" are in the superfluid condensate.

A second important theoretical advance came in the following year, when Fritz and Hans London set down their phenomenological theory of the electromagnetic properties of superconductors, in which the diamagnetic rather than electric aspects are assumed to be basic. They proposed that the electrical current density  $j_s$  carried by the superfluid is related to the magnetic vector potential  $A$  at each point in space by

$$\mathbf{j}_s = -\frac{1}{A c} \mathbf{A} \quad (2)$$

where  $A$  is a constant dependent on the material in question, which for a free electron gas model is given by  $A = m/\varrho_s e^2$ ,  $m$  and  $e$  being the electronic mass and charge, respectively.  $A$  is to be chosen such that  $\nabla \cdot A = 0$  to ensure current conservation. From (2) it follows that a magnetic field is excluded from a superconductor except within a distance

$$\lambda_L = \sqrt{A e^2 / 4\pi}$$

which is of order  $10^{-6}$  cm in typical superconductors for  $T$  well below  $T_c$ . Observed values of  $\lambda$  are generally several times the London value.

In the same year (1935) Fritz London (4) suggested how the diamagnetic

property (2) might follow from quantum mechanics, if there was a "rigidity" or stiffness of the wavefunction  $\psi$  of the superconducting state such that  $\psi$  was essentially unchanged by the presence of an externally applied magnetic field. This concept is basic to much of the theoretical development since that time, in that it sets the stage for the gap in the excitation spectrum of a superconductor which separates the energy of superfluid electrons from the energy of electrons in the normal fluid. As Leon Cooper will discuss, this gap plays a central role in the properties of superconductors.

In his book published in 1950, F. London extended his theoretical conjectures by suggesting that a superconductor is a "quantum structure on a macroscopic scale [which is a] kind of solidification or condensation of the average momentum distribution" of the electrons. This momentum space condensation locks the average momentum of each electron to a common value which extends over appreciable distance in space. A specific type of condensation in momentum space is central to the work Bardeen, Cooper and I did together. It is a great tribute to the insight of the early workers in this field that many of the important general concepts were correctly conceived before the microscopic theory was developed. Their insight was of significant aid in our own work.

The phenomenological London theory was extended in 1950 by Ginzburg and Landau (5) to include a spatial variation of  $\varrho_s$ . They suggested that  $\varrho_s/\varrho$  be written in terms of a phenomenological condensate wavefunction  $y(r)$  as  $\varrho_s(r)/\varrho = |\psi(r)|^2$  and that the free energy difference  $\Delta F$  between the superconducting and normal states at temperature  $T$  be given by

$$\Delta F = \int \left\{ \frac{\hbar^2}{2m} \left| (\nabla + \frac{\bar{e}}{c} A(r)) \psi(r) \right|^2 - a(T) |\psi(r)|^2 + \frac{b(T)}{2} \left| \psi(r) \right|^4 \right\} d^3r \quad (3)$$

where  $\bar{e}$ ,  $m$ ,  $a$  and  $b$  are phenomenological constants, with  $a(T_c) = 0$ .

They applied this approach to the calculation of boundary energies between normal and superconducting phases and to other problems.

As John Bardeen will discuss, a significant step in understanding which forces cause the condensation into the superfluid came with the experimental discovery of the isotope effect by E. Maxwell and, independently, by Reynolds, et al. (6). Their work indicated that superconductivity arises from the interaction of electrons with lattice vibrations, or phonons. Quite independently, Herbert Fröhlich (7) developed a theory based on electron-phonon interactions which yielded the isotope effect but failed to predict other superconducting properties. A somewhat similar approach by Bardeen (8) stimulated by the isotope effect experiments also ran into difficulties. N. Bohr, W. Heisenberg and other distinguished theorists had continuing interest in the general problem, but met with similar difficulties.

An important concept was introduced by A. B. Pippard (9) in 1953. On the basis of a broad range of experimental facts he concluded that a coherence length  $\xi$  is associated with the superconducting state such that a perturbation of the superconductor at a point necessarily influences the superfluid within a distance  $\xi$  of that point. For pure metals,  $\xi \sim 10^{-4}$  cm. for  $T \ll T_c$ . He gener-

alized the London equation (3) to a non-local form and accounted for the fact that the experimental value of the penetration depth is several times larger than the London value. Subsequently, Bardeen (10) showed that Pippard's non-local relation would likely follow from an energy gap model.

A major problem in constructing a first principles theory was the fact that the physically important condensation energy  $AF$  amounts typically to only 10W electron volts (e.V.) per electron, while the uncertainty in calculating the total energy of the electron-phonon system in even the normal state amounted to of order 1 e.V. per electron. Clearly, one had to isolate those correlations peculiar to the superconducting phase and treat them accurately, the remaining large effects presumably being the same in the two phases and therefore cancelling. Landau's theory of a Fermi liquid (11), developed to account for the properties of liquid He<sup>3</sup> formed a good starting point for such a scheme. Landau argued that as long as the interactions between the particles (He<sup>3</sup> atoms in his case, electrons in our case) do not lead to discontinuous changes in the microscopic properties of the system, a "quasi-particle" description of the low energy excitations is legitimate; that is, excitations of the fully interacting normal phase are in one-to-one correspondence with the excitations of a non-interacting fermi gas. The effective mass  $m$  and the Fermi velocity  $v_F$  of the quasi-particles differ from their free electron values, but aside from a weak decay rate which vanishes for states at the Fermi surface there is no essential change. It is the residual interaction between the quasi-particles which is responsible for the special correlations characterizing superconductivity. The ground state wavefunction of the superconductor  $\psi_0$  is then represented by a particular superposition of these normal state configurations,  $\Phi_n$ .

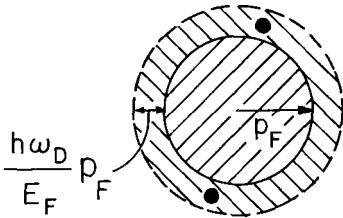
A clue to the nature of the states  $\Phi_n$  entering strongly in  $\psi_0$  is given by combining Pippard's coherence length  $\xi$  with Heisenberg's uncertainty principle

$$\Delta p \sim \hbar/\xi \sim 10^{-4} p_F \quad (4)$$

where  $p_F$  is the Fermi momentum. Thus,  $\Psi_0$  is made up of states with quasi-particles (electrons) being excited above the normal ground state by a momentum of order  $\Delta p$ . Since electrons can only be excited to states which are initially empty, it is plausible that only electronic states within a momentum  $10^4 p_F$  of the Fermi surface are involved significantly in the condensation, i.e., about  $10^4$  of the electrons are significantly affected. This view fits nicely with the fact that the condensation energy is observed to be of order  $10^4 k_B T_c$ . Thus, electrons within an energy  $\sim v_F \Delta p \simeq kT_c$  of the Fermi surface have their energies lowered by of order  $kT_c$  in the condensation. In summary, the problem was how to account for the phase transition in which a condensation of electrons occurs in momentum space for electrons very near the Fermi surface. A proper theory should automatically account for the perfect conductivity and diamagnetism, as well as for the energy gap in the excitation spectrum.

## II. THE PAIRING CONCEPT

In 1955, stimulated by writing a review article on the status of the theory of superconductivity, John Bardeen decided to renew the attack on the problem.



He invited Leon Cooper, whose background was in elementary particle physics and who was at that time working with C. N. Yang at the Institute for Advanced Study to join in the effort starting in the fall of 1955. I had the good fortune to be a graduate student of Bardeen at that time, and, having finished my graduate preliminary work, I was delighted to accept an invitation to join them.

We focused on trying to understand how to construct a ground state  $\Psi_0$  formed as a coherent superposition of normal state configurations  $\Phi_n$ ,

$$\Psi_0 = \sum_n a_n \Phi_n \quad (5)$$

such that the energy would be as low as possible. Since the energy is given in terms of the Hamiltonian  $H$  by

$$E_0 = (\Psi_0, H \Psi_0) = \sum_{n,n'} a_{n'}^* a_n (\Phi_{n'}, H \Phi_n) \quad (6)$$

we attempted to make  $E_0$  minimum by restricting the coefficients  $a_n$  so that only states which gave negative off-diagonal matrix elements would enter (6). In this case all terms would add in phase and  $E_0$  would be low.

By studying the eigenvalue spectrum of a class of matrices with off-diagonal elements all of one sign (negative), Cooper discovered that frequently a single eigenvalue is split off from the bottom of the spectrum. He worked out the problem of two electrons interacting via an attractive potential- $V$  above a quiescent Fermi sea, i.e., the electrons in the sea were not influenced by  $V$  and the extra pair was restricted to states within an energy  $\hbar\omega_D$  above the Fermi surface, as illustrated in Fig. 2. As a consequence of the non-zero density of quasi-particle states  $N(0)$  at the Fermi surface, he found the energy eigenvalue spectrum for two electrons having zero total momentum had a bound state split off from the continuum of scattering states, the binding energy being

$$E_B \cong \hbar\omega_D e^{-\frac{2}{N(0)V}} \quad (7)$$

if the matrix elements of the potential are constant equal to  $V$  in the region of interaction. This important result, published in 1956 (12), showed that, regardless of how weak the residual interaction between quasi-particles is, if the interaction is attractive the system is unstable with respect to the formation of bound pairs of electrons. Further, if  $E_B$  is taken to be of order  $k_B T_c$ , the uncertainty principle shows the average separation between electrons in the bound state is of order  $10^4 \text{ cm}$ .

While Cooper's result was highly suggestive, a major problem arose. If, as we discussed above, a fraction  $10^4$  of the electrons is significantly involved in the condensation, the average spacing between these condensed electrons

is roughly  $10^{-6}$  cm. Therefore, within the volume occupied by the bound state of a given pair, the centers of approximately  $(10^4/10^{-6})^3 \cong 10^6$  other pairs will be found, on the average. Thus, rather than a picture of a dilute gas of strongly bound pairs, quite the opposite picture is true. The pairs overlap so strongly in space that the mechanism of condensation would appear to be destroyed due to the numerous pair-pair collisions interrupting the binding process of a given pair.

Returning to the variational approach, we noted that the matrix elements  $\langle \Phi_{n'}, H \Phi_n \rangle$  in (6) alternate randomly in sign as one randomly varies  $n$  and  $n'$  over the normal state configurations. Clearly this cannot be corrected to obtain a low value of  $E_o$ , by adjusting the sign of the  $a_n$ 's since there are  $N^2$  matrix elements to be corrected with only  $N$  parameters  $a_n$ . We noticed that if the sum in (6) is restricted to include only configurations in which, if any quasi-particle state, say  $k, s$ , is occupied ( $s = \uparrow$  or  $\downarrow$  is the spin index), its "mate" state  $\bar{k}, \bar{s}$  is also occupied, then the matrix elements of  $H$  between such states would have a unique sign and a coherent lowering of the energy would be obtained. This correlated occupancy of pairs of states in momentum space is consonant with London's concept of a condensation in momentum.

In choosing the state  $\bar{k}, \bar{s}$  to be paired with a given state  $k, s$ , it is important to note that in a perfect crystal lattice, the interaction between quasi-particles conserves total (crystal) momentum. Thus, as a given pair of quasi-particles interact, their center of mass momentum is conserved. To obtain the largest number of non-zero matrix elements, and hence the lowest energy, one must choose the total momentum of each pair to be the same, that is

$$k + \bar{k} = q. \quad (8)$$

States with  $q \neq 0$  represent states with net current flow. The lowest energy state is for  $q = 0$ , that is, the pairing is such that if any state  $k\uparrow$  is occupied in an admissible  $\Phi_n$ , so is  $-k\downarrow$  occupied. The choice of  $\uparrow\downarrow$  spin pairing is not restrictive since it encompasses triplet and singlet paired states.

Through this reasoning, the problem was reduced to finding the ground state of the reduced Hamiltonian

$$H_{\text{red}} = \sum_{ks} \epsilon_k n_{ks} - \sum_{kk'} V_{k'k} b_{k'}^\dagger b_k. \quad (9)$$

The first term in this equation gives the unperturbed energy of the quasi-particles forming the pairs, while the second term is the pairing interaction in which a pair of quasi-particles in  $(k\uparrow, -k\downarrow)$  scatter to  $(k'\uparrow, -k'\downarrow)$ . The operators  $b_k^\dagger = c_{k\uparrow}^\dagger c_{-k\downarrow}^\dagger$ , being a product of two fermion (quasi-particle) creation operators, do not satisfy Bose statistics, since  $b_k^{+2} = 0$ . This point is essential to the theory and leads to the energy gap being present not only for dissociating a pair but also for making a pair move with a total momentum different from the common momentum of the rest of the pairs. It is this feature which enforces long range order in the superfluid over macroscopic distances.

### III. THE GROUND STATE

In constructing the ground state wavefunction, it seemed clear that the average occupancy of a pair state  $(k\uparrow, -k\downarrow)$  should be unity for  $k$  far below the Fermi

surface and 0 for  $k$  far above it, the fall off occurring symmetrically about  $kF$  over a range of momenta of order

$$\Delta k \sim \frac{1}{\xi} \sim 10^4 \text{ cm}^{-1}.$$

One could not use a trial  $\Psi_0$  as one in which each pair state is definitely occupied or definitely empty since the pairs could not scatter and lower the energy in this case. Rather there had to be an amplitude, say  $v_k$ , that  $(k\uparrow, -k\downarrow)$  is occupied in  $\Psi_0$  and consequently an amplitude  $u_k = \sqrt{1-v_k^2}$  that the pair state is empty. After we had made a number of unsuccessful attempts to construct a wavefunction sufficiently simple to allow calculations to be carried out, it occurred to me that since an enormous number ( $\sim 10^{19}$ ) of pair states  $(k'\uparrow, -k'\downarrow)$  are involved in scattering into and out of a given pair state  $(k\uparrow, -k\downarrow)$ , the "instantaneous" occupancy of this pair state should be essentially uncorrelated with the occupancy of the other pair states at that "instant". Rather, only the average occupancies of these pair states are related.

On this basis, I wrote down the trial ground state as a product of operators -one for each pair state-acting on the vacuum (state of no electrons),

$$\Psi_0 = \pi \prod_k (u_k + v_k b_k) |0\rangle, \quad (10)$$

where  $u_k = \sqrt{1-v_k^2}$ . Since the pair creation operators  $b_k^+$  commute for different  $k$ 's, it is clear that  $\Psi_0$  represents uncorrelated occupancy of the various pair states. I recall being quite concerned at the time that  $\Psi_0$  was an admixture of states with different numbers of electrons, a wholly new concept to me, and as I later learned to others as well. Since by varying  $v_k$  the mean number of electrons varied, I used a Lagrange multiplier  $\mu$  (the chemical potential) to make sure that the mean number of electrons ( $N_{\text{op}}$ ) represented by  $\Psi_0$  was the desired number  $N$ . Thus by minimizing

$$E_0 - \mu N = (\Psi_0, [H_{\text{red}} - \mu N_{\text{op}}] \Psi_0)$$

with respect to  $v_k$ , I found that  $v_k$  was given by

$$v_k^2 = \frac{1}{2} \left[ 1 - \frac{(\epsilon_k - \mu)}{E_k} \right] \quad (11)$$

where

$$E_k = \sqrt{(\epsilon_k - \mu)^2 + \Delta_k^2} \quad (12)$$

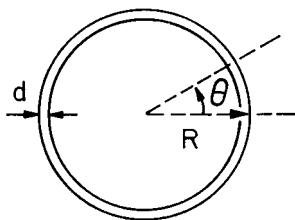
and the parameter  $\Delta_k$  satisfied what is now called the energy gap equation:

$$\Delta_k = - \sum_{k'k} V_{k'k} \frac{\Delta_{k'}}{2E_{k'}} \quad (13)$$

From this expression, it followed that for the simple model

$$V_{k'k} = \begin{cases} V, & |\epsilon_k - \mu| < \hbar\omega_D \text{ and } |\epsilon_{k'} - \mu| < \hbar\omega_D \\ 0, & \text{otherwise} \end{cases}$$

$$\Delta = \hbar \omega_{\text{DC}} - \frac{1}{N(0)V} \quad (14)$$



and the condensation energy at zero temperature is

$$\Delta F = \frac{1}{2} N(0) \Delta^2 \quad (15)$$

The idea occurred to me while I was in New York at the end of January, 1957, and I returned to Urbana a few days later where John Bardeen quickly recognized what he believed to be the essential validity of the scheme, much to my pleasure and amazement. Leon Cooper will pick up the story from here to describe our excitement in the weeks that followed, and our pleasure in unfolding the properties of the excited states.

#### IV. QUANTUM PHENOMENA ON A MACROSCOPIC SCALE

Superconductors are remarkable in that they exhibit quantum effects on a broad range of scales. The persistence of current flow in a loop of wire many meters in diameter illustrates that the pairing condensation makes the superfluid wavefunction coherent over macroscopic distances. On the other hand, the absorption of short wavelength sound and light by a superconductor is sharply reduced from the normal state value, as Leon Cooper will discuss. I will concentrate on the large scale quantum effects here.

The stability of persistent currents is best understood by considering a circular loop of superconducting wire as shown in Fig. 3. For an ideal small diameter wire, one would use the eigenstates  $e^{im\theta}$ , ( $m = 0, \pm 1, \pm 2, \dots$ ), of the angular momentum  $L_z$  about the symmetry axis to form the pairing. In the ground state no net current flows and one pairs  $m\uparrow$  with  $-m\downarrow$ , instead of  $k\uparrow$  with  $-k\downarrow$  as in a bulk superconductor. In both cases, the paired states are time reversed conjugates, a general feature of the ground state. In a current carrying state, one pairs  $(m+\nu)\uparrow$  with  $(-m+\nu)\downarrow$ , ( $\nu = 0, \pm 1, \pm 2, \dots$ ), so that the total angular momentum of each pair is identical,  $2\hbar\nu$ . It is this commonality of the center of mass angular momentum of each pair which preserves the condensation energy and long range order even in states with current flow. Another set of flow states which interweave with these states is formed by pairing  $(m+\nu)\uparrow$  with  $(-m+\nu+1)\downarrow$ , ( $\nu = 0, \pm 1, \pm 2, \dots$ ), with the pair angular momentum being  $(2\nu+1)\hbar$ . The totality of states forms a set with all integer multiples  $n$  of  $\hbar$  for allowed total angular momentum of pairs. Thus, even though the pairs greatly overlap in space, the system exhibits quantization effects as if the pairs were well defined.

There are two important consequences of the above discussion. First, the fact that the coherent condensate continues to exist in flow states shows that to scatter a pair out of the (rotating) condensate requires an increase of energy.

Crudely speaking, slowing down a given pair requires it to give up its binding energy and hence this process will occur only as a fluctuation. These fluctuations average out to zero. The only way in which the flow can stop is if all pairs simultaneously change their pairing condition from, say,  $v$  to  $v - 1$ . In this process the system must fluctuate to the normal state, at least in a section of the wire, in order to change the pairing. This requires an energy of order the condensation energy  $\Delta F$ . A thermal fluctuation of this size is an exceedingly rare event and therefore the current persists.

The second striking consequence of the pair angular momentum quantization is that the magnetic flux  $\Phi$  trapped within the loop is also quantized,

$$\Phi_n = n \cdot \frac{hc}{2e} \quad (n = 0, \pm 1, \pm 2 \dots). \quad (16)$$

This result follows from the fact that if the wire diameter  $d$  is large compared to the penetration depth  $\lambda$ , the electric current in the center of the wire is essentially zero, so that the canonical angular momentum of a pair is

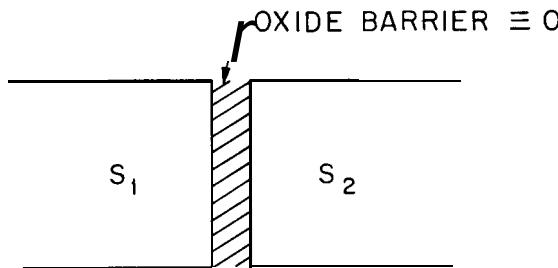
$$L_{\text{pair}} = \frac{2e}{c} r_{\text{pair}} \times A \quad (17)$$

where  $r_{\text{pair}}$  is the center of mass coordinate of a pair and  $A$  is the magnetic vector potential. If one integrates  $L_{\text{pair}}$  around the loop along a path in the center of the wire, the integral is  $nh$ , while the integral of the right hand side of

$$(17) \text{ is } \frac{2e}{c} \Phi.$$

A similar argument was given by F. London (4b) except that he considered only states in which the superfluid flows as a whole without a change in its internal structure, i.e., states analogous to the  $(m+v)\uparrow, (-m+v)\downarrow$  set. He found  $\Phi_z = n \cdot hc/e$ . The pairing  $(m+v)\uparrow, (m+v+1)\downarrow$  cannot be obtained by adding  $v$  to each state, yet this type of pairing gives an energy as low as the more conventional flow states and these states enter experimentally on the same basis as those considered by London. Experiments by Deaver and Fairbank (13), and independently by Doll and Nähauer (14) confirmed the flux quantization phenomenon and provided support for the pairing concept by showing that  $2e$  rather than  $e$  enters the flux quantum. Following these experiments a clear discussion of flux quantization in the pairing scheme was given by Beyers and Yang (15).

The idea that electron pairs were somehow important in superconductivity has been considered for some time (16, 17). Since the superfluidity of liquid  $\text{He}^4$  is qualitatively accounted for by Bose condensation, and since pairs of electrons behave in some respects as a boson, the idea is attractive. The essential point is that while a dilute gas of tightly bound pairs of electrons might behave like a Bose gas (18) this is not the case when the mean spacing between pairs is very small compared to the size of a given pair. In this case the inner structure of the pair, i.e., the fact that it is made of fermions, is essential; it is this which distinguishes the pairing condensation, with its energy gap for single pair translation as well as dissociation, from the spectrum of a Bose con-



densate, in which the low energy excitations are Bose-like rather than Fermi-like as occurs in actual superconductors. As London emphasized, the condensation is an ordering in occupying momentum space, and not a space-like condensation of clusters which then undergo Bose condensation.

In 1960, Ivar Giaever (19) carried out pioneering experiments in which electrons in one superconductor (\$S\_1\$) tunneled through a thin oxide layer (\$\sim 20\$-\$30\$ Å) to a second superconductor (\$S\_2\$) as shown in Fig. 4. Giaever's experiments were dramatic evidence of the energy gap for quasi-particle excitations. Subsequently, Brian Josephson made a highly significant contribution by showing theoretically that a superfluid current could flow between \$S\_1\$ and \$S\_2\$ with zero applied bias. Thus, the superfluid wavefunction is coherent not only in \$S\_1\$ and \$S\_2\$ separately, but throughout the entire system, \$S\_1\$-\$0\$-\$S\_2\$, under suitable circumstances. While the condensate amplitude is small in the oxide, it is sufficient to lock the phases of \$S\_1\$ and \$S\_2\$ together, as has been discussed in detail by Josephson (20) and by P. W. Anderson (21).

To understand the meaning of phase in this context, it is useful to go back to the ground state wavefunction \$\Psi\_0\$, (10). Suppose we write the parameter \$v\_k\$ as \$|v\_k| \exp i\varphi\$ and choose \$v\_k\$ to be real. If we expand out the \$k\$-product in \$\Psi\_0\$, we note that the terms containing \$N\$ pairs will have a phase factor \$\exp(iN\varphi)\$, that is, each occupied pair state contributes a phase \$\varphi\$ to \$\Psi\_0\$. Let this wavefunction, say \$\Psi\_0^{(1)}\$ represent \$S\_1\$, and have phase \$\varphi\_1\$. Similarly, let \$\Psi\_0^{(2)}\$ represent \$S\_2\$ and have phase angle \$\varphi\_2\$. If we write the state of the combined system as a product

$$\Psi_0^{(1,2)} = \Psi_0^{(1)} \Psi_0^{(2)} \quad (18)$$

then by expanding out the double product we see that the phase of that part of \$\Psi\_0^{(1,2)}\$ which has \$N\_1\$ pairs in \$S\_1\$ and \$N\_2\$ pairs in \$S\_2\$ is \$N\_1\varphi\_1 + N\_2\varphi\_2\$. For a truly isolated system, \$2(N\_1 + N\_2) = 2N\$ is a fixed number of electrons; however \$N\_1\$ and \$N\_2\$ are not separately fixed and, as Josephson showed, the energy of the combined system is minimized when \$\varphi\_1 = \varphi\_2\$ due to tunneling of electrons between the superconductors. Furthermore, if \$\varphi\_1 = \varphi\_2\$, a current flows between \$S\_1\$ and \$S\_2\$

$$j = j_1 \sin(\varphi_1 - \varphi_2) \quad (19)$$

If \$\varphi\_1 - \varphi\_2 = \varphi\$ is constant in time, a constant current flows with no voltage applied across the junction. If a bias voltage is \$V\$ applied between \$S\_1\$ and \$S\_2\$, then, according to quantum mechanics, the phase changes as

$$\frac{2eV}{\hbar} = \frac{d\varphi}{dt} \quad (20)$$

Hence a constant voltage applied across such a junction produces an alternating current of frequency

$$\nu = \frac{2eV}{\hbar} = 483 \text{ THz/V}. \quad (21)$$

These effects predicted by Josephson were observed experimentally in a series of beautiful experiments (22) by many scientists, which I cannot discuss in detail here for lack of time. I would mention, as an example, the work of Langenberg and his collaborators (23) at the University of Pennsylvania on the precision determination of the fundamental constant  $e/h$  using the frequency-voltage relation obeyed by the alternating Josephson supercurrent. These experiments have decreased the uncertainty in our experimental knowledge of this constant by several orders of magnitude and provide, in combination with other experiments, the most accurate available value of the Sommerfeld fine structure constant. They have resulted in the resolution of several discrepancies between theory and experiment in quantum electrodynamics and in the development of an "atomic" voltage standard which is now being used by the United States National Bureau of Standards to maintain the U.S. legal volt.

## V. CONCLUSION

As I have attempted to sketch, the development of the theory of superconductivity was truly a collaborative effort, involving not only John Bardeen, Leon Cooper and myself, but also a host of outstanding scientists working over a period of half a century. As my colleagues will discuss, the theory opened up the field for many exciting new developments, both scientific and technological, many of which no doubt lie in the future. I feel highly honored to have played a role in this work and I deeply appreciate the honor you have bestowed on me in awarding us the Nobel prize.

## REFERENCES

1. Kamerlingh Onnes, H., Nobel Lectures, Vol. 1, pp. 306-336.
2. Meissner, W. and Ochsenfeld, R., Naturwiss. 21, 787 (1933).
3. Garter, C. J. and Casimir, H. B. G., Phys. Z. 35, 963 (1934); Z. Techn. Phys. 15, 539 (1934).
4. London, F., [a] Phys. Rev. 24, 562 (1948); [b] Superfluids, Vol. 1 (John Wiley & Sons, New York, 1950).
5. Ginzburg, V. L. and Landau, L. D., J. Exp. Theor. Phys. (U.S.S.R.) 20, 1064 (1950).
6. Maxwell, E., Phys. Rev. 78, 477 (1950); Reynolds, C. A., Serin, B., Wright W. H. and Nesbitt, L. B., Phys. Rev. 78, 487 (1950).
7. Fröhlich, H., Phys. Rev. 79, 845 (1950).

8. Bardeen, J., Rev. Mod. Phys. 23, 261 (1951).
9. Pippard, A. B., Proc. Royal Soc. (London) A216, 547 (1953).
10. Bardeen, J., [a] Phys. Rev. 97, 1724 (1955); [b] Encyclopedia of Physics, Vol. 15 (Springer-Verlag, Berlin, 1956), p. 274.
11. Landau, L. D., J. Exp. Theor. Phys. (U.S.S.R.) 30 (3), 1058 (920) (1956); 32 (5), 59 (101) (1957).
12. Cooper, L. N., Phys. Rev. 104, 1189 (1956).
13. Deaver, B. S. Jr., and Fairbank, W. M., Phys. Rev. Letters 7, 43 (1961).
14. Doll, R. and NBbauer, M., Phys. Rev. Letters 7, 51 (1961).
15. Beyers, N. and Yang, C. N., Phys. Rev. Letters 7, 46 (1961).
16. Ginzburg, V. L., Usp. Fiz. Nauk 48, 25 (1952); transl. Fortsch. d. Phys. I, 101 (1953).
17. Schafroth, M. R., Phys. Rev. 96, 1442 (1954); 100, 463 (1955).
18. Schafroth, M. R., Blatt, J. M. and Butler, S. T., Helv. Phys. Acta 30, 93 (1957).
19. Giaever, I., Phys. Rev. Letters 5, 147 (1960).
20. Josephson, B. D., Phys. Letters I, 251 (1962); Advan. Phys. 14, 419 (1965).
21. Anderson, P. W., in Lectures on the Many-body Problem, edited by E. R. Caianiello (Academic Press, Inc. New York, 1964), Vol. II.
22. See Superconductivity, Parks, R. D., ed. (Dekker New York, 1969).
23. See, for example, Parker, W. H. Taylor B. N. and Langenberg, D. N. Phys. Rev. Letters 18, 287 (1967); Finnegan, T. F. Denenstein A. and Langenberg, D. N. Phys. Rev. B4, 1487 (1971).

**Physics 1973**

**LEO ESAKI, IVAR GIAEVER**

*for their experimental discoveries regarding tunneling phenomena in semiconductors and superconductors respectively*

**and**

**BRIAN D JOSEPHSON**

*for his theoretical predictions of the properties of a supercurrent through a tunnel barrier, in particular those phenomena which are generally known as Josephson effects*

# THE NOBEL PRIZE FOR PHYSICS

Speech by professor **STIG LUNDQVIST** of the Royal Academy of Sciences  
Translation from the Swedish text

Your Majesty, Your Royal Highnesses, Ladies and Gentlemen,  
The 1973 Nobel Prize for physics has been awarded to Drs. Leo Esaki, Ivar Giaever and Brian Josephson for their discoveries of tunnelling phenomena in solids.

The tunnelling phenomena belong to the most direct consequences of the laws of modern physics and have no analogy in classical mechanics. Elementary particles such as electrons cannot be treated as classical particles but show both wave and particle properties. Electrons are described mathematically by the solutions of a wave equation, the Schrödinger equation. An electron and its motion can be described by a superposition of simple waves, which forms a wave packet with a finite extension in space. The waves can penetrate a thin barrier, which would be a forbidden region if we treat the electron as a classical particle. The term tunnelling refers to this wave-like property - the particle "tunnels" through the forbidden region. In order to get a notion of this kind of phenomenon let us assume that you are throwing balls against a wall. In general the ball bounces back but occasionally the ball disappears straight through the wall. In principle this could happen, but the probability for such an event is negligibly small.

On the atomic level, on the other hand, tunnelling is a rather common phenomenon. Let us instead of balls consider electrons in a metal moving with high velocities towards a forbidden region, for example a thin insulating barrier. In this case we cannot neglect the probability of tunneling. A certain fraction of the electrons will penetrate the barrier by tunnelling and we may obtain a weak tunnel current through the barrier.

The interest for tunnelling phenomena goes back to the early years of quantum mechanics, i.e. the late twenties. The best known early application of the ideas came in the model of alpha-decay of heavy atomic nuclei. Some phenomena in solids were explained by tunnelling in the early years. However, theory and experiments often gave conflicting results, no further progress was made and physicists lost interest in solid state tunnelling in the early thirties.

With the discovery of the transistor effect in 1947 came a renewed interest in the tunnelling process. Many attempts were made to observe tunnelling in semiconductors, but the results were controversial and inconclusive.

It was the young Japanese physicist Leo Esaki, who made the initial pioneering discovery that opened the field of tunnelling phenomena for research. He was at the time with the Sony Corporation, where he performed some deceptively simple experiments, which gave convincing experimental evidence

for tunnelling of electrons in solids, a phenomenon which had been clouded by questions for decades. Not only was the existence of tunnelling in semiconductors established, but he also showed and explained an unforeseen aspect of tunnelling in semiconductor junctions. This new aspect led to the development of an important device, called the tunnel diode or the Esaki diode.

Esaki's discovery, published in 1958, opened a new field of research based on tunnelling in semiconductors. The method soon became of great importance in solid state physics because of its simplicity in principle and the high sensitivity of tunnelling to many finer details.

The next major advance in the field of tunnelling came in the field of superconductivity through the work of Ivar Giaever in 1960. In 1957, Bardeen, Cooper and Schrieffer had published their theory of superconductivity, which was awarded the 1972 Nobel Prize in physics. A crucial part of their theory is that an energy gap appears in the electron spectrum when a metal becomes superconducting. Giaever speculated that the energy gap should be reflected in the current-voltage relation in a tunnelling experiment. He studied tunnelling of electrons through a thin sandwich of evaporated metal films insulated by the natural oxide of the film first evaporated. The experiments showed that his conjecture was correct and his tunnelling method soon became the dominating method to study the energy gap in superconductors. Giaever also observed a characteristic fine structure in the tunnel current, which depends on the coupling of the electrons to the vibrations of the lattice. Through later work by Giaever and others the tunnelling method has developed into a new spectroscopy of high accuracy to study in detail the properties of superconductors, and the experiments have in a striking way confirmed the validity of the theory of superconductivity.

Giaver's experiments left certain theoretical questions open and this inspired the young Brian Josephson to make a penetrating theoretical analysis of tunnelling between two superconductors. In addition to the Giaever current he found a weak current due to tunnelling of coupled electron pairs, called Coopers pairs. This implies that we get a supercurrent through the barrier. He predicted two remarkable effects. The first effect is that a supercurrent may flow even if no voltage is applied. The second effect is that a high frequency alternating current will pass through the barrier if a constant voltage is applied.

Josephson's theoretical discoveries showed how one can influence supercurrents by applying electric and magnetic fields and thereby control, study and exploit quantum phenomena on a macroscopic scale. His discoveries have led to the development of an entirely new method called quantum interferometry. This method has led to the development of a rich variety of instruments of extraordinary sensitivity and precision with application in wide areas of science and technology.

Esaki, Giaever and Josephson have through their discoveries opened up new fields of research in physics. They are closely related because the pioneering work by Esaki provided the foundation and direct impetus for Giaever's discovery and Giaever's work in turn provided the stimulus which led to Jo-

sephson's theoretical predictions. The close relation between the abstract concepts and sophisticated tools of modern physics and the practical applications to science and technology is strongly emphasized in these discoveries. The applications of solid state tunnelling already cover a wide range. Many devices based on tunneling are now used in electronics. The new quantum interferometry has already been used in such different applications as measurements of temperatures near the absolute zero, to detect gravitational waves, for ore prospecting, for communication through water and through mountains, to study the electromagnetic field around the heart or brain, to mention a few examples.

Drs. Esaki, Giaever and Josephson,

In a series of brilliant experiments and calculations you have explored different aspects of tunnelling phenomena in solids. Your discoveries have opened up new fields of research and have given new fundamental insight about electrons in semiconductors and superconductors and about macroscopic quantum phenomena in superconductors.

On behalf of the Royal Academy of Sciences I wish to express our admiration and convey to you our warmest congratulations. I now ask you to proceed to receive your prizes from the hands of his Majesty the King.



Leo Eshai / 奈玲珍于山工之江

## LEO ESAKI

Leo Esaki was born in Osaka, Japan in 1925. Esaki completed work for a B.S. in Physics in 1947 and received his Ph.D in 1959, both from the University of Tokyo. Esaki is an IBM Fellow and has been engaged in semiconductor research at the IBM Thomas J. Watson Research Center, Yorktown Heights, New York, since 1960. Prior to joining IBM, he worked at the Sony Corp. where his research on heavily-doped Ge and Si resulted in the discovery of the Esaki tunnel diode; this device constitutes the first quantum electron device. Since 1969, Esaki has, with his colleagues, pioneered "designed semiconductor quantum structures" such as man-made superlattices, exploring a new quantum regime in the frontier of semiconductor physics.

The Nobel Prize in Physics (1973) was awarded in recognition of his pioneering work on electron tunneling in solids. Other awards include the Nishina Memorial Award (1959), the Asahi Press Award (1960), the Toyo Rayon Foundation Award for the Promotion of Science and Technology (1960), the Morris N. Liebmann Memorial Prize from IRE (1961), the Stuart Ballantine Medal from the Franklin Institute (1961), the Japan Academy Award (1965), the Order of Culture from the Japanese Government (1974), the American Physical Society 1985 International Prize for New Materials for his pioneering work in artificial semiconductor superlattices, the IEEE Medal of Honor in 1991 for contributions to and leadership in tunneling, semiconductor superlattices, and quantum wells. Dr. Esaki holds honorary degrees from Doshisha School, Japan, the Universidad Politecnica de Madrid, Spain, the University of Montpellier, France, Kwansei Gakuin University, Japan and the University of Athens, Greece. Dr. Esaki is a Director of IBM-Japan, Ltd., on the Governing Board of the IBM-Tokyo Research Laboratory, a Director of the Yamada Science Foundation and the Science and Technology Foundation of Japan. He serves on numerous international scientific advisory boards and committees, and is an Adjunct Professor of Waseda University, Japan. Currently he is a Guest Editorial writer for the Yomiuri Press. Dr. Esaki was elected a Fellow of the American Academy of Arts and Sciences in May 1974, a member of the Japan Academy on November 12, 1975, a Foreign Associate of the National Academy of Engineering (USA) on April 1, 1977, a member of the Max-Planck-Gesellschaft on March 17, 1989, and a foreign member of the American Philosophical Society in April of 1991.

# LONG JOURNEY INTO TUNNELING

Nobel Lecture, December 12, 1973

by

**LEO ESAKI**

IBM Thomas J. Watson Research Center, Yorktown Heights, N.Y., USA

## I. HISTORICAL BACKGROUND

In 1923, during the infancy of the quantum theory, de Broglie (1) introduced a new fundamental hypothesis that matter may be endowed with a dualistic nature-particles may also have the characteristics of waves. This hypothesis, in the hands of Schrodinger (2) found expression in the definite form now known as the Schrödinger wave equation, whereby an electron or a particle is assumed to be represented by a solution to this equation. The continuous nonzero nature of such solutions, even in classically forbidden regions of negative kinetic energy, implies an ability to penetrate such forbidden regions and a probability of tunneling from one classically allowed region to another. The concept of tunneling, indeed, arises from this quantum-mechanical result. The subsequent experimental manifestations of this concept can be regarded as one of the early triumphs of the quantum theory.

In 1928, theoretical physicists believed that tunneling could occur by the distortion, lowering or thinning, of a potential barrier under an externally applied high electric field. Oppenheimer (3) attributed the autoionization of excited states of atomic hydrogen to the tunnel effect: The coulombic potential well which binds an atomic electron could be distorted by a strong electric field so that the electron would see a finite potential barrier through which it could tunnel.

Fowler and Nordheim (4) explained, on the basis of electron tunneling, the main features of the phenomenon of electron emission from cold metals by high external electric fields, which had been unexplained since its observation by Lilienfeld (5) in 1922. They proposed a one-dimensional model. Metal electrons are confined by a potential wall whose height is determined by the work function  $\psi$  plus the fermi energy  $E_F$ , and the wall thickness is substantially decreased with an externally applied high electric field, allowing electrons to tunnel through the potential wall, as shown in Fig. 1. They successfully derived the well-known Fowler-Nordheim formula for the current as a function of electric field  $F$ :

$$J = AF^2 \exp[-4(2m)^{1/2} \Phi^{3/2} / 3\hbar F].$$

An application of these ideas which followed almost immediately came in the model for a decay as a tunneling process put forth by Gamow (6) and Gurney and Condon. (7) Subsequently, Rice (8) extended this theory to the description of molecular dissociation.

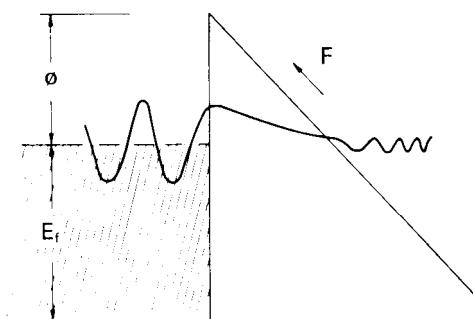


Fig. 1. Fowler-Nordheim tunneling.

The next important development was an attempt to invoke tunneling in order to understand transport properties of electrical contacts between two solid conductors. The problems of metal-to-metal and semiconductor-to-metal contacts are important technically, because they are directly related to electrical switches and rectifiers or detectors.

In 1930, Frenkel (9) proposed that the anomalous temperature independence of contact resistance between metals could be explained in terms of tunneling across a narrow vacuum separation. Holm and Meissner (10) then did careful measurements of contact resistances and showed that the magnitude and temperature independence of the resistance of insulating surface layers were in agreement with an explanation based on tunneling through a vacuum-like space. These measurements probably constitute the first correctly interpreted observations of tunneling currents in solids, (11) since the vacuum-like space was a solid insulating oxide layer.

In 1932, Wilson, (12) Frenkel and Joffe, (13) and Nordheim (14) applied quantum mechanical tunneling to the interpretation of metal-semiconductor contacts-rectifiers such as those made from selenium or cuprous oxide. From a most simplified energy diagram, shown in Fig. 2, the following well-known current-voltage relationship was derived:

$$J = J_s [\exp(eV/kT) - 1]$$

Apparently, this theory was accepted for a number of years until it was finally discarded after it was realized that it predicted rectification in the wrong direction for the ordinary practical diodes. It is now clear that, in the usual circumstance, the surface barriers found by the semiconductors in contact with metals, as illustrated in Fig. 2, are much too thick to observe tunneling current. There existed a general tendency in those early days of quantum mechanics to try to explain any unusual effects in terms of tunneling. In many cases, however, conclusive experimental evidence of tunneling was lacking, primarily because of the rudimentary stage of material science.

In 1934, the development of the energy-band theory of solids prompted Zener (15) to propose interband tunneling, or internal field emission, as an explanation for dielectric breakdown. He calculated the rate of transitions

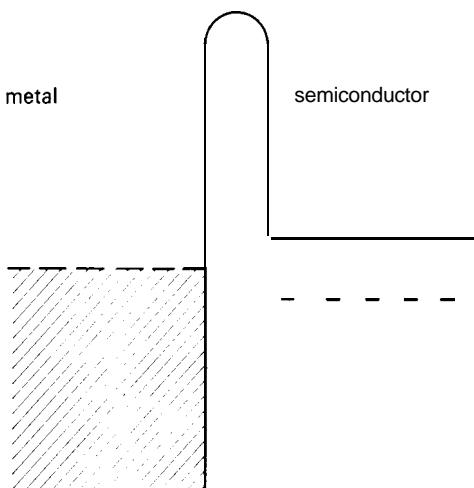


Fig. 2. Early model of metal-semiconductor rectifiers.

from a filled band to a next-higher unfilled band by the application of an electric field. In effect, he showed that an energy gap could be treated in the manner of a potential barrier. This approach was refined by Houston (16) in 1940. The Zener mechanism in dielectric breakdown, however, has never been proved to be important in reality. If a high electric field is applied to the bulk crystal of a dielectric or a semiconductor, avalanche breakdown (electron-hole pair generation) generally precedes tunneling, and thus the field never reaches a critical value for tunneling.

## II. TUNNEL DIODE

Around 1950, the technology of Ge p-n junction diodes, being basic to transistors, was developed, and efforts were made to understand the junction properties. In explaining the reverse-bias characteristic, McAfee et al. (17) applied a modified Zener theory and asserted that low-voltage breakdown in Ge diodes (specifically, they showed a 10-V breakdown) resulted from interband tunneling from the valence band in the p-type region to the empty conduction band in the n-type region. The work of McAfee et al. inspired a number of other investigations of breakdown in p-n junctions. Results of those later studies (18) indicated that most Ge junctions broke down by avalanche, but by that time the name "Zener diodes" had already been given to the low-breakdown Si diodes. Actually, these diodes are almost always avalanche diodes. In 1957, Chynoweth and McKay (19) examined Si junctions of low-voltage breakdown and claimed that they had finally observed tunneling. In this circumstance, in 1956, I initiated the investigation of interband tunneling or internal field emission in semiconductor diodes primarily to scrutinize the electronic structure of narrow (width) p-n junctions. This information, at the time, was also important from a technological point of view.

The built-in field distribution in p-n junctions is determined by the profile of impurities-donors and acceptors. If both the impurity distributions are

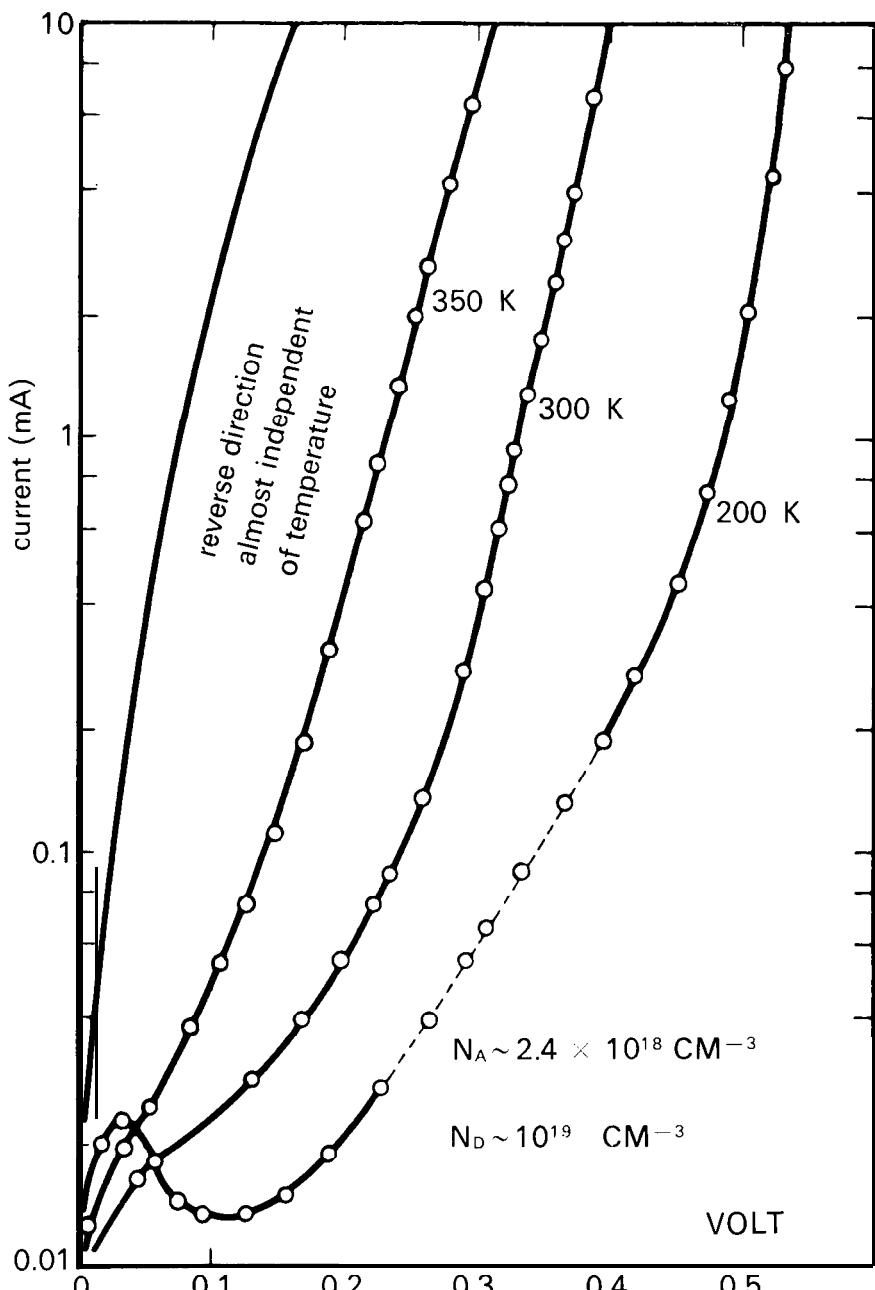


Fig. 3. Semilog plots of current-voltage characteristics in a tunnel diode, where  $N_A \sim 2.4 \times 10^{18} \text{ cm}^{-3}$  and  $N_D \sim 10^{19} \text{ cm}^{-3}$ .

assumed to be step functions, the junction width  $W$  is given by  $W = \text{const} [(N_A + N_D)/N_A N_D]^{1/2} \sim 1/N^{1/2}$ , where  $N_A$  and  $N_D$  are the acceptor and donor concentrations and  $N < N_A$  or  $N_D$ . Thus, first of all, we attempted to prepare heavily-doped Ge p-n junctions. Both the donor and acceptor concentrations are sufficiently high so that the respective sides of the junctions

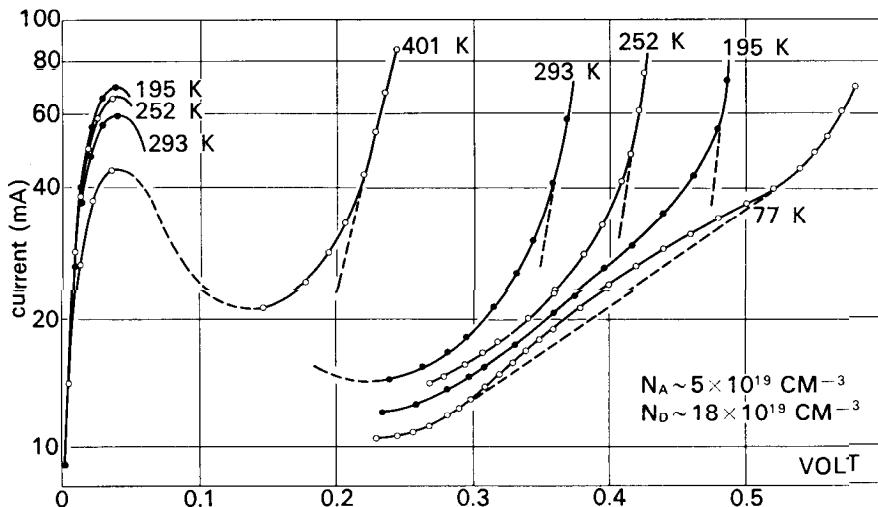


Fig. 4. Semilog plots of current-voltage characteristics in a tunnel diode, where  $N_A = 5 \times 10^{19} \text{ cm}^{-3}$  and  $N_D = 1.8 \times 10^{19} \text{ cm}^{-3}$ .

are degenerate, that is, the fermi energies are located well inside the conduction or valence band.

In this study, we first obtained a backward diode which was more conductive in the reverse direction than in the forward direction. In this respect it agreed with the rectification direction predicted by the previously-mentioned old tunneling rectifier theory. The calculated junction width at zero bias was approximately 200 Å, which was confirmed by capacitance measurements. In this junction, the possibility of an avalanche was completely excluded because the breakdown occurs at much less than the threshold voltage for electron-hole pair production. The current-voltage characteristic at room temperature indicated not only that the major current-flow mechanism was convincingly tunneling in the reverse direction but also that tunneling might be responsible for current flow even in the low-voltage range of the forward direction. When the unit was cooled, we saw, for the first time, a negative resistance, appearing, as shown in Fig. 3. By further narrowing the junction width (thereby further decreasing the tunneling path), through a further increase in the doping level, the negative resistance was clearly seen at all temperatures, as shown in Fig. 4. (20)

The characteristic was analyzed in terms of interband tunneling. In the tunneling process, if it is elastic, the electron energy will be conserved. Figures 5 (a), (b), (c), and (d) show the energy diagrams of the tunnel diode at zero bias and with applied voltages,  $V_p$ ,  $E$ 's, and  $V_s$ , respectively. As the bias is increased up to the voltage  $V_i$ , the interband tunnel current continues to increase, as shown by an arrow in Fig. 5 (b). However, as the conduction band in the n-type side becomes uncrossed with the valence band in the p-type side, with further increase in applied voltages, as shown in Fig. 5 (c), the current decreases because of the lack of allowed states of corresponding ener-

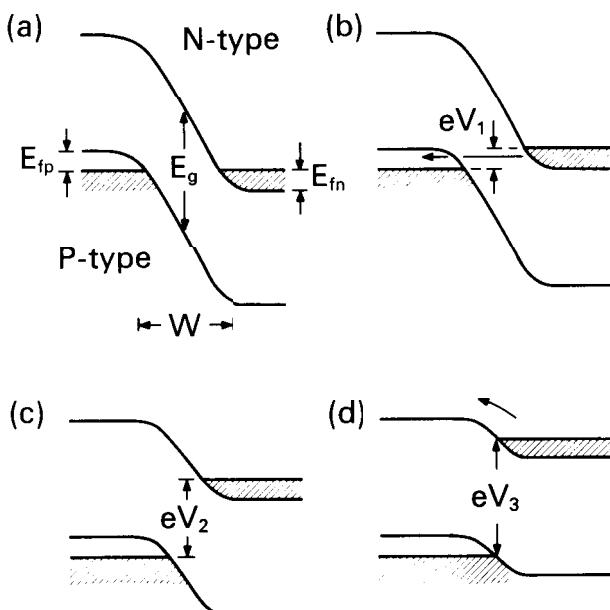


Fig. 5. Energy diagrams at varying bias-conditions in the tunnel diode.

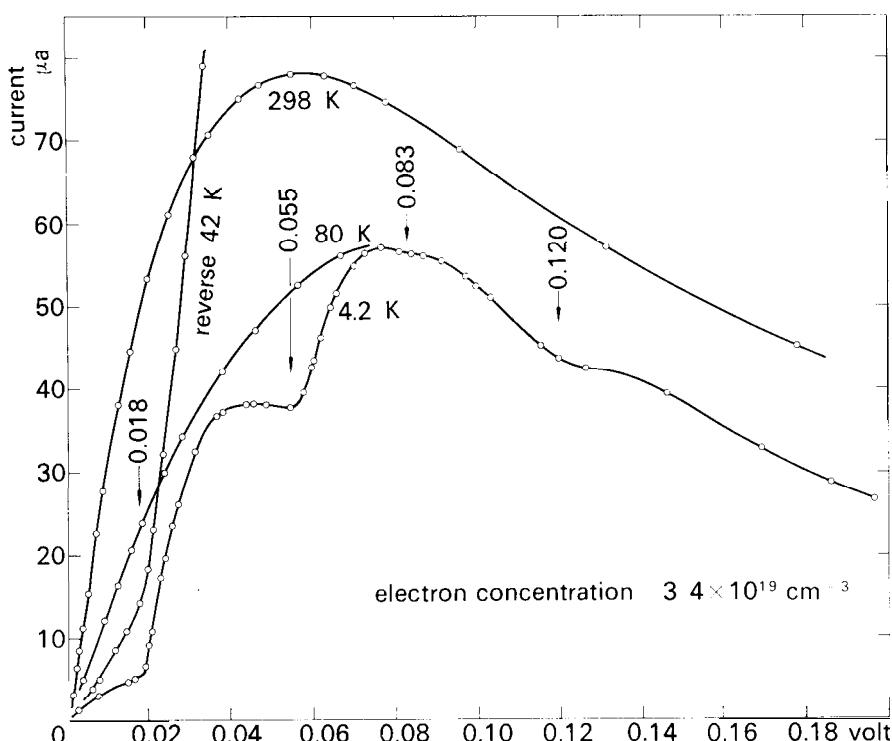


Fig. 6. Current-voltage characteristics in a Si tunnel diode at 4.2, 80 and 298 K.

gies for tunneling. When the voltage reaches  $V_2$ , or higher, the normal diffusion (or thermal) current will dominate as in the case of the usual p-n diode. Semiconductor materials other than Ge were quickly explored to obtain tunnel diodes: Si, InSb, GaAs, InAs, PbTe, GaSb, SiC, etc.

In our early study of the Si tunnel diode, (21) a surprisingly fine structure was found in the current-voltage curve at 4.2 K, indicating the existence of inelastic tunneling, as shown in Fig. 6. We were impressed with the fact that four voltages at the singularities shown in the figure agreed almost exactly with four characteristic energies due to acoustic and optical phonons, obtained from the optical absorption spectra (22) and also derived from the analysis of intrinsic recombination radiation (23) in pure silicon. The analysis of tunneling current in detail reveals not only the electronic states in the systems involved, but also the interactions of tunneling electrons with phonons, photons, plasmons, or even vibrational modes of molecular species in barriers. (24) As a result of the rich amount of information which can be obtained from a study of tunneling processes, a field called tunneling spectroscopy has emerged.

### III. NEGATIVE RESISTANCE IN METAL-OXIDE-SEMICONDUCTOR JUNCTIONS

This talk, however, is not intended as a comprehensive review of the many theoretical and experimental investigations of tunneling, which is available elsewhere. (25) Instead, I would like to focus on only one aspect for the rest of the talk: negative resistance phenomena in semiconductors which can be observed in novel tunnel structures.

Differential negative resistance occurs only in particular circumstances, where the total number of tunneling electrons transmitted across a barrier structure per second decreases, rather than increases as in the usual case, with an increase in applied voltage. The negative resistance phenomena themselves are not only important in solid-state electronics because of possible signal amplification, but also shed light on some fundamental aspects of tunneling.

Before proceeding to the main subject, I would like to briefly outline the independent-electron theory of tunneling. (26) In tunneling, we usually deal with a *one-dimensional potential barrier*  $V(x)$ . The transmission coefficient  $D$  for such a barrier is defined as the ratio of the intensity of the transmitted electron wave to that of the incident wave. The most common approximation for  $D$  is the use of the semiclassical WKB form

$$D(E_x) = \exp \left[ -2/\hbar \int_{x_1}^{x_2} (2m(V-E_x))^{1/2} dx \right] \quad (1)$$

where  $E_x$  is the kinetic energy in the direction normal to the barrier, and the quantities  $x_1$  and  $x_2$  are the classical turning points of an electron of energy  $E_x$  at the edges of the potential barrier. If the boundary regions are sharp, we first construct wave functions by matching values of functions as well as

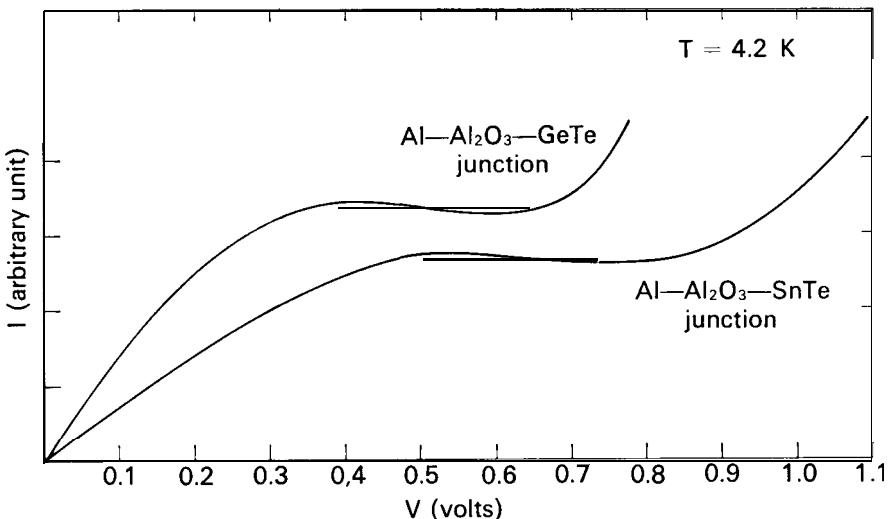


Fig. 7. Current-voltage characteristics in SnTe and GeTe tunnel junctions at 4.2 K.

their derivatives at each boundary, then calculate the transmission coefficient  $D$ .

The tunneling expression should include two basic conservation laws: 1) Conservation of the total electron energy; and 2) Conservation of the component of the electron wave vector parallel to the plane of the junction. The velocity of an incident electron associated with a state of wave number  $k_x$  is given by  $1/\hbar \partial E / \partial k_x$  in a one-particle approximation. Then, the tunneling current per unit area is written by

$$J = 2e/(2\pi)^3 \int \int \int D(E_x) (f(E) - f(E')) 1/\hbar \partial E / \partial k_x dk_x dk_y dk_z \quad (2)$$

where  $f$  is the fermi distribution function or occupation probability, and  $E$  and  $E'$  are the energy of the incident electron and that of the transmitted one, respectively. The front factor  $2/(2\pi)^3$  comes from the fact that the volume of a state occupied by two electrons of the opposite spin is  $(2\pi)^3$  in the wave-vector space for a unit volume crystal.

The previously-mentioned tunnel diode is probably the first structure in which the negative resistance effect was observed. But, now, I will demonstrate that a similar characteristic can be obtained in a metal-oxide-semiconductor tunnel junction, (27) where the origin of the negative resistance is quite different from that in the tunnel diode. The semiconductors involved here (SnTe and GeTe) are rather unusual-more metallic than semiconducting; both of them are nonstoichiometric and highly p-type owing to high concentrations of Sn or Ge vacancies with typical carrier concentrations about  $8 \times 10^{20}$  and  $2 \times 10^{20} \text{ cm}^{-3}$ , respectively. The tunnel junctions were prepared by evaporating SnTe or GeTe onto an oxidized evaporated stripe of Al on quartz or sapphire substrates. In contrast to the p-n junction diodes, all ma-

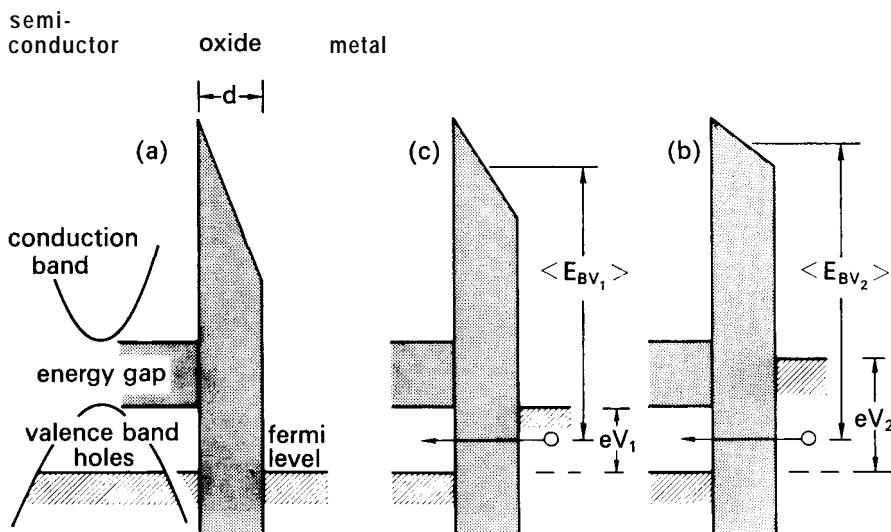


Fig. 8. Energy diagrams at varying bias-conditions in  $\text{Al}-\text{Al}_2\text{O}_3\text{SnTe}$  or - $\text{GeTe}$  tunnel junctions.

terials involved in these junctions are polycrystalline, although the Al oxide is possibly amorphous.

Figure 7 illustrates the current-voltage curves at 4.2 K of typical SnTe and GeTe junctions and Fig. 8 shows their energy diagrams at zero bias, and at applied voltages  $V_1$  and  $V_2$  from the left to the right. As is the case in the tunnel diode, until the bias voltage is increased such that the fermi level in the metal side coincides with the top of the valence band in the semiconductor side (Fig. 8 (b)), the tunnel current continues to increase. When the bias voltage is further increased (Fig. 8 (c)), however, the total number of empty allowed states or holes in the degenerate p-type semiconductor is unchanged, whereas the tunneling barrier height is raised, for instance from  $E_{BV_1}$  to  $E_{BV_2}$ , resulting in a decrease in tunneling probability determined by the exponential term,  $e^{-\lambda}$ , where  $\lambda \sim 2d(2mE_{BV})^{1/2}/\hbar$ , and  $E_{BV}$  and  $d$  are the barrier height and width, respectively. Thus a negative resistance is exhibited in the current-voltage curve. When the bias voltage becomes higher than the level corresponding to the bottom of the conduction band in the semiconductor, a new tunneling path from the metal to the conduction band is opened and one sees the current again increasing with the voltage. The rectification direction in this junction is again backward as is the case in the tunnel diode.

We might add that, in this treatment, the tunneling exponent is assumed to be determined only by the energy difference between the bottom of the conduction band in the oxide and the metal fermi energy. This assumption should be valid because this energy difference is probably much smaller than that between the top of the valence band in the oxide and the metal fermi energy.

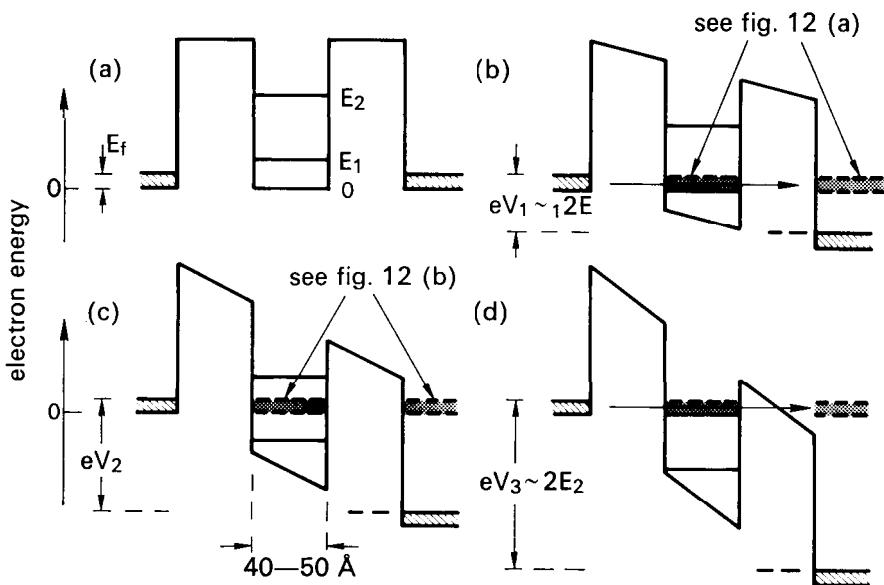


Fig. 9. Energy diagrams at varying bias-conditions in a double-barrier tunnel junction, indicating the resonant transmission in (b) and (d).

#### IV. NEGATIVE RESISTANCE DUE TO RESONANT TRANSMISSION

It has been known that there is a phenomenon called the resonant transmission. Historically, resonant transmission was first demonstrated in the scattering of electrons by atoms of noble gases and is known as the Ramsauer effect. In many textbooks (28) on quantum mechanics, the resonant transmission in tunneling or scattering is one of the more favored topics. In a one-dimensional double potential barrier, (29) the narrow central potential well has weakly-quantized (or quasi-stationary) bound states, of which the energies are denoted by  $E_1$  and  $E_2$  in Fig. 9 (a). If the energy of incident electrons coincides with these energies, the electrons may tunnel through both barriers without any attenuation. As seen in Fig. 10 (two curves at  $V = 0$ ), the transmission coefficient reaches unity at the electron energy  $E = E_1$  or  $E_2$ . Since  $E_1$  is a more strongly quantized state than  $E_2$ , the resonance peak at  $E_1$  is much sharper than that at  $E_2$ . Although this sharpness depends upon the barrier thickness, one can achieve at some energy a resonance condition of 100% transmission, whatever thickness is given to the two barriers.

This effect is quite intriguing because the transmission coefficient (or the attenuation factor) for two barriers is usually thought of as the product of two transmission coefficients, one for each barrier, resulting in a very small value for overall transmission. The situation, however, is somewhat analogous to the Fabry-Perot type interference filter in optics. The high transmissivity arises because, for certain wavelengths, the reflected waves from inside interfere destructively with the incident waves, so that only a transmitted wave remains.

This resonating condition can be extended to a periodic barrier structure. In the Kronig-Penney model of a one-dimensional crystal which consists of a series of equally-spaced potential barriers, it is well known that allowed bands of perfect transmission are separated by forbidden bands of attenuation. These one-dimensional mathematical problems can often be elegantly treated, leading to exact analytical solutions in textbooks of quantum mechanics. Many of these problems, however, are considered to be pure mathematical fantasy, far from reality.

We recently initiated an experimental project to materialize one-dimensional potential barriers in monocrystalline semiconductors in order to observe the predicted quantum-mechanical effects. (30) We choose n-type GaAs as a host semiconductor or a matrix in which potential barriers with the height of a fraction of one electron volt are made by inserting thin layers of  $\text{Ga}_{1-x}\text{Al}_x\text{As}$  or AlAs. Because of the similar properties of the chemical bond of Ga and Al together with their almost equal ion size, the introduction of AlAs into GaAs makes the least disturbance to the quality of single crystals. And yet the difference in the electronic structure between the two materials makes a sharp potential barrier inside the host semiconductor. We prepare the multi-layer structure with the technique of molecular beam epitaxy in ultra-high vacuum environment. Precise control of thickness and composition has been achieved by using a process control computer. (31)

With this facility, a double potential barrier structure has been prepared, (32) in which the barrier height and width are about 0.4 eV and a few tens of angstroms, respectively, and the width of the central well is as narrow as 40-50 Å. From these data, the first two energies,  $E_1$  and  $E_2$ , of the weakly-quantized states in the well are estimated to be 0.08 and 0.30 eV.

We have measured the current-voltage characteristic as well as the conductance  $dZ/dV$  as a function of applied voltages in this double tunnel junction. The results at 77 K are shown in Fig. 11, and they clearly indicate two singularities in each polarity and even show a negative resistance around +0.8 volt or -0.55 volt. The applied voltages at the singularities, averaged in both polarities, are roughly twice as much as the calculated bound-state energies. This general feature is not much different at 4.2 K, although no structure is seen at room temperature.

The energy diagrams at zero bias and at applied voltages  $V_1$ ,  $V_2$  and  $V_3$  are shown in Fig. 9. The electron densities on both the left and right GaAs sides are about  $10^{18} \text{ cm}^{-3}$  which gives a fermi energy of 0.04 eV at zero temperature. These electrons are considered to be classical free carriers with the effective mass,  $m^*$ , of which the kinetic energy  $E$  is given by

$$E = \frac{\hbar^2}{2m^*} (k_x^2 + k_y^2 + k_z^2).$$

On the other hand, the electrons in the central well have the weakly-quantized levels,  $E_1$ ,  $E_2$ , . . . , for motion in the x direction perpendicular to the walls with a continuum for motion in the y-z plane parallel to the walls. These

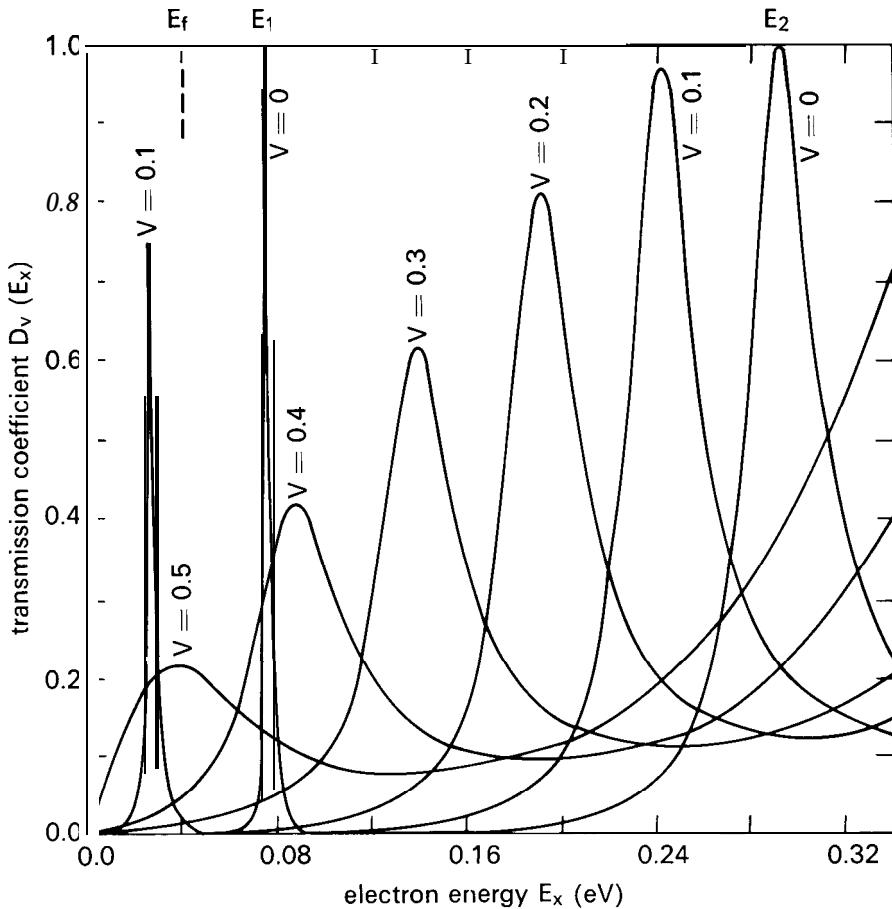


Fig. 10. Transmission coefficient versus electron energy, indicating the resonant transmission.

electrons are nearly two-dimensional, which is to say the kinetic energy  $E$  is given by

$$E = E_t + \frac{\hbar^2}{2m^*} (k_y^2 + k_z^2)$$

An approximation is made that the same electron effective mass,  $m^*$ , exists throughout the structure. Then an expression for the tunneling current in this structure (33) can be derived in the framework of the previously-described tunneling formalism in Eq. 2. Using  $\partial E / \partial k_x = \partial E_x / \partial k_x$ ,  $2\pi k_t dk_t = dk_y dk_z$  and  $T$  (temperature) = 0, the current is given by

$$\begin{aligned} J &= e/2\pi^2\hbar \int_0^{E_f} D(E_x) \int_0^{(2m^*(E_f - E_x))^{1/2}\hbar} k_t dk_t dE_x \\ &\quad - e/2\pi^2\hbar \int_0^{E_f - eV} D(E_x) \int_0^{(2m^*(E_f - E_x - eV))^{1/2}/\hbar} k_t dk_t dE_x, \end{aligned} \quad (3)$$

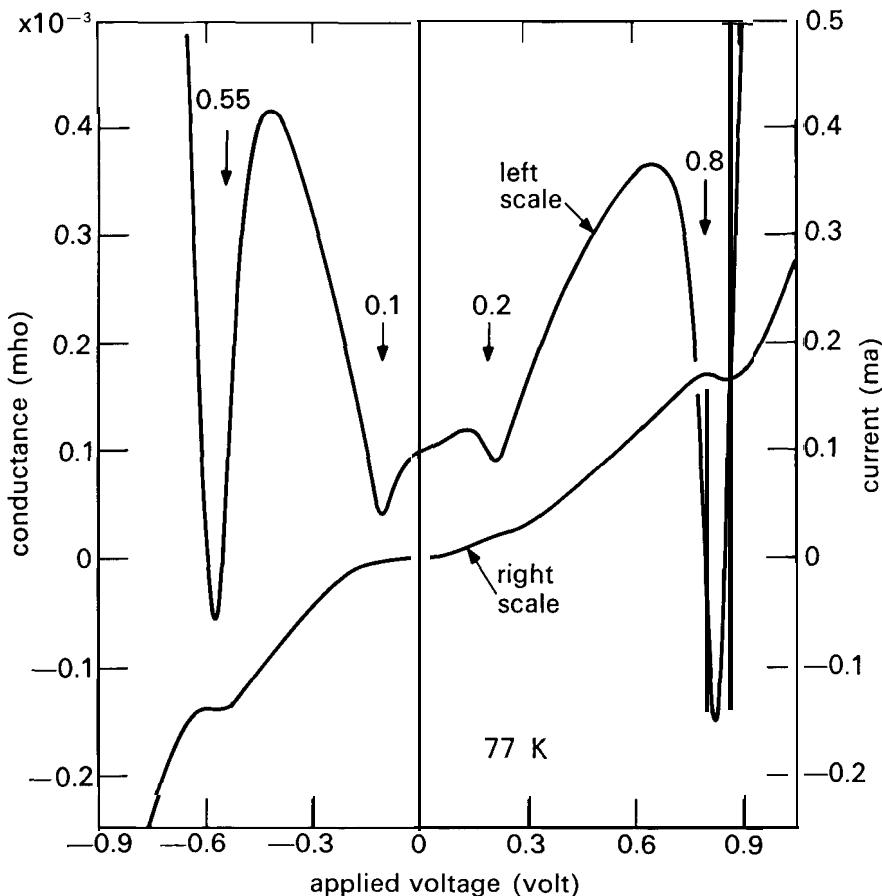


Fig. 11. Current,  $I$ , and conductance,  $dI/dV$ , versus voltage curves in a double barrier tunnel junction.

where  $V$  is the applied voltage, on which the transmission coefficient  $D(E_x)$  depends. The above expression can be integrated over the transverse wave number  $k_x$ , giving

$$\begin{aligned} \mathcal{J} = & em^*/2\pi^2\hbar^3 \int_0^{E_f} D_V(E_x)(E_f - E_x) dE_x \\ & - em^*/2\pi^2\hbar^3 \int_0^{E_f - eV} D_V(E_x)(E_f - E_x - eV) dE_x \end{aligned} \quad (4)$$

In both Eqs. 3 and 4, the second term is nonzero only for  $eV < E_f = 0.04$  eV.

Now, the transmission coefficient  $D_{v_i}(E_x)$  can be derived for each applied voltage from wave functions which are constructed by matching their values as well as derivatives at each boundary. Figure 10 shows one example of calculated  $D$  as a function of the electron energy for applied voltages

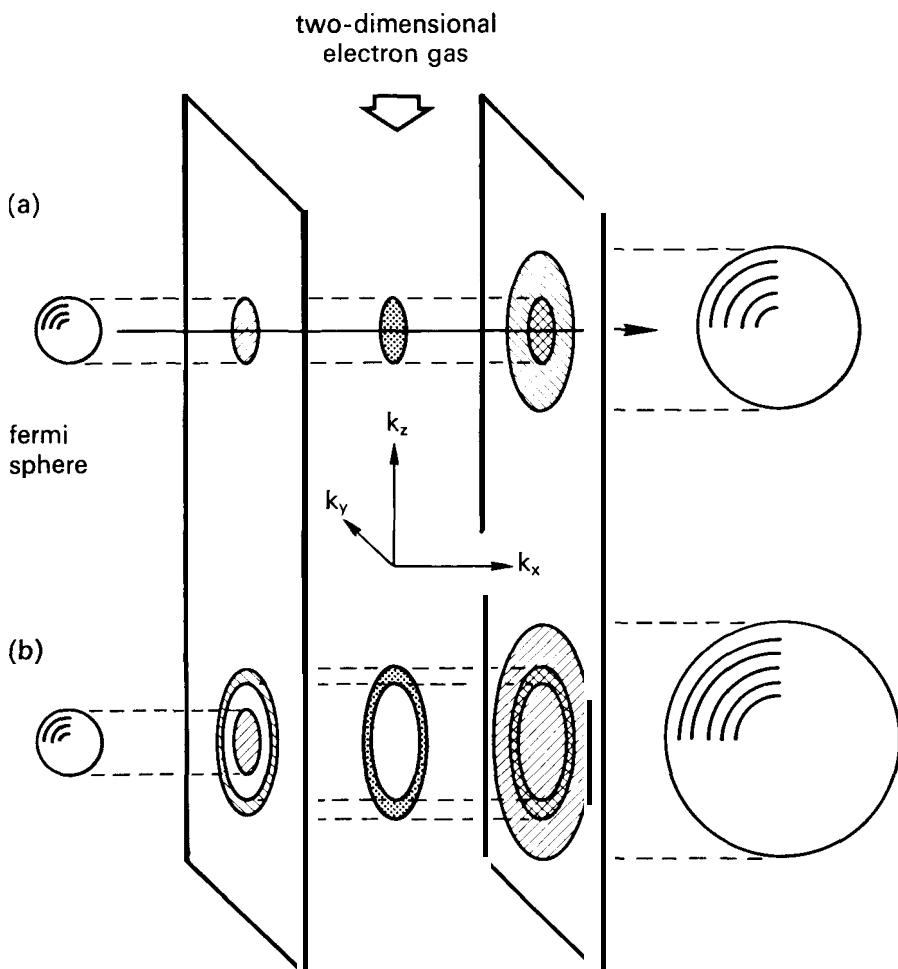


Fig. 12. Construction of shadows of energy surfaces on two  $k_y$ - $k_z$  planes corresponding to two barriers.

between zero and 0.5 volt. The energy zero is taken at the bottom of the conduction band on the left as shown in Fig. 9. In this example, the well width is taken to be 45 Å and the barrier height 0.4 eV at zero bias. The square shape for barriers and well is assumed for simplicity of calculation, although they are actually trapezoidal at any applied voltage.

Referring to Figs. 9 and 10, both the absolute values and the positions in energy for the maxima of the transmission coefficient decrease with increasing applied voltages, the origin of energy being the conduction band edge for the left outer GaAs layer. The current maxima occur at applied voltages such that the electron energies on the left coincide with the bound-state energies, as illustrated in Figs. 9 (b) and (d). This resonant transmission has been experimentally verified as shown in Fig. 11. The transmission coefficient itself at this resonance, however, is appreciably less than unity as indicated in

Fig. 10, primarily because of the asymmetric nature of the potential profile at applied voltages.

To gain an insight into this tunneling problem, particularly in view of the transverse wave-vector conservation (specular tunneling), a representation in the wave-vector space is useful and is shown in Fig. 12. Two  $k_y$ - $k_z$  planes are shown parallel to the junction plane, corresponding to the two barriers. Figures 12 (a) and (b) show two different bias-voltage conditions. First, the Fermi sphere on the left is projected on the first screen, making a circle. A similar projection, of the two-dimensional electrons in the central well which have the same total energies as electrons in the Fermi sphere on the left at the particular applied voltage, will form a circle (Fig. 12 (a)), or a ring (Fig. 12 (b)), depending upon the value of applied voltage. If the two projected patterns have no overlap, there will be no specular tunneling current. The situation on the right screen is slightly different, since an energy sphere on the right, in which electrons have the same total energies as electrons in the Fermi sphere on the left, is rather large; moreover, its size will be increased as the applied voltage increases. Thus in this case the two projected patterns always overlap. Figures 12 (a) and (b) correspond to the bias conditions in Figs. 9 (b) and (c), respectively. With an increase in applied voltage from  $V_1$  to  $V_2$ , the current will decrease because of a disappearance of overlapping regions, thereby causing a negative resistance. Since the current density is dependent upon the half-width of the resonant peaks shown in Fig. 10, we have observed a clear negative resistance associated with the second bound-state which is not swamped by possible excess currents arising for a variety of reasons.

## V. PERIODIC STRUCTURE-SUPERLATTICE

The natural extension of double barriers will be to construct a series of tunnel junctions by a periodic variation of alloy composition. (30) By using the same facilities for computer-controlled molecular beam epitaxy, we tried to prepare a Kronig-Penney type one-dimensional periodic structure-a man-made superlattice with a period of 100Å. (31) The materials used here are again GaAs and AlAs or  $\text{Ga}_{1-x}\text{Al}_x\text{As}$ .

The composition profile of such a structure (34) has been verified by the simultaneous use of ion sputter-etching of the specimen surface and Auger electron spectroscopy and is shown in Fig. 13. The amplitudes of the Al Auger signals serve as a measure of Al concentration near the surface within a sampling depth of only 10Å or so. The damping of the oscillatory behavior evident in the experimental data is not due to thermal diffusion or other reasons but due to a surface-roughening effect or non-uniformity in the sputter-etching process. The actual profile, therefore, is believed to be one which is illustrated by the solid line in Fig. 13. This is certainly one of the highest resolution structures ever built in monocrystalline semiconductors.

It should be recognized that the period of this superlattice is - 100Å--still large in comparison with the crystal lattice constant. If this period  $t$ , how-

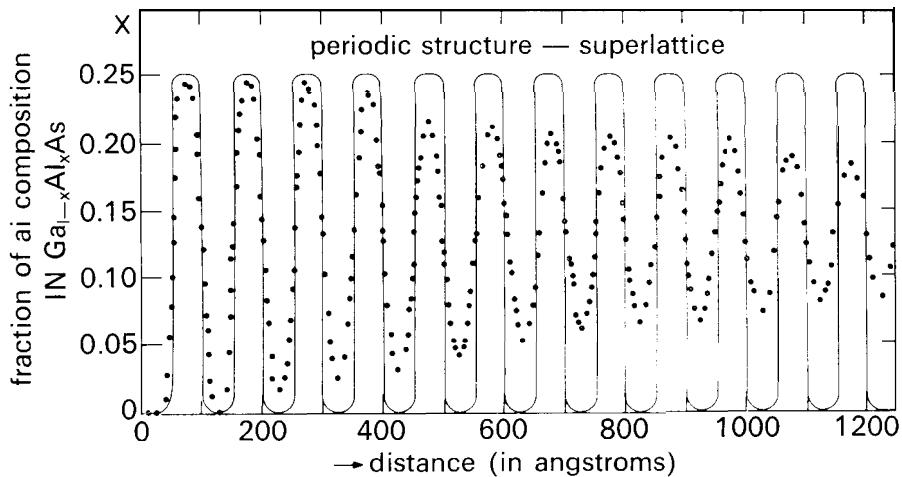


Fig. 13. Composition profile of a superlattice structure measured by a combination of ion sputter-etching and Auger electron spectroscopy.

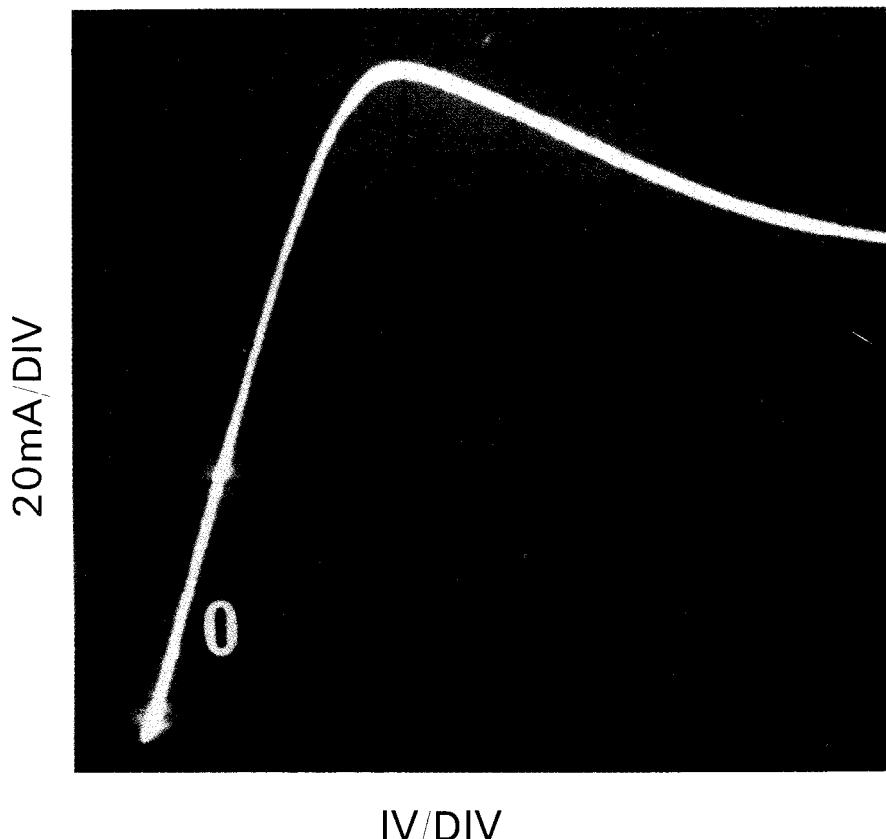


Fig. 14. Current-voltage characteristic at room temperature of a 70Å-period, GaAs-Ga<sub>0.9</sub>Al<sub>0.1</sub>As superlattice.

ever, is still shorter than the electron mean free path, a series of narrow allowed and forbidden bands is expected, due to the subdivision of the original Brillouin zone into a series of minizones. If the electron scattering time  $\tau$ , and an applied electric field  $F$ , meet a threshold condition:  $eF\tau/\hbar > 1$ , the combined effect of the narrow energy band and the narrow wave-vector zone makes it possible for electrons to be excited beyond an inflection point in the energy-wave vector relation. This would result in a negative resistance for electrical transport in the direction of the superlattice. This can be seen in another way. The de Broglie wavelength of conduction electrons having an energy of, for instance, 0.03 eV in n-type GaAs (the effective mass  $\sim 0.1$  m), is of the order of 200 Å. Therefore, an interaction of these electron waves with the Kronig-Penney type potential with a period of 100 Å can be expected, and will give rise to a nonlinear transport property.

We have begun to observe such current-voltage characteristic as shown in Fig. 14. The observed negative resistance may be interesting not only from the scientific aspect but also from a practical viewpoint because one can expect, at least theoretically, that the upper limit of operating frequencies would be higher than that for any known semiconductor devices.

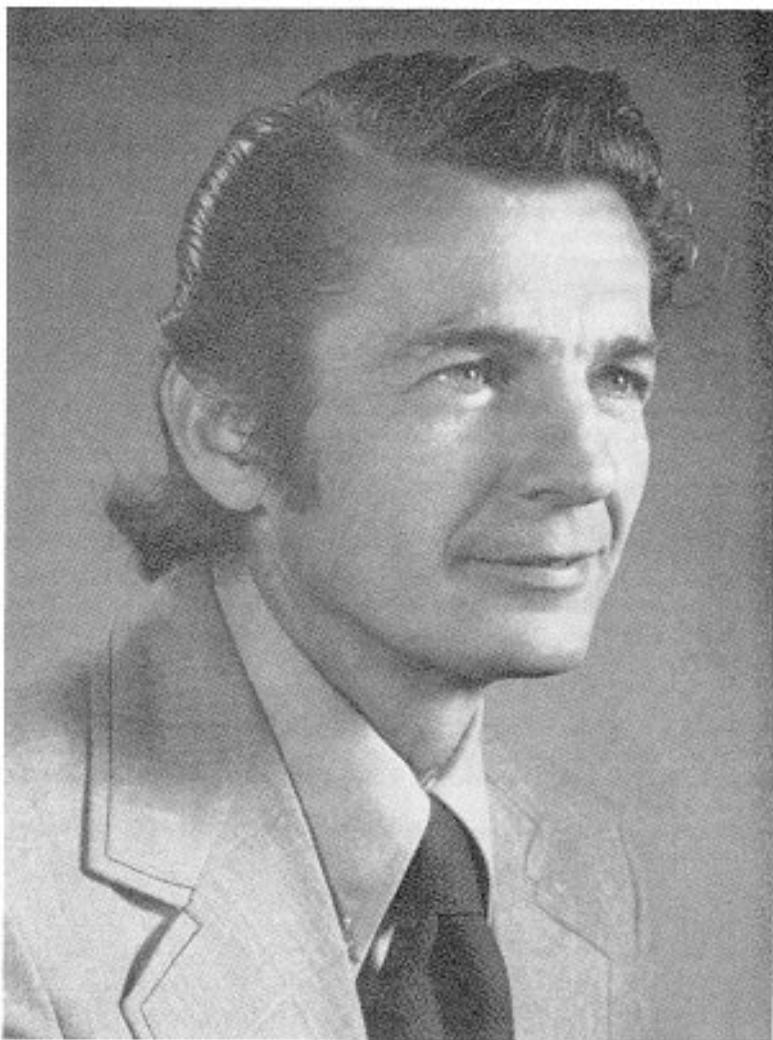
## VI. CONCLUSION

I am, of course, deeply aware of important contributions made by many colleagues and my friends throughout this long journey. The subject of Section II was carried out when I was in Japan and all the rest (35) has been performed in the United States of America. Since my journey into tunneling is still continuing, I do not come to any conclusions in this talk. However, I would like to point out that many high barriers exist in this world: Barriers between nations, races and creeds. Unfortunately, some barriers are thick and strong. But I hope, with determination, we will find a way to tunnel through these barriers easily and freely, to bring the world together so that everyone can share in the legacy of Alfred Nobel.

## REFERENCES

1. de Broglie, L., Nature 112, 540 (1923); Ann. de Physique (10 Ser.) 3, 22 (1925).
2. Schrödinger, E., Ann. d. Physik (4. Folge) 79, 361, 489 (1926).
3. Oppenheimer, J. R., Phys. Rev. 31, 66 (1928); Proc. Nat'l. Acad. Sci. U.S. 14, 363 (1928).
4. Fowler, R. H. and Nordheim, L., Proc. Roy. Soc. (London) A 119, 173 (1928).
5. Lilienfeld, J. E. Physik. Z. 23, 506 (1922).
6. Gamow, G., Physik, Z., 51, 204 (1928).
7. Gurney, R.W. and Condon, E. U., Nature 222, 439 (1928).
8. Rice, O. K. Phys. Rev. 34, 1451 (1929).
9. Frenkel, J., Phys. Rev. 36, 1604 (1930).
10. Helm, R. and Meissner, W., Z. Physik 74, 715 (1932), 86, 787 (1933).
11. Holm, R., Electric Contacts, Springer-Verlag, New York, 1967 p.118.
12. Wilson, A. H., Proc. Roy. Soc. (London) A 136, 487 (1932).
13. Frenkel, J. and Joffe, A., Physik. Z. Sowjetunion I, 60 (1932).
14. Nordheim, L., Z. Physik 7.5, 434 (1932).

15. Zener, C., Proc. Roy. Soc. (London) 14.5, 523 (1934).
16. Houston, W. V., Phys. Rev. 57, 184 (1940).
17. McAfee, K. B., Ryder, E. J., Shockley, W. and Sparks, M., Phys. Rev. 83, 650 (1951).
18. McKay, K. G. and McAfee, K. B. Phys. Rev. 91, 1079 (1953);  
McKay, K. G., Phys. Rev. 94, 877 (1954) ;  
Miller, S. L., Phys. 99, 1234 (1955).
19. Chynoweth, A. G. and McKay., K. G. Phys., Rev. 106, 418 (1957).
20. Esaki, L., Phys. Rev. 109, 603 (1958);  
Esaki, L., Solid State Physics in Electronics and Telecommunications, Proc. of Int. Conf. Brussels, 1958 (Desirant, M. and Michels, J. L., ed.), Vol. 1, Semiconductors, Part I, Academic Press, New York, 1960, p. 514.
21. Esaki, L. and Miyahara, Y., unpublished;  
Esaki, L. and Miyahara, Y., Solid-State Electron. 1, 13 (1960) ;  
Holonyak, N., Lesk, I. A., Hall, R. N., Tiemann, J. J. and Ehrenreich, H., Phys. Rev. Letters 3, 167 (1959).
22. Macfarlane, G. G., McLean, T. P., Quarrington, J. E. and Roberts, V., J. Phys. Chem. Solids 8, 388 ( 1959).
23. Haynes, J. R., Lax, M. and Flood, W. F., J. Phys. Chem. Solids 8, 392 (1959).
24. See, for instance, Tunneling Phenomena in Solids edited by Burstein, E. and Lundqvist, S., Plenum Press, New York, 1969.
25. Duke, C. B., Tunneling in Solids, Academic Press, New York, 1969.
26. Harrison, W. A., Phys. Rev. 123, 85 (1961).
27. Esaki, L. and Stiles, P. J., Phys. Rev. Letters 16, 1108 (1966) ; Chang, L. L., Stiles, P. J. and Esaki, L., J. Appl. Phys. 38, 4440 (1967) ; Esaki, L., J. Phys. Soc. Japan, Suppl. 21, 589 (1966).
28. See, for instance, Bohm, D., Quantum Theory, Prentice-Hall, New Jersey, 1951.
29. Kane, E. O., page 9-l 1 in reference 24.
30. Esaki, L. and Tsu, R., IBM J. Res. Develop. 24, 61 ( 1970); Esaki, L., Chang, L. L., Howard, W. E. and Rideout, V. L., Proc. 11th Int. Conf. Phys. Semicond., Warsaw, Poland, 1972, p. 431.
31. Chang, L. L., Esaki, L., Howard, W. E. and Ludeke, R., J. Vac. Sci. Technol. 10, 11 (1973);  
Chang, L. L., Esaki, L., Howard, W. E., Ludeke, R. and Schul, G., J. Vac. Sci. Technol. 10, 655 (1973).
32. Chang, L. L., Esaki, L. and Tsu, R., to be published.
33. Tsu, R. and Esaki, L., Appl. Phys. Letters 22, 562 (1973).
34. Ludeke, R., Esaki, L. and Chang, L. L., to be published.
35. Partly supported by Army Research Office, Durham, N. C.



Ivan Adair

## IVAR GIAEVER

Ivar Giaever was born in Bergen, Norway, April 5, 1929, the second of three children. He grew up in Toten where his father, John A. Giaever, was a pharmacist. He attended elementary school in Toten but received his secondary education in the city of Hamar. Next he worked one year at the Raufoss Munition Factories before entering the Norwegian Institute of Technology in 1948. He graduated in 1952 with a degree in mechanical engineering.

In 1953, Giaever completed his military duty as a corporal in the Norwegian Army, and thereafter he was employed for a year as a patent examiner for the Norwegian Government.

Giaever emigrated to Canada in 1954 and after a short period as an architect's aide he joined Canadian General Electric's Advanced Engineering Program. In 1956, he emigrated to the USA where he completed the General Electric Company's A, B and C engineering courses. In these he worked in various assignments as an applied mathematician. He joined the General Electric Research and Development Center in 1958 and concurrently started to study physics at Rensselaer Polytechnical Institute where he obtained a Ph.D. degree in 1964.

From 1958 to 1969 Dr. Giaever worked in the fields of thin films, tunneling and superconductivity. In 1965 he was awarded the Oliver E. Buckley Prize for some pioneering work combining tunneling and superconductivity. In 1969 he received a Guggenheim Fellowship and thereupon spent one year in Cambridge, England studying biophysics. Since returning to the Research and Development Center in 1970, Dr. Giaever has spent most of his effort studying the behavior of protein molecules at solid surfaces. In recognition of his work he was elected a Coolidge fellow at General Electric in May, 1973.

Dr Giaever is a member of the Institute of Electrical and Electronic Engineers, and the Biophysical Society, and he is a Fellow of the American Physical Society. Dr. Giaever has served on committees for several international conferences and presently he is a member of the Executive Committee of the Solid State division in the American Physical Society.

Ivar Giaever married Inger Skramstad in 1952 and they have four children. He became a naturalized US citizen in 1964.

*Notes added*

Linus Pauling is reported to have said that the Nobel Prize did not change his life - he was already famous! That was not true for me. The Nobel Prize opened a lot of doors, but also provided me with many distractions. I have, however, continued to work in biophysics, attempting to use physical methods and thoughts to solve biological problems. At the present time, I am studying the motion of mammalian cells in tissue culture by growing both normal and cancerous cells on small electrodes.

I left General Electric in 1988 to become an Institute Professor at Rensselaer (RPI) in Troy, New York 12180-3590, and concurrently I am also a Professor at the University of Oslo, Norway, sponsored by STATOIL.

On a personal note my wife and I are now the proud grandparents of almost four grandchildren.

# ELECTRON TUNNELING AND SUPERCONDUCTIVITY

Nobel Lecture, December 12, 1973

by

**IVAR GLÆVER**

General Electric Research and Development Center, Schenectady, N.Y., USA.

In my laboratory notebook dated May 2, 1960 is the entry: "Friday, April 22, I performed the following experiment aimed at measuring the forbidden gap in a superconductor." This was obviously an extraordinary event not only because I rarely write in my notebook, but because the success of that experiment is the reason I have the great honor and pleasure of addressing you today. I shall try in this lecture, as best I can, to recollect some of the events and thoughts that led to this notebook entry, though it is difficult to describe what now appears to me as fortuitous. I hope that this personal and subjective recollection will be more interesting to you than a strictly technical lecture, particularly since there are now so many good review articles dealing with superconductive tunneling.<sup>1,2</sup>

A recent headline in an Oslo paper read approximately as follows: "Master in billiards and bridge, almost flunked physics-gets Nobel Prize." The paper refers to my student days in Trondheim. I have to admit that the reporting is reasonably accurate, therefore I shall not attempt a "cover up", but confess that I almost flunked in mathematics as well. In those days I was not very interested in mechanical engineering and school in general, but I did manage to graduate with an average degree in 1952. Mainly because of the housing shortage which existed in Norway, my wife and I finally decided to emigrate to Canada where I soon found employment with Canadian General Electric. A three year Company course in engineering and applied mathematics known as the A, B and C course was offered to me. I realized this time that school was for real, and since it probably would be my last chance, I really studied hard for a few years.

When I was 28 years old I found myself in Schenectady, New York where I discovered that it was possible for some people to make a good living as physicists. I had worked on various Company assignments in applied mathematics, and had developed the feeling that the mathematics was much more advanced than the actual knowledge of the physical systems that we applied it to. Thus, I thought perhaps I should learn some physics and, even though I was still an engineer, I was given the opportunity to try it at the General Electric Research Laboratory.

The assignment I was given was to work with thin films and to me films meant photography. However I was fortunate to be associated with John Fisher who obviously had other things in mind. Fisher had started out as a mechanical engineer as well, but had lately turned his atten-

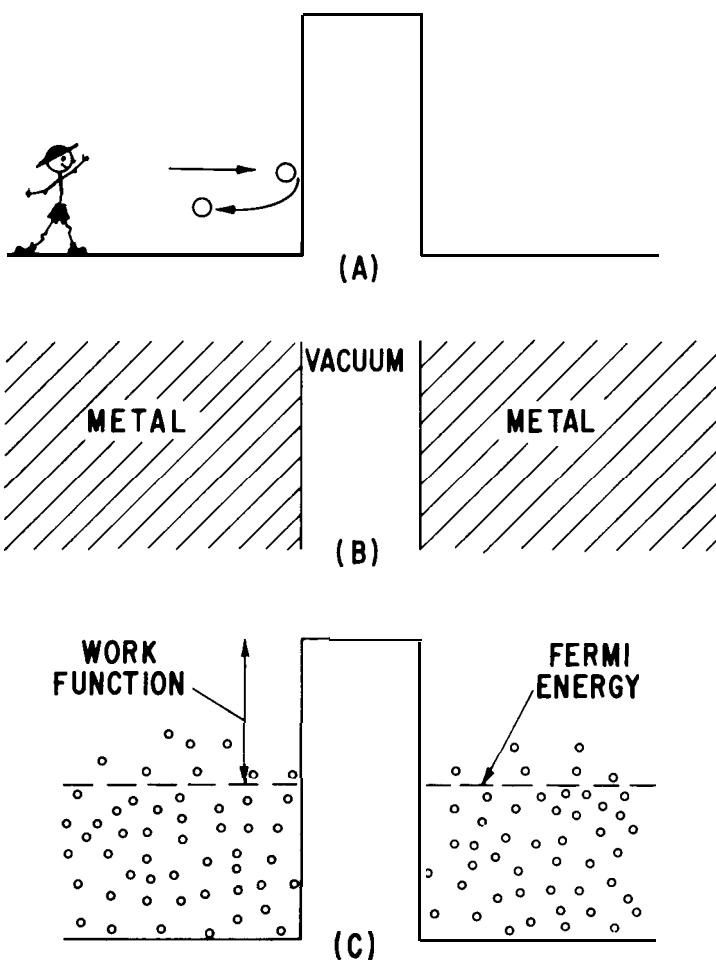


Fig. 1.

A. If a man throws a ball against a wall the ball bounces back. The laws of physics allow the ball to penetrate or tunnel through the wall but the chance is infinitesimally small because the ball is a macroscopic object. B. Two metals separated by a vacuum will approximate the above situation. The electrons in the metals are the "balls", the vacuum represents the wall. C. A pictorial energy diagram of the two metals. The electrons do not have enough energy to escape into the vacuum. The two metals can, however, exchange electrons by tunneling. If the metals are spaced close together the probability for tunneling is large because the electron is a microscopic particle.

tion towards theoretical physics. He had the notion that useful electronic devices could be made using thin film technology and before long I was working with metal films separated by thin insulating layers trying to do tunneling experiments. I have no doubt that Fisher knew about Leo Esaki's tunneling experiments at that time, but I certainly did not. The concept that a particle can go through a barrier seemed sort of strange to me, just struggling with quantum mechanics at Rensselaer Polytechnic Institute in Troy, where I took formal courses in Physics. For an engineer it sounds rather strange that if you

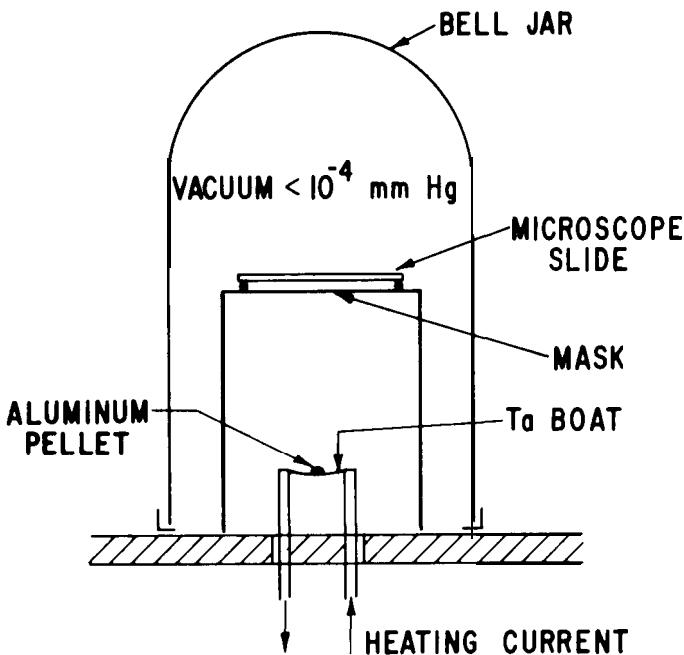


Fig. 2.

A schematic drawing of a vacuum system for depositing metal films. For example, if aluminum is heated resistively in a tantalum boat, the aluminum first melts, then boils and evaporates. The aluminum vapor will solidify on any cold substrate placed in the vapor stream. The most common substrates are ordinary microscope glass slides. Patterns can be formed on the slides by suitably shielding them with a metal mask.

throw a tennis ball against a wall enough times it will eventually go through without damaging either the wall or itself. That must be the hard way to a Nobel Prize! The trick, of course, is to use very tiny balls, and lots of them. Thus if we could place two metals very close together without making a short, the electrons in the metals can be considered as the balls and the wall is represented by the spacing between the metals. These concepts are shown in Figure 1. While classical mechanics correctly predicts the behavior of large objects such as tennis balls, to predict the behavior of small objects such as electrons we must use quantum mechanics. Physical insight relates to everyday experiences with large objects, thus we should not be too surprised that electrons sometimes behave in strange and unexpected ways.

Neither Fisher nor I had much background in experimental physics, none to be exact, and we made several false starts. To be able to measure a tunneling current the two metals must be spaced no more than about 100 Å apart, and we decided early in the game not to attempt to use air or vacuum between the two metals because of problems with vibration. After all, we both had training in mechanical engineering! We tried instead to keep the two metals apart by using a variety of thin insulators made from Langmuir films and from Formvar. Invariably, these films had pinholes and the mercury

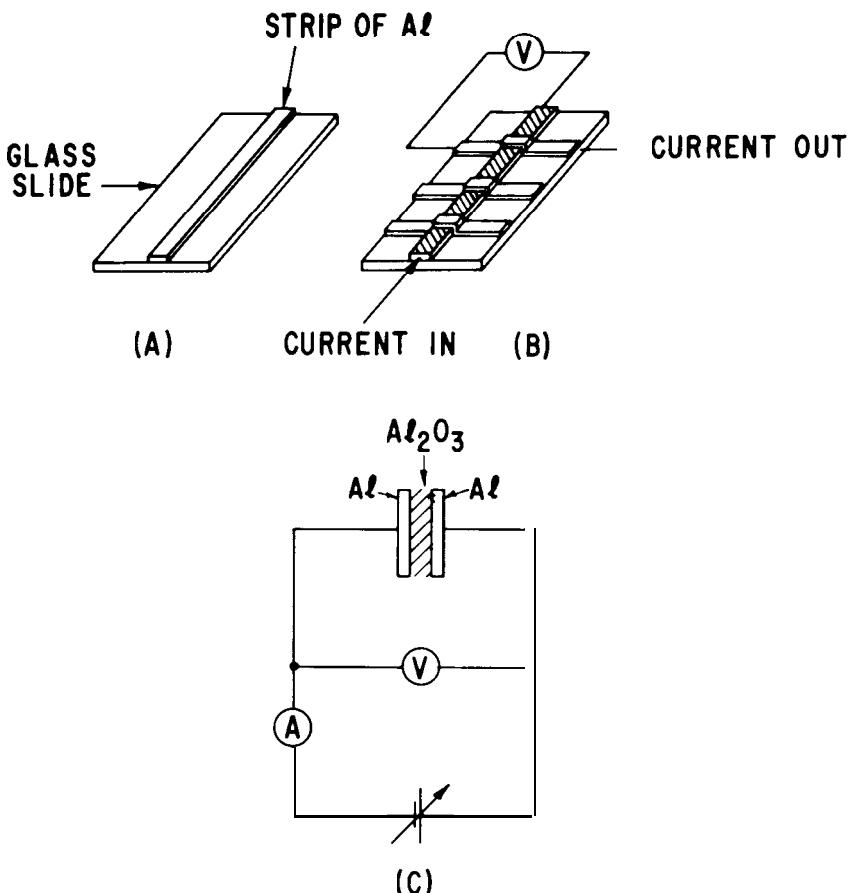


Fig. 3.

A. A microscope glass slide with a vapor deposited aluminum strip down the middle. As soon as the aluminum film is exposed to air, a protective insulating oxide forms on the surface. The thickness of the oxide depends upon such factors as time, temperature and humidity. B. After a suitable oxide has formed, cross strips of aluminum are evaporated over the first film, sandwiching the oxide between the two metal films. Current is passed along one aluminum film up through the oxide and out through the other film, while the voltage drop is monitored across the oxide. C. A schematic circuit diagram. We are measuring the current-voltage characteristics of the capacitor-like arrangement formed by the two aluminum films and the oxide. When the oxide thickness is less than  $50\text{\AA}$  or so, an appreciable dc current will flow through the oxide.

counter electrode which we used would short the films. Thus we spent some time measuring very interesting but always non-reproducible current-voltage characteristics which we referred to as miracles since each occurred only once. After a few months we hit on the correct idea: to use evaporated metal films and to separate them by a naturally grown oxide layer.

To carry out our ideas we needed an evaporator, thus I purchased my first piece of experimental equipment. While waiting for the evaporator to arrive I worried a lot-1 was afraid I would get stuck in experimental physics tied

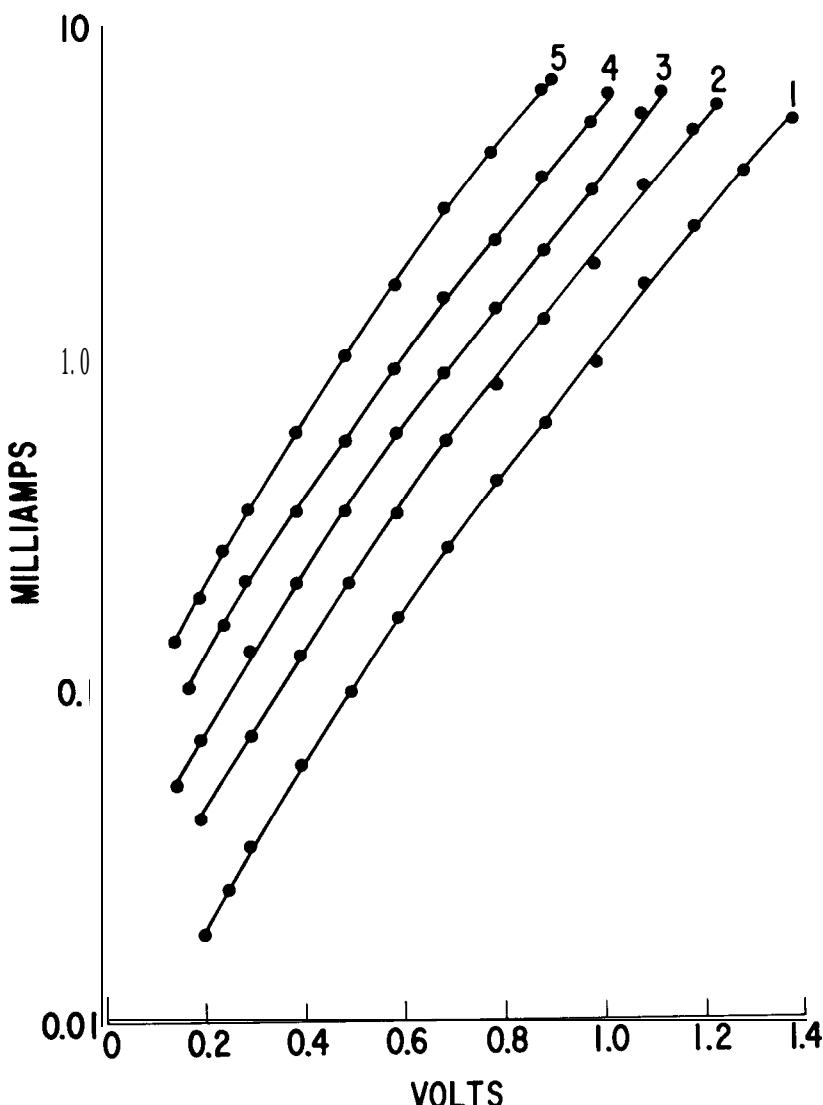


Fig. 4.

Current-voltage characteristics of five different tunnel junctions all with the same thickness, but with five different areas. The current is proportional to the area of the junction. This was one of the first clues that we were dealing with tunneling rather than shorts. In the early experiments we used a relatively thick oxide, thus very little current would flow at low voltages.

down to this expensive machine. My plans at the time were to switch into theory as soon as I had acquired enough knowledge. The premonition was correct; I did get stuck with the evaporator, not because it was expensive but because it fascinated me. Figure 2 shows a schematic diagram of an evaporator. To prepare a tunnel junction we first evaporated a strip of aluminum onto a glass slide. This film was removed from the vacuum system and heated

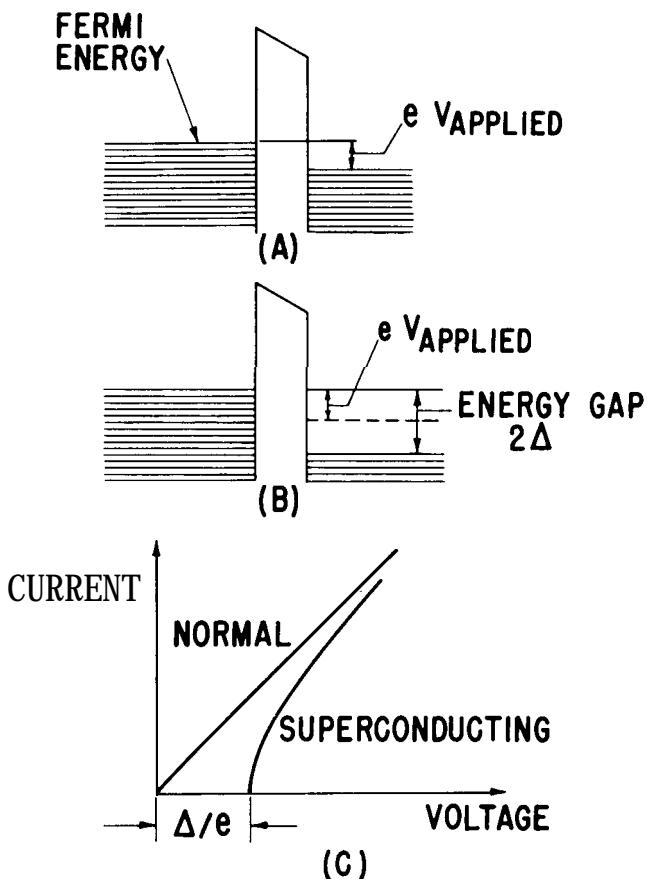


Fig. 5.

A. An energy diagram of two metals separated by a barrier. The Fermi energies in the two metals are at different levels because of the voltage difference applied between the metals. Only the left metal electrons in the energy range  $e \cdot V_{App}$  can make a transition to the metal on the right, because only these electrons face empty energy states. The Pauli Principle allows only one electron in each quantum state. B. The right-hand metal is now superconducting, and an energy gap  $2\Delta$  has opened up in the electron spectrum. No single electron in a superconductor can have an energy such that it will appear inside the gap. The electrons from the metal on the left can still tunnel through the barrier, but they cannot enter into the metal on the right as long as the applied voltage is less than  $\Delta/e$ , because the electrons either face a filled state or a forbidden energy range. When the applied voltage exceeds  $\Delta/e$ , current will begin to flow. C. A schematic current-voltage characteristic. When both metals are in the normal state the current is simply proportional to the voltage. When one metal is superconducting the current-voltage characteristic is drastically altered. The exact shape of the curve depends on the electronic energy spectrum in the superconductor.

to oxidize the surface rapidly. Several cross strips of aluminum were then deposited over the first film making several junctions at the same time. The steps in the sample preparation are illustrated in Figure 3. This procedure solved two problems, first there were no pinholes in the oxide because it is

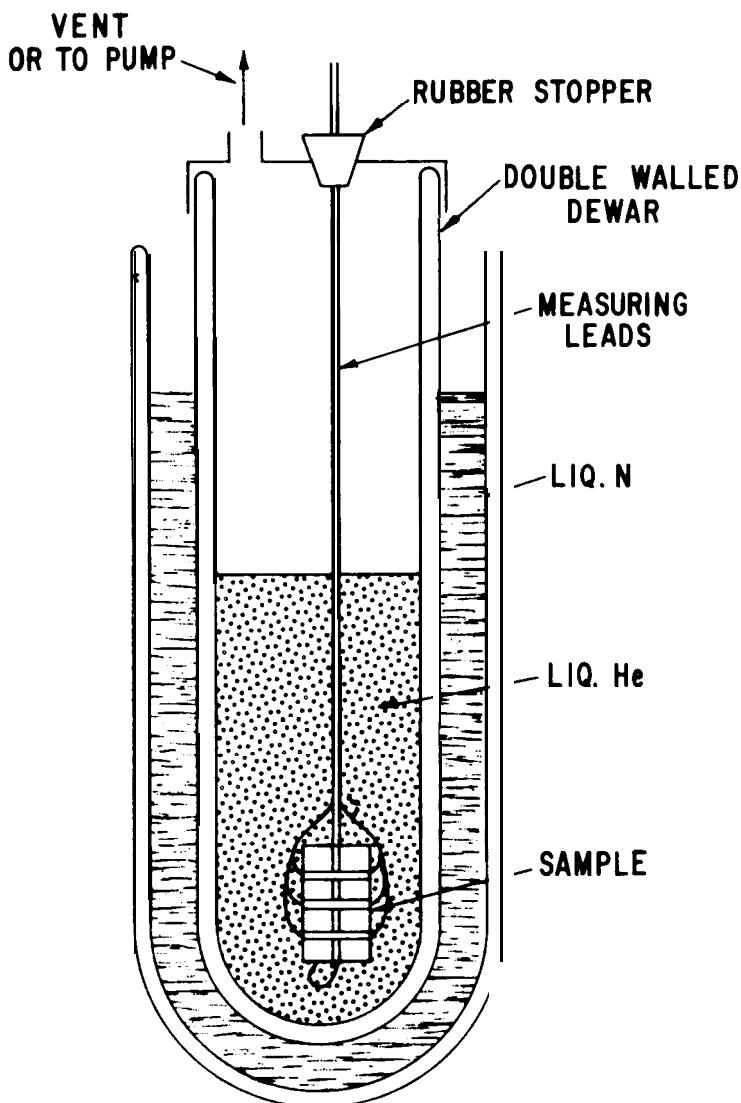


Fig. 6.

A standard experimental arrangement used for low temperature experiments. It consists of two dewars, the outer one contains liquid nitrogen, the inner one, liquid helium. Helium boils at  $4.2^\circ\text{K}$  at atmospheric pressure. The temperature can be lowered to about  $1^\circ\text{K}$  by reducing the pressure. The sample simply hangs into the liquid helium supported by the measuring leads.

self-healing, and second we got rid of mechanical problems that arose with the mercury counter electrode.

By about April, 1959, we had performed several successful tunneling experiments. The current-voltage characteristics of our samples were reasonably reproducible, and conformed well to theory. A typical result is shown in Figure 4. Several checks were done, such as varying the area and the oxide

thickness of the junction as well as changing the temperature. Everything looked OK, and I even gave a seminar at the Laboratory. By this time, I had solved Schrodinger's equation enough times to believe that electrons sometimes behave as waves, and I did not worry much about that part anymore.

However: there were many real physicists at the Laboratory and they properly questioned my experiment. How did I know I did not have metallic shorts? Ionic current? Semiconduction rather than tunneling? Of course, I did not know, and even though theory and experiments agreed well, doubts about the validity were always in my mind. I spent a lot of time inventing impossible schemes such as a tunnel triode or a cold cathode, both to try to prove conclusively that I dealt with tunneling and to perhaps make my work useful. It was rather strange for me at that time to get paid for doing what I considered having fun, and my conscience bothered me. But just like quantum mechanics, you get used to it, and now I often argue the opposite point; we should pay more people to do pure research.

I continued to try out my ideas on John Fisher who was now looking into the problems of fundamental particles with his characteristic optimism and enthusiasm; in addition, I received more and more advice and guidance from Charles Bean and Walter Harrison, both physicists with the uncanny ability of making things clear as long as a piece of chalk and a blackboard were available. I continued to take formal courses at RPI, and one day in a solid state physics course taught by Professor Huntington we got to superconductivity. Well, I didn't believe that the resistance drops to exactly zero—but what really caught my attention was the mention of the energy gap in a superconductor, central to the new Bardeen-Cooper-Schrieffer theory. If the theory was any good and if my tunneling experiments were any good, it was obvious to me that by combining the two, some pretty interesting things should happen, as illustrated in Figure 5. When I got back to the GE Laboratory I tried this simple idea out on my friends, and as I remember, it did not look as good to them. The energy gap was really a many body effect and could not be interpreted literally the way I had done. But even though there was considerable skepticism, everyone urged me to go ahead and make a try. Then I realized that I did not know what the size of the gap was in units I understood—electron volts. This was easily solved by my usual method: first asking Bean and then Harrison, and, when they agreed on a few millielectron volts: I was happy because that is in a easily measured voltage range.

I had never done an experiment requiring low temperatures and liquid helium—that seemed like complicated business. However one great advantage of being associated with a large laboratory like General Electric is that there are always people around who are knowledgeable in almost any field, and better still they are willing to lend you a hand. In my case, all I had to do was go to the end of the hall where Warren DeSorbo was already doing experiments with superconductors. I no longer remember how long it took me to set up the helium dewars I borrowed, but probably no longer than a day or two. People unfamiliar with low temperature work believe that the whole field of low temperature is pretty esoteric, but all it really requires is access

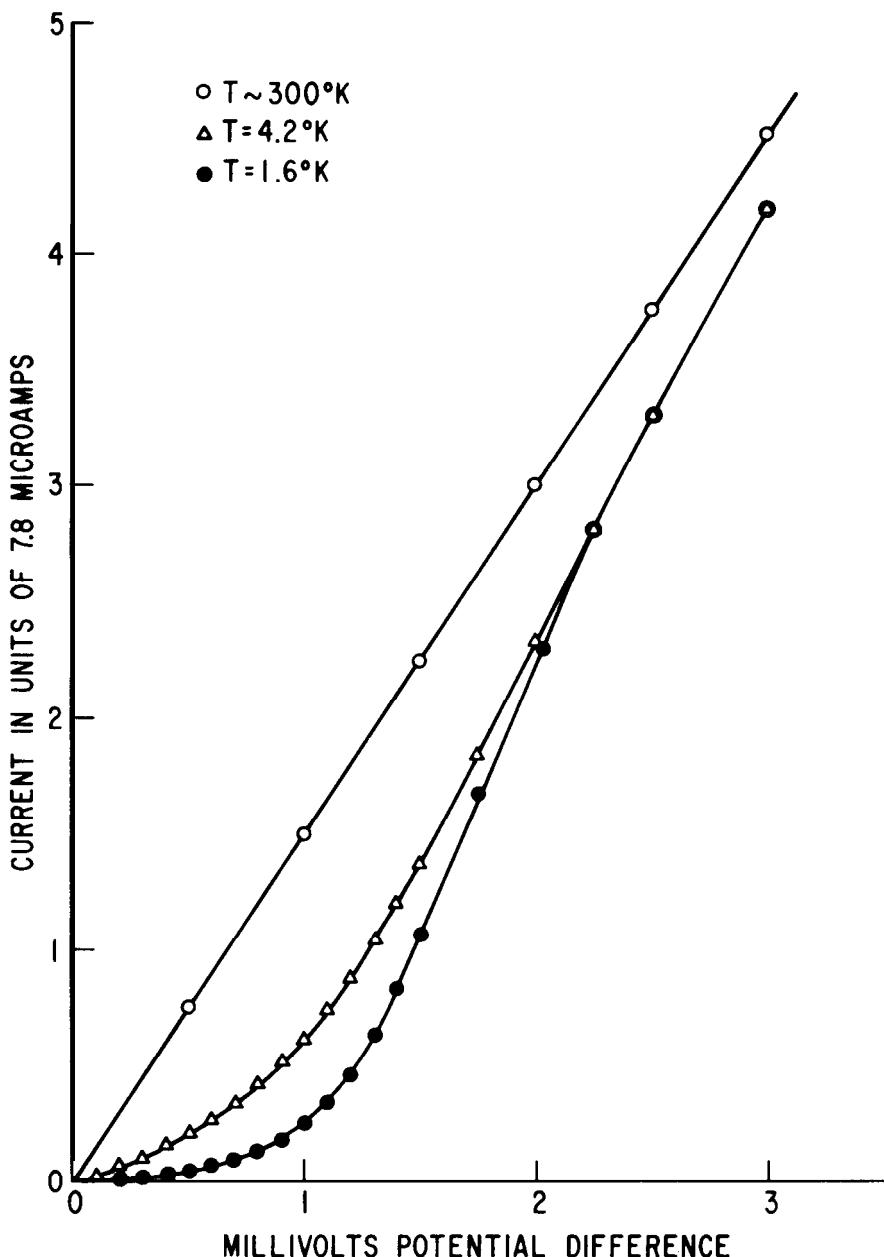


Fig. 7.

The current-voltage characteristic of an aluminum-aluminum oxide-lead sample. As soon as the lead becomes superconducting the current ceases to be proportional to the voltage. The large change between  $4.2^\circ\text{K}$  and  $1.6^\circ\text{K}$  is due to the change in the energy gap with temperature. Some current also flows at voltages less than  $\Delta/e$  because of thermally excited electrons in the conductors.

to liquid helium, which was readily available at the Laboratory. The experimental setup is shown in Figure 6. Then I made my samples using the familiar aluminum-aluminum oxide, but I put lead strips on top. Both lead and alu-

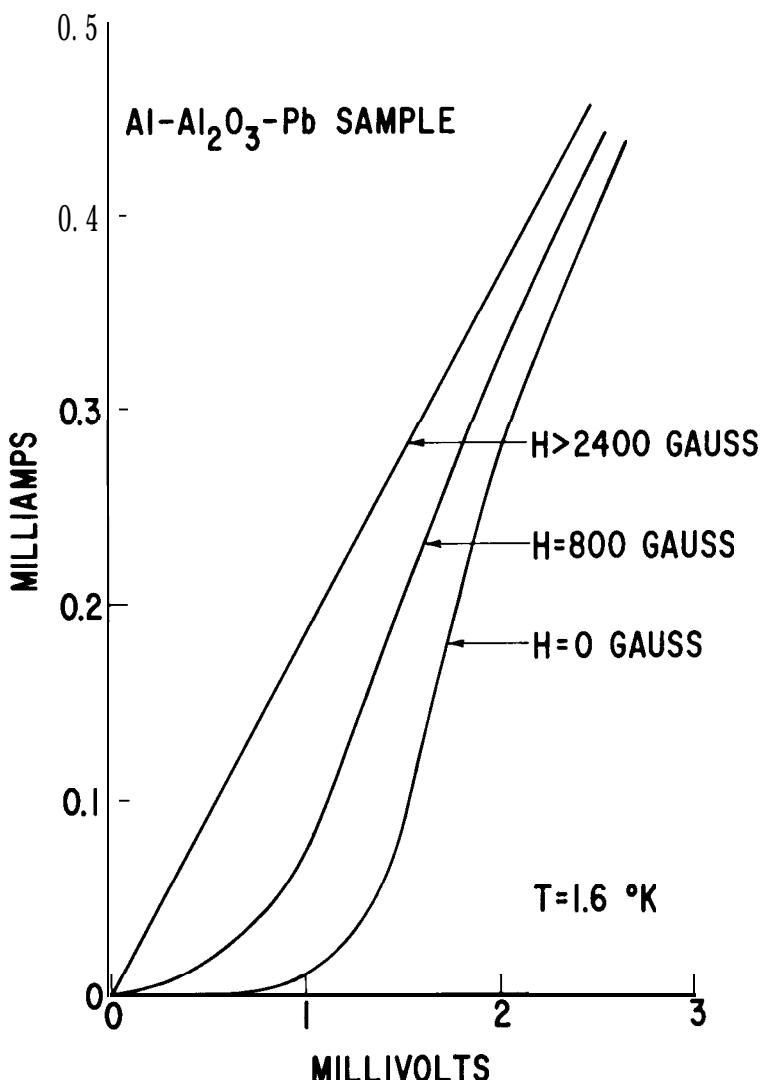


Fig. 8.

The current-voltage characteristic at  $1.6^{\circ} \text{K}$  as a function of the applied magnetic field. At 2400 gauss the films are normal, at 0 gauss the lead film is superconducting. The reason for the change in the characteristics between 800 gauss and 0 gauss is that thin films have an energy gap that is a function of the magnetic field.

minimum are superconductors, lead is superconducting at  $7.2^{\circ} \text{K}$  and thus all you need to make it superconducting is liquid helium which boils at  $4.2^{\circ} \text{K}$ . Aluminum becomes superconducting only below  $1.2^{\circ} \text{K}$ , and to reach this temperature a more complicated experimental setup is required.

The first two experiments I tried were failures because I used oxide layers which were too thick. I did not get enough current through the thick oxide to measure it reliably with the instruments I used, which were simply a standard voltmeter and a standard ammeter. It is strange to think about that



Fig. 9.

Informal discussion over a cup of coffee. From left: Ivar Giaever, Walter Harrison, Charles Bean, and John Fisher.

now, only 13 years later, when the Laboratory is full of sophisticated x-y recorders. Of course, we had plenty of oscilloscopes at that time but I was not very familiar with their use. In the third attempt rather than deliberately oxidizing the first aluminum strip, I simply exposed it to air for only a few minutes, and put it back in the evaporator to deposit the cross strips of lead. This way the oxide was no more than about  $30\text{\AA}$  thick, and I could readily measure the current-voltage characteristic with the available equipment. To me the greatest moment in an experiment is always just before I learn whether the particular idea is a good or a bad one. Thus even a failure is exciting, and most of my ideas have of course been wrong. But this time it worked! The current-voltage characteristic changed markedly when the lead changed from the normal state to the superconducting state as shown in Figure 7. That was exciting! I immediately repeated the experiment using a different sample -everything looked good! But how to make certain? It was well-known that superconductivity is destroyed by a magnetic field, but my simple setup of dewars made that experiment impossible. This time I had to go all the way across the hall where Israel Jacobs studied magnetism at low temperatures. Again I was lucky enough to go right into an experimental rig where both the temperature and the magnetic field could be controlled and I could quickly do all the proper experiments. The basic result is shown in Figure 8. Every-

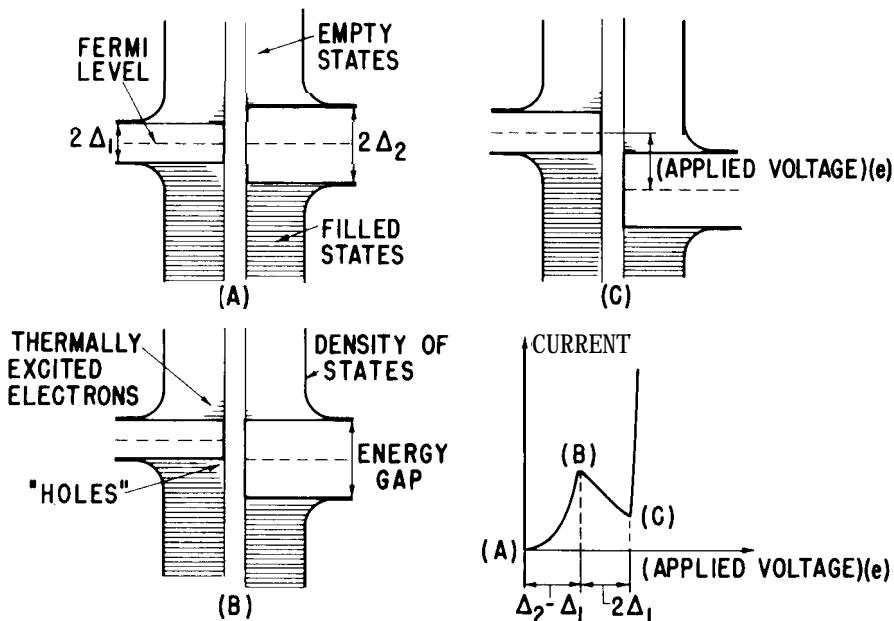


Fig. 10.

Tunneling between two superconductors with different energy gaps at a temperature larger than 0° K. A. No voltage is applied between the two conductors. B. As a voltage is applied it becomes energetically possible for more and more of the thermally excited electrons to flow from the superconductor with the smaller gap into the superconductor with the larger gap. At the voltage shown all the excited electrons can find empty states on the right. C. As the voltage is further increased, no more electrons come into play, and since the number of states the electrons can tunnel into decreases, the current will decrease as the voltage is increased. When the voltage is increased sufficiently the electrons below the gap in the superconductor on the left face empty states on the right, and a rapid increase in current will occur. D. A schematic picture of the expected current-voltage characteristic.

thing held together and the whole group, as I remember it, was very excited. In particular, I can remember Bean enthusiastically spreading the news up and down the halls in our Laboratory, and also patiently explaining to me the significance of the experiment.

I was, of course, not the first person to measure the energy gap in a superconductor, and I soon became aware of the nice experiments done by M. Tinkham and his students using infrared transmission. I can remember that I was worried that the size of the gap that I measured did not quite agree with those previous measurements. Bean set me straight with words to the effect that from then on other people would have to agree with me; my experiment would set the standard, and I felt pleased and like a physicist for the first time.

That was a very exciting time in my life; we had several great ideas to improve and extend the experiment to all sorts of materials like normal metals, magnetic materials and semiconductors. I remember many informal dis-

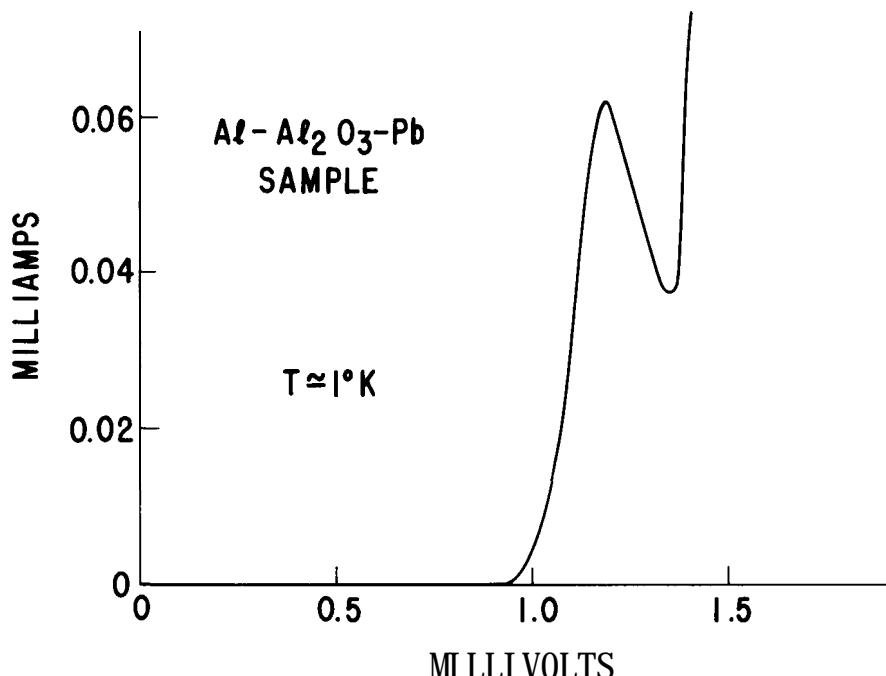


Fig. 11.

A negative resistance characteristic obtained experimentally in tunneling between two different superconductors.

cussions over coffee about what to try next and one of these sessions is in a photograph taken in 1960 which is shown in Figure 9. To be honest the picture was staged, we weren't normally so dressed up, and rarely did I find myself in charge at the blackboard! Most of the ideas we had did not work very well and Harrison soon published a theory showing that life is really complicated after all. But the superconducting experiment was charmed and always worked. It looked like the tunneling probability was directly proportional to the density of states in a superconductor. Now if this were strictly true, it did not take much imagination to realize that tunneling between two superconductors should display a negative resistance characteristic as illustrated in Figure 10. A negative resistance characteristic meant, of course, amplifiers, oscillators and other devices. But nobody around me had facilities to pump on the helium sufficiently to make aluminum become superconducting. This time I had to leave the building and reactivate an old low temperature setup in an adjacent building. Sure enough, as soon as the aluminum went superconducting a negative resistance appeared, and, indeed, the notion that the tunneling probability was directly proportional to the density of states was experimentally correct. A typical characteristic is shown in Figure 11.

Now things looked very good because all sorts of electronic devices could be made using this effect, but, of course, they would only be operative at low temperatures. We should remember that the semiconducting devices were not so advanced in 1960 and we thought that the superconducting junction

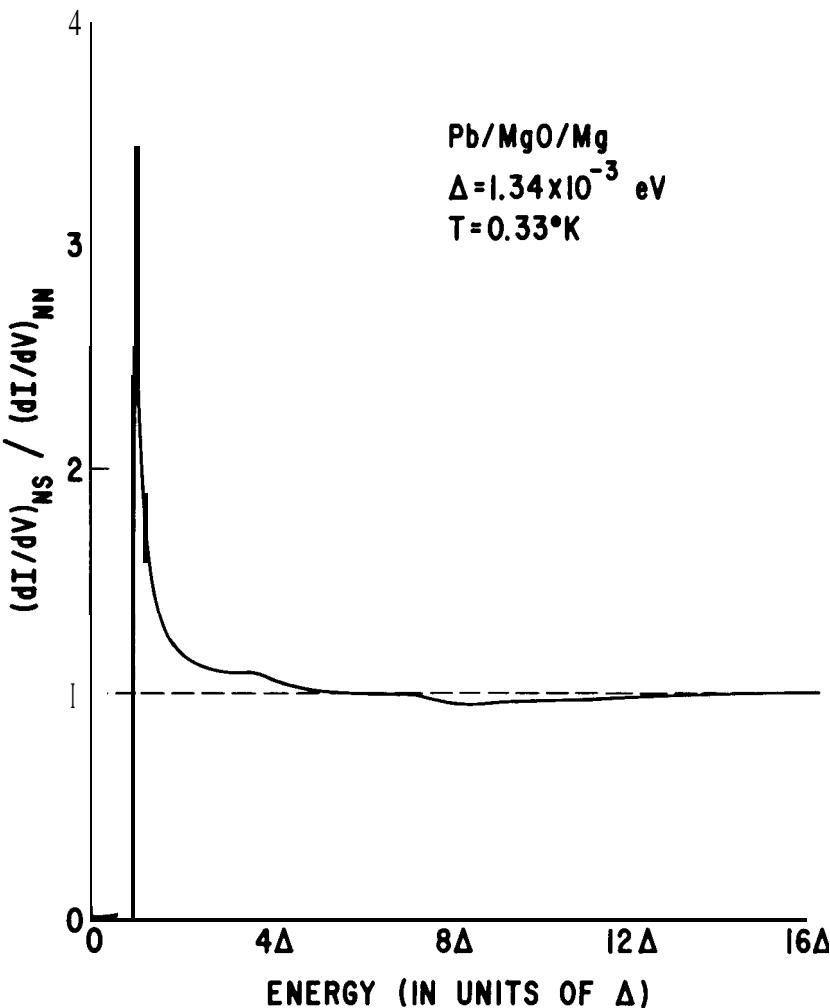


Fig. 12.

A normalized derivative of the current with respect to voltage of a lead junction at low temperature. The simple BCS-theory predicts that the derivative should approach unity asymptotically as the energy increases. Instead several wiggles are observed in the range between  $4\Delta$  and  $8\Delta$ . These wiggles are related to the phonon spectrum in lead.

would have a good chance of competing with, for example, the Esaki diode. The basic question I faced was which way to go: engineering or science? I decided that I should do the science first, and received full support from my immediate manager, Roland Schmitt.

In retrospect I realize how tempting it must have been for Schmitt to encourage other people to work in the new area, and for the much more experienced physicists around me to do so as well. Instead, at the right time, Schmitt provided me with a co-worker, Karl Megerle, who joined our Laboratory as a Research Training Fellow. Megerle and I worked well together and before long we published a paper dealing with most of the basic effects.

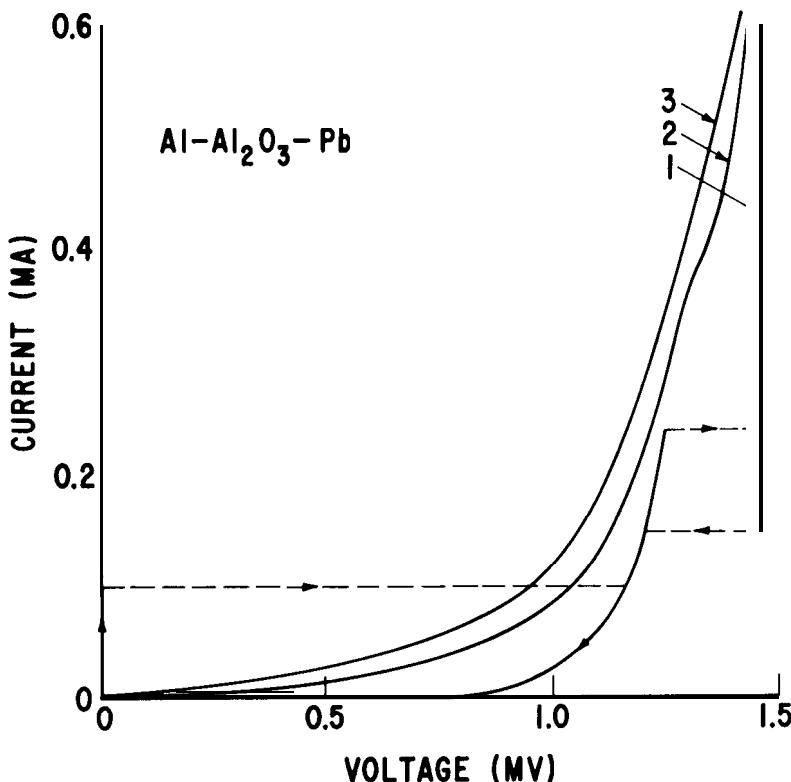


Fig. 13.

Effect of trapped magnetic field on a tunneling characteristic. Curve 1 is a virgin curve, while curve 3 is in a moderate magnetic field, and in curve 2 the magnetic field has been removed. In curve 1 we also have a small resistance-less current which we interpreted as caused by metallic shorts. In retrospect, it was actually due to the Josephson effect.

As always in physics, it is important to extend experiments to a higher energy, a greater magnetic field, or, in our case, to a lower temperature. Therefore, we joined forces with Howard Hart, who had just completed a helium 3 refrigerator that was capable of getting down to about 0.3° K. At the same time, Megerle finished a lock-in amplifier which we could use to measure directly the derivative of the current with respect to the voltage. That was really a nice looking machine with a magnet rotating past a pickup coil at eight cycles per second, but, of course, vastly inferior to the modern lock-in amplifier. We had known for some time that there were anomalies in the current-voltage characteristics of lead, and now we finally pinned them down by finding some extra wiggles in the derivative curve. This is shown in Figure 12. That made us happy because all that the tunneling experiments had done up till now was to confirm the BCS theory, and that is not what an experimentalist would really like to do. The dream is to show that a famous theory is incorrect, and now we had finally poked a hole in the theory. We speculated at the time that these wiggles were somehow associated with the phonons

thought to be the cause of the attractive electron-electron interaction in a superconductor. As often happens, the theorists turned the tables on us and cleverly used these wiggles to properly extend the theory and to prove that the BCS theory indeed was correct. Professor Bardeen gave a detailed account of this in his most recent Nobel Prize lecture.

I have, so far, talked mainly about what went on at General Electric at that time; sometimes it is difficult for me to realize that Schenectady is not the center of the world. Several other people began to do tunneling work, and to mention just a few: J. M. Rowell and W. L. McMillan were really the ones who unraveled the phonon structure in a superconductor; W. J. Tomaszch, of course, insisted on discovering his own effect; S. Shapiro and colleagues did tunneling between two superconductors at the same time we did; and J. Bardeen, and later M. H. Cohen et al., took care of most of the theory.

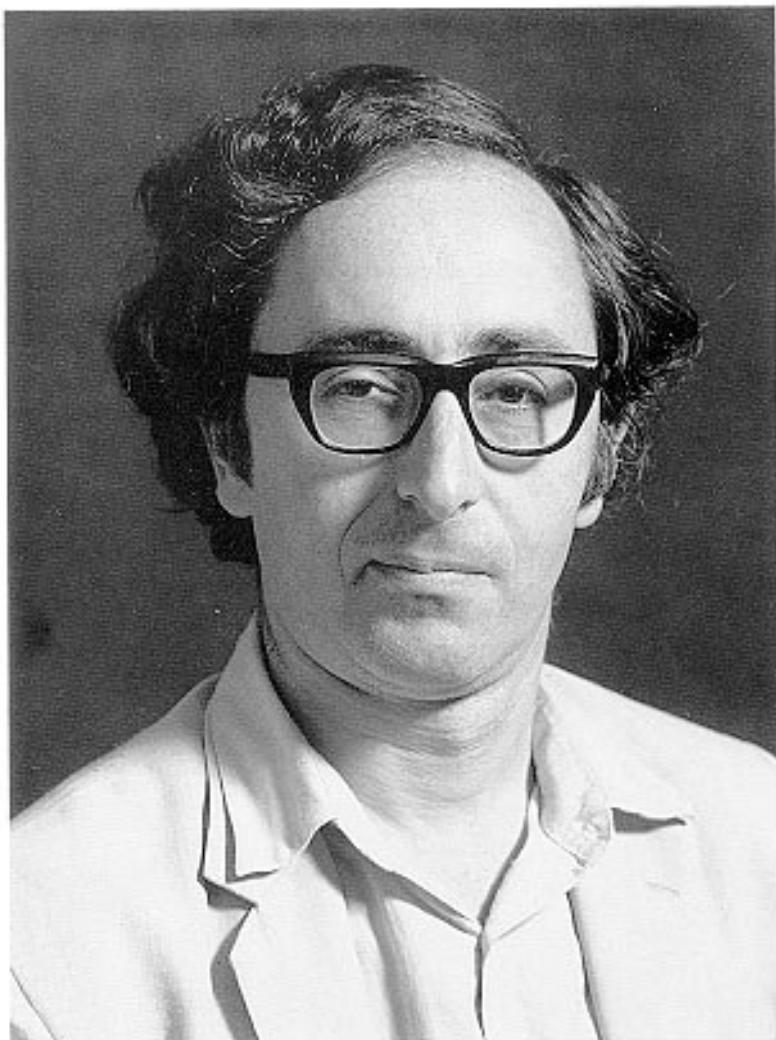
Meanwhile, back at RPI, I had finished my course work and decided to do a theoretical thesis on ordered-disordered alloys with Professor Huntington because tunneling in superconductors was mainly understood. Then someone made me aware of a short paper by Brian Josephson in *Physics Letters*-what did I think? Well, I did not understand the paper, but shortly after I had the chance to meet Josephson at Cambridge and I came away impressed. One of the effects Josephson predicted was that it should be possible to pass a supercurrent with zero voltage drop through the oxide barrier when the metals on both sides were superconducting; this is now called the dc Josephson effect. We had observed this behavior many times; matter-of-fact, it is difficult not to see this current when junctions are made of tin-tin oxide-tin or lead-lead oxide-lead. The early tunnel junctions were usually made with aluminum oxide which generally is thicker and therefore thermal fluctuations suppress the dc current. In our first paper Megerle and I published a curve, which is shown in Figure 13, demonstrating such a supercurrent and also that it depended strongly on a magnetic field. However, I had a ready-made explanation for this supercurrent-it came from a metallic short or bridge. I was puzzled at the time because of the sensitivity to the magnetic field which is unexpected for a small bridge, but no one knew how a 20 $\text{\AA}$  long and 20 $\text{\AA}$ , wide bridge would behave anyway. If I have learned anything as a scientist it is that one should not make things complicated when a simple explanation will do. Thus all the samples we made showing the Josephson effect were discarded as having shorts. This time I was too simple-minded! Later I have been asked many times if I feel bad for missing the effect? The answer is clearly no, because to make an experimental discovery it is not enough to observe something, one must also realize the significance of the observation, and in this instance I was not even close. Even after I learned about the dc Josephson effect, I felt that it could not be distinguished from real shorts, therefore I erroneously believed that only the observation of the so-called ac effect would prove or disprove Josephson's theory.

In conclusion I hope that this rather personal account may provide some slight insight into the nature of scientific discovery. My own beliefs are that the road to a scientific discovery is seldom direct, and that it does not neces-

sarily require great expertise. In fact, I am convinced that often a newcomer to a field has a great advantage because he is ignorant and does not know all the complicated reasons why a particular experiment should not be attempted. However, it is essential to be able to get advice and help from experts in the various sciences when you need it. For me the most important ingredients were that I was at the right place at the right time and that I found so many friends both inside and outside General Electric who unselfishly supported me.

#### REFERENCES

1. *Tunneling Phenomena in Solids* edited by Burstein, E. and Lundquist, S. Plenum Press, New York, 1969.
2. Superconductivity edited by Parks, R. D. Marcel Dekker, Inc., New York. 1969.



Brian Josephson

## BRIAN D. JOSEPHSON

Date of birth: 4 January 1940

Place of birth: Cardiff High School

<i>Education:</i>	Cardiff High School	
	University of Cambridge, B.A.	1960
	University of Cambridge, M.A., Ph.D	1964

### Academic Career

Fellow of Trinity College, Cambridge	1962
Research Assistant Professor, University of Illinois	1964-65
Assistant Director of Research, University of Cambridge	1967-72
NSF Senior Foreign Scientist Fellow, Cornell University	1971
Reader in Physics, University of Cambridge	1972-74
Professor of Physics, University of Cambridge	1974
Visiting Professor - Computer Science Department, Wayne State University, Detroit	1983
Visiting Professor, Indian Institute of Science, Bangalore	1984
Visiting Professor, University of Missouri-Rolla	1987

### Awards

New Scientist	1969
Research Corporation	1969
Fritz London	1970

### Medals

Guthrie (Institute of Physics)	1972
van der Pol	1972
Elliott Cresson (Franklin Institute)	1972
Hughes (Royal Society)	1972
Holweck (Institute of Physics and French Institute of Physics)	1972
Faraday (Institution of Electrical Engineers)	1982
Sir George Thomson (Institute of Measurement and Control)	1984

*Other Information*

Fellow of the Institute of Physics	
Honorary D.Sc., University of Wales	1974
Honorary Member, American Academy of Arts and Sciences	1974
Honorary Member, Institute of Electrical and Electronic Engineers	1982
Honorary D.Sc., University of Exeter	1983
Invited presentation on subject of 'Higher States of Consciousness', to US Congressional Committee	1983

# THE DISCOVERY OF TUNNELLING SUPERCURRENTS

Nobel Lecture, December 12, 1973

by

**BRIAN D. JOSEPHSON**

Cavendish Laboratory, Cambridge, England

The events leading to the discovery of tunnelling supercurrents took place while I was working as a research student at the Royal Society Mond Laboratory, Cambridge, under the supervision of Professor Brian Pippard. During my second year as a research student, in 1961-2, we were fortunate to have as a visitor to the laboratory Professor Phil Anderson, who has made numerous contributions to the subject of tunnelling supercurrents, including a number of unpublished results derived independently of myself. His lecture course in Cambridge introduced the new concept of 'broken symmetry' in superconductors, (1) which was already inherent in his 1958 pseudospin formulation of superconductivity theory, (2) which I shall now describe.

As discussed by Cooper in his Nobel lecture last year (3) according to the Bardeen-Cooper-Schrieffer theory there is a strong positive correlation in a superconductor between the occupation of two electron states of equal and opposite momentum and spin. Anderson showed that in the idealized case where the correlation is perfect the system can be represented by a set of interacting 'pseudospins', with one pseudospin for each pair of electron states. The situation in which both states are unoccupied is represented by a pseudospin in the positive z direction, while occupation of both states is represented by a pseudospin in the negative z direction; other pseudospin orientations correspond to a superposition of the two possibilities.

The effective Hamiltonian for the system is given by

$$H = -2 \sum_k (\varepsilon_k - \mu) s_{kz} - \sum_{k \neq k'} V_{kk'} (s_{kx} s_{k'x} + s_{ky} s_{k'y}) \quad (1)$$

the first term being the kinetic energy and the second term the interaction energy. In this equation  $s_{kx}, s_{ky}$  and  $s_{kz}$  are the three components of the  $k^{\text{th}}$  pseudospin,  $\varepsilon_k$  is the single-particle kinetic energy,  $\mu$  the chemical potential and  $V_{kk'}$  the matrix element for the scattering of a pair of electrons of equal and opposite momentum and spin. The  $k^{\text{th}}$  pseudospin sees an effective field

$$\mathbf{H}_k = 2(\varepsilon_k - \mu) \hat{\mathbf{z}} + 2 \sum_{k' \neq k} V_{kk'} \mathbf{s}_{k'} \perp \quad (2)$$

where  $\hat{\mathbf{z}}$  is a unit vector in the z direction and  $\perp$  indicates the component of the pseudospin in the xy plane.

One possible configuration of pseudospins consistent with (2) is shown in Fig. 1 (a). All the pseudospins lie in the positive or negative z direction, and

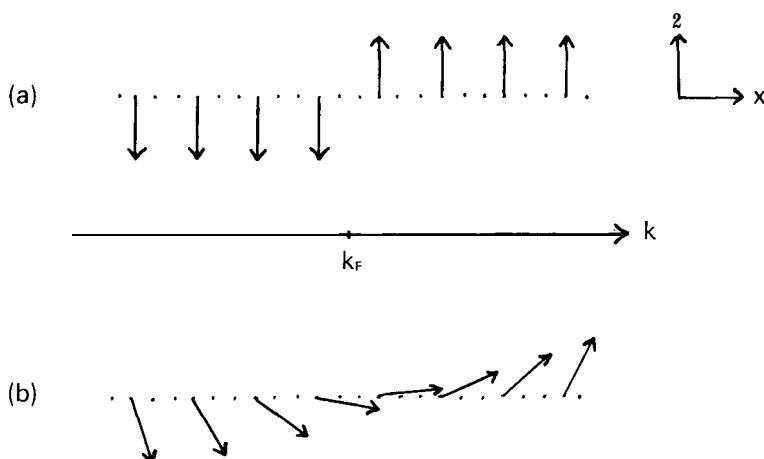


Fig. 1. Pseudospin configurations in (a) a normal metal (b) a superconductor.  $k_F$  is the Fermi momentum.

the direction reverses as one goes through the Fermi surface, since  $\epsilon_k - \mu$  changes sign there. If the interaction is attractive, however (corresponding to negative  $V_{kk'}$ ), a configuration of lower energy exists, in which the pseudospins are tilted out of the  $z$  direction into a plane containing the  $z$  axis, and the pseudospin direction changes continuously as one goes through the Fermi surface, as in Fig. 1 (b).

The ground state of Fig. 1 (b) breaks the symmetry of the pseudospin Hamiltonian (1) with respect to rotation about the  $z$  axis, which is itself a consequence of conservation of number of electrons in the original Hamiltonian. Because of this symmetry a degenerate set of ground states exists, in which the pseudospins can lie in any plane through the  $z$  axis. The angle  $\Phi$  which this plane makes with the  $Oxz$  plane will play an important role in what follows. Anderson made the observation that with a suitable interpretation of the Gor'kov theory, (4)  $\Phi$  is also the phase of the complex quantity  $F$  which occurs in that theory.

I was fascinated by the idea of broken symmetry, and wondered whether there could be any way of observing it experimentally. The existence of the original symmetry implies that the absolute phase angle  $\Phi$  would be unobservable, but the possibility of observing phase differences between the  $F$  functions in two separate superconductors was not ruled out. However, consideration of the number-phase uncertainty relation suggested that the phase difference  $\Delta\Phi$  could be observed only if the two superconductors were able to exchange electrons. When I learnt of observations suggesting that a supercurrent could flow through a sufficiently thin normal region between two superconductors (5, 6), I realized that such a supercurrent should be a function of  $\Delta\Phi$ . I could see in principle how to calculate the supercurrent, but considered the calculation to be too difficult to be worth attempting.

I then learnt of the tunnelling experiments of Giaever, (7) described in the

preceding lecture (8). Pippard (9) had considered the possibility that a Cooper pair could tunnel through an insulating barrier such as that which Giaever used, but argued that the probability of two electrons tunnelling simultaneously would be very small, so that any effects might be unobservable. This plausible argument is now known not to be valid. However, in view of it I turned my attention to a different possibility, that the normal currents through the barrier might be modified by the phase difference. An argument in favour of the existence of such an effect was the fact that matrix elements for processes in a superconductor are modified from those for the corresponding processes in a normal metal by the so-called coherence factors, (3) which are in turn dependent on  $\Delta\Phi$  (through the  $u_k$ 's and  $v_k$ 's of the BCS theory). At this time there was no theory available to calculate the tunnelling current, apart from the heuristic formula of Giaever, (7) which was in agreement with experiment but could not be derived from basic theory. I was able, however, to make a qualitative prediction concerning the time dependence of the current. Gor'kov (4) had noted that the  $F$  function in his theory should be time-dependent, being proportional to  $e^{-2i\mu t/\hbar}$ , where  $\mu$  is the chemical potential as before. (10) The phase  $\Phi$  should thus obey the relation

$$\partial\Phi/\partial t = -2\mu/\hbar, \quad (3)$$

while in a two-superconductor system the phase difference obeys the relation

$$\frac{\partial}{\partial t} (\Delta\Phi) = 2eV/\hbar, \quad (4)$$

where  $V$  is the potential difference between the two superconducting regions. so that

$$\Delta\Phi = 2eVt/\hbar + \text{const.} \quad (5)$$

Since nothing changes physically if  $\Delta\Phi$  is changed by a multiple of  $2\pi$ , I was led to expect a periodically varying current at a frequency  $2eV/\hbar$ .

The problem of how to calculate the barrier current was resolved when one day Anderson showed me a preprint he had just received from Chicago, (11) in which Cohen, Falicov and Phillips calculated the current flowing in a superconductor-barrier-normal metal system, confirming Giaever's formula. They introduced a new and very simple way to calculate the barrier current—they simply used conservation of charge to equate it to the time derivative of the amount of charge on one side of the barrier. They evaluated this time derivative by perturbation theory, treating the tunnelling of electrons through the barrier as a perturbation on a system consisting of two isolated subsystems between which tunnelling does not take place.

I immediately set to work to extend the calculation to a situation in which both sides of the barrier were superconducting. The expression obtained was of the form

$$I = I_0(V) + I_1'(V) \cos(\Delta\Phi) + I_1(V) \sin(\Delta\Phi). \quad (6)$$

At finite voltages the linear increase with time of  $\Delta\Phi$  implies that the only contribution to the dc current comes from the first term, which is the same as Giaever's prediction, thus extending the results of Cohen et al. to the two-superconductor case. The second term had a form consistent with my expectations of a  $\Delta\Phi$  dependence of the current due to tunnelling of quasi-particles. The third term, however, was completely unexpected, as the coefficient  $I_1(V)$ , unlike  $I_0(V)$ , was an even function of  $V$  and would not be expected to vanish when  $V$  was put equal to zero. The  $\Delta\Phi$  dependent current at zero voltage had the obvious interpretation of a supercurrent, but in view of the qualitative argument mentioned earlier I had not expected a contribution to appear to the same order of magnitude as the quasiparticle current, and it was some days before I was able to convince myself that I had not made an error in the calculation.

Since  $\sin(\Delta\Phi)$  can take any value from  $-1$  to  $+1$ , the theory predicted a value of the critical supercurrent of  $I_1(0)$ . At a finite voltage  $V$  an 'ac supercurrent' of amplitude

$$\sqrt{[I_1(V)]^2 + [I_1'(V)]^2}$$

and frequency  $2eV/h$  was expected. As mentioned earlier, the only contribution to the dc current in this situation ( $V \neq 0$ ) comes from the  $I_0(V)$  term, so that a two-section current-voltage relation of the form indicated in Fig. 2 is expected.

I next considered the effect of superimposing an oscillatory voltage at frequency  $v$  on to a steady voltage  $V$ . By assuming the effect of the oscillatory voltage to be to modulate the frequency of the ac supercurrent I concluded that constant-voltage steps would appear at voltages  $V$  for which the frequency of the unmodulated ac supercurrent was an integral multiple of  $v$ , i.e. when  $V = nhv/2e$  for some integer  $n$ .

The embarrassing feature of the theory at this point was that the effects predicted were too large! The magnitude of the predicted supercurrent was proportional to the normal state conductivity of the barrier, and of the same order of magnitude as the jump in current occurring as the voltage passes through that at which production of pairs of quasi-particles becomes possible. Examination of the literature showed that possibly dc supercurrents of this magnitude had been observed, for example in the first published observation of tunnelling between two evaporated-film superconductors by Nicol, Shapiro and Smith (12) (fig. 3). Giaever (13) had made a similar observation, but ascribed the supercurrents seen to conduction through metallic shorts through the barrier layer. As supercurrents were not always seen, it seemed that the explanation in terms of shorts might be the correct one, and the whole theory might have been of mathematical interest only (as was indeed suggested in the literature soon after).

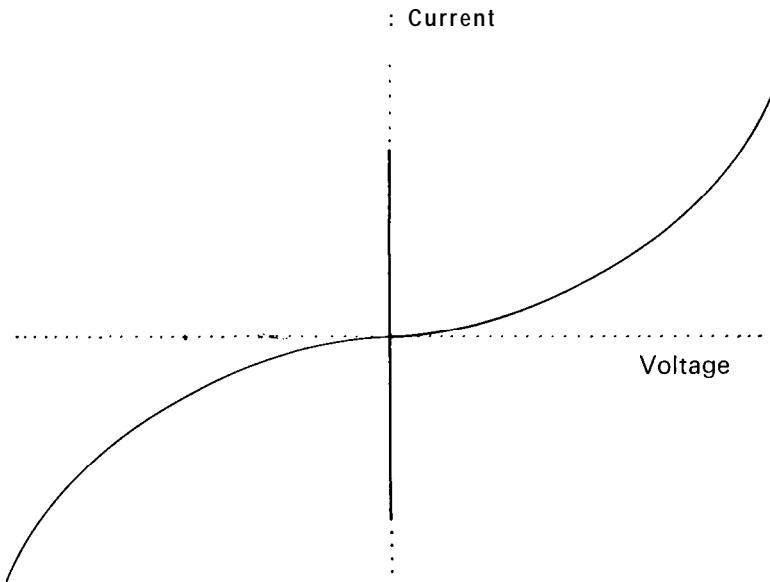


Fig. 2. Predicted two-part current-voltage characteristic of a superconducting tunnel junction.

Then a few days later Phil Anderson walked in with an explanation for the missing supercurrents, which was sufficiently convincing for me to decide to go ahead and publish my calculation, (14) although it turned out later not to have been the correct explanation. He pointed out that my relation between the critical supercurrent and the normal state resistivity depended on the assumption of time-reversal symmetry, which would be violated if a magnetic field were present. I was able to calculate the magnitude of the effect by using the Ginzburg-Landau theory to find the effect of the field on the phase of the  $F$  functions, and concluded that the Earth's field could have a drastic effect on the supercurrents.

Brian Pippard then suggested that I should try to observe tunneling supercurrents myself, by measuring the characteristics of a junction in a compensated field. The result was negative-a current less than a thousandth of the predicted critical current was sufficient to produce a detectable voltage across the junction. This experiment was at one time to be written up in a chapter of my thesis entitled 'Two Unsuccessful Experiments in Electron Tunnelling between Superconductors'.

Eventually Anderson realized that the reason for the non-observation of dc supercurrents in some specimens was that electrical noise transmitted down the measuring leads to the specimen could be sufficient in high-resistance

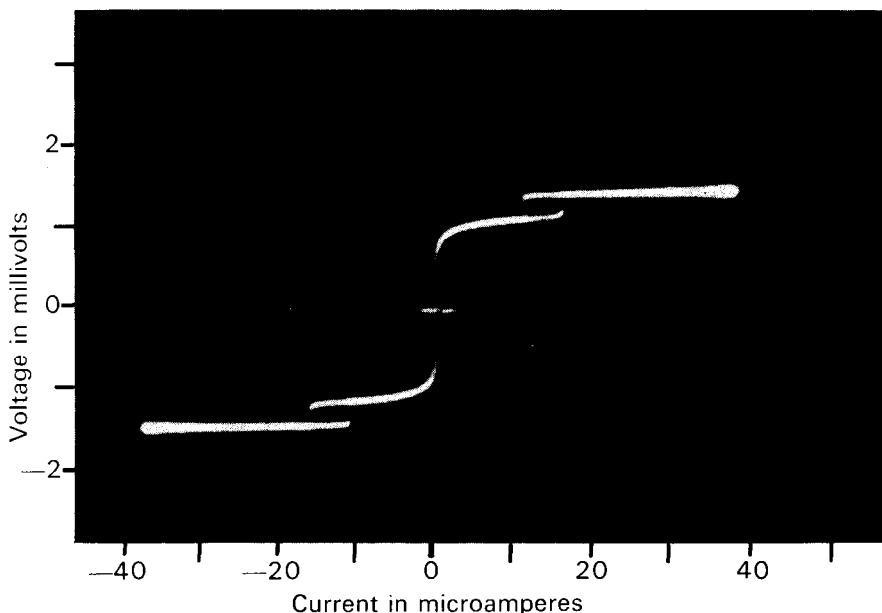


Fig. 3. The first published observation of tunnelling between two evaporated-film superconductors (Nicol, Shapiro and Smith, reference 6). A zero-voltage supercurrent is clearly visible. It was not until the experiments of Anderson and Rowe<sup>11</sup> (reference 15) that such supercurrents could be definitely ascribed to the tunnelling process.

specimens to produce a current exceeding the critical current. Together with John Rowe<sup>11</sup> he made some low resistance specimens and soon obtained convincing evidence (15) for the existence of tunnelling supercurrents, shown particularly by the sensitivity to magnetic fields, which would not be present in the case of conduction through a metallic short. In one specimen they found a critical current of 0.30 mA in the Earth's magnetic field. When the field was compensated, the critical current increased by more than a factor of two, to 0.65 mA, while a field of 2mT was sufficient to destroy the zero-voltage supercurrents completely. Later Rowe<sup>11</sup> (16) investigated the field dependence of the critical current in detail, and obtained results related to the diffraction pattern of a single slit, a connection first suggested by J. C. Phillips (unpublished). This work was advanced by Jaklevic, Lambe, Silver and Mercereau, (17) who connected two junctions in parallel and were able to observe the analogue of the Young's slit interference experiment. The sensitivity of the critical current to applied magnetic field can be increased by increasing the area enclosed between the two branches of the circuit, and Zimmerman and Silver (18) were able to achieve a sensitivity of  $10^{-13}$  T.

Indirect evidence for the ac supercurrents come soon after. Shapiro (19) shone microwaves on to a junction and observed the predicted appearance of steps in the current-voltage characteristics. The voltages at which the steps occurred changed as the frequency of the microwaves was changed, in the manner expected. In 1966, Langenberg, Parker and Taylor (20) measured

the ratio of voltage to frequency to 60 parts per million and found agreement with the value of  $h/2e$  then accepted. Later they increased their accuracy sufficiently to be able to discover errors in the previously accepted values of the fundamental constants and derive more accurate estimates, (21, 22), thus carrying out to fruition an early suggestion of Pippard (unpublished). The ac supercurrent is now used to compare voltages in different standards laboratories without the necessity for the interchange of banks of standard cells. If two laboratories irradiate specimens with radiation of the same frequency, constant-voltage steps appear at identical voltages. The intercomparison of frequencies can be carried out in a straightforward manner by transmission of radio signals.

At the end of 1963, the evidence for the existence of the ac supercurrent was only indirect. John Adkins and I tried to observe the effect by coupling together two junctions by a short ( $\sim 0.2$  mm.) thin-film transmission line. The idea was that radiation emitted by one junction would modify the characteristics of the other. The experiment, planned to form the second part of the thesis chapter referred to above, was unsuccessful, for reasons which are still unclear. Later, Giaever (23) was able to observe the ac supercurrent by a similar method to the one we had considered, and then Yanson, Svistunov and Dmitrenko (24) succeeded in observing radiation emitted by the ac supercurrent with a conventional detector.

Finally, I should like to describe the SLUG, (25) developed in the Royal Society Mond Laboratory by John Clarke while he was a research student. John was attempting to make a high-sensitivity galvanometer using the previously described magnetic interferometers with two junctions connected in parallel. One day Paul Wraight, who shared a room with John, observed that the fact that one cannot solder niobium using ordinary solder must mean that if one allows a molten blob of solder to solidify in contact with niobium there must be an intermediate layer of oxide, which might have a suitable thickness to act as a tunnelling barrier. This proved to be the case. However, in John's specimens, in which a niobium wire was completely surrounded by a blob of solder, the critical current through the barrier proved to be completely insensitive to externally applied magnetic fields. It was, however, found to be sensitive to the magnetic field produced by passing a current through the central wire. This fact led to the development of a galvanometer with sensitivity of  $10^{-14}$  volts at a time constant of 1 s.

There have been many other developments which I have not had time to describe here. I should like to conclude by saying how fascinating it has been for me to watch over the years the many developments in laboratories over the world, which followed from asking one simple question, namely what is the physical significance of broken symmetry in superconductors?

## REFERENCES AND FOOTNOTES

1. Anderson, P. W., *Concepts in Solids*, W. A. Benjamin, Inc., New York, 1963.
2. Anderson, P. W. Phys. Rev. 112, 1900 (1958).
3. Cooper, L. N., *Les Prix Nobel en 1972*, Nobel Foundation 1973, p. 64.
4. Gor'kov, L. P., J. Exptl. Theoret. Phys. (USSR) 36, 1918 (1959); translation: Soviet Phys. JETP 9, 1364 (1959).
5. Meissner, H. Phys. Rev. 117, 672 (1960).
6. Smith, P. H., Shapiro, S., Miles, J. L. and Nicol, J. Phys. Rev. Letters 6, 686 (1961).
7. Giaever, I., Phys. Rev. Letters 5, 464 (1960).
8. Giaever, I., *Les Prix Nobel en 1973*, Nobel Foundation 1974, p.
9. Pippard, A. B., *Proceedings of the VIIth International Conference on Low Temperature Physics*, ed. Graham, G. M. and Hollis Hallett, A. C., North-Holland, Amsterdam 1961, p. 320.
10. Gor'kov's result may be extended to finite temperatures by the following argument. The density operator of a system in equilibrium has the form  $Z^{-1} \exp\{-\beta(H-\mu N)\}$  where  $\mu$  must contain a small symmetrybreaking term in order to set up an ensemble in which  $\Phi$  has a definite value. Since this operator commutes with  $H-\mu N$ , all quantities are time-independent if the Hamiltonian of the system is taken to be  $H-\mu H$ . Transition to a situation where the time dependence is given by the true Hamiltonian  $H$  can be accomplished by means of a gauge transformation, and consideration of the effect of this transformation on the electron operators gives immediately Gor'kov's result  $F \propto \exp(-2i\mu t/\hbar)$ .
11. Cohen, M. H., Falicov, L. M. and Phillips, J. C., Phys. Rev. Letters 8, 316 (1962).
12. Nicol, J., Shapiro, S. and Smith, P. H., Phys. Rev. Letters 5, 461 (1960).
13. Giaever, I., Phys. Rev. Letters 5, 464 (1960).
14. Josephson, B. D., Phys. Letters 1, 251 (1962).
15. Anderson, P. W. and Rowell, J. M., Phys. Rev. Letters 10, 230 (1963).
16. Rowell, J. M., Phys. Rev. Letters 11, 200 (1963).
17. Jaklevic, R. C., Lambe, J., Silver, A. H. and Mercereau, J. E., Phys. Rev. Letters 12, 159 (1964).
18. Zimmerman, J. E., and Silver, A. H., Phys. Rev. 141, 367 (1966).
19. Shapiro, S. Phys. Rev. Letters 11, 80 (1963).
20. Langenberg, D. N., Parker, W. H. and Taylor, B. N., Phys. Rev. 150, 186 (1966).
21. Parker, W. H., Taylor, B. N. and Langenberg, D. N., Phys. Rev. Letters 18, 287 (1967).
22. Taylor, B. N., Parker, W. H. and Langenberg, D. N., *The Fundamental Constants and Quantum Electrodynamics*, Academic Press, New York and London, 1969.
23. Giaever, I., Phys. Rev. Letters 14, 904 (1965).
24. Yanson, I. K., Svistunov, V. M. and Dmitrenko, I. M., J. Exptl. Theoret. Phys. (USSR) 21, 650 (1965); translation: Soviet Phys. JETP 48, 976 (1965).
25. Clarke, J., Phil. Mag. 13, 115 (1966).

Physics 1974

**MARTIN RYLE and ANTONY HEWISH**

*for their pioneering research in radio astrophysics: Ryle for his observations and inventions, in particular of the aperture synthesis technique, and Hewish for his decisive role in the discovery of pulsars*

## THE NOBEL PRIZE FOR PHYSICS

Speech by professor **HANS WILHELMSSON** of the Royal Academy of Sciences  
Translation from the Swedish text

Your Majesty, Your Royal Highnesses, Ladies and Gentlemen,  
The subject of the Nobel Prize in Physics this year is the science of Astrophysics, the Physics of the stars and galactic systems.

Problems concerning our Universe on a large scale, its constitution and evolution, play an essential role in present day scientific discussions.

We are curious about the behaviour of our Universe. In order to draw reliable conclusions regarding cosmological models it is necessary to gather detailed information about conditions in the remote parts of the Cosmos.

Radio-astronomy offers unique possibilities for studying what is taking place, or in reality what occurred very long ago, at enormous distances from Earth, as far out as thousands of millions of lightyears from us. The radio waves now reaching us have been travelling for thousands of millions of years at the speed of light to reach our Earth from those very remote sources.

It is indeed a thrilling fact that the radio signals we record today here on Earth left their cosmic sources at a time when hardly any flowers or living creatures, and certainly no physicists, existed on Earth.

New and epoch-making discoveries have been made in the field of Radio-astrophysics during the last decade, discoveries that are also exceedingly important contributions to modern Physics, for example in establishing through radio-astronomical observations the presence of matter in a superdense state.

One single cubic centimeter of this superdense matter has a weight of thousands of millions of tons. It consists of tightly-packed neutrons. A neutron star appears as a consequence of a star explosion, a so-called supernova event. Neutron stars, with a diameter of about 10 kilometers, are from a cosmic point of view extremely small objects. They represent the final state in the evolution of certain stars.

This year's Nobel Prize winners in Physics, Martin Ryle and Antony Hewish, developed new radio-astronomical techniques. Their observations of cosmic radio sources represent extremely noteworthy research results.

In order to collect radio waves from cosmic radio sources one utilizes radio-telescopes. It is important that a radio-telescope should have a large area, both for highest possible sensitivity and for the high angular resolution that is needed to discriminate among the various cosmic sources of radio radiation.

For observation of exceedingly small sources it is, however, no longer possible to build a single radio-telescope of sufficient size. Ryle and his collaborators therefore developed the method of aperture synthesis. Instead of making one huge aerial, a number of small aerials are used in this method,

and the signals received by them are combined in such a way as to provide the necessary extreme accuracy.

Instead of many small aerials, Ryle in fact made use of a few aerials that could be moved successively to different positions on the ground. Ryle also invented the extremely elegant and powerful technique utilizing the rotation of the Earth to move his radio-telescopes. With this technique he obtained a resolution in his observations that corresponded to an aerial of enormous size.

Ryle's measurements enable us to conclude that a steady-state model of the Universe can not be accepted. The Cosmos on a large scale has to be described by dynamic, evolutionary models.

In his latest construction in Cambridge, Ryle obtained an angular resolution permitting the mapping of cosmic radio sources with an error less than one second of arc!

The radio-astronomical instruments invented and developed by Martin Ryle, and utilized so successfully by him and his collaborators in their observations, have been one of the most important elements of the latest discoveries in Astrophysics.

Antony Hewish and his collaborators in Cambridge, in the Autumn of 1967, made a unique and unexpected discovery that has revolutionized Astrophysics. They had constructed new aerials and instruments to study the influence of the outer corona of the Sun on the radiation detected from remote point sources. A special receiver capable of extremely rapid response had been built.

The fast receiver provided a result quite different from its intended purpose. By chance the receiver detected short pulses of radio signals that were repeated periodically about every second, and with exceedingly high precision in the pulse repetition rates.

It was concluded that the radiation originated from cosmic sources of previously unknown type. These sources were subsequently named pulsars.

One has come to the conclusion that the central part of a pulsar consists of a neutron star. The pulsars are also accompanied by magnetic fields many millions of times stronger than those found in laboratories on Earth. The neutron star is surrounded by an electrically-conducting gas or plasma. Each pulsar rotates and emits beams of radiation in the Universe, resembling those from a light-house. The beams strike the Earth periodically with high precision.

These pulsars are indeed the world clocks which our Nobel Prize winner Harry Martinson mentions in his poetry.

Allow me to quote this poet of space:

"World clocks tick and space gleams  
everything changes place and order".

Early in the history of pulsar research it was suspected that neutron star matter existed in the centres of supernovas. Radio-telescopes were aimed towards the centre of the Crab nebula, a magnificent glaring gaseous remnant of a supernova event that is known, from Chinese annals, to have occurred in 1054 A.D., and indeed, they detected a pulsar! This pulsar emits not only

radio pulses, as expected from a pulsar, but pulses of light and x-rays as well. It is comparatively young, rotates rapidly and is in fact exceptional among pulsars.

Antony Hewish played a decisive role in the discovery of pulsars. This discovery, which is of extraordinary scientific interest, opens the way to new methods for studying matter under extreme physical conditions.

The contributions of Ryle and Hewish represent an important step forward in our knowledge of the Universe. Thanks to their work new fields of research have become part of Astrophysics. The gigantic laboratory of the Universe offers rich possibilities for future research.

Sir Martin,

Some of the most fundamental questions in Physics have been elucidated as a result of your brilliant research. Your inventions and observations have brought new foundations for our conception of the Universe.

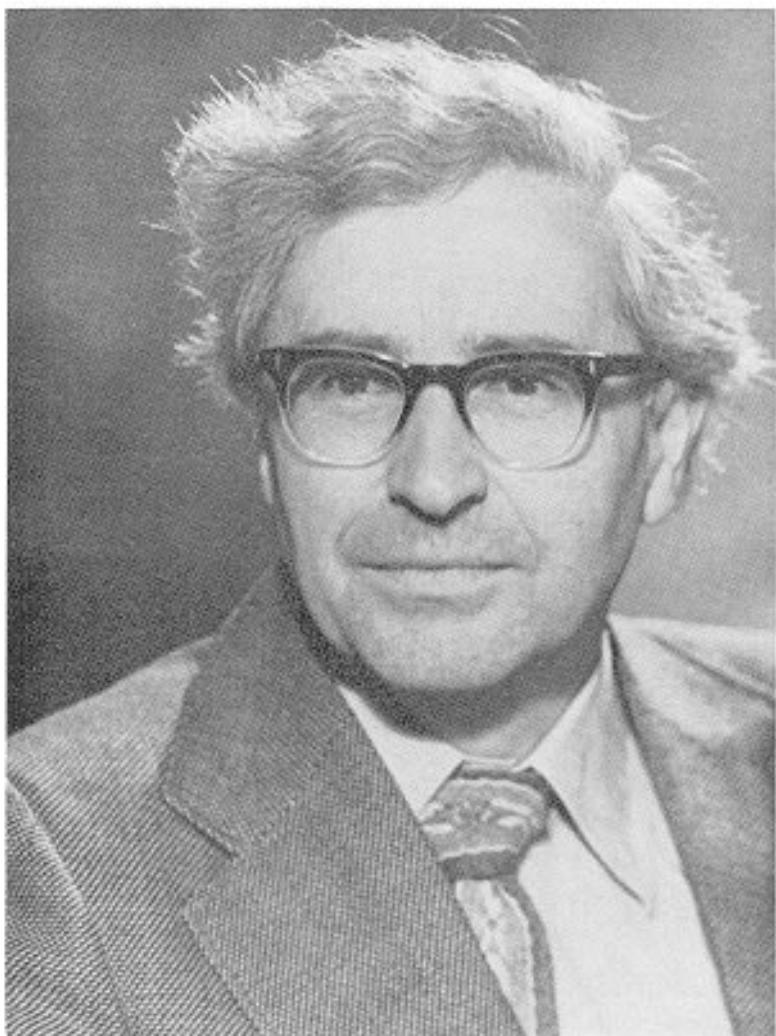
Professor Antony Hewish,

The discovery of pulsars, for which you played a decisive role, is a most outstanding example of how in recent years our knowledge of the Universe has been dramatically extended. Your research has contributed greatly to Astrophysics and to Physics in general.

On behalf of the Royal Academy of Sciences I wish to express our admiration and to convey to you our warmest congratulations.

The Royal Academy of Sciences regrets that Sir Martin Ryle is not here today.

May I now ask you, Professor Hewish, to receive your prize and also the prize awarded to Sir Martin Ryle from the hands of His Majesty the King.



Anthony Hewish.

## ANTHONY HEWISH

I was born in Fowey, Cornwall, on 11 May 1924, the youngest of three sons and my father was a banker. I grew up in Newquay, on the Atlantic coast and there developed a love of the sea and boats. I was educated at Ring's College, Taunton and went to the University of Cambridge in 1942. From 1943-46 I was engaged in war service at the Royal Aircraft Establishment, Farnborough and also at the Telecommunications Research Establishment, Malvern. I was involved with airborne radar-counter-measure devices and during this period I also worked with Martin Ryle.

Returning to Cambridge in 1946 I graduated in 1948 and immediately joined Ryle's research team at the Cavendish Laboratory. I obtained my Ph.D. in 1952, became a Research Fellow at Gonville and Caius College where I had been an undergraduate, and in 1961 transferred to Churchill College as Director of Studies in Physics. I was University Lecturer during 1961-69, Reader during 1969-71 and Professor of Radio Astronomy from 1971 until my retirement in 1989. Following Ryle's illness in 1977 I assumed leadership of the Cambridge radio astronomy group and was head of the Mullard Radio Astronomy Observatory from 1982-88.

My decision to begin research in radio astronomy was influenced both by my wartime experience with electronics and antennas and by one of my teachers, Jack Ratcliffe, who had given an excellent course on electromagnetic theory during my final undergraduate year and whom I had also encountered at Malvern. He was head of radiophysics at the Cavendish Laboratory at that time.

My first research was concerned with propagation of radiation through inhomogeneous transparent media and this has remained a lifelong interest. The first two radio "stars" had just been discovered and I realised that their scintillation, or "twinkling", could be used to probe conditions in the ionosphere. I developed the theory of diffraction by phase-modulating screens and set up radio interferometers to exploit my ideas. Thus I was able to make pioneering measurements of the height and physical scale of plasma clouds in the ionosphere and also to estimate wind speeds in this region. Following our Cambridge discovery of interplanetary scintillation in 1964 I developed similar methods to make the first ground-based measurements of the solar wind and these were later adopted in the USA, Japan and India for long term observations. I also showed how interplanetary scintillation could be used to obtain very high angular resolution in radio astronomy, equivalent to an interferometer with a baseline of 1000 km - something which had not then been achieved in this field. It was to exploit

this technique on a large sample of radio galaxies that I conceived the idea of a giant phased-array antenna for a major sky survey. This required instrumental capabilities quite different from those of any existing radio telescope, namely very high sensitivity at long wavelengths, and a multi-beam capability for repeated whole-sky surveys on a day to day basis.

I obtained funds to construct the antenna in 1965 and it was completed in 1967. The sky survey to detect all scintillating sources down to the sensitivity threshold began in July. By a stroke of good fortune the observational requirements were precisely those needed to detect pulsars. Jocelyn Bell joined the project as a graduate student in 1965, helping as a member of the construction team and then analysing the paper charts of the sky survey. She was quick to spot the week to week variability of one scintillating source which I thought might be a radio flare star, but our more detailed observations subsequently revealed the pulsed nature of the signal.

Surprisingly, the phased array is still a useful research instrument. It has been doubled in area and considerably improved over the years and one of my present interests is the way our daily observations of scintillation over the whole sky can be used to map large-scale disturbances in the solar wind. At present this is the only means of seeing the shape of interplanetary weather patterns so our observations make an useful addition to in-situ measurements from spacecraft such as Ulysses, now (1992) on its way to Jupiter.

Looking back over my forty years in radio astronomy I feel extremely privileged to have been in at the beginning as a member of Martin Ryle's group at the Cavendish. We were a closely-knit team and besides my own research programmes I was also involved in the design and construction of Ryle's first antennas employing the novel principle of aperture synthesis.

Teaching physics at the University, and more general lecturing to wider audiences has been a major concern. I developed an association with the Royal Institution in London when it was directed by Sir Lawrence Bragg, giving one of the well known Christmas Lectures and subsequently several Friday Evening Discourses. I believe scientists have a duty to share the excitement and pleasure of their work with the general public, and I enjoy the challenge of presenting difficult ideas in an understandable way.

I have been happily married since 1950. My son is a physicist and obtained his Ph.D. for neutron scattering in liquids, while my daughter is a language teacher.

### *Honours and Awards*

Hamilton Prize, Cambridge (1952); Eddington Medal, Royal Astronomical Society (1969); Charles Vernon Boys Prize, Institute of Physics (1970); Dellinger Medal, International Union of Radio Science (1972); Michelson Medal, Franklin Institute (1973); Hopkins Prize, Cambridge Philosophical Society (1973); Holwech Medal and Prize, Société Francaise de Physique (1974); Nobel Prize in Physics (1974); Hughes Medal, Royal Society (1976).

Honorary ScD.s from the Universities of Leicester (1976), Exeter (1977), Manchester (1989) and Santa Maria, Brazil (1989).

Fellow of the Royal Society (1968), Foreign Honorary Member, American Academy of Arts and Sciences (1977), Foreign Fellow, Indian National Science Academy (1982) Honorary Fellow, Indian Institution of Electronics and Telecommunication Engineers (1985) Associate Member, Belgian Royal Academy (1989).

# PULSARS AND HIGH DENSITY PHYSICS

Nobel Lecture, December 12, 1974

bY

**ANTONY HEWISH**

University of Cambridge, Cavendish Laboratory, Cambridge, England

## DISCOVERY OF PULSARS

The trail which ultimately led to the first pulsar began in 1948 when I joined Ryle's small research team and became interested in the general problem of the propagation of radiation through irregular transparent media. We are all familiar with the twinkling of visible stars and my task was to understand why radio stars also twinkled. I was fortunate to have been taught by Ratcliffe, who first showed me the power of Fourier techniques in dealing with such diffraction phenomena. By a modest extension of existing theory I was able to show that our radio stars twinkled because of plasma clouds in the ionosphere at heights around 300 km, and I was also able to measure the speed of ionospheric winds in this region (1).

My fascination in using extra-terrestrial radio sources for studying the intervening plasma next brought me to the solar corona. From observations of the angular scattering of radiation passing through the corona, using simple radio interferometers, I was eventually able to trace the solar atmosphere out to one half the radius of the Earth's orbit (2).

In my notebook for 1954 there is a comment that, if radio sources were of small enough angular size, they would illuminate the solar atmosphere with sufficient coherence to produce interference patterns at the Earth which would be detectable as a very rapid fluctuation of intensity. Unfortunately the information then available showed that the few sources known were more than one hundred times too large to produce this effect, and I did not pursue the idea. This was sad because the phenomenon was discovered by chance, about eight years later, by Margaret Clarke long after I had forgotten all about my comment. She was involved with a survey of radio sources at Cambridge and noticed that three particular sources showed variations of intensity. She pointed out that two of the sources were known to have angular sizes of less than 2" and estimated that a scintillation mechanism required plasma irregularities at distances of thousands of km but she concluded that the fluctuations were an unsolved mystery (3). During a group discussion I suddenly remembered my earlier conclusion and realised that, if the radio sources subtended an angle of less than 1", they might show the predicted intensity scintillation caused by plasma clouds in the interplanetary medium. With the assistance of Scott and Collins special observations of 3C 48 and other quasi-stellar radio sources were made and the scintillation phenomenon was immediately confirmed (4).

Since interplanetary scintillation, as we called this new effect, could be

detected in any direction in space I used it to study the solar wind, which had by then been discovered by space probes launched into orbits far beyond the magnetosphere. It was interesting to track the interplanetary diffraction patterns as they raced across England at speeds in excess of  $300 \text{ km s}^{-1}$ , and to sample the behaviour of the solar wind far outside the plane of the ecliptic where spacecraft have yet to venture (5).

The scintillation technique also provided an extremely simple and useful means of showing which radio sources had angular sizes in the range  $0''.1\text{--}1''.0$ . The first really unusual source to be uncovered by this method turned up in 1965 when, with my student Okoye, I was studying radio emission from the Crab Nebula. We found a prominent scintillating component within the nebula which was far too small to be explained by conventional synchotron radiation and we suggested that this might be the remains of the original star which had exploded and which still showed activity in the form of flare-type radio emission (6). This source later turned out to be none other than the famous Crab Nebula Pulsar.

In 1965 I drew up plans for a radio telescope with which I intended to carry out a large-scale survey of more than 1000 radio galaxies using interplanetary scintillation to provide high angular resolution. To achieve the required sensitivity it was necessary to cover an area of  $18,000 \text{ m}^2$  and, because scintillation due to plasmas is most pronounced at long wavelengths, I used a wavelength of 3.7 m. The final design was an array containing 2048 dipole antennas. Later that year I was joined by a new graduate student, Jocelyn Bell, and she became responsible for the network of cables connecting the dipoles. The entire system was built with local effort and we relied heavily upon the willing assistance of many members of the Cambridge team.

The radio telescope was complete, and tested, by July 1967 and we immediately commenced a survey of the sky. Our method of utilising scintillation for the quantitative measurement of angular sizes demanded repeated observations so that every source could be studied at many different solar elongations. In fact we surveyed the entire range of accessible sky at intervals of one week. To maintain a continuous assessment of the survey we arranged to plot the positions of scintillating radio sources on a sky-chart, as each record was analysed, and to add points as the observations were repeated at weekly intervals. In this way genuine sources could be distinguished from electrical interference since the latter would be unlikely to recur with the same celestial coordinates. It is greatly to Jocelyn Bell's credit that she was able to keep up with the flow of paper from the four recorders.

One day around the middle of August 1967 Jocelyn showed me a record indicating fluctuating signals that could have been a faint source undergoing scintillation when observed in the antisolar direction. This was unusual since strong scintillation rarely occurs in this direction and we first thought that the signals might be electrical interference. So we continued the routine survey. By the end of September the source had been detected on several occasions, although it was not always present, and I suspected that we had located a flare star, perhaps similar to the M-type dwarfs under investigation by Lovell.

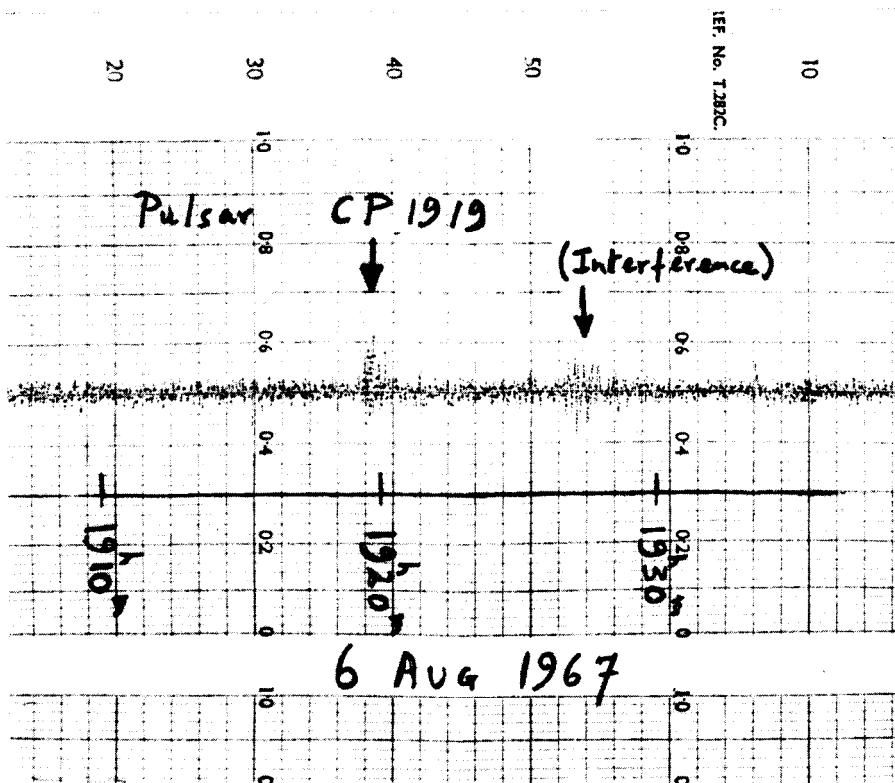


Fig. 1. The first signals from CP 1919.

But the source also exhibited apparent shifts of right ascension of up to 90 seconds which was evidence against a celestial origin. We installed a high-speed recorder to study the nature of the fluctuating signals but met with no success as the source intensity faded below our detection limit. During October

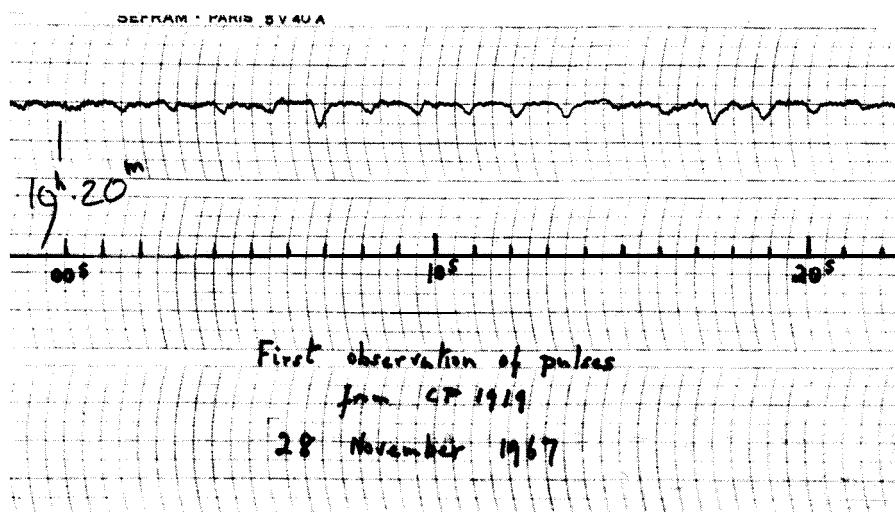


Fig. 2. The first indication of pulsed radio emission from CP 1919.

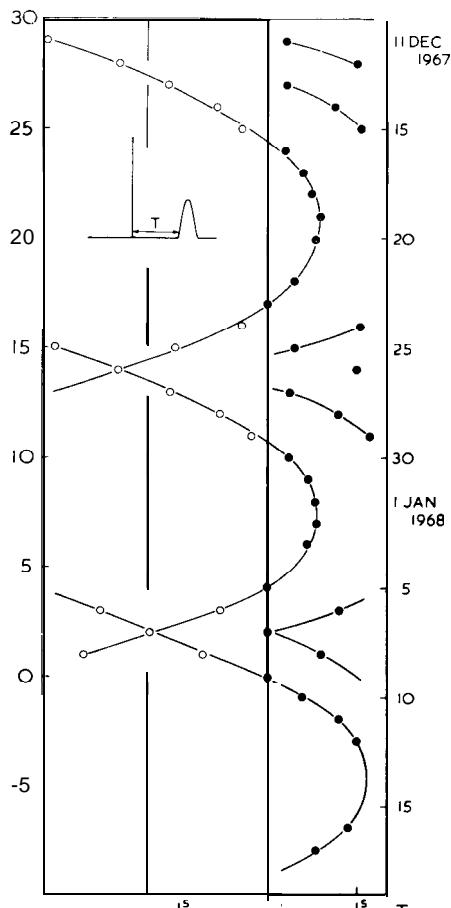


Fig. 3. Timing measurements showing Doppler shift due to the orbital motion of the Earth.

this recorder was required for pre-arranged observations of another source, 3C 273, to check certain aspects of scintillation theory, and it was not until November 28th that we obtained the first evidence that our mysterious source was emitting regular pulses of radiation at intervals of just greater than one second. I could not believe that any natural source would radiate in this fashion and I immediately consulted astronomical colleagues at other observatories to enquire whether they had any equipment in operation which might possibly generate electrical interference at a sidereal time near  $19^{\text{h}}19^{\text{m}}$ .

In early December the source increased in intensity and the pulses were clearly visible above the noise. Knowing that the signals were pulsed enabled me to ascertain their electrical phase and I reanalysed the routine survey records. This showed that the right ascension was constant. The apparent variations had been caused by the changing intensity of the source. Still sceptical, I arranged a device to display accurate time marks at one second intervals broadcast from the MSF Rugby time service and on December 11th began daily timing measurements. To my astonishment the readings fell in a regular pattern, to within the observational uncertainty of 0.1s, showing that the pulsed source kept time to better than 1 part in  $10^6$ . Meanwhile my col-

leagues Pilkington, and Scott and Collins, found by quite independent methods that the signal exhibited a rapidly sweeping frequency of about  $-5 \text{ MHz s}^{-1}$ . This showed that the duration of each pulse, at one particular radio frequency, was approximately 16 ms.

Having found no satisfactory terrestrial explanation for the pulses we now began to believe that they could only be generated by some source far beyond the solar system, and the short duration of each pulse suggested that the radiator could not be larger than a small planet. We had to face the possibility that the signals were, indeed, generated on a planet circling some distant star, and that they were artificial. I knew that timing measurements, if continued for a few weeks, would reveal any orbital motion of the source as a Doppler shift, and I felt compelled to maintain a curtain of silence until this result was known with some certainty. Without doubt, those weeks in December 1967 were the most exciting in my life.

It turned out that the Doppler shift was precisely that due to the motion of the Earth alone, and we began to seek explanations involving dwarf stars, or the hypothetical neutron stars. My friends in the library at the optical observatory were surprised to see a radio astronomer taking so keen an interest in books on stellar evolution. I finally decided that the gravitational oscillation of an entire star provided a possible mechanism for explaining the periodic emission of radio pulses, and that the fundamental frequency obtainable from white dwarf stars was too low. I suggested that a higher order mode was needed in the case of a white dwarf, or that a neutron star of the lowest allowed density, vibrating in the fundamental mode, might give the required periodicity. We also estimated the distance of the source on the assumption that the frequency sweep was caused by pulse dispersion in the interstellar plasma, and obtained a value of 65 parsec, a typical stellar distance.

While I was preparing a coherent account of this rather hectic research, in January 1968, Jocelyn Bell was scrutinising all our sky-survey recordings with her typical persistence and diligence and she produced a list of possible additional pulsar positions. These were observed again for evidence of pulsed radiation and before submitting our paper for publication, on February 8th, we were confident that three additional pulsars existed although their parameters were then only crudely known. I well remember the morning when Jocelyn came into my room with a recording of a possible pulsar that she had made during the previous night at a right ascension  $09^{\text{h}}50^{\text{m}}$ . When we spread the chart over the floor and placed a metre rule against it a periodicity of 0.25s was just discernible. This was confirmed later when the receiver was adjusted to a narrower bandwidth, and the rapidity of this pulsar made explanations involving white dwarf stars increasingly difficult.

The months that followed the announcement (7) of our discovery were busy ones for observers and theoreticians alike, as radio telescopes all over the world turned towards the first pulsars and information flooded in at a phenomenal rate. It was Gold (8) who first suggested that the rotation of neutron stars provided the simplest and most flexible mechanism to explain the pulsar clock, and his prediction that the pulse period should increase with

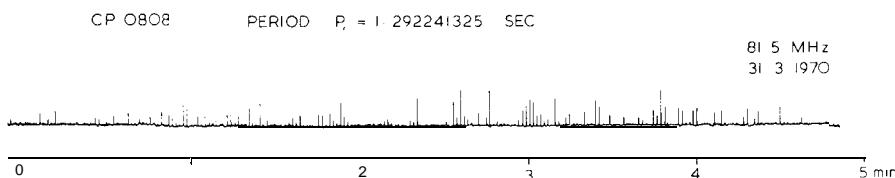


Fig. 4. Radiation from a typical pulsar.

time soon received dramatic confirmation with the discovery of the pulsar in the Crab Nebula (9, 10). Further impressive support for the neutron star hypothesis was the detection of pulsed light from the star which had previously been identified as the remnant of the original explosion. This, according to theories of stellar evolution, is precisely where a young neutron star should be created. Gold also showed that the loss of rotational energy, calculated from the increase of period for a neutron star model, was exactly that needed to power the observed synchrotron light from the nebula.

Now, in 1974, with more than 130 pulsars charted in the heavens, there is overwhelming evidence that the neutron star "lighthouse" model is correct. No other star could spin fast enough, without fragmenting, to account for the most rapid pulsars yet periods ranging from 33 ms to 3.5 s are readily accommodated by the rotation theory. At the same time there is unfortunately no satisfactory theory to account for the radio emission generated by these tiny stars which have radii of only 10 km.

#### HIGH DENSITY PHYSICS INSIDE NEUTRON STARS

The prediction that matter at the almost unimaginable density of  $10^{18} \text{ kg m}^{-3}$  might be formed under gravitational compression inside stars was first made by Baade and Zwicky (11) in 1934, soon after Chadwick's discovery of the neutron. At this density only a small fraction of the original protons and electrons could exist and matter would consist predominantly of neutrons. It is the degeneracy pressure arising from the neutrons, which obey Fermi statistics, that balances further gravitational compression, although finally the Fermi energy becomes relativistic and further gravitational collapse ensues. Since complex nuclei are generated by nuclear fusion inside hot stars, where there is a large thermal pressure, the degenerate neutron state can only be found when fusion ceases and we deal with the cooling "ashes" of stellar evolution. The stars that give rise to neutron stars are more massive than the Sun, and it is believed that the formation of neutron stars is associated with supernova explosions.

Since the discovery of pulsars there has been great activity amongst solid-state physicists around the world because neutron matter, at any temperature less than about  $10^9 \text{ K}$ , behaves rather like ordinary matter close to the absolute zero of temperature. The generally agreed model of a neutron star consists of concentric shells with very different physical properties as reviewed by Ruderman (12).

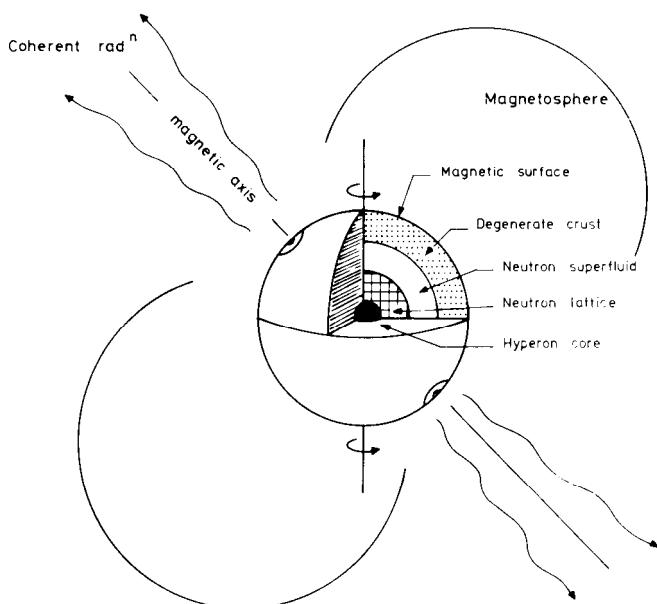


Fig. 5. Model of a neutron star.

At the surface of the star it is likely that there exists a shell of iron since  $^{56}\text{Fe}$  is the most stable nucleus. The atoms would be normal if no magnetic field were present. In astrophysics it is unwise to ignore magnetic phenomena and gravitational collapse following a supernova explosion probably compresses the original stellar magnetic flux to produce surface field strengths of  $10^8\text{T}$  or more. In fields of this magnitude the radius of gyration of electrons in atomic energy levels becomes smaller than the Bohr radius and the electronic wave functions adopt a cylindrical shape. It is far harder to ionize distorted atoms of this kind and this is of importance when considering the generation of a magnetosphere surrounding the neutron star.

Beneath the iron skin the increasing compression forces electrons into higher energy states until they are entirely freed from the positive nuclei. The unscreened nuclei then settle into a rigid lattice having a melting temperature of about  $^8\text{K}$ . At greater depths the electron energies become relativistic and they begin to combine with protons in the nuclei, thus adding to the neutron population. This is the process of inverse  $\beta$  decay. At a sufficient depth nearly all the electrons and protons have disappeared and the nuclei have been converted to a sea of neutrons.

The energy gap for neutron pairing is of the order of several MeV, corresponding to a superfluid transition temperature of  $10^8\text{-}10^{10}\text{K}$ , and since young neutron stars cool rapidly to temperatures below  $10^9\text{K}$ , the neutron sea is expected to behave like a quantum superfluid. The few remaining protons will similarly pair and enter a superconducting state, while the residual electrons will behave normally. The bulk motion of the neutron superfluid must be irrotational, but an effective solid body rotation can be simulated with a

distribution of quantised vortex lines containing a small fraction of normal fluid neutrons.

At yet deeper levels the neutron-neutron interaction may result in the creation of a solid neutron lattice, although this possibility is under debate, and finally there is the question of a material composed of stable hyperons.

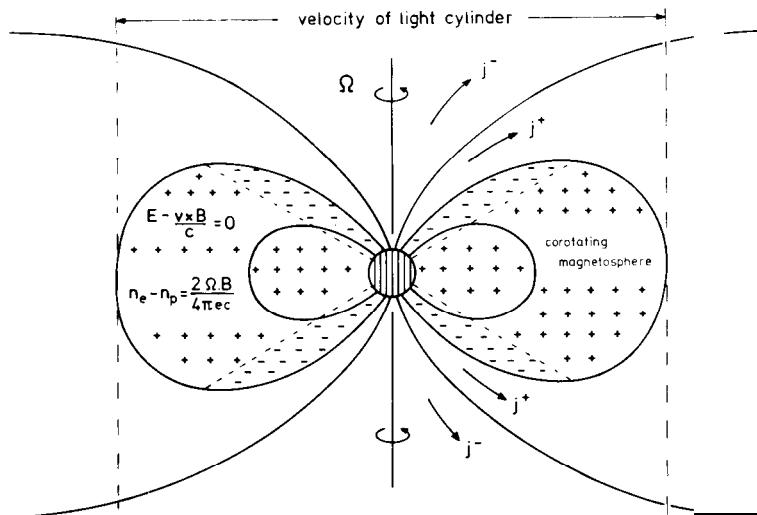
Evidence that neutron stars do indeed have a structure similar to the predicted models has been obtained from extended timing observations of pulsars (13). These show that the systematic increase of period, corresponding to a steady loss of rotational energy from the spinning star, is sometimes interrupted by discontinuous changes. Most pulsars are observed to be slowing down on a typical timescale of  $10^6$ - $10^7$  years, although the most rapid pulsars, in the Crab and Vela supernovae, have timescales of only  $10^3$  and  $10^4$  years respectively. The discontinuities often show an abrupt decrease of period, followed by a recovery to a slightly reduced value with a characteristic relaxation time.

For the Crab pulsar this effect can be explained by a rigid crust-liquid core model. Young neutron stars are likely to be spinning rapidly at birth, with angular velocities up to  $10^4$  radian s $^{-1}$ , and they will therefore have a spheroidal shape. As a star slows down it will tend to become less spheroidal and the rigid crust will fracture at irregular intervals as the increasing strain overcomes rigidity. When this occurs the crust will momentarily spin more rapidly, but later the increased angular momentum will be coupled into the fluid interior, where the bulk of the mass resides. The observed time constant for coupling is in good agreement with the superfluid model, and would be far smaller in the case of a normal fluid interior. It is remarkable that a crust shrinkage of only 10  $\mu\text{m}$  is sufficient to explain the period anomalies for the Crab pulsar. When similar reasoning is applied to the Vela pulsar, for which the anomalies are larger, it is found necessary to invoke a solid neutron lattice core in which strains imposed when the star was young are intermittently relaxed.

#### PLASMA PHYSICS OUTSIDE NEUTRON STARS

It is strange that there appears to be more understanding of the interior of neutron stars, than of their atmospheres wherein is generated the radiation which makes them detectable. Ginzburg and Zheleznyakov (14) have summarised the electrodynamic problems in detail. The model upon which theorists are concentrating most attention is that of an oblique magnetic rotator, in which the pulsar may be regarded as a dynamo, powered by the initial store of rotational kinetic energy, and converting this into radiation together with a flux of relativistic particles by means of the large magnetic field. The oblique rotator model was first considered by Pacini (15) before pulsars had been found, and it was Gold (8) who suggested that an extended corotating magnetosphere played a vital role.

Goldreich and Julian (16) showed that electrical forces arising from unipolar induction would be sufficient to drag charges from the stellar surface



**Fig. 6.** The neutron star magnetosphere for an aligned magnetic rotator.

and then distribute them in a corotating magnetosphere. It is not yet known whether such a distribution is stable, and the plasma differs from laboratory plasmas in that almost complete charge separation occurs. Inertial forces must dominate when the corotation velocity approaches  $c$ , and beyond the velocity of light cylinder the plasma breaks away to create a stellar wind. In such models the polar regions are believed to play a crucial role since particles can escape along 'open' field lines.

Within such an overall framework exists the ordered motion of the charges which generate the beamed radio waves that we observe, and also those regions which emit light and X-rays for the youngest pulsar in the Crab. The fascinating richness of the phenomena involving polarisation, pulse shapes, radio spectra, intensity variations, and complex secondary periodicities, must eventually provide vital evidence to resolve our present uncertainties. There is good reason to believe that the general outline is correct. Simple dynamics shows that the surface magnetic field strength  $B_0$  is proportional to  $\left(P \frac{dP}{dt}\right)^{-1/2}$ , where  $P$  is the pulsar period, and observations of many pulsars give  $B_0 \sim 10^8$  T when conventional neutron star models are assumed. Further evidence comes from pulsar ages which are approximately  $P\left(\frac{dP}{dt}\right)^{-1}$ .

Typical ages are  $10^6$ - $10^7$  years although  $10^3$  years is obtained for the Crab pulsar, in good agreement with the known age of the supernova.

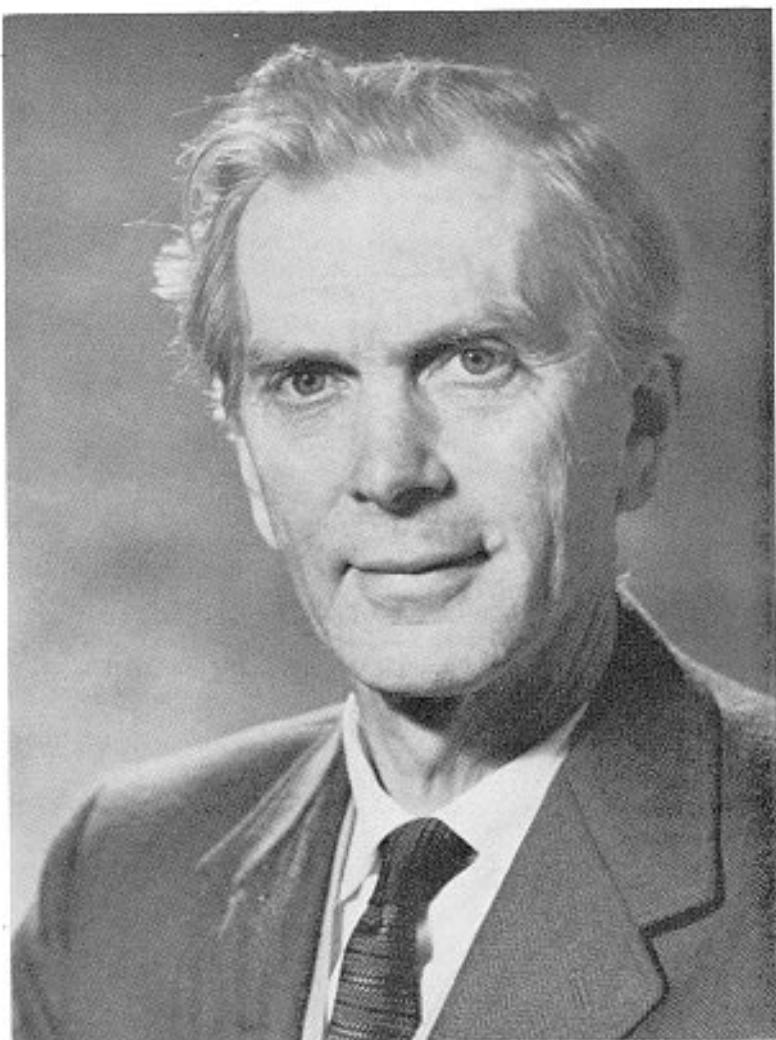
#### CONCLUSION

In outlining the physics of neutron stars, and my good fortune in stumbling upon them, I hope I have given some idea of the interest and rewards of

extending physics beyond the confines of laboratories. These are good times in which to be an astrophysicist. I am also deeply aware of my debt to all my colleagues in the Cavendish Laboratory. Firstly to Sir Martin Ryle for his unique flair in creating so congenial and stimulating a team in which to work. Secondly to Jocelyn Bell for the care, diligence and persistence that led to our discovery so early in the scintillation programme, and finally to my friends who contributed so generously in many aspects of the work.

## REFERENCES

1. Hewish, A., Proc. Roy. Soc. (London), 214, 494, (1952).
2. Hewish, A. and Wyndham, J. D., Mon. Not. R. astr. Soc., 126, 469, (1963).
3. Clarke, M. E., Ph.D. Thesis. Cambridge (1964).
4. Hewish, A., Scott, P. F. and Wills, D., Nature, 203, 1214, (1964).
5. Dennison, P. A. and Hewish, A., Nature, 213, 343, (1967).
6. Hewish, A. and Okoye, S. E., Nature, 207, 59, (1965).
7. Hewish, A., Bell, S. J., Pilkington, J. D. H., Scott, P. F. and Collins, R. A., Nature 217, 709, (1968).
8. Gold, T., Nature, 218, 731, (1968).
9. Staelin, D. H. and Reifenstein, E. C., Science, 162, 1481, (1968).
10. Comella, J. M., Craft, H. D., Lovelace, R. V. E., Sutton, J. M. and Tyler, G. L., Nature, 221, 453, (1969).
11. Baade, W., and Zwicky, F., Proc. Nat. Acad. Sci. 20, 255, (1934).
12. Ruder-man, M., Ann. Rev. Astron. Astrophys., 20, 427, (1972).
13. Pines, D., Shaham, J., and Ruder-man, M., I.Au.c Symposium No 53, (1973).
14. Ginzburg, V. L. and Zheleznyakov, V. V., Ann. Rev. Astron. Astrophys., 13, (1975) in press.
15. Pacini, F., Nature, 219, 145, (1968).
16. Goldreich, P. and Julian, W. H., Astrophys. J., 157, 869, (1969).



Martin Rydberg

## MARTIN RYLE

I was born on September 27, 1918, the second of five children. My father John A. Ryle was a doctor who, after the war, was appointed to the first Chair of Social Medicine at Oxford University.

I was educated at Bradfield College and Oxford, where I graduated in 1939. During the war years I worked on the development of radar and other radio systems for the R.A.F. and, though gaining much in engineering experience and in understanding people, rapidly forgot most of the physics I had learned.

In 1945 J. A. Ratcliffe, who had been leading the ionospheric work in the Cavendish Laboratory, Cambridge before the war, suggested that I apply for a fellowship to join his group to start an investigation of the radio emission from the Sun, which had recently been discovered accidentally with radar equipment.

During these early months, and for many years afterwards both Ratcliffe and Sir Lawrence Bragg, then Cavendish Professor, gave enormous support and encouragement to me. Bragg's own work on X-ray crystallography involved techniques very similar to those we were developing for "aperture synthesis", and he always showed a delighted interest in the way our work progressed.

In 1948 I was appointed to a Lectureship in Physics and in 1949 elected to a Fellowship at Trinity College. At this time Tony Hewish joined me, and in fact four other members of our present team started their research during the period 1948-52.

In 1959 the University recognized our work by appointing me to a new Chair of Radio Astronomy.

During 1964-7 I was president of Commission 40 of the International Astronomical Union, and in 1972 was appointed Astronomer Royal.

In 1947 I married Rowena Palmer, and we have two daughters, Alison and Claire, and a son, John. We enjoy sailing small boats, two of which I have designed and built myself.

### *Awards*

- 1952 Fellow of Royal Society of London.
- 1954 Hughes Medal, Royal Society of London.
- 1955 Halley Lecturer, University of Oxford.
- 1958 Bakerian Lecturer, Royal Society of London.

- 1963 Van der Pol Medal, U.R.S.I.
- 1964 Gold Medal, Royal Astronomical Society, London.
- 1965 Henry Draper Medal, U.S. National Academy of Sciences; Holweck Prize, Société Francaise de Physique.
- 1968 Elected Foreign Member of the Royal Danish Academy of Sciences and Letters.
- 1970 Elected Foreign Honorary Member of American Academy of Arts and Sciences.
- 1971 Elected Foreign Member of U.S.S.R. Academy of Sciences.  
Morris N. Liebmann Award; Institution of Electrical & Electronic Engineers.  
Faraday Medal, Institution of Electrical Engineers.  
Popov Medal, USSR Academy of Sciences.  
Michelson Medal, Franklin Institute, U.S.A.
- 1973 Royal Medal, Royal Society of London.
- 1974 Bruce Medal, Astronomical Society of the Pacific.  
Foreign Member, Deutsche Akademie der Naturforscher, Leopoldina.  
Honorary DSc. of the Universities of Strathclyde (1968), Oxford (1969) and Nicholas Copernicus University of Torun (1973).

*Sir Martin Ryle died in 1984.*

# RADIO TELESCOPES OF LARGE RESOLVING POWER

Nobel Lecture, December 12, 1974

by **MARTIN RYLE**

University of Cambridge, Cavendish Laboratory, Cambridge, England

I think that the event which, more than anything else, led me to the search for ways of making more powerful radio telescopes, was the recognition, in 1952, that the intense source in the constellation of Cygnus was a distant galaxy-1000 million light years away. This discovery showed that some galaxies were capable of producing radio emission about a million times more intense than that from our own Galaxy or the Andromeda nebula, and the mechanisms responsible were quite unknown. It seemed quite likely that some of the weaker sources already detected with the small radio telescopes then available might be similar in character; if so they would be at distances comparable with the limits of observation of the largest optical telescopes. It was therefore possible that more powerful radio telescopes might eventually provide the best way of distinguishing between different cosmological models. It was not until 1958 (1) that it could be shown with some certainty that most of the sources were indeed powerful extragalactic objects, but the possibilities were so exciting even in 1952 that my colleagues and I set about the task of designing instruments capable of extending the observations to weaker and weaker sources, and of exploring their internal structure.

The early observations were severely limited both by the poor angular resolution and by the limited sensitivity. It was usually impossible to obtain any information about the structure of a source, and adjacent sources could often not be properly separated, whilst attempts to relate the radio sources to optically visible objects were often prevented by the poor positional accuracy. The use of interferometers allowed better positions to be obtained, and sometimes made it possible to derive simple models for the source structure. Few of the sources were found to have an angular size greater than 2-3 minutes of arc.

The problem of making detailed maps of such sources arises simply from the fact that the wavelengths used are some million times greater than optical wavelengths-so that even to obtain a radio picture with the same resolution as that of the unaided human eye ( $\sim 1'$  arc) we would need a telescope having a diameter of about 1 km operating at a wavelength of 50 cm. At the same time the instrument will be effective only if the surface accuracy is good enough to make a proper image, corresponding to errors of  $\leq \lambda/20$  or a few cm; the engineering problems of building such an instrument are clearly enormous.

With the development, around 1960, of masers and parametric amplifiers capable of providing receiving systems of good sensitivity at wavelengths of a few cm, it became possible to build telescopes of diameter 10-100 m with

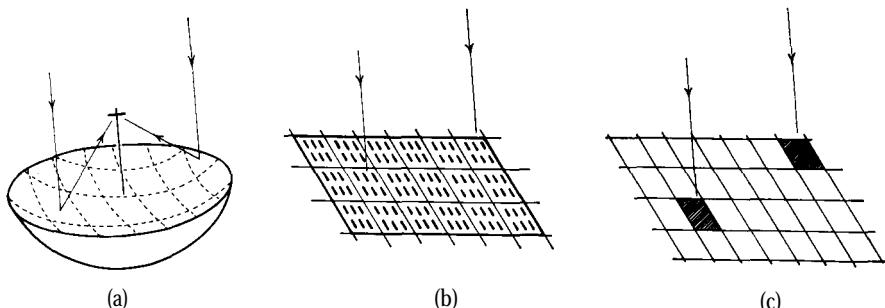


Fig. 1. The use of (a) a paraboloid, (b) an array of dipoles or (c) the sequential sampling of the wavefront by small aerial elements to achieve a high resolving power by combining the signals from a large part of the incident wavefront.

sufficient sensitivity and with angular resolutions of  $b 1'$  arc; even with such instruments the engineering problems of constructing a rigid enough surface are considerable, and it is likely to be difficult to build a conventional paraboloid capable of angular resolutions much better than  $1'$  arc.

I would like now to describe an entirely different approach to the problem in which small aerial elements are moved to occupy successively the whole of a much larger aperture plane. The development and use of "aperture synthesis" systems has occupied much of our team in Cambridge over the past 20 years.

The principle of the method is extremely simple. In all methods used to obtain a large resolving power, that is to distinguish the wavefront from a particular direction and ignore those from adjacent directions, we arrange to combine the field measured over as large an area as possible of the wavefront. In a paraboloid we do this by providing a suitably shaped reflecting surface, so that the fields incident on different parts of the sampled wavefront are combined at the focus (Fig. 1 (a)); the voltage produced in the receiving dipole represents the sum of these fields. We can achieve the same result if we use an array of dipoles connected together through equal lengths of cable (Fig. 1 (b)).

Suppose now that only a small part of the wavefront is sampled, but that different parts are sampled in turn (Fig. 1 (c)); could we combine these signals to produce the same effect? Since in general, we do not know the phase of the incident field at different times this would not normally be possible but if we continue to measure one of the samples while we measure the others we can use the signal from this one as a phase reference to correct the values measured in other parts of the wavefront. In this way, by using two small aerial elements, we can again add the fields over the wavefront—the area of which is now determined by the range of relative positions taken by the two aerial elements.

It might be thought that this method would be extremely slow, for if we are to sample an area of side  $D$  using elements of side  $d$ , it is necessary to

observe with  $\frac{2D^3}{d^3}$  different relative positions of the two aerial elements. In

practice, however, the method is not significantly slower than the use of the large equivalent instrument for although a large number of observations must be made, the results may be combined in a computer using additional phase differences, which correspond to many different wave directions (as in a phased array or dipoles), so that with the one set of observations an area of sky may be mapped which is limited only by the diffraction pattern of the small elements themselves; there are in fact some  $\frac{D^2}{d^2}$  different directions which can be scanned in this way, and which would have had to be explored sequentially by a conventional instrument, so that the total observing time of the two methods is nearly the same.

It can also be seen that the sensitivity of the system is much better than would be associated with the small elements, for the signal from a particular point in the sky is contributing to that point on the map for the whole observing period; the resulting signal-to-noise is in fact equivalent to the use of

an instrument having a collecting area  $2d^2 \frac{2D^2}{d^2}$

$\sim 3Dd$ , a figure which may be much greater than that of the elements themselves, and although it is not as great as if the full instrument of area  $D^2$  had been built, it may exceed that of any instrument which can be built.

Unlike a paraboloid or array, in which both the sensitivity and resolving power are fixed as soon as the wavelength is decided, the value of  $d$  may be chosen so that the sensitivity, for any particular wavelength and type of observation, is matched to the resolution.

The method of aperture synthesis avoids the severe structural problems of building very large and accurate paraboloids or arrays, and allows both high resolving power and large effective collecting area to be obtained with a minimum of engineering structure and therefore cost. Provision must be made for the relative movement of the small elements, and their relative positions and electrical connecting paths must be known with an accuracy equal to the surface accuracy of the equivalent instrument ( $< i/20$ ). Automatic computing is needed to carry out the Fourier inversion involved in combining the observations to provide a map of the sky.

Historically, the forerunners of this type of instrument were realized in the early days when observations in both Australia and England with aerial elements having a range of separations were used to determine the distribution of radio brightness across the solar disc. In the earliest observations the Sun was assumed to show spherical symmetry, and no measurements of phase were necessary so that a precise knowledge of the relative positions of the elements, and of the electrical path lengths to the receiver were unnecessary. A similar technique was used to establish the profile of radio brightness across the plane of the Galaxy (2).

The first synthesis instrument capable of mapping an arbitrary distribution of sources was built at Cambridge in 1954 by John Blythe (3); it consisted of a long thin element covering, in effect, a whole row of Fig. 1 (c) used in con-

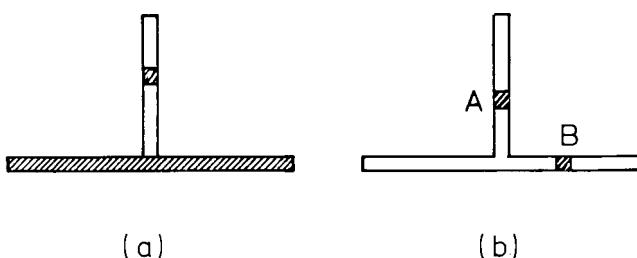


Fig. 2. (a) The arrangement used in the instrument built in 1954 by J. H. Bythe.  
 (b) The equivalent instrument using two small elements.

junction with a smaller element moved to 38 different positions along a perpendicular line (Fig. 2 (a) ) to synthesise a square instrument giving a resolution of  $2^{\circ}.2$ . This instrument provided the first detailed maps of the galactic emission at a long radio wavelength (7.9 m).

Larger instruments using this same configuration were built at Cambridge during the succeeding years, including an instrument of high sensitivity and  $45'$  arc resolution also at  $\lambda = 7.9$  m (4) and a second operating at  $\lambda = 1.7$  m with  $25'$  arc resolution which was used by Paul Scott and others to locate nearly 5000 sources in the northern sky (5, 6).

These instruments used a very cheap form of construction; for  $\lambda > 1$  m an efficient reflecting surface may be provided by thin ( $\sim 1$  mm diameter) wires 5-10 cm apart. In the case of the  $\lambda = 1.7$  m instrument, wires stretched across simple parabolic frames of welded steel tube provided a cylindrical paraboloid 450 m long and 20 m wide (Fig. 3) at a cost of about £2 per m<sup>2</sup>.

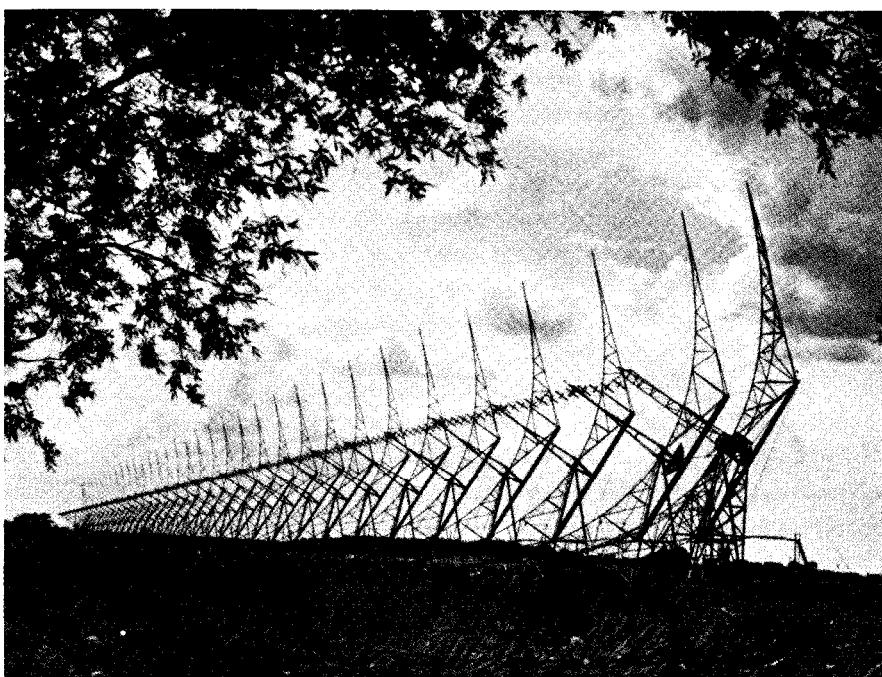


Fig. 3. Photograph of the east-west arm of the  $\lambda = 1.7$  m instrument built in 1957 with which nearly 5000 sources were located.

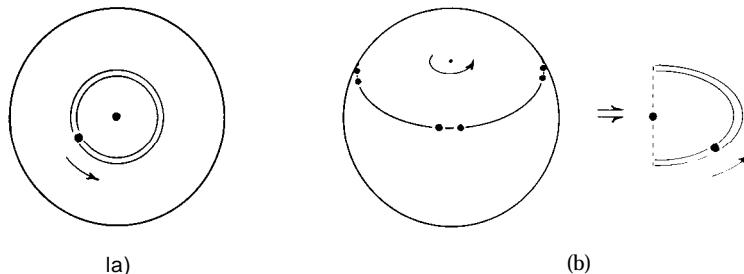


Fig. 4. (a) Two aerial elements mounted near the North Pole observing throughout the day are equivalent to one ring of a much larger instrument.  
 (b) The elements may be used at other latitudes if arranged on an east-west line and used to track the chosen point for 12<sup>h</sup>.

With the need for still greater resolving power, we realized that physically larger systems operating at metre wavelengths would no longer prove successful, because of the limitation imposed by irregularities of electron density in the ionosphere. But at shorter wavelengths where these are unimportant it becomes difficult to make efficient reflectors by using stretched wires, both because of their deflection by the wind, and because with the closer spacing needed there is difficulty with them twisting together. For operating wavelengths of < 50 cm a much more rigid supporting structure must be used, and the engineering costs of building a long element become very great.

The obvious solution is to use the system illustrated in Fig. 1 (c), in which the engineering structure is confined to two small elements-where much higher costs per m<sup>2</sup> are acceptable. The method for altering the relative positions of the two elements presents some practical problems; suppose that the elements are mounted on two railway tracks at right angles (Fig. 2 (b) ), so that for each position of A on the N-S track B is moved to every position along the E-W track. For values of  $\frac{D}{d} \sim 50$ , there are then 5000 different arrangements, and if B is moved each day, the observations will take 5000 days and although a map will then be available for the whole strip of sky, the period is too long for a graduate student's thesis!

Alternatively B could be moved rapidly-so that several positions could be fitted into the time during which the area of sky remains in the beam of the small elements. This will reduce the total time of the observations, at the expense of observing only parts of the strip of sky. We can clearly extend this period, and so allow more relative positions of A and B each day, if we arrange for the elements to track the chosen point in the sky for an extended period.

As soon as we do this, we realize that the rotation of the earth is itself providing us with a relative motion of A and B as seen from the source, without our having to move them on the surface of the earth at all. Suppose, for example, we have our two elements mounted near the North Pole and we use them to observe an area of sky centred on the Celestial Pole; in this case we do not even have to arrange for them to track. Over a 24<sup>h</sup> period, one will have traced out a circular path about the other (Fig. 4 (a) ), and the signals

recorded during this time can be combined to provide the same response as that of the equivalent ring aerial; by simply altering the separation along a line on say 50 successive days a complete aperture can then be synthesized. Miss Ann Neville and I set up an experimental system in 1960-61 to test the method and develop the computing; we used it to map a region  $8^\circ$  in diameter round the North Celestial Pole at a wavelength of 1.7 m (7). We connected up different 14 m sections of the long cylindrical paraboloid (Fig. 3) with some other small aerials to simulate the use of two 14 m diameter elements at different spacings. The effective diameter of the synthesized instrument was 1 km and it provided an angular resolution of  $4'.5$  arc.

As well as showing that the method really worked, it provided some interesting astronomical results-in particular by allowing the detection of sources some 8 times weaker than had been observed before; even though the area of sky covered was only some 50 square degrees the results were useful in our cosmological investigations.

In practice only  $12^\circ$  observations are needed because of the symmetry of the system and observations need not be made from the North Pole or limited to the Celestial Pole, provided that the elements are situated on an East-West axis, and each is able to track the required region of sky for  $12^\circ$  (Fig. 4 (b)). At low declinations the synthesized instrument becomes elliptical with the north-south aperture reduced by  $\sin \delta$ . The engineering simplicity of moving the elements along a line, and the consequent great saving in the area of land needed are, however, such great advantages that we eventually built three large instruments in Cambridge with equivalent instrumental diameters of 0.8, 1.6 and nearly 5 km.

These instruments are known as the Half-Mile, the One-Mile and (because its construction coincided with the early negotiations for the entry of Britain into the European Community), the 5 km Telescopes! The One-Mile telescope was the first to be built, and this started observations in 1964.

It is interesting that as early as 1954 we had discussed the possibility of building a high resolution instrument on exactly these principles, and I have recently found two entries in an old note-book:-

"8.6.1954 Possible research student and other projects.

... 3(f). North Polar Survey on 81.5 Mc/s. Effective gain area  $\sim 25 \times 1500 = 37,500$  sq. ft. Effective resolving power area  $\simeq 10^6$  sq. ft."

(The entry included a diagram of the proposed aerial element)

"29.6.1954

Do 3(f) in all directions where  $180^\circ$  rotation available? above about  $20^\circ$  might be possible by directing aerials in successively different directions-i.e. observation not on meridian."

A third entry on 22.7.1954 discusses the east-west rail track to be used for the latter programme with two 30 ft aerials mounted on it, the arrangement of cabling needed to compensate for the different path lengths to the two aerials when observing off the meridian, and the selection of directions of observation "to give uniform cover of Fourier terms".

Why then, with its obvious simplicity and economy, did we not build this instrument in 1954? The answer is that at this time there were no computers with sufficient speed and storage capacity to do the Fourier inversion of the data. EDSAC I, which was the first stored-programme computer, was built by Dr. M. V. Wilkes of the Cambridge University Mathematical Laboratory, and came into operation in 1949. It was used for reducing John Blythe's observations and took some 15<sup>h</sup> of computing to do the 38-point transform for every 4<sup>m</sup> of the 24<sup>h</sup> scan of the sky. It would not have been practicable to use it for the 2-dimensional inversion needed for the earth-rotation synthesis. By 1958 the completion of the much faster EDSAC II, and the development by Dr. David Wheeler of the Mathematical Laboratory of the fast Fourier transform (incidentally some six years before these methods came into general use) made possible the efficient reduction of the 7.9 m and 1.7 m surveys, and also enabled the trials of the 1.7 m earth-rotation synthesis to be made in 1961; even with EDSAC II, however, the reduction for the small area of sky covered in the latter survey took the whole night.

During the early stages in the design of the One-Mile telescope in 1961, I discussed with Maurice Wilkes the considerably greater problems of reducing the data from this instrument, but by then the replacement of EDSAC II was planned and the new TITAN computer, which came into operation in 1963, was easily capable of dealing with the output of the One-Mile telescope. The development of aperture synthesis has therefore been very closely linked to the development of more and more powerful computers, and it is interesting to speculate how our work in Cambridge would have proceeded if, for example, computer development had been five years behind its actual course.

The two programmes in my 1954 note-book subsequently formed the basis of two Ph.D. theses in 1964 and 1965.

Now I return to the design of the large instruments whose layout is shown in Fig. 5. The One-Mile telescope consists of three 18 m dishes, two fixed at 0.8 km spacing, the third mounted on a 0.8 km rail-track (Fig. 6); this ar-

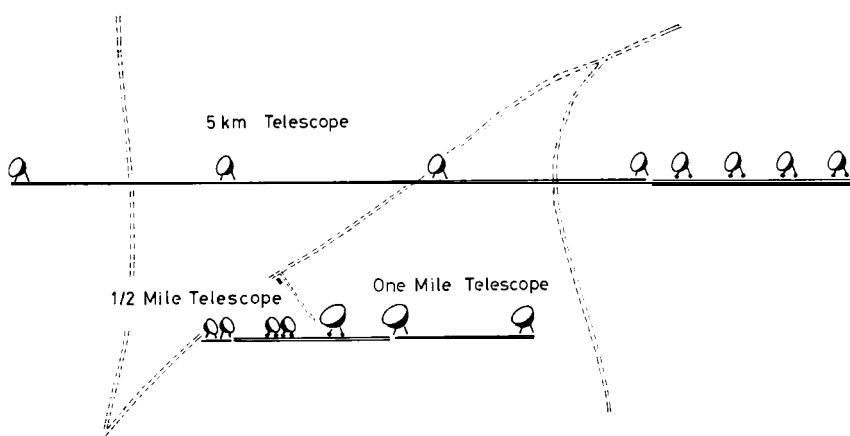


Fig. 5. Sketch map showing the arrangement of One-Mile, Half-Mile and 5 km telescopes.

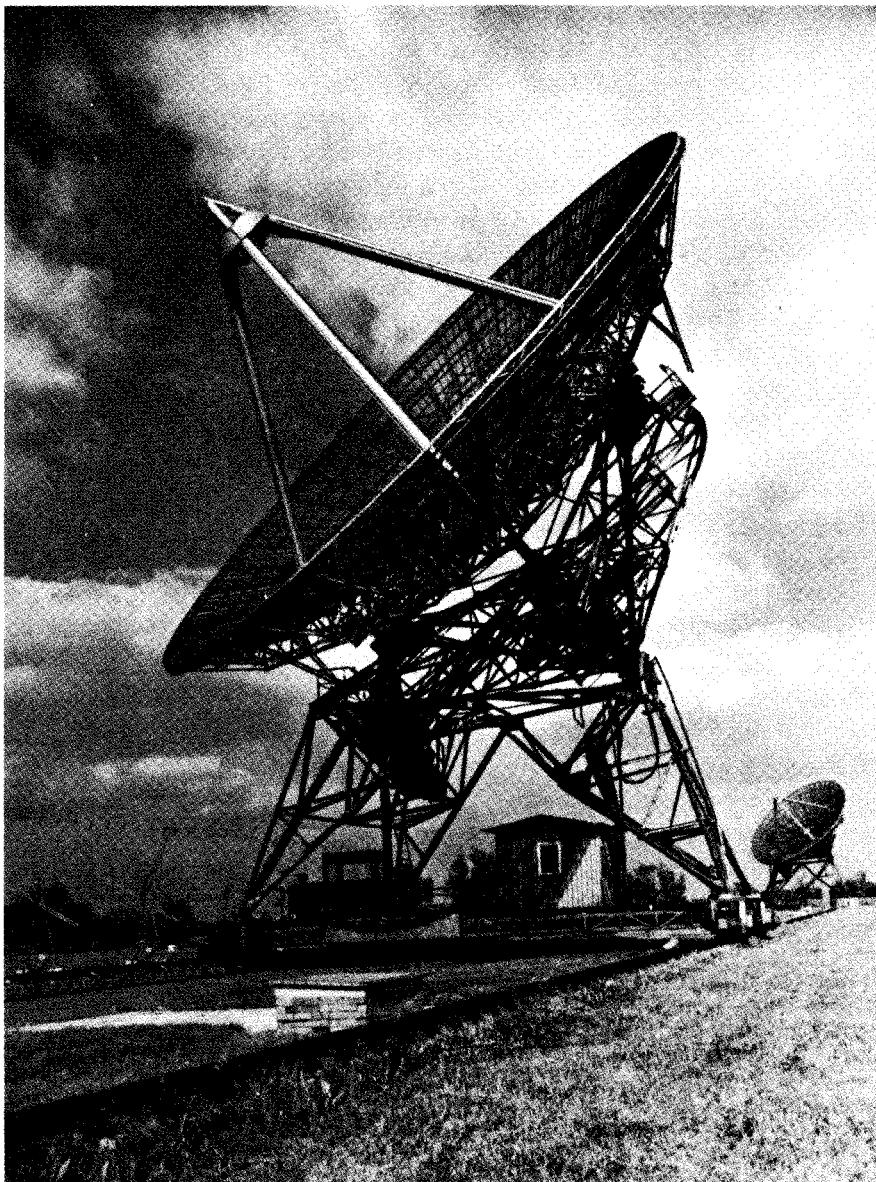


Fig. 6. The One-Mile telescope, showing the west, railmounted, dish in the foreground, with the two fixed dishes behind.

rangement was cheaper than building the longer rail track and it also provided two spacings at a time. It was designed for two main programmes: (a) The detection of much fainter and therefore more distant sources (see Fig. 7) in order to explore the early history of the Universe, and so try and distinguish between different cosmological models, and (b) To make radio maps of individual sources in an attempt to understand the physical mechanisms within them; most of the sources studied have been powerful extragalactic objects, but the remnants of supernova explosions are perhaps physically as important.

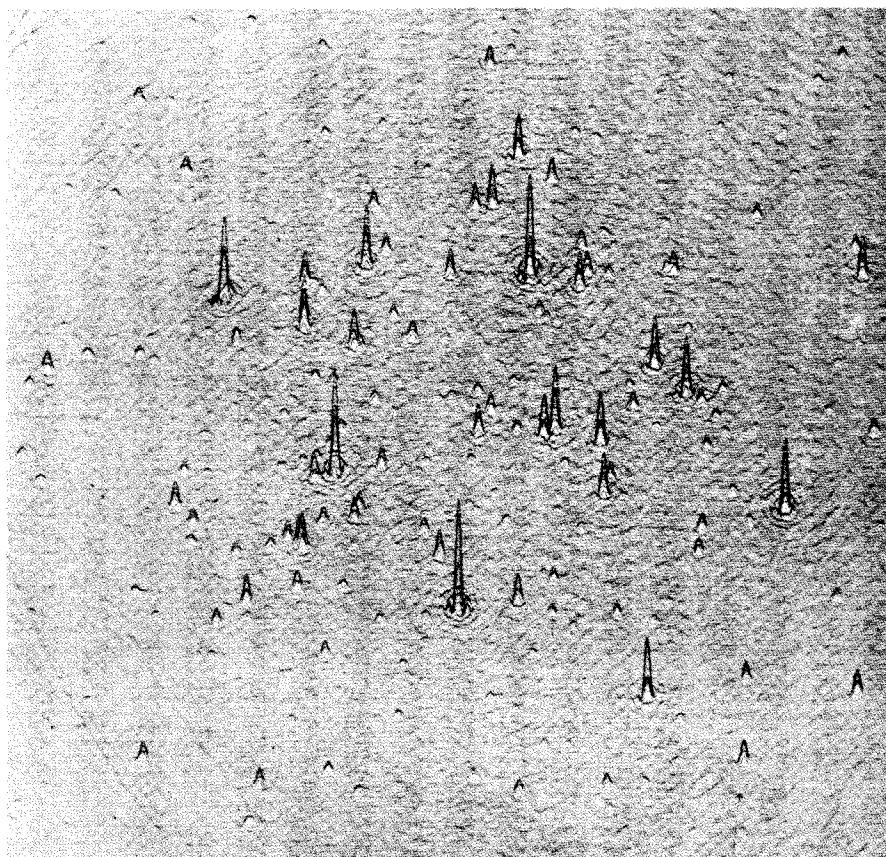


Fig. 7. Map obtained with the One-Mile telescope showing sources about 100 times fainter than had been observed before.

The problem of the physics of radio galaxies and quasars and the cosmological problem are strangely linked; we appear to be living in an evolving Universe, so that very distant sources which, due to the signal travel time, we observe as they were when the Universe was younger, may be systematically different from a sample of nearby sources. But the intrinsically most powerful sources are so rare that there are no nearby ones, whilst the weak sources cannot be detected at great distances. If we are to understand *how* the Universe is evolving, we may first have to solve the physical problem of the individual source-so that we can infer the differences in its evolution at earlier cosmological epochs.

The Half-Mile telescope was built later by John Shakeshaft and John Baldwin. It was actually built very cheaply because as can be seen from Fig. 5, it made use of the same rail track, and we were able to get the four 9 m dishes at scrap-metal prices from a discontinued radio link service, and only the mounts had to be built. It has been used mainly with a radio spectrometer covering the 21 cm wavelength band of neutral hydrogen to map the distribution of density and velocity of the hydrogen in a number of nearby galaxies,

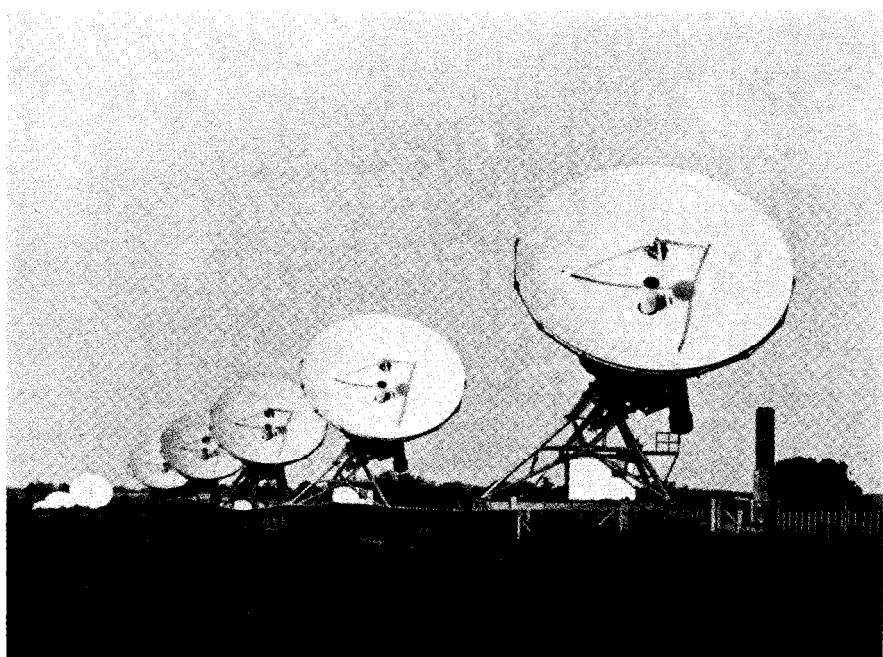


Fig. 8. The 5 km telescope, with the movable dishes in the foreground.

and forms part of a programme concerning the formation and evolution of galaxies.

The 5 km telescope was completed in 1971, and because it represents a rather more advanced design I will describe it in more detail. It was designed solely for the purpose of mapping individual sources, and besides its larger overall size, the individual dishes are more accurate to allow operation at wavelengths as short as 2 cm. As a result the angular resolution is  $\sim 1''$  arc, a figure comparable with the resolution attained by large optical telescopes on good mountain sites. It is at present being used on a wavelength of 6 cm, where the resolution is  $2''$  arc.

In order to improve the speed of observation, four fixed and four movable elements mounted on a rail-track are used, as shown in Fig. 5; this arrangement provides 16 spacings simultaneously, and a single  $12''$  observation produces a  $2''$  arc main response with circular grating responses separated by  $42''$  arc. Sources of smaller extent than  $42''$  arc can therefore be mapped with a single  $12''$  observation; more extensive fields of view require further observations with intermediate positions of the movable elements to suppress the grating responses.

For operation at these short wavelengths the positioning of the elements, and the electrical cable connections, must be stable and measured with an accuracy better than 1 mm. Conventional surveying methods allowed each element to be located to  $\pm 10$  mm, and the final alignment had to be based entirely on radio observations; the distance between the two outer fixed elements (on which the scale of declination is based) was found in this way

to be  $3430828.7 \pm 0.25$  mm, and no changes outside this error have been found over a 2-year period. The combination of azimuth and longitude, on which the measurement of right-ascension depends, was established by observing the bright fundamental star Algol, which is a weak and variable radio source.

The telescope is controlled by an on-line computer which continually updates the position of the selected map-centre for precession, aberration etc., and uses this to compute the path differences (corrected for atmospheric refraction) to each pair of elements; these values are then used to control electrical delays in the signals from each element before they are combined in the receivers. The outputs of the receivers are sampled by the computer and stored on a magnetic disc, so that at the completion of the observation they may be combined to form a map of the area observed. The map is then drawn on a curve-plotter controlled by the computer.

This instrument has been used in a wide range of astronomical programmes from the study of ionized hydrogen clouds in our Galaxy to distant quasars. Following the accurate calibration survey it became evident that as an astrometric instrument-to establish a coordinate system across the sky-its measuring accuracy was comparable with the best optical methods, whilst overcoming some of the difficulties in optical work such as the measurement of large angles. Bruce Elsmore is involved in a collaborative programme with optical observers to relate the positions of quasars-(some of which are compact sources at both optical and radio wavelengths)-as measured by radio means, to those derived from the fundamental stars, in order to determine any large-scale non-uniformities which may exist in the present astrometric systems. He also showed how this type of instrument may be used for the direct measurement of astronomical time-without the need for collaborative observations at different longitudes to correct for polar motion-again with an accuracy comparable with optical methods ( $\sim 5$  mS in a  $12^{\text{h}}$  observation).

Another programme is concerned with a study of the birth of stars; when a cloud of gas condenses to form a star, the dust which it contains provides such an effective screen that newly-formed stars, with their surrounding regions of ionized hydrogen, can never be seen optically; only after this dust cloud has dispersed does the star appear. The dust introduces no appreciable absorption at radio wavelengths, so that radio observations allow these regions to be studied at the earliest stages.

NGC 7538 is an example of such a region, and the upper part of Fig. 9 shows the radio emission as mapped with the One-Mile telescope. The large diffuse component corresponds almost exactly with the optical nebulosity, and represents the cloud of gas ionized by one or more O-stars formed about a million years ago, with the dust sufficiently dispersed to allow the light to be seen. The compact lower component corresponds to gas ionized by much younger stars, which are still embedded in dust too dense for any optical emission to escape, and it is invisible on the photograph. When this southern component was mapped with the higher resolution of the 5 km telescope, the

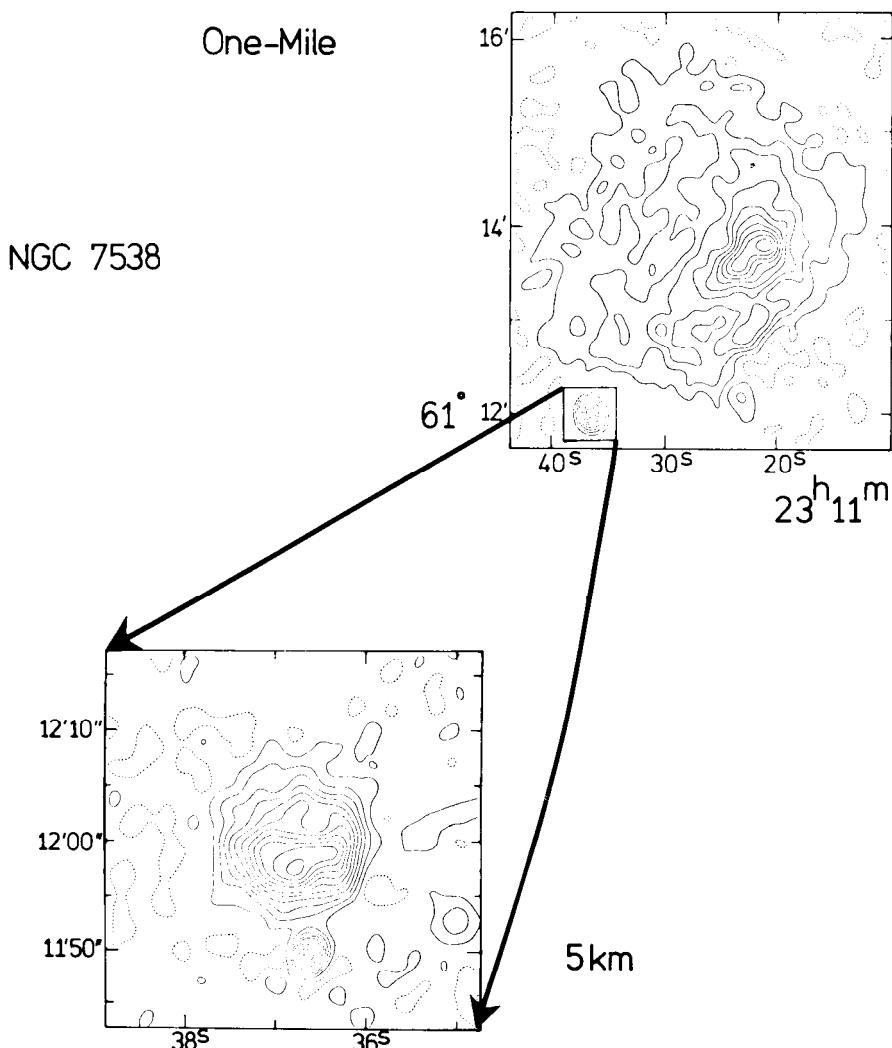


Fig. 9. The ionized hydrogen cloud NGC 7533. The upper radio map shows the large cloud associated with the optical emission, and another, compact, component to the south. This compact component is shown with greater resolution below.

lower map was obtained, showing that there is an ionized cloud some 10" arc in diameter, probably produced by the radiation from a star of spectral type O8, and an even more compact cloud to the south of this, produced by a still younger star, only a few thousand years old. The dust surrounding these two compact regions is heated by the young stars they contain, and both have been detected by their infra-red emission (8).

But the most extensive programme has been the mapping of extragalactic sources—the radio galaxies and quasars; galaxies which, during a brief fraction of their lives, produce some  $10^{80}$  ergs of energy, equivalent to the total annihilation of the matter in about a million suns, by a mechanism which is not understood.

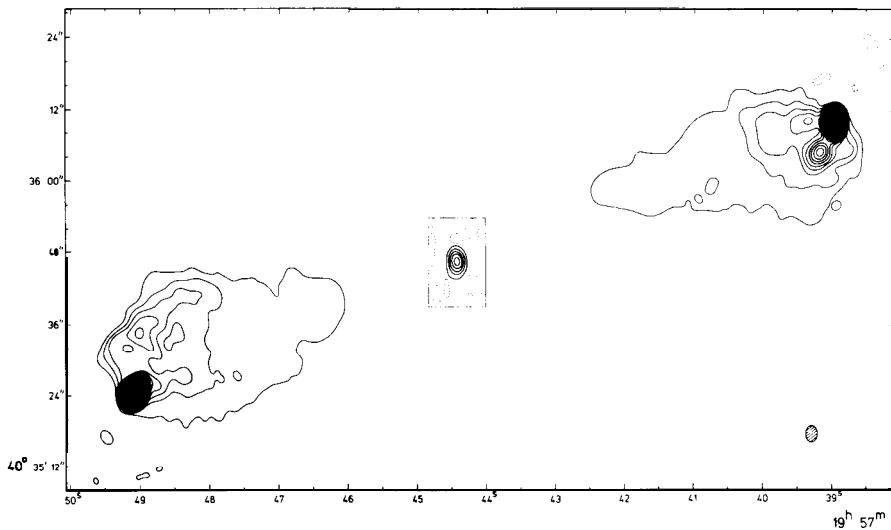


Fig. 10. The powerful radio **galaxy** in the constellation of Cygnus mapped with the 5 km telescope. The compact outer components are exceedingly bright-(31 and 41 contours). The central component—which corresponds to the nucleus of the optical galaxy is very weak and is drawn with contours spaced at 1/5 the interval.

Fig. 10 shows the new radio map of the source in the constellation of Cygnus—the first powerful radio galaxy to be recognized. The distribution of polarized emission from the north component is shown in Fig. 11, giving in-

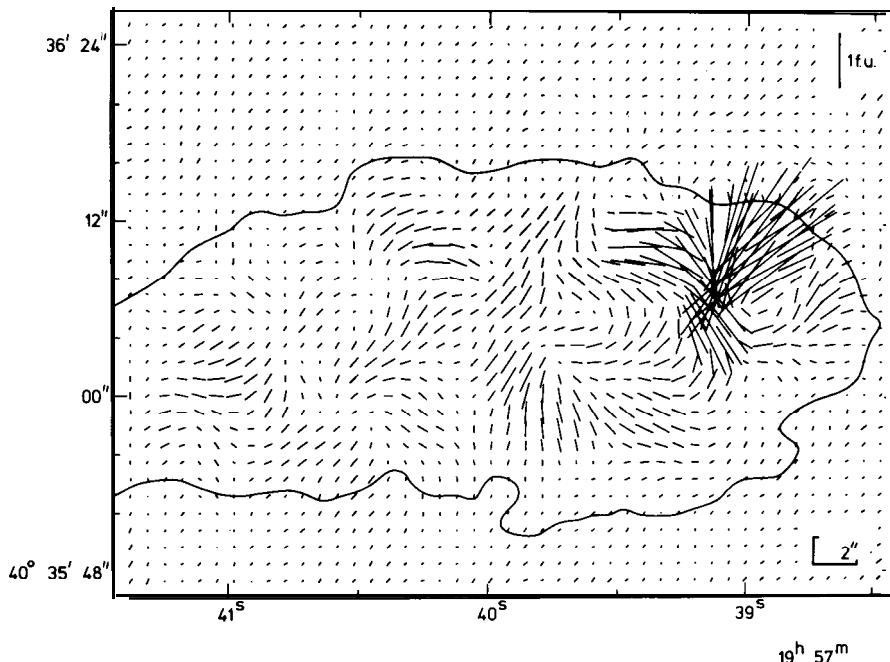


Fig. 11. The polarization of the emission from the north component of the Cygnus source which shows the magnetic field to be turbulent on a scale  $\sim 10^4$  light-years.

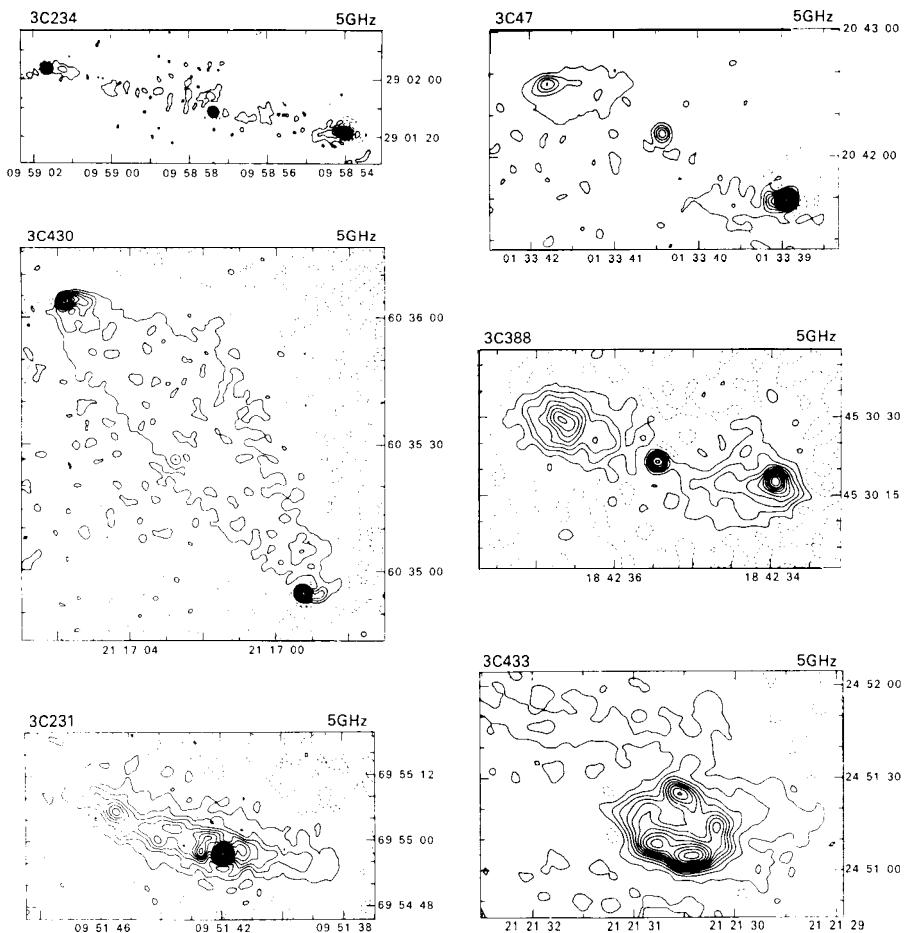


Fig. 12. Maps of six extragalactic radio sources.

formation on the magnetic field. Maps of a number of other sources made with the 5 km telescope are shown in Fig. 12.

In most cases the radio emission originates mainly in two huge regions disposed far outside the associated galaxy—although weak emission may also be detectable from a very compact central source coincident with the nucleus of the galaxy. In some cases much more extensive components or a bridge linking the components occur.

The finer detail provided by the 5 km telescope has already enabled some important conclusions to be drawn; the energy is probably being produced more or less continuously over a period of  $10^7$ – $10^8$  years in a very compact nucleus and not, as was originally thought, in some single explosive event. The source of this energy may be associated with the gravitational collapse of large numbers of stars, in the manner which Tony Hewish describes in his lecture, or by material falling into a much more massive collapsed object at the nucleus of the galaxy. The mechanism for transmitting this energy to the compact heads of the main components (e.g. Fig. 10) is not understood, but

may involve a narrow beam of low frequency electromagnetic waves or relativistic particles (9, 10). The interaction of this beam with the surrounding intergalactic medium might then accelerate the electrons responsible for the radio emission from the compact heads, and their subsequent diffusion into the region behind the heads can probably explain the general shape of the extensive components.

While much remains unanswered, the present conclusions were only reached when detailed maps became available; the physical processes relating the nucleus, the compact heads, and the extensive tails or bridges can clearly only be investigated when the relationship between these structural components is known.

What can we expect in the future? In 1954, the first aperture synthesis telescope provided maps with a resolution of  $2^{\circ}.2$ ; today we have maps with a resolution of  $2''$  arc. Can we foresee a continuing development with radio pictures having much better resolution than the optical ones? The technical problems of increasing the aperture or decreasing the operating wavelength are severe, but they do not appear to be as serious as the limitations imposed by the earth's atmosphere; in optical observations atmospheric turbulence on a scale of  $\sim 10$  cm in the lower atmosphere introduces irregularities in the incident wavefront which normally limits the resolution to  $\sim 1''$  arc. At radio wavelengths the contribution of these small-scale irregularities is not important, but there are also irregularities of refractive index on a much larger scale in the troposphere. Two distinct types have been found in a series of observations with the One-Mile and 5 km telescopes; neither can be attributed to variations of air density, and both are probably due to non-uniformity in the partial pressure of water-vapour, which makes an important contribution to the refractive index at radio wavelengths. One class has a typical scale size of  $\sim 0.7$  km and is attributed to turbulence in the troposphere due to solar heating of the ground in the same way that fair-weather cumulus clouds develop. These irregularities, however, are often detected in clear air conditions without the formation of cumulus clouds; they only occur during day-time and are more severe during summer months. The second class, which shows only slight diurnal or annual variation, has a much larger scale size, typically 10-20 km, and there may be still larger scales which have not yet been recognized. The origin of these disturbances is not known, and it is therefore not possible to predict how they might depend on geographical position.

Under very good conditions-representing about 1% of the total time, the atmospheric irregularities are extremely small and correspond to a distortion of the incident wavefront by  $< 0.2$  mm over 5 km; under these conditions, operation at a wavelength of 4 mm or less would be possible and should provide maps with a resolution better than  $0''.2$  arc. These excellent observing conditions have only been encountered during periods of widespread winter fog when the atmosphere is extremely stable, a result which illustrates the differing requirements in seeking good sites for optical and radio observatories!

For most of the time the atmospheric irregularities are considerably worse, and although there is insufficient information on scale sizes  $> 20$  km, the use of instruments much larger than this will introduce difficulties associated with the curvature of the atmosphere. One might guess that it should be possible to build instruments which would give a resolution better than  $0''.5$  arc for perhaps 50% of the winter months.

To reach a greater resolution new techniques capable of correcting for the atmospheric effects will be necessary. One simple, though expensive, solution would be to build a second dish alongside each element, so that observations of a reference point source close to the area to be mapped, could be made simultaneously at every spacing; the observed phase errors for this reference source could then be used to provide a continuous correction for the signals from the area being mapped.

Such techniques can clearly be extended to the interferometers having baselines of many thousands of km (VLBI) which have been made possible by the development of atomic frequency standards. These instruments have shown the existence of very small components,  $\sim 0''.001$  arc in some sources. The use of a comparison source for eliminating both atmospheric and instrumental phase was first used at Jodrell Bank in the special case of sources of the OH maser line at  $\lambda = 18$  cm, where different components within the primary beam can be distinguished by their frequency; if one is used as a phase reference the relative positions of the others can be found (11).

For continuum sources a reference outside the primary beam of the instrument must, in general, be used and two elements at each location are needed. This technique has been used in the U.S.A. to reduce both instrumental and atmospheric phase variations in measurements of the gravitational deflection of radio waves by the sun (12); one pair of elements was used to observe a source close to the Sun, while the other pair observed a reference source about  $10^\circ$  away.

The accuracy of the correction, and hence the shortest wavelength at which mapping could be achieved, would depend on the angular separation between the area to be mapped and a reference source sufficiently intense and of sufficiently small angular size. But even if adequate phase stability can be attained in this way, there is a serious practical difficulty in making maps with resolution  $\sim 0''.001$  arc, due to the inevitable poor sampling of the aperture plane. Even with 5 or 6 stations distributed across one hemisphere of the world, and using every possible combination of the signals from them, with observing periods lasting several hours, the fraction of the aperture plane which can be filled is still very small, so that the field of view which can be mapped without ambiguity from secondary responses is unlikely to exceed  $\sim 0''.02$  arc. Whilst there seems little hope of deriving complete maps of most sources with this resolution, there are certainly some central components where such a map could provide very important information.

But I think it may also be important for our understanding of the mechanisms operating in the main components of radio sources, to obtain complete maps with intermediate resolution; for this work extensions of the

present synthesis techniques, while retaining good filling of the aperture plane, are needed.

The last 25 years have seen a remarkable improvement in the performance of radio telescopes, which has in turn led to a much greater understanding of the strange sources of "high-energy astrophysics" and of the nature of the Universe as a whole.

I feel very fortunate to have started my research at a time which allowed me and my colleagues to play a part in these exciting developments.

#### REFERENCES

1. Ryle, M. (1958) Proc. Roy. Soc. A., 248, 289.
2. Scheuer, P. A. G. & Ryle, hi. (1953) Mon. Not. R. astr. Soc., **113**, 3.
3. Blythe, J. H. (1957) Mon. Not. R. astr. Soc., **117**, 644.
4. Costain, C. H. & Smith, F. G. (1960) Mon. Not. R. astr. Soc., 121, 405.
5. Pilkington, J. D. H. & Scott, P. F. ( 1965) Mem. R. astr. Soc., 69, 183.
6. Gower, J. F. R., Scott, P. F. & Wills, D. (1967) Mem. R. astr. Soc., 71, 49
7. Ryle, M. & Neville, A. C. (1962) Mon. Not. R. astr. Soc., 125, 39.
8. Wynn-Williams, C. G., Becklin, E. E. & Neugebauer, G. (1974) Ap. J., 187, 473.
9. Rees, M. (1971) Nature, 229, 312.
10. Scheuer, P. A. G. (1974) Mon. Not. R. astr. Soc., **166**, 513.
11. Cooper, A. J., Davies, R. D. & Booth, R. S. (1971) Mon. Not. R. astr. Soc., 152, 383
12. Counselmann, C. C., Kent, S. M., Knight, C. A., Shapiro, I. I., Clarke, T. A., Hinteregger, H. F., Rogers, A. E. E. & Whitney, A. R. (1974) Phys. Rev. Lett.. **33**, 1621.

Physics 1975

**AAGE BOHR, BEN R MOTTELSON and  
JAMES RAINWATER**

*for the discovery of the connection between collective motion and particle motion in atomic nuclei and the development of the theory of the structure of the atomic nucleus based on this connection*

## THE NOBEL PRIZE FOR PHYSICS

Speech by professor SVEN JOHANSSON of the Royal Academy of Sciences  
Translation from the Swedish text

Your Majesties, Your Royal Highnesses, Ladies and Gentlemen,

At the end of the 1940's, nuclear physics had advanced to a stage where a more detailed picture of the structure of the atomic nucleus was beginning to emerge and it was becoming possible to calculate its properties in a quantitative way. One knew that the nucleus consists of protons and neutrons, the so-called nucleons. They are kept together by nuclear forces, which give rise to a potential well, in which the nucleons move. The details of the nuclear structure were, however, unknown and one had to a great extent to rely upon models. These models were rather incomplete and partly contradictory. The oldest is the drop model in which the nucleus is regarded as a liquid drop, the nucleons corresponding to the molecules of the liquid. This model could be used with a certain success for a description of the mechanism of nuclear reactions, in particular for fission. On the other hand, one could not find any excited states of the nucleus corresponding to rotations or vibrations of the drop. Neither could certain other properties of the nucleus, particularly those associated with the "magic numbers", be explained by means of the drop model. These show that individual nucleons in a decisive way affect the behaviour of the nucleus. This discovery, which is systematized in the shell model, was awarded the 1963 Nobel Prize for Physics.

It was soon found that the nucleus has properties, which cannot be explained by these models. perhaps the most striking one was the very marked deviation of the charge distribution from spherical symmetry, which was observed in several cases. It was also pointed out that this might indicate that certain nuclei are not spherical but are deformed as an ellipsoid, but no one could give a reasonable explanation of this phenomenon.

The solution of the problem was first presented by James Rainwater of Columbia University, New York, in a short paper submitted for publication in April 1950. In this, he considers the interaction between the main part of the nucleons, which form an inner core, and the outer, the valence nucleons. He points out that the valence nucleons can influence the shape of the core. Since the valence nucleons move in a field which is determined by the distribution of the inner nucleons, this influence is mutual. If several valence nucleons move in similar orbits, this polarizing effect on the core can be so great that the nucleus as a whole becomes permanently deformed. Expressed very simply, it can be said that as a result of their motion, certain nucleons expose the "walls" of the nucleus to such high centrifugal pressure that it becomes deformed. Rainwater also attempted to calculate this effect and got results that agreed with experimental data on the charge distributions.

Aage Bohr, working in Copenhagen, but at this time on a visit to Columbia University, had, independently of Rainwater, been thinking along the same lines. In a paper, submitted for publication about a month after Rainwater's, he formulates the problem of the interaction of a valence nucleon with the core in a general way.

These relatively vague ideas were further developed by Bohr in a famous work from 1951, in which he gives a comprehensive study of the coupling of oscillations of the nuclear surface to the motion of the individual nucleons. By analysing the theoretical formula for the kinetic energy of the nucleus, he could predict the different types of collective excitations: vibration, consisting of a periodic change of the shape of the nucleus around a certain mean value, and rotation of the whole nucleus around an axis perpendicular to the symmetry axis. In the latter case, the nucleus does not rotate as a rigid body, but the motion consists of a surface wave propagating around the nucleus.

Up to this point, the progress made had been purely theoretical and the new ideas to a great extent lacked experimental support. The very important comparison with experimental data was done in three papers, written jointly by Aage Bohr and Ben Mottelson and published in the years 1952-53. The most spectacular finding was the discovery that the position of energy levels in certain nuclei could be explained by the assumption that they form a rotational spectrum. The agreement between theory and experiment was so complete that there could be no doubt of the correctness of the theory. This gave stimulus to new theoretical studies, but, above all, to many experiments to verify the theoretical predictions.

This dynamic progress very soon led to a deepened understanding of the structure of the atomic nucleus. Even this further development towards a more refined theory was inspired and influenced in a decisive way by Bohr and Mottelson. For example, they showed together with Pines that the nucleons have a tendency to form pairs. A consequence of this is that nuclear matter has properties reminiscent of superconductors.

Drs Bohr, Mottelson and Rainwater,

In your pioneering works you have laid the foundation of a theory of the collective properties of atomic nuclei. This has been an inspiration to an intensive research activity in nuclear structure physics. The further development in this field has in a striking way confirmed the validity and great importance of your fundamental investigations.

On behalf of the Royal Academy of Sciences I wish to convey to you our warmest congratulations and I now ask you to receive your prize from the hands of His Majesty the King.



George Bohm

## AAGE BOHR

I was born in Copenhagen on June 19, 1922, as the fourth son of Niels Bohr and Margrethe Bohr (née Norlund). During my early childhood, my parents lived at the Institute for Theoretical Physics (now the Niels Bohr Institute), and the remarkable generation of scientists who came to join my father in his work became for us children Uncle Kramers, Uncle Klein, Uncle Nishina, Uncle Heisenberg, Uncle Pauli, etc. When I was about ten years old, my parents moved to the mansion at Carlsberg, where they were hosts for widening circles of scholars, artists, and persons in public life.

I went to school for twelve years at Sortedam Gymnasium (H. Adler's Fallesskole) and am indebted to many of my teachers, both in the humanities and in the sciences, for inspiration and encouragement.

I began studying physics at the University of Copenhagen in 1940 (a few months after the German occupation of Denmark). By that time, I had already begun to assist my father with correspondence, with his writing of articles of a general epistemological character, and gradually also in connection with his work in physics. In those years, he was concerned partly with problems of nuclear physics and partly with problems relating to the penetration of atomic particles through matter.

In October 1943, my father had to flee Denmark to avoid arrest by the Nazis, and the whole family managed to escape to Sweden, where we were warmly received. Shortly afterwards, my father proceeded to England, and I followed after him. He became associated with the atomic energy project and, during the two years until we returned to Denmark, in August 1945, we travelled together spending extensive periods in London, Washington, and Los Alamos. I was acting as his assistant and secretary and had the opportunity daily to share in his work and thoughts. We were members of the British team, and my official position was that of a junior scientific officer employed by the Department of Scientific and Industrial Research in London. In another context, I have attempted to describe some of the events of those years and my father's efforts relating to the prospects raised by the atomic weapons<sup>1</sup>.

On my return to Denmark, I resumed my studies at the University and obtained a master's degree in 1946. My thesis was concerned with some aspects of atomic stopping problems.

For the spring term of 1948, I was a member of the Institute for Advanced Study in Princeton. On a visit during that period to Columbia University

<sup>1</sup>Niels Bohr. His life and work as seen by his friends and colleagues, p. 191. Ed. by S. Rosenthal, North-Holland Publishing Company, Amsterdam 1967.

and through discussions with professor I. I. Rabi, I became interested in a newly discovered effect in the hyperfine structure in deuterium. This led on to my association with Columbia University from January 1949 to August 1950. As described in my lecture, this was for me a very fruitful association.

Soon after my return to Copenhagen, I began the close cooperation with Ben Mottelson which has continued ever since. The main direction of our work is described in the lectures included in the present volume. During the last fifteen years, a major part of our efforts has been connected with the attempt to present the status of our understanding of nuclear structure in a monograph, of which Volume I (Single-Particle Motion) appeared in 1969, and Volume II (Nuclear Deformations) in 1975. We feel that in our cooperation, we have been able to exploit possibilities that lie in a dialogue between kindred spirits that have been attuned through a long period of common experience and jointly developed understanding. It has been our good fortune to work closely together with colleagues at the Niels Bohr Institute and Nordita, including the many outstanding scientists who have come from all parts of the world and have so greatly enriched the scientific atmosphere and personal contacts.

I have been connected with the Niels Bohr Institute since the completion of my university studies, first as a research fellow and from 1956 as a professor of physics at the University of Copenhagen. After the death of my father in 1962, I followed him as director of the Institute until 1970.

For our whole circle, it has been a challenge to exploit the opportunities provided by the traditions of the Institute, of which I would like especially to mention two aspects. One concerns the fruitful interplay between experimental and theoretical investigations. The other concerns the promotion of international cooperation as a vital factor in the development of science itself and also as a means to strengthen the mutual knowledge and understanding between nations.

In 1957, Nordita (Nordisk Institut for Teoretisk Atomfysik) was founded on the premises of the Niels Bohr Institute, and the two institutes operate in close association. I have been a member of the Board of Nordita from 1957 until 1975, and since then director of this institute.

In March 1950, in New York City, I was married to Marietta Soffer. We have three children, Vilhelm, Tomas, and Margrethe. Both for my wife and myself, the personal friendships that have grown out of scientific contacts with colleagues from many different countries have been an important part of our lives, and the travels we have made together in connection with the world-wide scientific co-operation have given us rich treasures of experiences.

# ROTATIONAL MOTION IN NUCLEI

Nobel Lecture, December 11, 1975

*bY*

AAGE BOHR

The Niels Bohr Institute and Nordita  
Copenhagen, Denmark

The exploration of nuclear structure over the last quarter century has been a rich experience for those who have had the privilege to participate. As the nucleus has been subjected to more and more penetrating probes, it has continued to reveal unexpected facets and to open new perspectives. The preparation of our talks today has been an occasion for Ben Mottelson and myself to relive the excitement of this period and to recall the interplay of so many ideas and discoveries coming from the worldwide community of nuclear physicists, as well as the warmth of the personal relations that have been involved.

In this development, the study of rotational motion has had a special role. Because of the simplicity of this mode of excitation and the many quantitative relations it implies, it has been an important testing ground for many of the general ideas on nuclear dynamics. Indeed, the response to rotational motion has played a prominent role in the development of dynamical concepts ranging from celestial mechanics to the spectra of elementary particles.

## EARLY IDEAS ON NUCLEAR ROTATION

The question of whether nuclei can rotate became an issue already in the very early days of nuclear spectroscopy (1, 2). Quantized rotational motion had been encountered in molecular spectra (3), but atoms provide examples of quantal systems that do not rotate collectively. The available data on nuclear excitation spectra, as obtained for example from the fine structure of a decay, appeared to provide evidence against the occurrence of low-lying rotational excitations, but the discussion was hampered by the expectation that rotational motion would either be a property of all nuclei or be generally excluded, as in atoms, and by the assumption that the moment of inertia would have the rigid-body value, as in molecular rotations. The issue, however, took a totally new form with the establishment of the nuclear shell model (4).

Just at that time, in early 1949, I came to Columbia University as a research fellow and had the good fortune of working in the stimulating atmosphere of the Pupin Laboratory where so many great discoveries were being made under the inspiring leadership of I.I. Rabi. One of the areas of great activity was the study of nuclear moments, which was playing such a crucial role in the development of the new ideas on nuclear structure.

To-day, it is difficult to fully imagine the great impact of the evidence for nuclear shell structure on the physicists brought up with the concepts of the liquid-drop and compound-nucleus models, which had provided the basis for

interpreting nuclear phenomena over the previous decade (5)<sup>1</sup>. I would like also to recall my father's reaction to the new evidence, which presented the sort of dilemma that he would respond to as a welcome opportunity for deeper understanding. In the summer of 1949, he was in contact with John Wheeler on the continuation of their work on the fission process, and in this connection, in order to "clear his thoughts", he wrote some tentative comments on the incorporation of the contrasting evidence into a more general picture of nuclear constitution and the implications for nuclear reactions (7). These comments helped to stimulate my own thinking on the subject, which was primarily concerned with the interpretation of nuclear moments<sup>2</sup>.

The evidence on magnetic moments, which at the time constituted one of the most extensive quantitative bodies of data on nuclear properties, presented a special challenge. The moments showed a striking correlation with the predictions of the one-particle model (9, 4), but at the same time exhibited major deviations indicative of an important missing element. The incomparable precision that had been achieved in the determination of the magnetic moments, as well as in the measurement of the hyperfine structure following the pioneering work of Rabi, Bloch, and Purcell, was even able to provide information on the distribution of magnetism inside the nucleus (10, 11).

A clue for understanding the deviations in the nuclear coupling scheme from that of the single-particle model was provided by the fact that many nuclei have quadrupole moments that are more than an order of magnitude larger than could be attributed to a single particle<sup>3</sup>. This finding directly implied a sharing of angular momentum with many particles, and might seem to imply a break-down of the one-particle model. However, essential features of the single-particle model could be retained by assuming that the average nuclear field in which a nucleon moves deviates from spherical symmetry (15). This picture leads to a nuclear model resembling that of a molecule, in which the nuclear core possesses vibrational and rotational degrees

1 The struggle involved in facing up to the new evidence is vividly described by Jensen (6). Our discussions with Hans Jensen over the years concerning many of the crucial issues in the development provided for us a special challenge and inspiration.

2 The interplay between individual-particle and collective motion was also at that time taken up by John Wheeler. Together with David Hill, he later published the extensive article on "Nuclear Constitution and the Interpretation of Fission Phenomena" (8), which has continued over the years to provide inspiration for the understanding of new features of nuclear phenomena.

<sup>3</sup>The first evidence for a non-spherical nuclear shape came from the observation of a quadrupole component in the hyperfine structure of optical spectra (12). The analysis showed that the electric quadrupole moments of the nuclei concerned were more than an order of magnitude greater than the maximum value that could be attributed to a single proton and suggested a deformation of the nucleus as a whole (13). The problem of the large quadrupole moments came into focus with the rapid accumulation of evidence on nuclear quadrupole moments in the years after the war and the analysis of these moments on the basis of the shell model (14).

of freedom. For the rotational motion there seemed no reason to expect the classical rigid-body value; however, the large number of nucleons participating in the deformation suggested that the rotational frequency would be small compared with those associated with the motion of the individual particles. In such a situation, one obtains definite limiting coupling schemes (see Fig. 1) which could be compared with the empirical magnetic moments and the evidence on the distribution of nuclear magnetism, with encouraging results (15, 17)<sup>4</sup>.

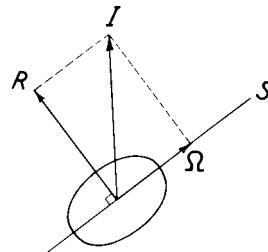


Fig. 1. Coupling scheme for particle in slowly rotating spheroidal nucleus. The intrinsic quantum number represents the projection of the particle angular momentum along the nuclear symmetry axis  $S$ , while  $R$  is the collective angular momentum of the nuclear core and is directed perpendicular to the symmetry axis, since the component along  $S$  which is a constant of the motion, vanishes in the nuclear ground state. The total angular momentum is denoted by  $I$ . The figure is from (16).

In the meantime and, in fact, at nearly the same point in space, James Rainwater had been thinking about the origin of the large nuclear quadrupole moments and conceived an idea that was to play a crucial role in the following development. He realized that a non-spherical equilibrium shape would arise as a direct consequence of single-particle motion in anisotropic orbits, when one takes into account the deformability of the nucleus as a whole, as in the liquid-drop model (19).

On my return to Copenhagen in the autumn of 1950, I took up the problem of incorporating the coupling suggested by Rainwater into a consistent dynamical system describing the motion of a particle in a deformable core. For this coupled system, the rotational motion emerges as a low-frequency component of the vibrational degrees of freedom, for sufficiently strong coupling. The rotational motion resembles a wave travelling across the nuclear surface and the moment of inertia is much smaller than for rigid rotation (see Fig. 2).

Soon, I was joined by Ben Mottelson in pursuing the consequences of the interplay of individual-particle and collective motion for the great variety of nuclear phenomena that was then coming within the range of experimental

<sup>4</sup>The effect on the magnetic moments of a sharing of angular momentum between the single particle and oscillations of the nuclear surface was considered at the same time by Foldy and Milford (18).

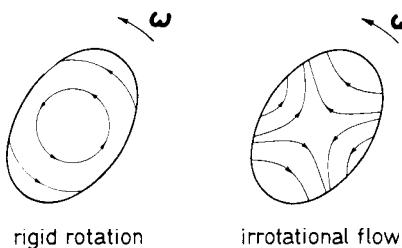


Fig. 2. Velocity fields for rotational motion. For the rotation generated by irrotational flow, the velocity is proportional to the nuclear deformation (amplitude of the travelling wave). Thus, for a spheroidal shape, the moment of inertia is  $\mathcal{J} = \mathcal{J}_{\text{rig}}(\Delta R/R)^2$ , where  $\mathcal{J}_{\text{rig}}$  is the moment for rigid rotation, while  $R$  is the mean radius and  $\Delta R$  (assumed small compared with  $R$ ) is the difference between major and minor semi-axes. The figure is from (16).

studies (20). In addition to the nuclear moments, important new evidence had come from the classification of the nuclear isomers (21) and beta decay (22) as well as from the discovery of single-particle motion in nuclear reactions (23, 24). It appeared that one had a framework for bringing together most of the available evidence, but in the quantitative confrontation with experiment, one faced the uncertainty in the parameters describing the collective properties of the nucleus. It was already clear that the liquid-drop description was inadequate, and one lacked a basis for evaluating the effect of the shell structure on the collective parameters.

#### THE DISCOVERY OF ROTATIONAL SPECTRA

At this point, one obtained a foothold through the discovery that the coupling scheme characteristic of strongly deformed nuclei with the striking rotational band structure was in fact realized for an extensive class of nuclei. The first indication had come from the realization by Goldhaber and Sunyar that the electric quadrupole transition rates for the decay of low-lying excited states in even-even nuclei were, in some cases, much greater than could be accounted for by a single-particle transition and thus suggested a collective mode of excitation (21). A rotational interpretation (25) yielded values for the nuclear eccentricity in promising agreement with those deduced from the spectroscopic quadrupole moments.

Soon after, the evidence began to accumulate that these excitations were part of a level sequence with angular momenta  $I = 0, 2, 4 \dots$  and energies proportional to  $I(I+1)$  (26, 27); examples of the first such spectra are shown in Fig. 3. For ourselves, it was a thrilling experience to receive a prepublication copy of the 1953 compilation by Hollander, Perlman, and Seaborg (29) with its wealth of information on radioactive transitions, which made it possible to identify so many rotational sequences.

The exciting spring of 1953 culminated with the discovery of the Coulomb excitation process (30, 31) which opened the possibility for a systematic study of rotational excitations (30, 32). Already the very first experiments by

Huus and Zupančič (see Fig. 4) provided a decisive quantitative test of the rotational coupling scheme in an odd nucleus, involving the strong coupling between intrinsic and rotational angular momenta".

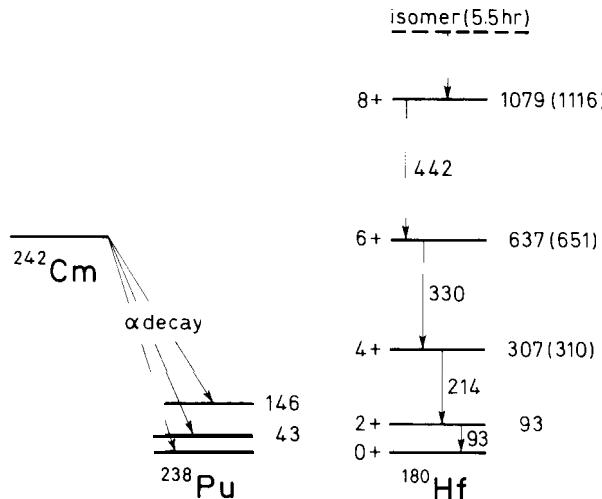


Fig. 3. Rotational spectra for  $^{238}\text{Pu}$  and  $^{180}\text{Hf}$ . The spectrum of  $^{180}\text{Hf}$  (from (26)) was deduced from the observed  $\gamma$  lines associated with the decay of the isomeric state (28). The energies are in keV, and the numbers in parenthesis are calculated from the energy of the first excited state, assuming the energies to be proportional to  $I(I+1)$ .

The spectrum of  $^{238}\text{Pu}$  was established by Asaro and Perlman (27) from measurements of the fine structure in the  $\alpha$  decay of  $^{242}\text{Cm}$ . Subsequent evidence showed the spin-parity sequence to be  $0+, 2+, 4+$ , and the energies are seen to be closely proportional to  $I(I+1)$ .

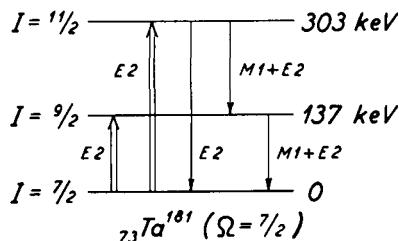


Fig. 4. Rotational excitations in  $^{181}\text{Ta}$  observed by Coulomb excitation. In an odd-A nucleus with intrinsic angular momentum  $\Omega$  (see Fig. 1), the rotational excitations involve these sequence  $z = \Omega, \Omega+1, \Omega+2, \dots$ , all with the same parity. In the Coulomb excitation process, the action of the electric field of the projectile on the nuclear quadrupole moment induces E2 (electric quadrupole) transitions and can thus populate the first two rotational excitations. The observed energies (30) are seen to be approximately proportional to  $I(I+1)$ .

The excited states decay by E2 and M1 (magnetic dipole) transitions, and the rotational interpretation implies simple intensity relations. For example, the reduced E2 matrix elements within the band are proportional to the Clebsch-Gordan coefficient  $\langle I_i \Omega | 20 | I_f \Omega \rangle$ , where  $I_i$  and  $I_f$  are the angular momenta of initial and final states. The figure is from (16).

<sup>a</sup>The quantitative interpretation of the cross sections could be based on the semi-classical theory of Coulomb excitation developed by Ter-Martirosyan (33) and Alder and Winther (34).

This was a period of almost explosive development in the power and versatility of nuclear spectroscopy, which rapidly led to a very extensive body of data on nuclear rotational spectra. The development went hand in hand with a clarification and expansion of the theoretical basis.

Fig. 5 shows the region of nuclei in which rotational band structure has so far been identified. The vertical and horizontal lines indicate neutron and proton numbers that form closed shells, and the strongly deformed nuclei are seen to occur in regions where there are many particles in unfilled shells that can contribute to the deformation.

The rotational coupling scheme could be tested not only by the sequence of spin values and regularities in the energy separations, but also by the intensity relations that govern transitions leading to different members of a rotational band (37, 38, 39). The leading order intensity rules are of a purely geometrical character depending only on the rotational quantum numbers and the multipolarity of the transitions (see the examples in Fig. 4 and Fig. 10).

The basis for the rotational coupling scheme and its predictive power were greatly strengthened by the recognition that the low-lying bands in odd-A nuclei could be associated with one-particle orbits in the deformed potential (40, 41, 42). The example in Fig. 6 shows the spectrum of  $^{235}\text{U}$  with its high level density and apparently great complexity. However, as indicated, the states can be grouped into rotational bands that correspond uniquely to those expected from the Nilsson diagram shown in Fig. 7.

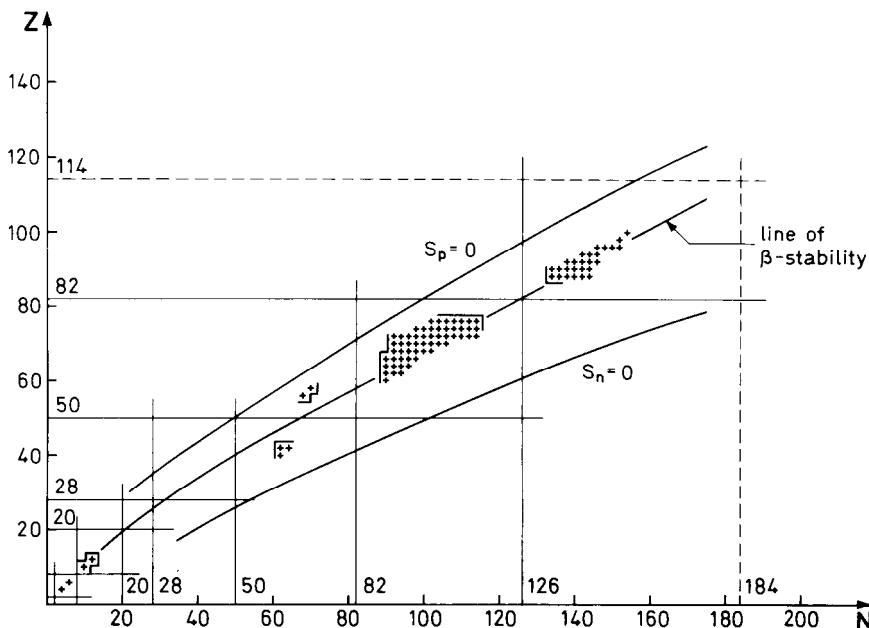


Fig. 5. Regions of deformed nuclei. The crosses represent even-even nuclei, whose excitation spectra exhibit an approximate  $I(I+1)$  dependence, indicating rotational band structure. The figure is from (35) and is based on the data in (36). The curves labelled  $S_n = 0$  and  $S_p = 0$  are the estimated borders of instability with respect to neutron and proton emission.

		<u>1053    7/2</u>		
		([743 7/2]0+) 7/2	<u>987.5    13/2</u>	
927	15/2		920.7    11/2	
850.5	13/2		([743 7/2]2+) 11/2	
777.7	11/2		804.9    25/2	A = 5.14
720.7	9/2		701.2    7/2	
671.1	7/2		664.6    5/2	
633.1	5/2		670.9    23/2	
	608.1    11/2		637.9    3/2	
[752 5/2]	533.3    9/2	([743 7/2]2+)^3/2	550.4    21/2	
A = 5.43		A = 5.34		
474.3    7/2			438.5    19/2	
426.7    5/2	414.8    9/2			
393.2    3/2	367.0    7/2		338.7    17/2	
[631 3/2]	332.8    5/2			
A = 6.79			357.2    15/2	368.8    13/2
B = -2.8 x 10^-3	[633 5/2]			
A_3 = -6.2 x 10^-3	B = 4.25		294.7    13/2	291.1    11/2
			249.1    15/2	
			170.7    13/2	197.1    11/2
			103.0    11/2	150.5    9/2
			46.2    9/2	171.3    7/2
			0    7/2	129.3    5/2
			[743 7/2]	
			B = 5.08	81.8    7/2
			B = 1.4 x 10^-3	[622 5/2]
				51.7    5/2
				13.0    3/2
				A = 6.03
				0.08    1/2
				B = -0.7 x 10^-3
				[631 1/2]
				A = 6.05
				A_1 = 1.75
				B = -2.2 x 10^-3
				B_1 = 2.0 x 10^-3
235    U	92			

Fig. 6. Spectrum of  $^{235}\text{U}$ . The figure is from (35) and is based on the experimental data from Coulomb excitation (43),  $^{239}\text{Pu}$   $\alpha$  decay (43a), one-particle transfer (44), and the  $^{234}\text{U}(\text{n}, \gamma)$  reaction (45). All energies are in keV. The levels are grouped into rotational bands characterized by the spin sequence, energy dependence, and intensity rules. The energies within a band can be represented by a power series expansion of the form  $E(Z) = AI(I+1) + BI^2(I+1)^2 + \dots + (-1)^{I+\Omega}(I+\Omega)!((I-\Omega)!)^{-1}(A_{2\Omega} + B_{2\Omega} Z(Z+1) + \dots)$ , with the parameters given in the figure. The low-lying bands are labelled by the quantum numbers of the available single-particle orbits (see Fig. 7), with particle-like states drawn to the right of the ground-state band and hole-like states to the left. The bands beginning at 638, 921, and 1053 keV represent quadrupole vibrational excitations of the ground-state configuration.

The regions of deformation in Fig. 5 refer to the nuclear ground-state configurations; another dimension is associated with the possibility of excited states with equilibrium shapes quite different from those of the ground state. For example, some of the closed-shell nuclei are found to have strongly deformed excited configurations<sup>6</sup>. Another example of sharp isomerism with associated rotational band structure is encountered in the metastable, very strongly deformed states that occur in heavy nuclei along the path to fission (50, 51).

<sup>6</sup> The fact that the first excited states in  $^{16}\text{O}$  and  $^{40}\text{Ca}$  have positive parity, while the low-lying single-particle excitations are restricted to negative parity, implies that these states involve the excitation of a larger number of particles. It was suggested (47) that the excited positive parity states might be associated with collective quadrupole deformations. The existence of a rotational band structure in  $^{16}\text{O}$  was convincingly established as a result of the  $^{12}\text{C}$  (aa) studies (48) and the observation of strongly enhanced E2-transition matrix elements (49).

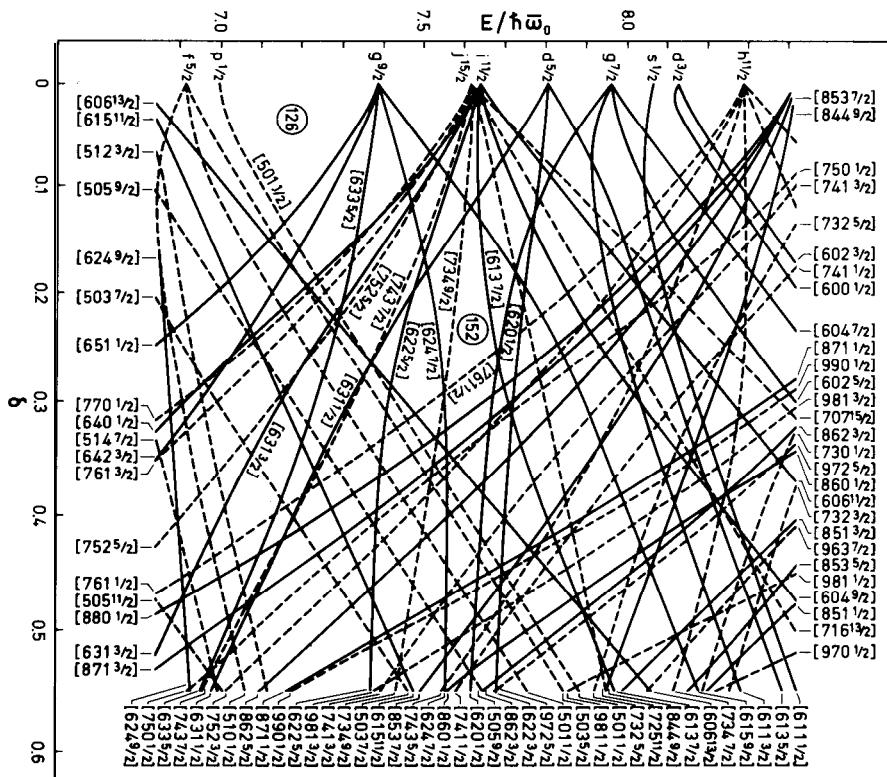


Fig. 7. Neutron orbits in prolate potential. The figure (from (35)) shows the energies of single-particle orbits calculated in an appropriate nuclear potential by Gustafson, Lamm, Nilsson, and Nilsson (46). The single-particle energies are given in units of  $\hbar\omega$ , which represents the separation between major shells and, for  $^{235}\text{U}$ , has the approximate value 6.6 MeV. The deformation parameter  $\delta$  is a measure of the nuclear eccentricity; the value determined for  $^{235}\text{U}$ , from the observed E2 transition moments, is  $\delta \approx 0.25$ . The single-particle states are labelled by the "asymptotic" quantum numbers  $[Nl_A \Omega]$ . The last quantum number  $\Omega$ , which represents the component  $j_z$  of the total angular momentum along the symmetry axis, is a constant of the motion for all values of  $\delta$ . The additional quantum numbers refer to the structure of the orbits in the limit of large deformations, where they represent the total number of nodal surfaces (N), the number of nodal surfaces perpendicular to the symmetry axis ( $n$ ), and the component of orbital angular momentum along the symmetry axis ( $A$ ). Each orbit is doubly degenerate ( $j_z = \pm \Omega$ ), and a pairwise filling of orbits contributes no net angular momentum along the symmetry axis. For  $^{235}\text{U}$ , with neutron number 143, it is seen that the lowest two configurations are expected to involve an odd neutron occupying the orbits  $[743 7/2]$  or  $[631 1/2]$ , in agreement with the observed spectrum (see Fig. 6). It is also seen that the other observed low-lying bands in  $^{235}\text{U}$  correspond to neighbouring orbits in the present figure.

New possibilities for studying nuclear rotational motion were opened by the discovery of marked anisotropies in the angular distribution of fission fragments (52), which could be interpreted in terms of the rotational quantum numbers labelling the individual channels through which the fissioning nucleus passes the saddle-point shape (53). Present developments in the ex-

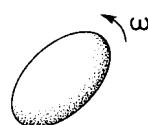
perimental tools hold promise of providing detailed information about band structure in the fission channels and thereby on rotational motion under circumstances radically different from those studied previously.

### CONNECTION BETWEEN ROTATIONAL AND SINGLE-PARTICLE MOTION

The detailed testing of the rotational coupling scheme and the successful classification of intrinsic spectra provided a firm starting point for the next step in the development, which concerned the dynamics underlying the rotational motion.

The basis for this development was the bold idea of Inglis (54) to derive the moment of inertia by simply summing the inertial effect of each particle as it is dragged around by a uniformly rotating potential (see Fig. 8). In this approach, the potential appears to be externally "cranked", and the problems concerning the self-consistent origin for the rotating potential and the limitations of such a semi-classical description have continued over the years to be hotly debated issues. The discussion has clarified many points concerning the connection between collective and single-particle motion, but the basic idea of the cranking model has stood its tests to a remarkable extent (55, 35).

The evaluation of the moments of inertia on the basis of the cranking model gave the unexpected result that, for independent-particle motion, the moment would have a value approximately corresponding to rigid rotation (56). The fact that the observed moments were appreciably smaller than the rigid-body values could be qualitatively understood from the effect of the residual interactions that tend to bind the particles into pairs with angular momentum zero. A few years later, a basis for a systematic treatment of the moment of inertia with the inclusion of the many-body correlations associated with the pairing effect was given by Migdal (57) and Belyaev (58),



$$H = H_0 - \hbar \omega J_x$$

$$J = 2\hbar^2 \sum_i \frac{\langle i | J_x | o \rangle^2}{E_i - E_0}$$

cranking model

Fig. 8. Nuclear moment of inertia from cranking model. The Hamiltonian  $\mathbf{H}$  describing particle motion in a potential rotating with frequency  $\omega$  about the  $x$  axis is obtained from the Hamiltonian  $\mathbf{H}_0$  for motion in a fixed potential by the addition of the term proportional to the component  $J_x$  of the total angular momentum, which represents the Coriolis and centrifugal forces acting in the rotating co-ordinate frame. The moment of inertia is obtained from a second-order perturbation treatment of this term and involves a sum over the excited states  $i$ . For independent-particle motion, the moment of inertia can be expressed as a sum of the contributions from the individual particles.

exploiting the new concepts that had in the meantime been developed for the treatment of electronic correlations in a superconductor (59); see also the following talk (60).

The nuclear moment of inertia is thus intermediate between the limiting values corresponding to rigid rotation and to the hydrodynamical picture of irrotational flow that was assumed in the early models of nuclear rotation. Indeed, the classical pictures involving a local flow provide too limited a framework for the description of nuclear rotation, since, in nuclear matter, the size of the pairs (the coherence length) is greater than the diameter of the largest existing nuclei. Macroscopic superflow of nuclear matter and quantized vortex lines may occur, however, in the interior of rotating neutron stars (61).

While these developments illuminated the many-body aspects of nuclear rotation, appropriate to systems with a very large number of nucleons, a parallel development took its starting point from the opposite side. Shell-model calculations exploiting the power of group-theoretical classification schemes and high-speed electronic computers could be extended to configurations with several particles outside of closed shells. It was quite a dramatic moment when it was realized that some of the spectra in the light nuclei that had been successfully analyzed by the shell-model approach could be given a very simple interpretation in terms of the rotational coupling scheme<sup>7</sup>.

The recognition that rotational features can manifest themselves already in configurations with very few particles provided the background for Elliott's discovery that the rotational coupling scheme can be given a precise significance in terms of the  $SU_3$  unitary symmetry classification, for particles moving in a harmonic oscillator potential (65). This elegant model had a great impact at the time and has continued to provide an invaluable testing ground for many ideas concerning nuclear rotation. Indeed, it has been a major inspiration to be able, even in this limiting case, to see through the entire correlation structure in the many-body wave function associated with the collective motion. Thus, for example, the model explicitly exhibits the separation between intrinsic and collective motion and implies an intrinsic excitation spectrum that differs from that of independent-particle motion in a deformed field by the removal of the "spurious" degrees of freedom that have gone into the collective spectrum.

This development also brought into focus the limitation to the concept of rotation arising from the finite number of particles in the nucleus. The rotational spectrum in the  $SU_3$  model is of finite dimension (compact symmetry group) corresponding to the existence of a maximum angular momentum that can be obtained from a specified shell-model configuration. For low-lying bands, this maximum angular momentum is of the order of magnitude

<sup>7</sup>In this connection, a special role was played by the spectrum of  $^{10}\text{F}$ . The shell-model analysis of this three-particle configuration had been given by Elliott and Flowers (62) and the rotational interpretation was recognized by Paul (63); the approximate identity of the wave functions derived by the two approaches was established by Redlich (64).

of the number of nucleons  $A$  and, in some of the light nuclei, one has, in fact, obtained evidence for such a limitation in the ground-state rotational bands\*. However, the proper place of this effect in nuclear rotations is still an open issue due to the major deviations from the schematized  $SU_3$  picture.

### GENERAL THEORY OF ROTATION

The increasing precision and richness of the spectroscopic data kept posing problems that called for a framework, in which one could clearly distinguish between the general relations characteristic of the rotational coupling scheme and the features that depend more specifically on the internal structure and the dynamics of the rotational motion<sup>9</sup>. For ourselves, an added incentive was provided by the challenge of presenting the theory of rotation as part of a broad view of nuclear structure. The view-points that I shall try to summarize gradually emerged in this prolonged labour (70, 71, 35).

In a general theory of rotation, symmetry plays a central role. Indeed, the very occurrence of collective rotational degrees of freedom may be said to originate in a breaking of rotational invariance, which introduces a "deformation" that makes it possible to specify an orientation of the system. Rotation represents the collective mode associated with such a spontaneous symmetry breaking (Goldstone boson) .

The full degrees of freedom associated with rotations in three-dimensional space come into play if the deformation completely breaks the rotational symmetry, thus permitting a unique specification of the orientation. If the deformation is invariant with respect to a subgroup of rotations, the corresponding elements are part of the intrinsic degrees of freedom, and the collective rotational modes of excitation are correspondingly reduced, disappearing entirely in the limit of spherical symmetry.

The symmetry of the deformation is thus reflected in the multitude of states that belong together in rotational families and the sequence of rotational quantum numbers labelling these states, in a similar manner as in the symmetry classification of molecular rotational spectra. The nuclear rotational spectra shown in Figs. 3, 4, and 6 imply a deformation with axial symmetry and invariance with respect to a rotation of  $180^\circ$  about an axis perpendicular to the symmetry axis ( $D_\infty$  symmetry group). It can also be inferred from the observed spectra that the deformation is invariant with respect to space and time reflection.

\*The evidence (66, 67) concerns the behaviour of the quadrupole transition rates, which are expected to vanish with the approach to the band termination (65). This behaviour reflects the gradual alignment of the angular momenta of the particles and the associated changes in the nuclear shape that lead eventually to a state with axial symmetry with respect to the angular momentum and hence no collective radiation (68), (35).

<sup>9</sup>In this development, a significant role was played by the high-resolution spectroscopic studies (69) which led to the establishment of a generalized intensity relation in the E2 decay of the  $\gamma$ -vibrational band in  $^{158}Gd$ .

The recognition of the deformation and its degree of symmetry breaking as the central element in defining rotational degrees of freedom opens new perspectives for generalized rotational spectra associated with deformations in many different dimensions including spin, isospin, and gauge spaces, in addition to the geometrical space of our classical world. The resulting rotational band structure may involve comprehensive families of states labelled by the different quantum numbers of the internally broken symmetries. Relations between quantum numbers belonging to different spaces may arise from invariance of the deformation with respect to a combination of operations in the different spaces<sup>10</sup>.

The Regge trajectories that have played a prominent role in the study of hadronic properties have features reminiscent of rotational spectra, but the symmetry and nature of possible internal deformations of hadrons remain to be established. Such deformations might be associated with boundaries for the regions of quark confinement.

The condensates in superfluid systems involve a deformation of the field that creates the condensed bosons or fermion pairs. Thus, the process of addition or removal of a correlated pair of electrons from a superconductor (as in a Josephson junction) or of a nucleon pair from a superfluid nucleus constitutes a rotational mode in the gauge space in which particle number plays the role of angular momentum (73). Such pair rotational spectra, involving families of states in different nuclei, appear as a prominent feature in the study of two-particle transfer processes (74). The gauge space is often felt as a rather abstract construction but, in the particle-transfer processes, it is experienced in a very real manner.

The relationship between the members of a rotational band manifests itself in the simple dependence of matrix elements on the rotational quantum numbers, as first encountered in the  $I(I + 1)$  dependence of the energy spectra and in the leading-order intensity rules that govern transitions leading to different members of a band. The underlying deformation is expressed by the occurrence of collective transitions within the band.

For sufficiently small values of the rotational quantum numbers, the analysis of matrix elements can be based on an expansion in powers of the angular momentum. The general structure of such an expansion depends on the symmetry of the deformation and takes an especially simple form for axially symmetric systems. As an example, Fig. 9 shows the two lowest bands observed in  $^{166}\text{Er}$ . The energies within each band have been measured with enormous precision and can be expressed as a power series that converges rather rapidly for the range of angular momentum values included in the figure. Similar expansions can be given for matrix elements of tensor operators representing

A well-known example is provided by the strong-coupling fixed-source model of the pion-nucleon system, in which the intrinsic deformation is invariant with respect to simultaneous rotations in geometrical and isospin spaces resulting in a band structure with  $I = T$  (72, 35).

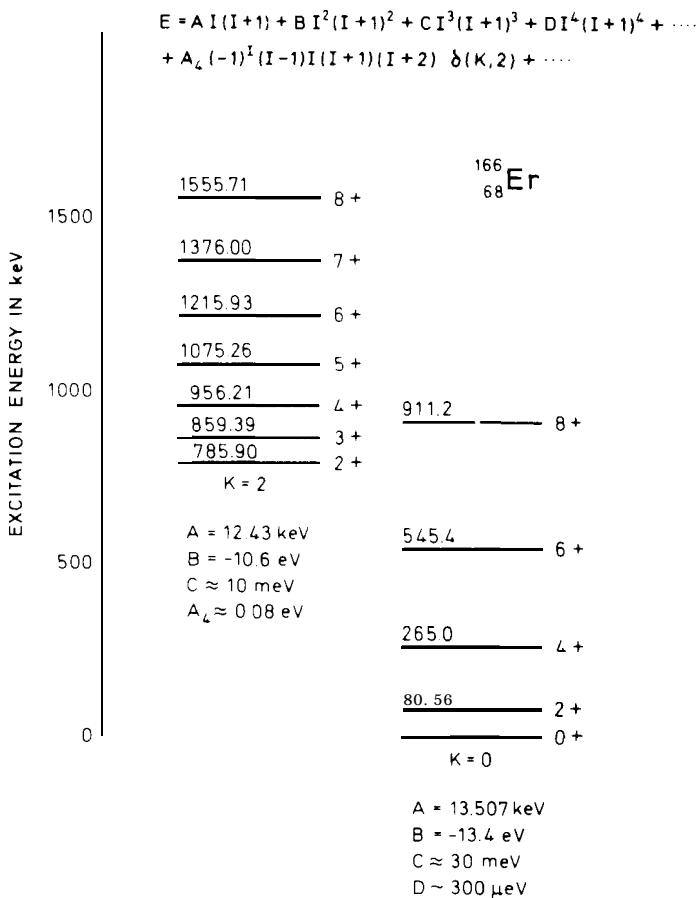


Fig. 9. Rotational bands in  $^{166}\text{Er}$ . The figure is from (35) and is based on the experimental data by Reich and Cline (75). The bands are labelled by the component  $K$  of the total angular momentum with respect to the symmetry axis. The  $K = 2$  band appears to represent the excitation of a mode of quadrupole vibrations involving deviations from axial symmetry in the nuclear shape.

electromagnetic transitions,  $\beta$  decay, particle transfer, etc. Thus, extensive measurements have been made of the E2 transitions between the two bands in  $^{166}\text{Er}$ , and Fig. 10 shows the analysis of the empirical transition matrix elements in terms of the expansion in the angular momentum quantum numbers of initial and final states.

Such an analysis of the experimental data provides a phenomenological description of the rotational spectra in terms of a set of physically significant parameters. These parameters characterize the internal structure of the system with inclusion of the renormalization effects arising from the coupling to the rotational motion.

A systematic analysis of these parameters may be based on the ideas of the cranking model, and this approach has yielded important qualitative insight into the variety of effects associated with the rotational motion. However, in

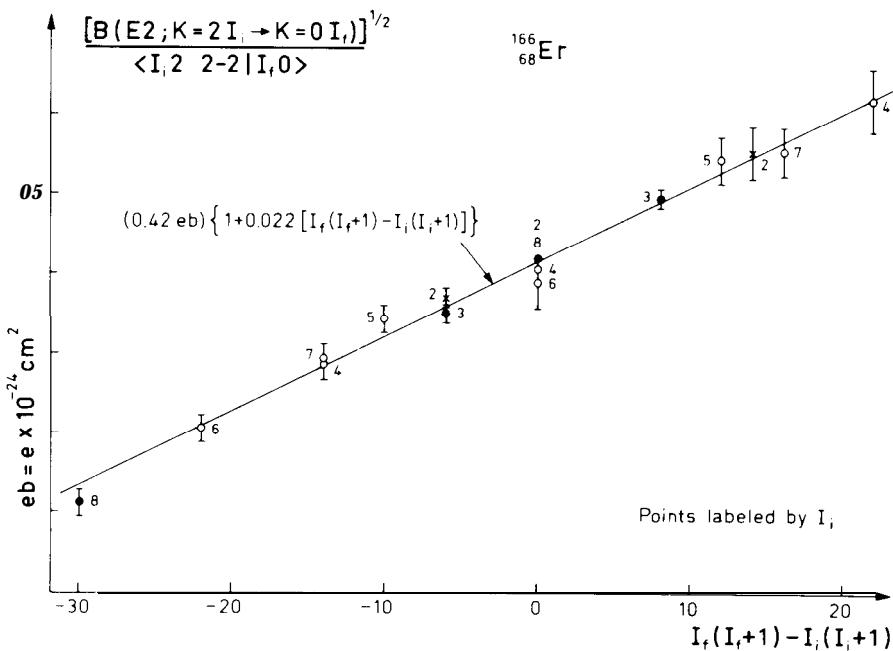


Fig. 10. Intensity relation for E2 transitions between rotational bands. The figure, which is from (35) and is based upon experimental data in (76), shows the measured reduced electric quadrupole transition probabilities  $B(E2)$  for transitions between members of the  $K = 2$  and  $K = 0$  bands in  $^{166}\text{Er}$  (see Fig. 9). An expansion similar to that of the energies in Fig. 9, but taking into account the tensor properties of the E2 operator, leads to an expression for  $(B(E2))^{1/2}$  which involves a Clebsch-Gordan coefficient  $\langle I_f K_f 2-2 | I_i K_i 0 \rangle$  (geometrical factor) multiplied by a power series in the angular momenta of  $I_i$  and  $I_f$  of the initial and final states. The leading term in this expansion is a constant, and the next term is linear in  $I_f(I_f+1) - I_i(I_i+1)$ ; the experimental data are seen to be rather well represented by these two terms.

this program, one faces significant unsolved problems. The basic coupling involved in the cranking model can be studied directly in the Coriolis coupling between rotational bands in **odd-A** nuclei associated with different orbits of the unpaired particle (77). The experiments have revealed, somewhat shockingly that, in many cases, this coupling is considerably smaller than the one directly experienced by the particles as a result of the nuclear rotation with respect to the distant galaxies (78). It is possible that this result may reflect an effect of the rotation on the nuclear potential itself (57, 79, 80, 35), but the problem stands as an open issue.

#### CURRENT PERSPECTIVES

In the years ahead, the study of nuclear rotation holds promising new perspectives. Not only are we faced with the problem already mentioned of a more deep-going probing of the rotational motion, which has become possible with the powerful modern tools of nuclear spectroscopy, but new frontiers are opening up through the possibility of studying nuclear states with very large

values of the angular momentum. In reactions induced by heavy ions, it is in fact now possible to produce nuclei with as much as a hundred units of angular momentum. We thus encounter nuclear matter under quite novel conditions, where centrifugal stresses may profoundly affect the structure of the nucleus. The challenge of this new frontier has strongly excited the imagination of the nuclear physics community.

A schematic phase diagram showing energy versus angular momentum for a nucleus with mass number  $A \approx 160$  is shown in Fig. 11. The lower curve representing the smallest energy, for given angular momentum, is referred to as the yrast line. The upper curve gives the fission barrier, as a function of angular momentum, estimated on the basis of the liquid-drop model (81). For  $I \approx 100$ , the nucleus is expected to become unstable with respect to fission, and the available data on cross sections for compound-nucleus formation in heavy ion collisions seem to confirm the approximate validity of this estimate of the limiting angular momentum (82).

Present information on nuclear spectra is confined almost exclusively to a small region in the left-hand corner of the phase diagram, and a vast extension of the field is therefore coming within range of exploration. Special interest attaches to the region just above the yrast line, where the nucleus, though highly excited, remains cold, since almost the entire excitation energy is concentrated in a single degree of freedom. One thus expects an excitation spectrum with a level density and a degree of order similar to that near the ground state. The extension of nuclear spectroscopy into this region may therefore offer the opportunity for a penetrating exploration of how the nuclear structure responds to the increasing angular momentum.

In recent years, it has been possible to identify quantal states in the yrast region up to  $I \approx 20$  to 25, and striking new phenomena have been observed.

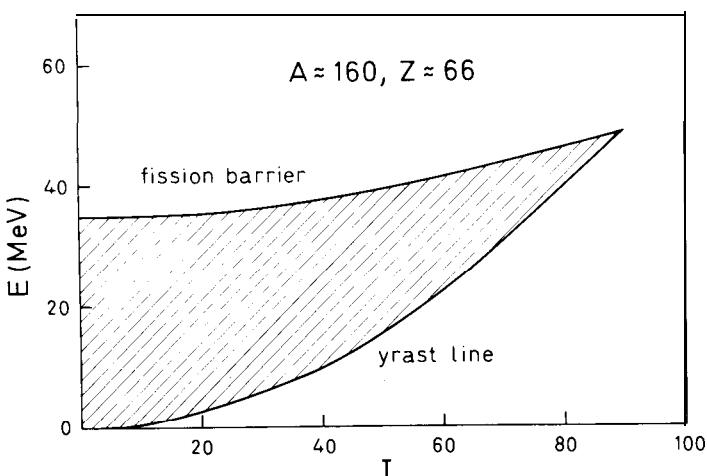


Fig. 11. Nuclear phase diagram for excitation energy versus angular momentum. The yrast line and the fission barrier represent estimates, due to Cohen, Plasil, and Swiatecki (81), based on the liquid-drop model, with the assumption of the rigid-body value for the moment of inertia.

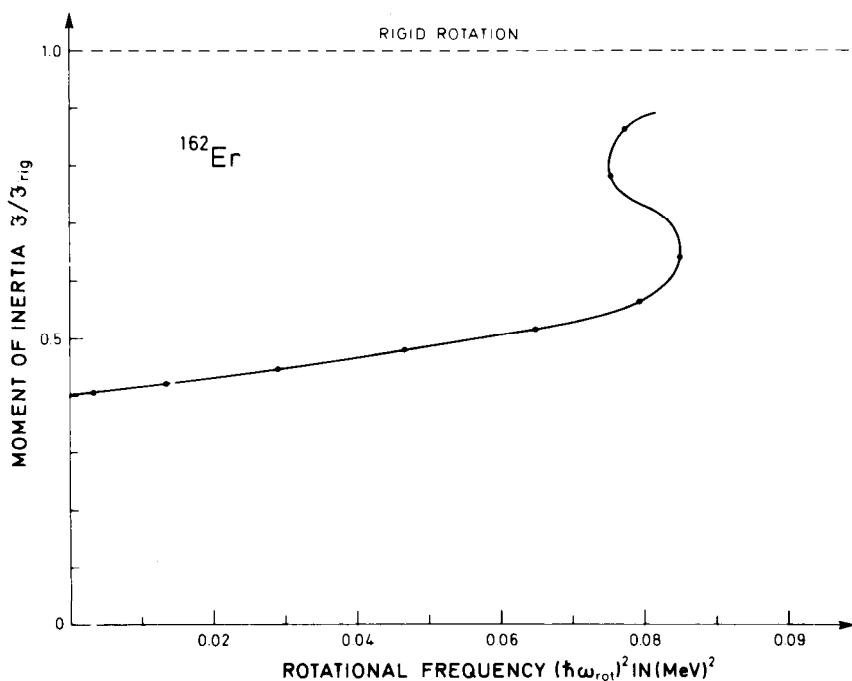


Fig. 12. Moment of inertia as function of rotational frequency. The figure is from (83) and is based on the experimental data of Johnson, Ryde, and Hjorth (84). The rotational frequency is defined as the derivative of the rotational energy with respect to the angular momentum and is obtained by a linear interpolation in the variable  $Z(Z+1)$  between the quantal states. The moment of inertia is defined in the usual manner as the ratio between the angular momentum and the rotational frequency.

An example is shown in Fig. 12, in which the moment of inertia is plotted against the rotational frequency. This "back-bending" effect was discovered here in Stockholm at the Research Institute for Atomic Physics, and has been found to be a rather general phenomenon.

In the region of angular momenta concerned, one is approaching the phase transition from superfluid to normal nuclear matter, which is expected to occur when the increase in rotational energy implied by the smaller moment of inertia of the superfluid phase upsets the gain in correlation energy (85). The transition is quite analogous to the destruction of superconductivity by a magnetic field and is expected to be associated with an approach of the moment of inertia to the rigid-body value characteristic of the normal phase.

The back-bending effect appears to be a manifestation of a band crossing, by which a new band with a larger moment of inertia and correspondingly smaller rotational frequency for given angular momentum, moves onto the yrast line. Such a band crossing may arise in connection with the phase transition, since the excitation energy for a quasiparticle in the rotating potential may vanish, even though the order parameter (the binding energy of the correlated pairs) remains finite, in rather close analogy to the situation in gapless superconductors (86). In fact, in the rotating potential, the angular

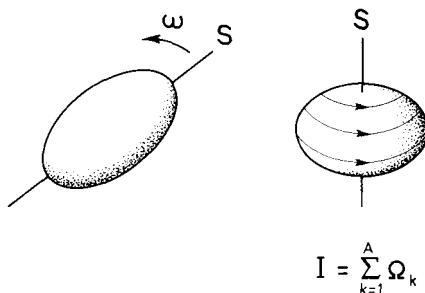


Fig. 13. Collective rotation contrasted with alignment of particle angular momenta along a symmetry axis.

momentum carried by the quasiparticle tends to become aligned in the direction of the axis of rotation. The excitation of the quasiparticle is thus associated with a reduction in the angular momentum and, hence, of the energy that is carried by the collective rotation (87).

It must be emphasized that, as yet, there is no quantitative interpretation of the striking new phenomena, as exemplified by Fig. 12. One is facing the challenge of analyzing a phase transition in terms of the individual quanta1 states.

For still larger values of the angular momentum, the centrifugal stresses are expected to produce major changes in the nuclear shape, until finally the system becomes unstable with respect to fission. The path that a given nucleus follows in deformation space will depend on the interplay of quanta1 effects associated with the shell structure and classical centrifugal effects similar to those in a rotating liquid drop. A richness of phenomena can be envisaged, but I shall mention only one of the intriguing possibilities.

The classical centrifugal effects tend to drive the rotating system into a shape that is oblate with respect to the axis of rotation, as is the case for the rotating earth. An oblate nucleus, with its angular momentum along the symmetry axis, will represent a form for rotation that is entirely different from that encountered in the low-energy spectrum, where the axis of rotation is perpendicular to the symmetry axis (see Fig. 13). For a nucleus spinning about its symmetry axis, the average density and potential are static, and the total angular momentum is the sum of the quantized contributions from the individual particles. In this special situation, we are therefore no longer dealing with a collective rotational motion characterized by enhanced radiative transitions, and the possibility arises of yrast states with relatively long lifetimes (88). If such high-spin metastable states (super-dizzy nuclei) do in fact occur, the study of their decay will provide quite new opportunities for exploring rotational motion in the nucleus at very high angular momenta.

Thus, the study of nuclear rotation has continued over the years to be alive and to reveal new, challenging dimensions. Yet, this is only a very special aspect of the broader field of nuclear dynamics that will be the subject of the following talk.

## REFERENCES

1. Thibaud, J., Comptes rendus 191, 656 ( 1930)
2. Teller, E., and Wheeler, J. A., Phys. Rev. 53, 778 (1938)
3. Bjerrum, N. in **Nernst** Festschrift, p. 90, Knapp, Halle 1912
4. Mayer, M. G., Phys. Rev. 75, 209 ( 1949) ; Haxel, O., Jensen, J. H. D., and Suess, H. E., Phys. Rev. 75, 1766 (1949)
5. Bohr, N., Nature 137, 344 ( 1936) ; Bohr, N., and Kalckar, F., Mat. Fys. Medd. Dan. Vid. Selsk. 14, no. 10 (1937); Weisskopf, V. F., Phys. Rev. 52, 295 (1937) ; Meitner, L., and Frisch, O. R., Nature 243, 239 (1939); Bohr, N., and Wheeler, J. A., Phys. Rev. 56, 426 (1939); Frenkel, J., J. Phys. (USSR) 1, 125 (1939); see also Phys. Rev. 55, 987 (1939)
6. Jensen, J. H. D., Les Prix Nobel en 1963, p. 153 Imprimerie Royale P. A. Norstedt & Shiner, Stockholm 1964
7. Bohr, N., **Tentative Comments on Atomic and Nuclear Constitution.** Manuscript dated August 1949. Niels Bohr Archives, The Niels Bohr Institute, Copenhagen
8. Hill, D. L., and Wheeler, J. A., Phys. Rev. 89, 1102 (1953)
9. Schmidt, T., Z Physik 106, 358 (1937)
10. Bitter, F., Phys. Rev. 75, 1326 (1949); 76, 150, (1949)
11. Bohr, A., and Weisskopf, V. F., Phys. Rev. 77, 94 (1950)
12. Schüler, H., and Schmidt, Th., Z. Physik 94, 457 (1935)
13. Casimir, H. B. G., **On the Interaction Between Atomic Nuclei and Electrons,** Prize Essay, Taylor's Tweede Genootschap, Haarlem (1936); see also Kopfermann, H., Kernmomente, Akademische Verlagsgesellschaft, Leipzig 1940
14. Townes, C. H., Foley, H. M., and Low, W., Phys. Rev. 76, 1415 (1949)
15. Bohr, A., Phys. Rev. 81, 134 (1951)
16. Bohr, A., **Rotational States of Atomic Nuclei,** Munksgård, Copenhagen 1954
17. Bohr, A., Phys. Rev. 81, 331 (1951)
18. Foldy, L. L., and Milford, F. J., Phys. Rev. 80, 751 ( 1950)
19. Rainwater, J., Phys. Rev. 79, 432 (1950)
20. Bohr, A., and Mottelson, B. R., Mat. Fys. Medd. Dan. Vid. Selsk. 27, no. 16 (1953)
21. Goldhaber, M., and Sunyar, A. W., Phys. Rev. 83, 906 ( 1951)
22. Mayer, M. G., Moszkowski, S. A., and Nordheim, L. W., Rev. Mod. Phys. 23, 315 (1951) Nordheim, L. W., Rev. Mod. Phys. 23, 322 (1951)
23. Barschall, H. H., Phys. Rev. 86, 431 ( 1952)
24. Weisskopf, V. F., Physica 18, 1083 (1952)
25. Bohr, A., and Mottelson, B. R., Phys. Rev. 89, 316 (1953)
26. Bohr, A., and Mottelson, B. R., Phys. Rev. 90, 717 ( 1953)
27. Asaro, F., and Perlman, I., Phys. Rev. 91, 763 (1953)
28. Burson, S. B., Blair, K. W., Keller, H. B., and Wexler, S., Phys. Rev. 83, 62 (1951)
29. Hollander, J. M., Perlman, I., and Seaborg, G. T., Rev. Mod. Phys. 25, 469 (1953)
30. Huus, T. and Zupančič, C., Mat. Fys. Medd. Dan. Vid. Selsk. 28, no. 1 (1953)
31. McClelland, C. L., and Goodman, C., Phys. Rev. 91, 760 (1953)
32. Heydenburg, N. P., and Temmer, G. M., Phys. Rev. 94, 1399 ( 1954) and 100, 150 (1955)
33. Ter-Martirosyan, K. A., Zh. Eksper. Teor. Fiz. 22, 284 ( 1952)
34. Alder, K., and Winther, A. Phys. Rev. 91, 1578 (1953)
35. Bohr, A., and Mottelson, B. R., **Nuclear Structure,** Vol. II. Benjamin, W. A. Inc., Reading, Mass. 1975
36. Sakai, M., Nuclear Data Tables **A8**, 323 (1970) and **A10**, 511 (1972)
37. Alaga, G., Alder, K., Bohr, A., and Mottelson, B. R., Mat. Fys. Medd. Dan. Vid. Selsk. 29, no. 9 (1955)
38. Bohr, A., Froman, P. O., and Mottelson, B. R., Mat. Fys. Medd. Dan. Vid. Selsk. 29, no. 10 (1955)

39. Satchler, G. R., Phys. Rev. 97, 1416 (1955)
40. Nilsson, S. G., Mat. Fys. Medd. Dan. Vid. selsk. 29, no. 16 ( 1955)
41. Mottelson, B. R., and Nilsson, S. G., Phys. Rev. 99, 1615 ( 1955)
42. Gottfried, K., Phys. Rev. 203, 1017 (1956)
43. Stephens, F. S., Holtz, M. D., Diamond, R. M., and Newton, J. O., Nuclear Phys. **A115**, 129 (1968)
- 43a. Cline, J. E., Nuclear Phys. **A106**, 481 (1968)
44. Elze, Th. W., and Huizenga, J. R., Nuclear Phys. **A133**, 10 (1969); Braid, T. H., Chasman, R. R., Erskine, J. R., and Friedman, A. M., Phys. Rev. C1, 275 (1970)
45. Jurney, E. T., **Neutron Capture Gamma-Ray Spectroscopy**, Proceedings of the International Symposium held in Studsvik, p. 431, International Atomic Energy Agency, Vienna 1969
46. Gustafson, C., Lamm, I. L., Nilsson, B., and Nilsson, S. G., Arkiv Fysik 36, 613 (1967)
47. Morinaga, H., Phys. Rev. **101**, 254 (1956)
48. Carter, E. B., Mitchell, G. E., and Davies, R. H., Phys. Rev. 133, B1421 ( 1964)
49. Gorodetzky, S., Mennrath, P., Benenson, W., Chevallier, P., and Scheibling, F., J. phys. radium 24, 887 (1963)
50. Polikanov, S. M., Druin, V. A., Karnaukhov, V. A., Mikheev, V. L., Pleve, A. A., Skobelev, N. K., Subbotin, V. G., Ter-Akop'yan, G. M., and Fomichev, V. A., J. Exptl. Theoret. Phys. (USSR) 42, 1464 (1962); transl. Soviet Physics JETP **15**, 1016
51. Specht, H. J., Weber, J., Konecny, E., and Heunemann, D., Phys. Letters 41B, 43 (1972)
52. Winhold, E. J., Demos, P. T., and Halpern, I., Phys. Rev. 87, 1139 ( 1952)
53. Bohr, A. in **Proc. Intern. Conf. on the Peaceful Uses of Atomic Energy**, vol. 2, p. 151, United Nations, New York, 1956
54. Inglis, D. R., Phys. Rev. 96, 1059 ( 1954)
55. Thouless, D. J., and Valatin, J. G., Nuclear Phys. 31, 211 (1962)
56. Bohr, A., and Mottelson, B. R., Mat. Fys. Medd. Dan. Vid. Selsk. 30, no. 1 ( 1955)
57. Migdal, A. B., Nuclear Phys. **13**, 655 (1959)
58. Belyaev, S. T., Mat. Fys. Medd. Dan. Selsk. 31, no. 11 (1959)
59. Bardeen, J., Cooper, L. N., and Schrieffer, J. R., Phys. Rev. 106, 162 and 108, 1175 (1957)
60. Mottelson, B. R., this volume, p. 82.
61. Ruderman, M., Ann. Rev. Astron. and Astrophys. 10, 427 (1972)
62. Elliott, J. P., and Flowers, B. H., Proc. Roy. Soc. (London) **A229**, 536, (1955)
63. Paul, E. B., Phil. Mag. 2, 311 (1957)
64. Redlich, M. G., Phys. Rev. **110**, 468 (1958)
65. Elliott, J. P., Proc. Roy. Soc. (London) **A245**, 128 and 562 (1958)
66. Jackson, K. P., Ram, K. B., Lawson, P. G., Chapman N. G., and Allen, K. W., Phys. Letters 30B, 162 ( 1969)
67. Alexander, T. K., Häusser, O., McDonald, A. B., Ferguson, A. J., Diamond, W. T., and Litherland, A. E., Nuclear Phys. **A179**, 477 (1972)
68. Bohr, A. in **Intern. Nuclear Phys. Conf.**, p. 489, Ed.-in-Chief R. L. Becker, Academic Press, New York 1967
69. Hansen, P. G., Nielsen, O. B., and Sheline, R. K., Nuclear Phys. 12, 389 ( 1959)
70. Bohr, A., and Mottelson, B. R., Atomnaya Energiya **14**, 41 (1963)
71. Bohr, A. in **Symmetry Properties of Nuclei**, Proc. of 15th Solvay Conf. on Phys. 1970, p. 187. Gordon and Breach Science Publ., London 1974
72. Henley, E. M., and Thirring, W., Elementary Quantum **Field Theory**, McGraw Hill, New York 1962
73. Anderson, P. W., Rev. Mod. Phys. 38, 298 (1966)
74. Middleton, R., and Pullen, D. J., Nuclear Phys. 52, 77 (1964); see also Broglia, R.

- A., Hansen, O., and Riedel, *C. Advances in Nuclear Physics* 6, 287, Plenum Press, New York 1973
75. Reich, C. W., and Cline, J. E., Nuclear Phys. **A159**, 181 (1970)
76. Gallagher, C. J., Jr., Nielsen, O. B., and Sunyar, A. W., Phys. Letters 16, 298 (1965); Günther, C., and Parsignault, D. R., Phys. Rev. 153, 1297 (1967); Domingos, J. M., Symons, G. D., and Douglas, A. C., Nuclear Phys. **A180**, 600 (1972)
77. Kerman, A. K., Mat. Fys. Medd. Dan. Vid. Selsk. 30, no. 15 ( 1956)
78. Stephens, F. (1960), quoted by Hyde, E., Perlman, I., and Seaborg, G. T. in **The Nuclear Properties of the Heavy Elements**, Vol. II, p. 732, Prentice Hall, Englewood Cliffs, N.J. 1964; Hjorth, S. A., Ryde, H., Hagemann, K. A., Løvhøiden, G., and Waddington, J. C., Nuclear Phys. **A144**, 513 ( 1970) ; see also the discussion in (35)
79. Belyaev, S. T., Nuclear Phys. 24, 322 (1961)
80. Hamamoto, I., Nuclear Phys. **A232**, 445, ( 1974)
81. Cohen, S., Plasil, F., and Swiatecki, W. J., Ann. Phys. 82, 557 (1974)
82. Britt, H. C., Erkilla, B. H., Stokes, R. H., Gutbrod, H. H., Plasil, F., Ferguson, R. L., and Blann, M., Phys. Rev. C in press; Gauvin, H., Guerrau, D., Le Beyec, Y., Lefort, M., Plasil, F., and Tarrago, X., Phys. Letters, 58B, 163 ( 1975)
83. Bohr, A., and, Mottelson, B. R., **The Many Facets of Nuclear Structure**. Ann. Rev. Nuclear Science 23, 363 (1973)
84. Johnson, A., Ryde, H., and Hjorth, S. A., Nuclear Phys. **A179**, 753 (1972)
85. Mottelson, B. R., and Valatin, J. G., Phys. Rev. Letters 5, 511 ( 1960)
86. Goswami, A., Lin, L., and Struble, G. L., Phys. Letters 25B, 451 ( 1967)
87. Stephens, F. S., and Simon, R. S., Nuclear Phys. **A183**, 257 (1972)
88. Bohr, A., and Mottelson, B. R., Physica Scripta **10A**, 13 ( 1974)



*Ben R. Motelson*

## BEN R. MOTTELSON

I was born in Chicago, Illinois, on July 9, '1926, the second of three children of Goodman Mottelson and Georgia Mottelson (née Blum). My father held a university degree in engineering. My childhood home was a place where scientific, political and moral issues were freely and vigorously discussed. I attended primary school and high school in the village of La Grange, Illinois.

Graduating from high school during the second world war, I was sent by the U.S. Navy to Purdue University for officers training (V12 program) and remained there to receive a Bachelor of Science degree in 1947. My graduate studies were at Harvard University and my PhD work on a problem in nuclear physics was directed by Professor Julian Schwinger and completed in 1950.

Receiving a Sheldon Traveling Fellowship from Harvard University I chose to spend the year ( 1950-51) at the Institute for Theoretical Physics in Copenhagen (later the Niels Bohr Institute) where so much of modern physics had been created and where there were such special traditions for international cooperation. A fellowship from the U.S. Atomic Energy Commission permitted me to continue my work in Copenhagen for two more years after which I held a research position in the CERN (European Organization for Nuclear Research) theoretical study group that was formed in Copenhagen. With the founding of the Nordic Institute for Theoretical Atomic Physics in Copenhagen (1957) I received a position as professor which I have held since. The spring term of 1959 was spent as visiting professor in the University of California at Berkeley.

The close scientific collaboration with Aage Bohr was begun in 1951 and has continued ever since. We feel that in this cooperation we have been able to exploit possibilities that lie in a dialogue between kindred spirits that have been attuned through a long period of common experience and jointly developed understanding. The lectures that are published in this volume attempt a discussion of the main influences that we have built on and the viewpoints that have been developed in this collaboration. It has been our good fortune to work closely together with colleagues at the Niels Bohr Institute and Nordita, including the many outstanding scientists who have come from all parts of the world and have so enriched the scientific atmosphere and personal contacts.

In 1948 I married Nancy Jane Reno and we had three children. The family became Danish citizens in 1971.

# ELEMENTARY MODES OF EXCITATION IN THE NUCLEUS

Nobel Lecture, December 11, 1975

by

BEN R. MOTTELSON

NORDITA, Copenhagen, Denmark

In the field of nuclear dynamics a central theme has been the struggle to find the proper place for the complementary concepts referring to the independent motion of the individual nucleons and the collective behaviour of the nucleus as a whole. This development has been a continuing process involving the interplay of ideas and discoveries relating to all different aspects of nuclear phenomena. The multi-dimensionality of this development makes it tempting to go directly to a description of our present understanding and to the problems and perspectives as they appear today. However, an attempt to follow the evolution of some of the principal ideas may be instructive in illustrating the struggle for understanding of many-body systems, which have continued to inspire the development of fundamental new concepts, even in cases where the basic equations of motion are well established. Concepts appropriate for describing the wealth of nuclear phenomena have been derived from a combination of many different approaches, including the exploration of general relations following from considerations of symmetry, the study of model systems, sometimes of a grossly oversimplified nature, and, of course, the clues provided by the experimental discoveries which have again and again given the development entirely new directions.<sup>1</sup>

The situation in 1950, when I first came to Copenhagen, was characterized by the inescapable fact that the nucleus sometimes exhibited phenomena characteristic of independent-particle motion, while other phenomena, such as the fission process and the large quadrupole moments, clearly involved a collective behaviour of the whole nucleus.

It was also clear from the work of Rainwater that there was an important coupling between the motion of the individual particles and the collective deformation, and one was thus faced with the problem of exploring the properties of a dynamical system involving such coupled degrees of freedom (1, 2, 3, 4).

$$\begin{aligned}
 H &= H_{\text{vib}} + H_{\text{part}} + H_{\text{coupl}} \\
 H_{\text{vib}} &= \frac{1}{2} C_\lambda \sum_\mu |\alpha_{\lambda\mu}|^2 + \frac{1}{2} D_\lambda \sum_\mu |\alpha_{\lambda\mu}|^2 \\
 H_{\text{coupl}} &= \sum_p k(r_p) \sum_\mu \alpha_{\lambda\mu} Y_{\lambda\mu}(\varphi_p \theta_p)
 \end{aligned} \tag{1}$$

<sup>1</sup>We would like to take this opportunity to pay tribute to the ingenuity and resourcefulness of the generation of experimentalists whose untiring efforts have created the basis for the development sketched in our reports today.

where  $a_{\lambda\mu}$  are the amplitudes of the nuclear deformation expanded in spherical harmonics and  $(r_p, \theta_p, \varphi_p)$  are the coordinates of the particles considered. The coupling term represents the effect of the deformation on the one-particle potential.

I remember vividly the many lively discussions in these years reflecting the feeling of unease, not to say total disbelief, of many of our colleagues concerning the simultaneous use of both collective and single-particle coordinates to describe a system that we all agreed was ultimately built out of the neutrons and protons themselves. Niels Bohr participated very actively in these discussions. Something of the flavour of this contribution can perhaps be gathered from the exchange recorded in the Proceedings of the CERN International Physics Conference in Copenhagen from June 1952; I had given a report on our work, and in the discussion Rosenfeld "asked how far this model is based on first principles". N. Bohr "answered that it appeared difficult to define what one should understand by first principles in a field of knowledge where our starting point is empirical evidence of different kinds, which is not directly combinable".

I would like to take this opportunity to acknowledge the tremendous inspiration it has been for me to have had the privilege to work for the entire period covered by this report within the unique scientific environment created by Niels Bohr.

#### INTERPRETATION OF LOW-ENERGY NUCLEAR EXCITATION SPECTRA

In the beginning of the 1950'ies, there existed very little evidence on nuclear spectra, which could be used to test these ideas. In the following years, however, a dramatic development of nuclear spectroscopy took place. The new data made possible the identification of the characteristic patterns of rotational spectra (5) and shortly afterwards the recognition by Scharff-Goldhaber and Wenner (6) that a significant class of spectra exhibit patterns corresponding to quadrupole vibrations about a spherical equilibrium<sup>2</sup>. The existence of the static deformations in certain classes of nuclei received further decisive confirmation in the successful classification of the intrinsic states of these spectra in terms of one-particle motion in an appropriately deformed potential (5).

A striking feature in the developing picture of nuclear excitation spectra was the distinction between a class of nuclei with spherical shape and others with large deformations. The clue to the origin of this distinction came, rather unexpectedly, from the analysis of the moments of inertia of the rotational spectra. The cranking model of Inglis (14) had provided a starting point for a microscopic interpretation of the rotational motion, and the analysis

<sup>2</sup>This step followed the recognition of striking regularities in the low-energy spectra of even-even nuclei, including the spins and parities (7, 8), energy systematics (8, 9, 10, 11) and selection rules (12).

showed that significant deviations from independent-particle motion were required to account for the observed magnitude of the moments of inertia. These correlations could be attributed to the residual interactions that tend to bind the nucleons into pairs with angular momentum zero. Such a pair is spherically symmetric, and this nucleonic correlation could therefore, at the same time, be seen to provide an effect tending to stabilize the spherical shape (15).

Thus, quite suddenly the way was opened to a qualitative understanding of the whole pattern of the low-energy excitation spectra in terms of a competition between the pairing effect and the tendency toward deformations implied by the anisotropy of the single-particle orbits. The outcome of this competition depends on the number of particles in unfilled shells; for few particles, the deformation in the absence of interactions is relatively small and can easily be dominated by the tendency to form spherical pairs; but with increasing number of particles, the spherical equilibrium shape becomes less stable, and eventually a transition takes place to a deformed equilibrium shape. These considerations are illustrated by the potential energy surfaces shown in Fig. 1.

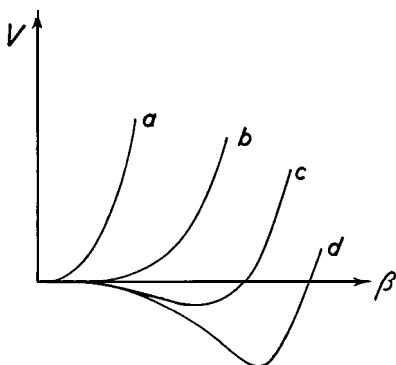


Fig. 1. Nuclear potential energy function. The figure, taken from (13), gives a schematic representation of the nuclear energy as function of the deformation  $\beta$ . The curve a represents a configuration with only relatively few particles outside of closed shells. As particles are added, the restoring force for the spherical shape ( $\beta = 0$ ) decreases (curve b). Still further from the closed shells, the spherical shape may become unstable (curves c and d) and the nucleus acquires a non-spherical equilibrium shape.

#### MICROSCOPIC DESCRIPTION OF COLLECTIVE MOTION

This qualitative interpretation of the nuclear coupling schemes could soon be given a firmer basis in terms of many-body wave functions that describe the correlation effects governing the low-energy nuclear spectra.

A step towards a microscopic understanding of the deformation effect resulted from the discovery of rotational spectra in light nuclei<sup>3</sup>. For these nuclei, even a few particles represent a significant fraction of the total and can

<sup>3</sup>Rotational band structure and the classification of the intrinsic states for (sd)-shell nuclei was first established in 1955 following the extensive series of experiments at Chalk River (see the survey by A. Litherland et al. (16)). For the classification of the p-shell nuclei in terms of the rotational coupling scheme, see Kurath and Pičman (17). For our own understanding of the special flavour of these very light nuclei, the discussions over the years with Tom Lauritsen were a continuing challenge and source of inspiration.

give rise to deformations that are among the largest observed. The spectra of some of these nuclei had previously been successfully analyzed in terms of shell-model configurations (5). Thus, for the first time one had a many-body wave function with rotational relationships and one could see explicitly that the main effect of the rather complicated finite range interactions employed in the shell-model calculations had been to generate a deformed average potential.

The essence of this development was brought into focus by Elliott's discovery that the  $SU_3$  classification scheme for particles in a harmonic oscillator potential leads to multiplets with rotational relations (18). The effective two-body interaction that is invariant under  $SU_3$  symmetry (when acting within the configurations of a major shell) and thus leads to the rotational coupling scheme, is given by the scalar product of the quadrupole moments of each pair of particles.

$$V_{\text{eff}} = \frac{1}{2} H \sum_{ij} (q(i)q(j))_0 \quad (2)$$

$$q_\mu(i) = r_i^2 Y_{2\mu}(\varphi_i \theta_i)$$

Such a two-body force is equivalent to the interaction of each particle with the total quadrupole moment of the system and thus to the effect of an ellipsoidal deformation in the average potential.

In retrospect, the important lesson of this development was the recognition that the aligned wave function

$$\Psi = A (\psi_{k1}(x_1) \psi_{k2}(x_2) \dots \psi_{kA}(x_A)) \quad (3)$$

obtained as a simple product of single-particle states in a self-consistent deformed potential provides a starting point for the full many-body wave function<sup>4</sup>. This view-point had indeed been implied by the establishment of the classification based on the Nilsson scheme, but the revelation of the exact  $SU_3$  solution, even in such an oversimplified model, contributed greatly to the confidence in this approach.

The second major development involved the many-particle interpretation of the nuclear pairing effect. As we have seen, this problem had become a crucial one for the quantitative analysis of collective motion in the nucleus, but the story of the pairing effect goes back much further, to the very earliest days of nuclear physics (21). The discovery of the neutron made it possible to

<sup>4</sup>The wave function given by Eq. (3) represents the intrinsic state in the absence of rotation, and can be directly employed in obtaining the leading-order intensity relations. The I-dependent terms, such as the rotational energy, are obtained by including the rotational perturbations in the intrinsic motion, as in the cranking model. The  $SU_3$  coupling scheme represents a special case in which the total function, with the inclusion of rotational effects can be expressed as a projection of the intrinsic wave function onto a state of specified angular momentum (18). (Such projected wave functions had been employed earlier (19); see also the discussion in (20).)

interpret the accumulated systematics concerning the differences in stability of odd and even nuclei in terms of an additional binding associated with even numbers of protons or neutrons (22). This effect later provided the basis for understanding the striking difference in the fission of the odd and even isotopes of uranium (23). The pairing effect also played an important role in the development of the shell model since it provided the basis for the interpretation of many of the properties of odd- $A$  nuclei in terms of the binding states of the last odd particle (24, 25, 26).

The key to understanding the correlation effect underlying the odd-even differences came from the discovery by Bardeen, Cooper, and Schrieffer of the profound new concepts for treating the electronic correlations in superconductors (27)<sup>5</sup>. It was a marvellous thing that the correlations, which might appear to be associated with such complexity, could be simply expressed in terms of a generalized one-body problem in which the particles move in a potential which creates and annihilates pairs of particles giving rise to the quasiparticles that are superpositions of particles and holes (30, 31). It could also be seen that the many-body wave function represented a generalization of Racah's seniority coupling scheme (32) which had been exploited in the interpretation of the one-particle model in nuclei.

One thus had available the basic tools for a microscopic analysis of the coupling schemes encountered in the low-energy nuclear spectra. These tools were rapidly exploited to treat the moments of inertia of rotating nuclei (33, 34, 35, 36), the potential energy surfaces and inertial parameters for the vibrations of spherical nuclei (33, 37), as well as the effects of pair correlations on a variety of nuclear processes (38, 39, 40, 41).

This was indeed a period of heady development in the understanding of many-body problems with a fruitful interplay of experience gained from the study of so many different systems that nature had provided, including the "elementary particles" that had stimulated the development of the powerful tools of relativistic field theory. An important clarification in the description of collective motion was the new way of viewing the normal modes of vibration as built out of correlated two-quasiparticle (or particle-hole) excitations. The significant part of the interactions creates and annihilates two such basic excitations, and the vibrations can thus be obtained from the solution of a generalized two-body problem (42). This approach not only comple-

<sup>5</sup>It was a fortunate circumstance for us that David Pines spent a period of several months in Copenhagen in the summer of 1957, during which he introduced us to the exciting new developments in the theory of superconductivity. Through the discussions with him, the relevance of these concepts to the problem of pair correlations in nuclei became apparent (28). An important component in these discussions was the fact that the experimental evidence had been accumulating for the existence of an energy gap in the excitation spectra of nuclei reminiscent of that observed in superconductors (15, 28). (For the recognition of the odd-even difference in nuclear excitation spectra, striking evidence had come from the high-resolution spectroscopic studies of  $^{182}\text{W}$  and  $^{183}\text{W}$  made possible by the bent crystal spectrometer (29).)

mented the previously applied adiabatic treatment of nuclear collective motion, but also gave a broader scope to the concept of vibration that was to be important for the subsequent development.

The whole picture of nuclear physics at this stage in the development is beautifully expressed by Weisskopf in his summary talk at the Kingston Conference in 1960, where the recurring theme is his comment again and again: "It works surprisingly well".<sup>6</sup>

### THE GREAT VARIETY OF COLLECTIVE MODES

While the low-frequency spectra are dominated by transitions of particles within the partly filled shells, new aspects of nuclear dynamics are associated with the excitation of the closed shells. The classic example of a collective excitation of this type is the "giant dipole resonance" which was discovered in the study of the photo-processes soon after the war (43), and which could be given an interpretation in terms of collective motion of the neutron and proton fluids with respect to each other (44,45).

After the development of the shell model, attempts were made to describe the photo-absorption in terms of single-particle excitations (46), but one encountered the problem that the one-particle excitations that should carry the main part of the dipole strength appeared in a part of the spectrum quite distinct from that in which the strong dipole absorption was observed (see Fig. 2). This led to a period of lively discussions, and for a time it was felt that the single-particle and collective descriptions represented opposite and mutually exclusive interpretations (47).

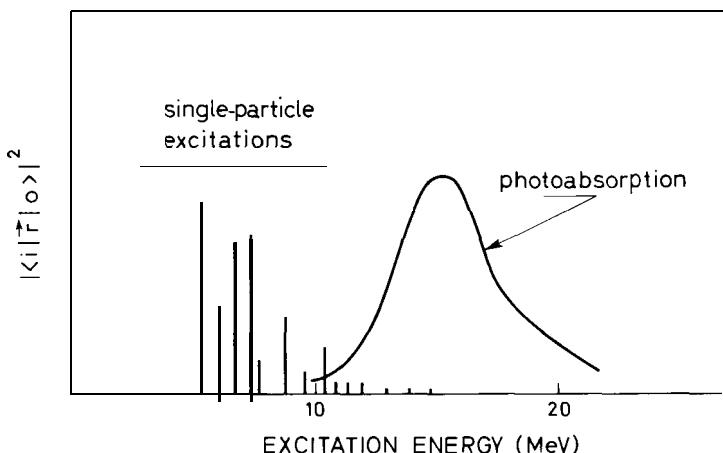


Fig. 2. Frequency distribution of nuclear electric dipole excitations. The figure is a schematic representation (for  $A \approx 100$ ) of the dipole strength for single-particle excitations as compared with the observed frequency distribution of the photo-absorption cross section.

<sup>6</sup>We would like to acknowledge the deep importance for us of the close personal contact with Viki Weisskopf who has provided inspiration for a whole generation of nuclear physicists.

A step in the resolution of the problem resulted from a study of the interaction effects in the single-particle excitations of the closed-shell configuration of  $^{16}\text{O}$ , which revealed a strong tendency towards the formation of linear combinations of different particle-hole configurations collecting the major part of the dipole strength and shifting it to higher energy (48). A highly simplified model based on degenerate single-particle excitations, as in the harmonic oscillator potential, again provided valuable insight by exhibiting exact solutions, in which the total dipole strength was collected into a single high-frequency excitation (49, 50). These schematic models could soon be seen in the more general framework of the normal modes treatment referred to above.

In carrying through this program, one faced the uncertainty in the effective forces to be employed, but it was found possible to represent the interactions by an oscillating average potential acting with opposite sign on neutrons and protons, the strength of which could be related to the isovector component in the static central potential that is present in nuclei with a neutron excess (51, 20).<sup>7</sup> Indeed, it appears that all the collective nuclear modes that have been identified can be traced back to average fields of specific symmetry generated by the effective interaction.

The new insight into the manner in which the vibrations are generated by the interactions in the various channels of particle excitations opened a whole new perspective, since one became liberated from the classical picture of vibrations and could begin to imagine the enormously greater variety of vibrational phenomena that are characteristic of quantal systems. This perspective became apparent 10 to 15 years ago, but there was at that time very little experimental evidence on which to build. The understanding of some of the features in this rich fabric of possibilities has been the result of a gradual process (which added a decade to the gestation of Vol. II of our work on Nuclear Structure) and which is still continuing. A few examples may give an impression of the scope of the new phenomena.

The dipole mode is of isovector character and each quantum of excitation carries unit isospin. It is thus a component of a triplet, which also includes excitations that turn neutrons into protons and vice versa. In a nucleus with equal numbers of neutrons and protons, and total isospin  $T_0 = 0$  in the ground state, the triplet of excitations represents an isobaric multiplet and the different states are therefore directly related in terms of rotations in isospace. However, in a nucleus with neutron excess and total isospin  $T_0 \neq 0$  in the ground state, the dipole excitations with charge exchange may be very different from those with zero component of the isospin (see Fig. 3). The resulting dipole excitation spectrum is schematically illustrated in Fig. 4 and

<sup>7</sup>The close similarity of the results of the hydrodynamic and the microscopic treatments is a special feature of the dipole mode (20), associated with the fact that the single-particle response function for this channel is concentrated in a single frequency region (see Fig. 2).

presents an example of symmetry breaking resulting from the lack of isobaric isotropy of the "vacuum" (the nuclear ground state). Some of the features in the pattern indicated in Fig. 4 have been experimentally confirmed, but the major part of this rich structure remains to be explored<sup>8</sup>.

Another dimension to the vibrational concept is associated with the possibility of collective fields that create or annihilate pairs of particles, in

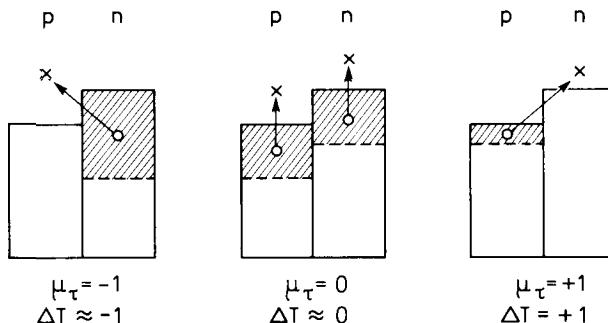


Fig. 3. Single-particle dipole excitations in a nucleus with neutron excess. The boxes represent the occupied proton (p) and neutron (n) orbits and the hatched domains correspond to the particle orbits that can be excited by the isovector dipole field with different components,  $\mu_\tau$ . For large values of the neutron excess, the excitations lead to a change  $\Delta T$  in the total isospin quantum number equal to  $\mu_\tau$ . The figure is from (20).

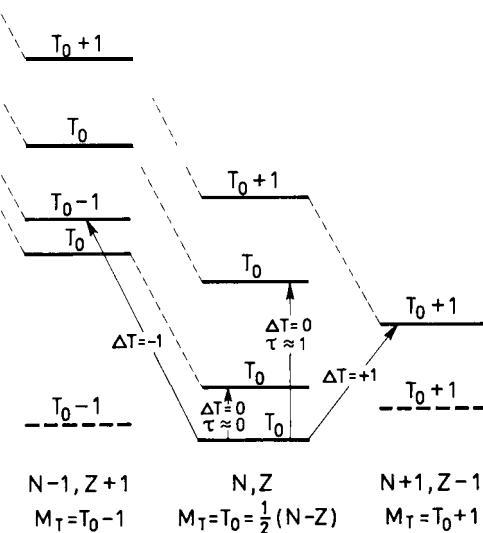


Fig. 4. Isospin of vibrational excitations in nucleus with neutron excess. The ground state of such a nucleus has a total isospin component  $M_T = \frac{1}{2}(N-Z)$  and total isospin  $T_0 = M_T$ . The figure gives a schematic illustration of the pattern of states formed by adding vibrational quanta with isospin  $\tau = 0$  and  $\tau = 1$ . Isobaric analogue states are connected by thin broken lines. The ground states of the isobaric nuclei with  $M_T = T_0 \pm 1$  are indicated by dashed lines. The figure is from (20).

<sup>8</sup>For a summary of this development, see Fallieros (52) and reference (20).

contrast to the field associated with the dipole mode that creates particle-hole pairs and therefore conserves particle number. The new fields are connected with the pairing component in the nuclear interactions which tend to bind pairs into a highly correlated state of angular momentum zero. The addition of such a pair to a closed shell constitutes an excitation that can be repeated and which can thus be viewed as a quantum of a vibrational mode. Fig. 5 shows the pair-vibrational spectrum with the two modes associated with addition and removal of the neutrons from the closed-shell configuration of  $^{208}\text{Pb}$ . One thus encounters a vibrational band in which the members belong to different nuclei. In systems with many particles outside closed shells the ground state can be viewed as a condensate of correlated pairs as in the superconductor (54). Such a condensate can be expressed as a static *deformation* in the magnitude of the pair field, and the addition and removal of pairs from the condensate constitute the associated rotational mode of excitation.

The clarification of the dynamical role of pair fields in the nucleus has resulted from a close interplay of experimental and theoretical work<sup>9</sup>. From the

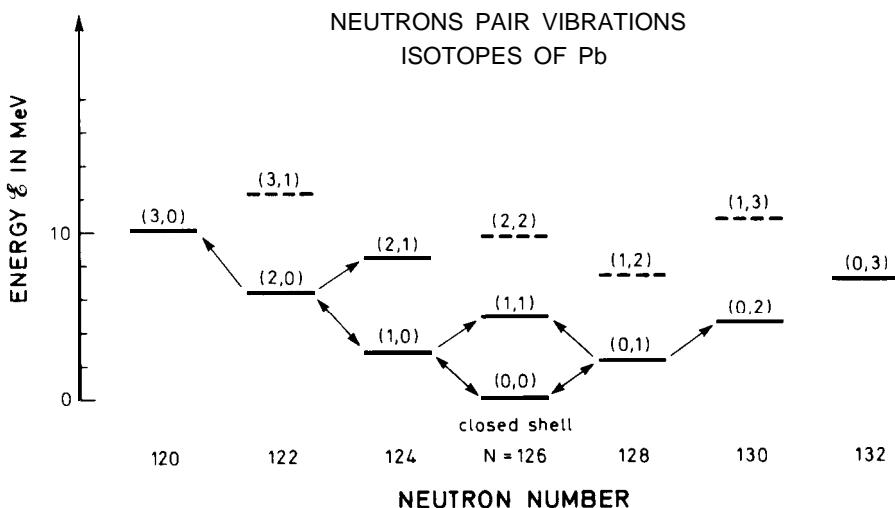


Fig. 5. Neutron monopole pair vibrations based on  $^{208}\text{Pb}$ . The levels in the pair-vibrational spectrum are labelled by the quantum numbers  $(n_-, n_+)$  where  $n_-$  corresponds to the number of correlated  $\Delta = 0$  pairs that have been added to or removed from the closed-shell configuration of  $^{208}\text{Pb}$ . Thus, the levels  $(n_-, 0)$  and  $(0, n_+)$  correspond to the ground states of the even Pb isotopes. The observed levels are indicated by solid lines, while the dashed lines indicate the predicted positions of additional levels. The strong two-neutron transfer processes ((pt) and (tp)) that have so far been observed are indicated by arrows. The figure is from (53).

<sup>9</sup>The concept of pair vibrations in nuclei evolved through the discussions of Högaasen-Feldman (55) early versions of ref. (20) (see, for example, (56)), and Bès and Broglia (57). Excited states of pair vibrational type were identified in the region of  $^{208}\text{Pb}$  by Bjerregaard, Hansen, Nathan, and Hinds (58).

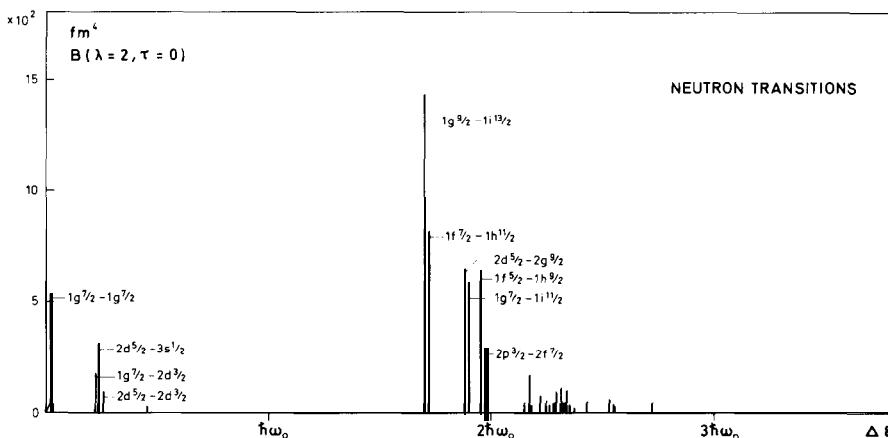


Fig. 6. Single-particle response function for quadrupole excitations. The figure gives the strength of the transitions produced by the quadrupole operator  $r^2Y_4$ , acting on a nucleus with neutron number  $N = 60$ . The single-particle spectrum has been obtained from a potential represented by a harmonic oscillator with the addition of spin-orbit coupling and anharmonic terms reflecting the flatter bottom and steeper sides of the nuclear potential. The excitation energies are plotted in terms of the oscillator frequency  $\omega_0$ , and for the nucleus considered  $h\nu \approx 8.7$  MeV. The figure is from (20).

experimental side, the decisive contribution came from the study of reactions in which a correlated pair of nucleons is added or removed from the nucleus as in the (tp) or (pt) reactions (59).

The new views of vibrations also lead to important insight concerning shape oscillations. While the early considerations were guided by the classical picture provided by the liquid-drop model (60, 4), the lesson of the microscopic theory has been that one must begin the analysis of the collective modes by studying the single-particle excitations produced by fields of the appropriate symmetry.

For quadrupole excitations, an example of such a single-particle response function is shown in Fig. 6 and reveals that the quadrupole excitations involve two very different frequency regions. The first is associated with transitions within the partially filled shells and gives rise to the low-frequency quadrupole mode discussed above. The second frequency region in the quadrupole response function is associated with transitions between orbits separated by two major shells and contains most of the oscillator strength. This group of transitions generates a high-frequency collective mode which has been eagerly expected for many years (61); a few years ago, the study of inelastic electron scattering led to the identification of this mode (62) (see Fig. 7), which has since been found as a systematic feature in a wide variety of inelastic scattering experiments (63). This discovery opens the possibility for a deeper probing of one of the fundamental degrees of freedom in the nucleus.

Returning of the quadrupole response function, the low-frequency excitations reflect a degeneracy in the single-particle spectrum, which is responsi-

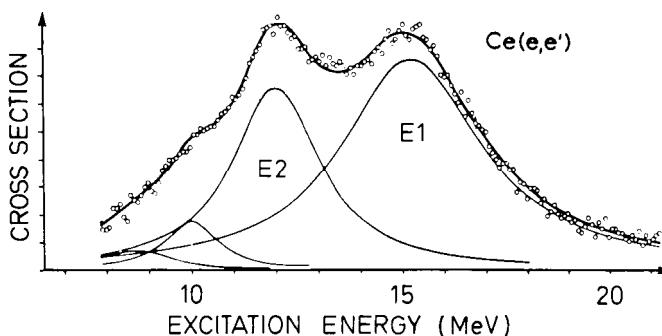
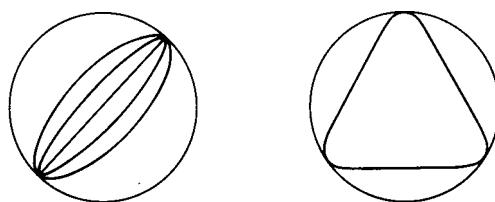


Fig. 7. Inelastic electron scattering on Ce. The highest energy resonance line corresponds with the well-known isovector dipole resonance observed in photo-absorption, while the resonance at an excitation energy of about 12 MeV is identified with the isoscalar quadrupole mode. The figure is from (62).

ble for the tendency to break away from spherical symmetry and to form a spheroidal equilibrium shape (Jahn-Teller effect). One may ask: What underlies this degeneracy in the single-particle spectrum? Apparently this question was never seriously asked until the discovery of the fission isomers (64) revealed the occurrence of important shell-structure effects in potentials that deviate in a major way from spherical symmetry (65) (saddle-point shape). These developments posed in an acute way the question of the general conditions for the occurrence of significant degeneracies in the eigenvalue spectrum for the wave equation. It has been possible to relate this question to the occurrence of degenerate families of periodic orbits in the corresponding classical problem (66, 67, 20) and the instabilities that arise for partially filled shells directly reflect the geometry of these classical orbits. Thus, the observed quadrupole deformations in nuclei can be associated with the elliptical orbits for particle motion in an harmonic oscillator potential (see Fig. 8). The nuclear potential in heavy nuclei also supports orbits of triangular symmetry, and indeed there is evidence for an incipient octupole instability in heavy nuclei.



Classical orbits in nuclear potential

Fig. 8. Periodic orbits in nuclear potential. For small values of angular momentum the motion resembles the elliptical orbits in the oscillator potential. For larger values of angular momentum the effects of the rather sharp nuclear surface can give rise to approximately triangular orbits.

## MODERN VIEW OF PARTICLE-VIBRATION COUPLING

The picture of nuclear dynamics that has emerged from these developments thus involves a great variety of different collective excitations that are as elementary as the single-particle excitations themselves, in the sense that they remain as approximately independent entities in the construction of the nuclear excitation spectrum. Examples of the superposition of elementary modes of excitation are given in Fig. 9 (see also Fig. 5).

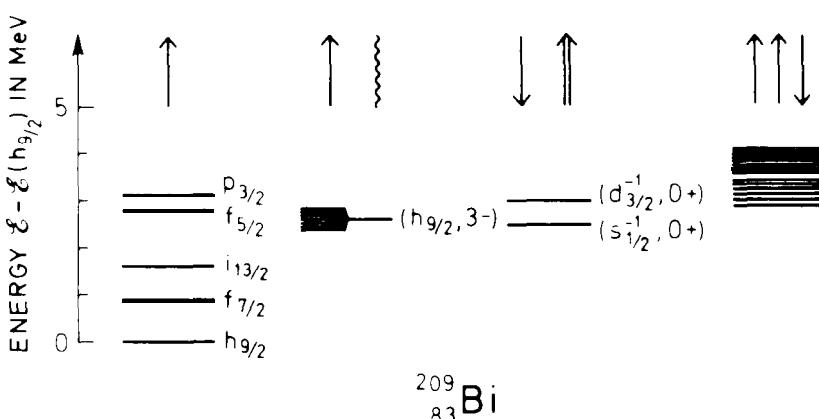
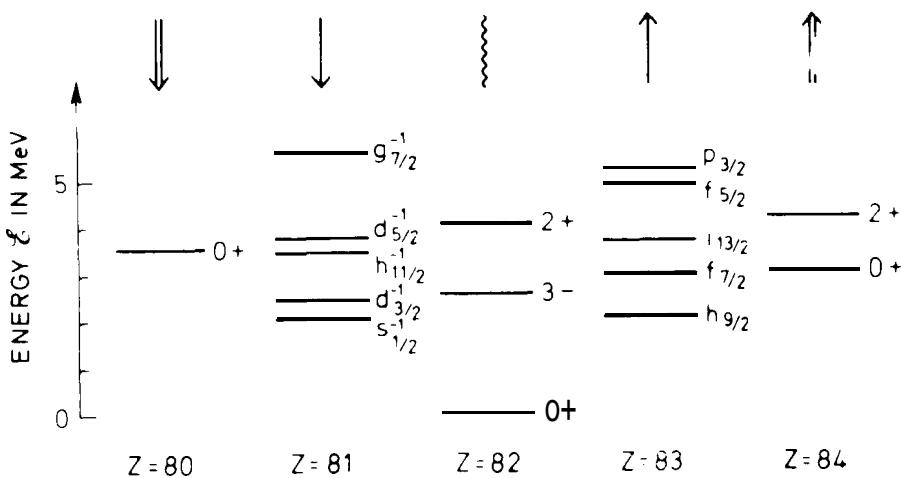


Fig. 9. Elementary excitations based on the closed shell of  $^{208}\text{Pb}$ . The upper part of the figure shows fermion excitations involving the addition or removal of a single proton ( $\Delta Z = +1$  or  $\Delta Z = -1$ ), and boson excitations involving correlated pairs of protons ( $\Delta Z = \pm 2$ ) as well as collective shape oscillations (particle-hole excitations) in  $^{208}\text{Pb}$  itself.

The lower part of the figure gives the observed spectrum of  $^{209}\text{Bi}$ , which comprises partly the single-proton states, and partly states involving the combinations of a single particle or a single hole with a collective boson. The configuration  $(h_{9/2}, 3^-)$  gives rise to a septuplet of states with  $I = 3/2, 5/2 \dots 15/2$  which have all been identified within an energy region of a few hundred keV (see figure 12). At an excitation energy of about 3 MeV, a rather dense spectrum of two-particle one-hole states sets in, as indicated to the right in the figure. The figure is from (53).

In the analysis of the elementary modes and their interactions, a central element is the particle-vibration coupling which expresses the variations in the average potential associated with the collective vibrational amplitude. This coupling is the organizing element that generates the self-consistent collective modes out of the particle excitations. At the same time it gives rise to interactions that provide the natural limitation to the analysis in terms of elementary modes.

Information about the particle-vibration coupling comes from a variety of sources. For some modes, such as the shape oscillations, the coupling can be related to observed static potentials. More generally, the couplings directly manifest themselves in inelastic scattering processes and indirectly in the properties of the modes and their interactions.

The average one-particle potentials appearing in the particle-vibration coupling are of course ultimately related to the underlying nucleonic interactions. Indeed, many of our colleagues would stress the incompleteness in a description that is not explicitly based on these interactions. However, we would emphasize that the potentials are physically significant quantities in terms of which one can establish relationships between a great variety of nuclear phenomena.<sup>10</sup>

It is of course a great challenge to exploit the extensive and precise information available on the two-body forces and the structure of hadrons in order to shed light on the average nuclear potentials. The problem is a classical one in nuclear physics and has continued to reveal new facets, not only because of the complexity of the nuclear forces, but also due to the many subtle correlations that may contribute to the effective interactions in the nuclear medium.

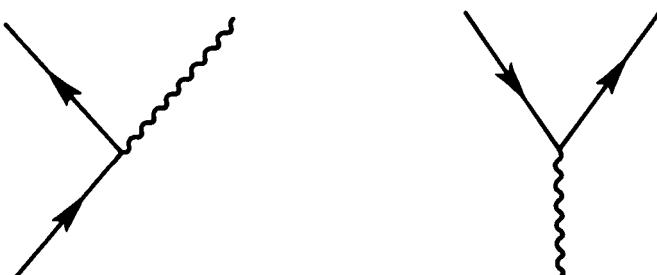


Fig. 10. Basic diagrams for particle-vibration coupling. The solid lines represent particles and the wavy line a phonon of a collective excitation. The particle-vibration coupling creates or annihilates vibrational quanta and at the same time either scatters a particle (or hole) or creates a particle-hole pair.

<sup>10</sup> This issue appears to be endemic in all strongly interacting many-body systems ranging from condensed matter to elementary particles. The approach described here is closely related to that of the Fermi-liquid theory as developed by Landau (68). This formulation operates with a phenomenological effective interaction between the quasiparticles from which the coupling between the particles and the collective modes can be derived. The description of nuclear dynamics in terms of the concepts employed in the theory of Fermi liquids has been developed by Migdal (69).

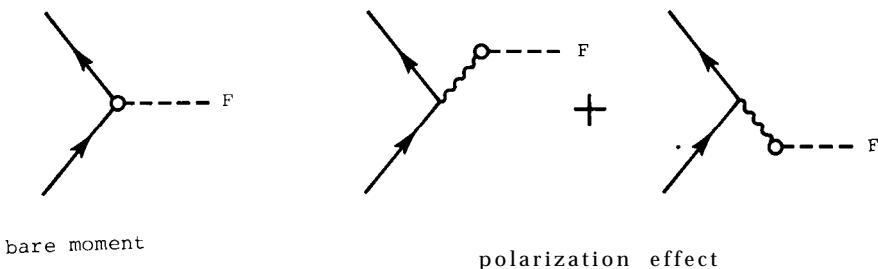


Fig. 11. Renormalization of the matrix elements of a single-particle moment resulting from particle-vibration coupling. The moment  $F$  may be any operator that acts on the degrees of freedom of a single particle, such as an electric or magnetic moment,  $\beta$ -decay transition moment, etc.

The basic matrix elements of the particle-vibration coupling can be represented by the diagrams in Fig. 10, which form the basis for a nuclear field theory based on the elementary modes of fermion and boson type. In lowest order, the coupling gives rise to a renormalization of the effective moments of a particle illustrated by the diagrams in Fig. 11. This renormalization is a major effect in the transitions between low-lying single-particle states and provides the answer to the old dilemma concerning the distribution of the strength between the particle excitations and the collective modes. Thus, for example, for the dipole mode, the one-particle excitations carry a very small admixture of the collective mode, which is sufficient to almost cancel the dipole moment of the bare particle<sup>11</sup>.

Acting in higher order, the particle-vibration coupling gives rise to a wealth of different effects, including interactions between the different elementary modes, anharmonicities in the vibrational motion, self-energy effects, etc. An example is provided by the interaction between a single particle and a phonon in  $^{209}\text{Bi}$  (see Fig. 12)<sup>12</sup>. The lowest single-proton state  $h\ 9/2$  can be superposed on the octupole excitation observed in  $^{208}\text{Pb}$  and gives rise to a septuplet with  $I = 3/2, \dots, 15/2$ . The splitting of the septuplet receives

<sup>11</sup>While the renormalization of the electric quadrupole operator followed directly from the coupling to the deformation of the nuclear surface (13, 20), the occurrence of large deviations in the magnetic moments for configurations with a single particle outside of closed shells was felt as an especially severe challenge to the description in terms of particles coupled to surface oscillations (see, for example, (4)). The clue to the understanding of this effect came from the recognition that special kinds of configuration mixings could give rise to large first-order effects in the magnetic moments (70, 71). Later, it was recognized that this was a manifestation of the particle-vibration coupling involving collective modes of spin-flip type ( $\lambda\pi = 1 +$ ) (61, 20). Experimental evidence for the occurrence of such modes in heavy nuclei came only at a much later time (72). The interpretation of the strong M1 transitions in light nuclei was discussed by Kurath (73) in terms of an intermediate situation between (LS) and (jj) coupling.

<sup>12</sup>The discovery of the weak-coupling multiplet in  $^{209}\text{Bi}$  (75) was a major incentive to the exploration of the scope of the particle-vibration coupling (76, 20).

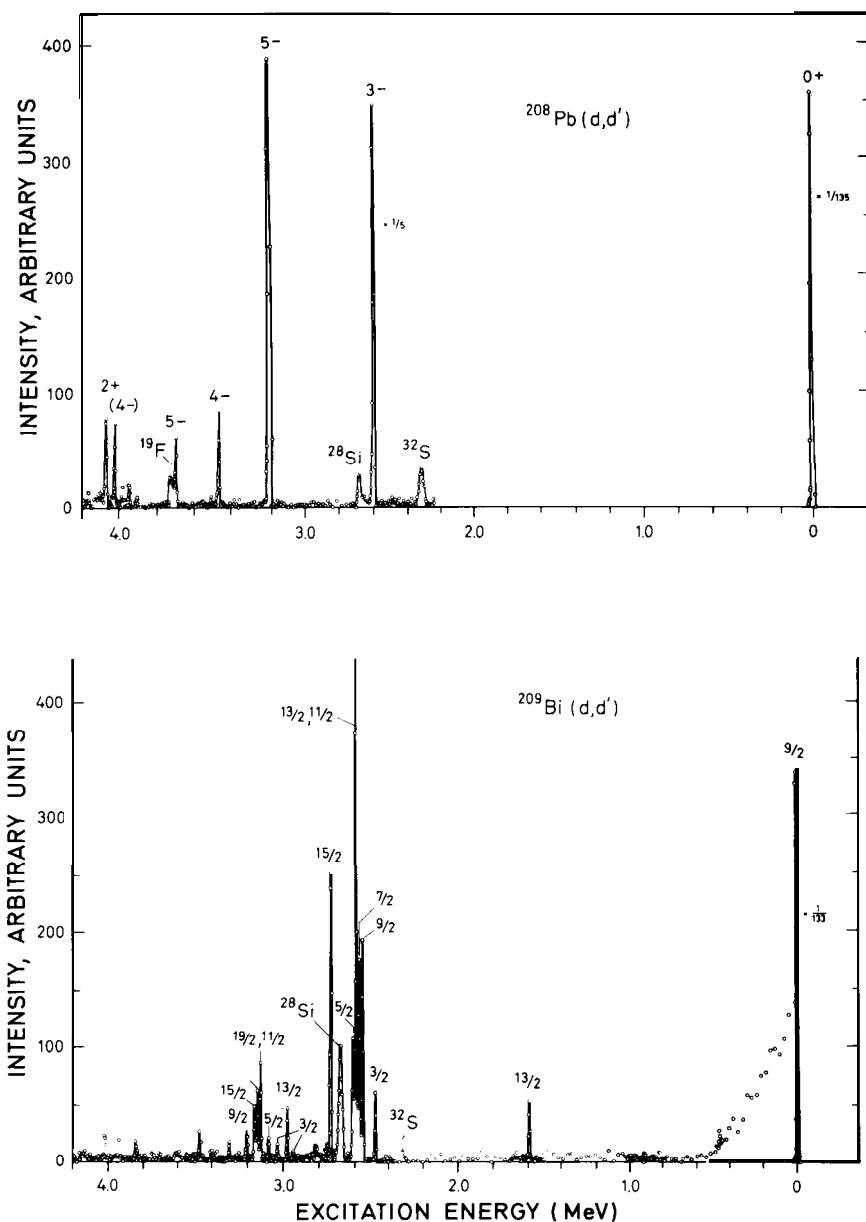


Fig. 12. Energy spectra of deuterons scattered from  $^{208}\text{Pb}$  and  $^{209}\text{Bi}$ . The prominent inelastic group in  $^{208}\text{Pb}$  corresponds to the excitation of an octupole vibrational phonon ( $I\pi = 3^-$ ;  $\hbar\omega_3 = 2.6$  MeV). In  $^{209}\text{Bi}$  the ground state has  $I = 9/2$ , corresponding to a single  $\hbar\pi$  proton outside the closed-shell configurations. The excitation of the octupole quantum in  $^{209}\text{Bi}$  leads to a septuplet of states in the neighbourhood of 2.6 MeV with  $Z = 3/2, 5/2 \dots 15/2$ . The figure is from (20) and is based on data from (74).

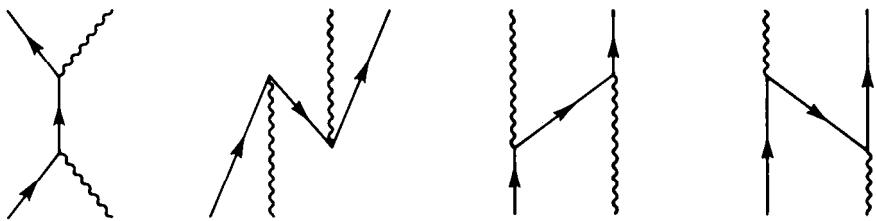


Fig. 13. Second-order diagrams contributing to the energy of a particle-phonon multiplet.

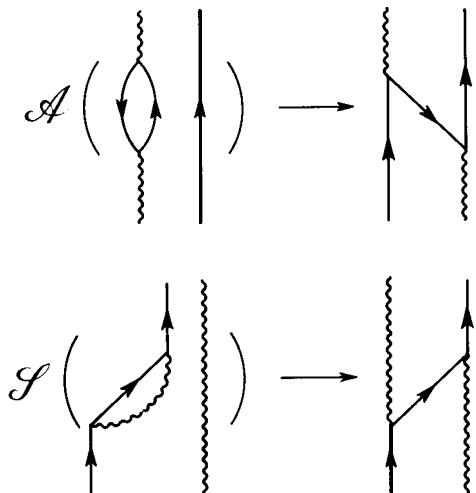


Fig. 14. Linked diagrams associated with symmetrization of particle plus phonon states.

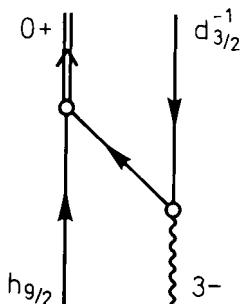


Fig. 15. Coupling between the configurations  $(h_{9/2} 3-)3/2+$  and  $(d_{3/2}^{-1} 0+)3/2+$ , based on particle-vibration vertices.

contributions from the octupole coupling, which can be estimated from the second-order diagrams shown in Fig. 13 (and which are seen to correspond to those of the Compton effect in electrodynamics).

It is an important feature of this calculation that the interactions contain the effect of the antisymmetry between the single particle considered and the particles out of which the vibration is built. This effect is contained in the last diagram in Fig. 13, as schematically indicated in Fig. 14. In a similar manner, the third diagram in Fig. 13 contains the effect of the Bose symmetry of the two identical octupole quanta.

The particle-vibration coupling also leads to the interaction between "crossed" channels, such as illustrated in Fig. 15, which exhibits the cou-

pling of the  $I = 3/2$  member of the septuplet in  $^{209}\text{Bi}$  to the state obtained by superposing a quantum consisting of a pair of protons coupled to angular momentum zero (as in the ground state of  $^{210}\text{Po}$ ) and a single-proton hole in the configuration  $d_{3/2}^{-1}$  (as observed in the spectrum of  $^{207}\text{Tl}$ ). The treatment of this diagram takes proper account of the fact that the two configurations considered are not mutually orthogonal, as must be expected quite generally in a description that exploits simultaneously the quanta of particle-hole type as well as those involving two particles or holes.

As illustrated by these examples, it appears that the nuclear field theory based upon the particle-vibration coupling provides a systematic method for treating the old problems of the overcompleteness of the degrees of freedom, as well as those arising from the identity of the particles appearing explicitly and the particles participating in the collective motion (20, 77). This development is one of the active frontiers in the current exploration of nuclear dynamics.

Looking back over this whole development one cannot help but be impressed by the enormous richness and variety of correlation effects exhibited by the nucleus. This lesson coincides with that learned in so many other domains of quanta1 physics and reflects the almost inexhaustible possibilities in the quanta1 many-body systems. The connections between the problems encountered in the different domains of quanta1 physics dealing with systems with many degrees of freedom have become increasingly apparent, and have been of inspiration, not least to the nuclear physicists who find themselves at an intermediate position on the quantum ladder. Looking forward, we feel that the efforts to view the various branches of quanta1 physics as a whole may to an even greater extent become a stimulus to a deeper understanding of the scope of this broad development.

## REFERENCES:

1. Foldy, L. L., and Milford, F. J., Phys Rev. 80, 751 ( 1950)
2. Bohr, A., Mat Fys. Medd. Dan. Vid. Selsk. 26 no. 14 (1952)
3. Hill, D. L., and Wheeler, J. A., Phys. Rev. 89, 1102 (1953)
4. Bohr, A., and Mottelson, B. R., Mat. Fysk. Medd. Dan. Vid. Selsk. 27, no. 16 (1953)
5. see A. Bohr, preceding lecture
6. Scharff-Goldhaber, G., and Weneser, J., Phys. Rev. 98, 212 ( 1955)
7. Goldhaber, M., and Sunyar, A. W., Phys. Rev. 83, 906 ( 1951)
8. Scharff-Goldhaber, G., Physica 18, 1105 (1952)
9. Stähelin, P., andr Preiswerk, P., Helv. Phys. Acta 24, 623 (1951)
10. Rosenblum, S., and Valadares, M., Compt. rend. 235, 711 ( 1952)
11. Asaro, F., and Perlman, I., Phys. Rev. 87, 393 ( 1953)
12. Kraushaar, J. J., and Goldhaber, M., Phys. Rev. 89, 1081 ( 1953)
13. Alder, K., Bohr, A., Huus, T., Mottelson, B., and Winther, A., Rev. Mod. Phys. 28, **432 (1956)**
14. Inglis, D. R., Phys. Rev. 96, 1059, (1954)
15. Bohr, A., and Mottelson, B. R., Mat. Fys. Medd. Dan. Vid. Selsk. 30, no. 1 ( 1955)
16. Litherland, A. E., McManus, H., Paul, E. B., Bromley, D. A., and Gove, H. E., Can. J. Phys. 36, 378 (1958)

17. Kurath D., and Pičman, L., Nuclear Phys. 10, 313 (1959)
18. Elliott, J. P., Proc. Roy. Soc. (London) **A245**, 128 and 562 (1958)
19. Hill, D. L., and Wheeler, J. A., Phys. Rev. 89, 1102 (1953); Peierls, R. E., and Yoccoz, J., Proc. Phys. Soc. (London) **A70**, 381 (1957); Yoccoz, J. Proc. Phys. Soc. (London) **A70**, 388 (1957) ;Villars, F., Ann. Rev. Nuclear Sci. 7, 185 (1957)
20. Bohr, A., and Mottelson, B. R., Nuclear Structure, Vol. II, Benjamin, W. A. Inc., Reading, Mass. 1975
21. Rutherford, E., Chadwick, J., and Ellis, C. D., **Radiations from Radiooactive Substances**, Cambridge 1930
22. Heisenberg, W., Z. Physik 78, 156 (1932)
23. Bohr N. and Wheeler, J. A. Phys. Rev. 56, 426 (1939)
24. Mayer, M. G. Phys. Rev. 78, 22 (1950)
25. Racah, G. in **Farkas Memorial Volume**, p. 294 eds. A. Farkas and E. P. Wigner, Research Council of Israel, Jerusalem 1952
26. Racah, G. and Talmi, I. Phys. Rev. 89, 913 (1953)
27. Bardeen, J., Cooper, L. N., Schrieffer, J. R., Phys. Rev. **106**, **162** and 108, 1175 (1957)
28. Bohr, A., Mottelson, B. R., and Pines, D., Phys. Rev. 110, 936 (1958)
29. Murray, J. J., Boehm, F., Marmier, P., and Du Mond, J. W., Phys. Rev. 97, 1007 (1955)
30. Bogoliubov, N. N., J. Exptl. Theoret. Phys. (USSR) 34, 58 (1958)
31. Valatin, J. G., Nuovo cimento 7, 843 (1958)
32. Racah, G., Phys. Rev. 63, 367 (1943)
33. Belyaev, S. T., Mat. Fys. Medd. Dan. Vid. Selsk. **31**, no. 11 (1959)
34. Migdal, A. B., Nuclear Phys. 13, 644 (1959)
35. Griffin, J. J., and Rich, M., Phys. Rev. 118, 850 ( 1960)
36. Nilsson, S. G., and Prior, O., Mat. Fys. Medd. Dan. Vid. Selsk. 32, no. 16 ( 1961)
37. Kisslinger, L. S., and Sorensen, R. A., Mat. Fys. Medd. Vid. Selsk. 32, no. 9 (1960)
38. Soloviev, V G., Mat. Fys. Skr. Dan. Vid. Selsk. 1, no. 11 ( 1961) and Phys. Letters 1, **202** (1962)
39. Yoshida, S., Phys. Rev. 123, 2122 (1961) and Nuclear Phys. 33, 685 (1962)
40. Mang, H. J., and Rasmussen, J. O., Mat. Fys. Skr. Dan. Vid. Selsk. 2, no. 3 (1962)
41. Cohen, B. L., and Price, R. E., Phys. Rev. 121, 1441 (1961)
42. Glassgold, A. E., Heckrotte, W., and Watson, K. M., Ann. Phys. 6, 1 (1959) ; Ferrell, R. A., and Fallieros, S., Phys. Rev. 116, 660 ( 1959) ; Goldstone, J., and Gottfried, K., Nuovo cimento 13, 849 (1959); Takagi, S., Progr. Theoret. Phys. (Kyoto) 21, 174 (1959); Ikeda, K., Kobayasi, M., Marumori, T., Shiozaki, T., and Takagi, S., Progr. Theoret. Phys. (Kyoto) 22, 663 (1959); Arvieu, R., and Vénéroni, M., Compt. rend. 250, 992 and 2155 ( 1960) and 252, 670 (1961); Baranger, M., Phys. Rev. 120, 957 ( 1960) ; Kobayasi, M., and Marumori, T., Progr. Theoret. Phys. (Kyoto) 23, 387 ( 1960) ; Marumori, T., Progr. Theoret. Phys. (Kyoto) 24, 331 (1960); Thouless, D. J., Nuclear Phys. 22, 78 (1961)
43. Baldwin, G. C., and Klaiber, G. S., Phys. Rev. 71, 3 ( 1947) and 73, 1156 ( 1948)
44. Goldhaber, M., and Teller, E., Phys. Rev. 74, 1046 ( 1948)
45. Steinwedel, H., and Jensen, J. H. D., Z. Naturforsch. 5a, 413 ( 1950)
46. see especially Wilkinson, D. H., Physica 22, 1039 (1956)
47. see, for example, the discussion in **Proc. Glasgow Conf. on Nuclear and Meson Physics**, eds. Bellamy, E. H. and Moorhouse, R. G., Pergamon Press, London 1955
48. Elliott, J. P., and Flowers, B. H., Proc. Roy. Soc. (London) **A242**, 57 ( 1957 )
49. Brink, D. M., Nuclear Phys, 4, 215 (1957)
50. Brown, G. E., and Bolsterli, M., Phys. Rev. Letters 3, 472 ( 1959)
51. Bohr, A., and Mottelson, B. R. in **Neutron Capture Gamma-Ray Spectroscopy**, p. 3, International Atomic Energy Agency, Vienna 1969

52. Fallieros, S. in ***Intern. Conf. on Photonuclear Reactions and Applications***, Vol. I, p. 401, ed. Berman, B. L., U.S. Atomic Energy Commission, Oak Ridge, Term. 1973
53. Bohr, A., and Mottelson, B., Ann. Rev. Nuclear Sci. 23, 363 (1973)
54. Anderson, P. W. in ***Lectures on the Many-Body Problem*** 2, 113, ed. Caianiello, E. R., Academic Press, New York 1964
55. Högaasen-Feldman, J., Nuclear Phys. 28, 258 (1961)
56. Bohr A. in Compt. rend. du Congrès Intern. de Physique Nucléaire, Vol. 1, p. 487, ed. Gugenberger, P., C.N.R.S. Paris 1964 and ***Nuclear Structure***, Dubna Symposium, p. 179, International Atomic Energy Agency, Vienne 1968
57. Bès, D. R., and Broglia, R. A., Nuclear Phys. 80, 289 (1966)
58. Bjerregaard, J. H., Hansen, O., Nathan, O., and Hinds, S., Nuclear Phys. 89, 337 (1966)
59. see the review by Broglia, R. A. Hansen, O. and Riedel, ***C. Advances in Nuclear Phys.*** 6, 287, Plenum Press, New York 1973
60. Bohr, N., and Kalckar, F., Mat. Fys. Medd. Dan Vid. Selsk. 14, no. 10 (1937)
61. Mottelson, B. in ***Proc. International Conf. on Nuclear Structure***, p. 525, eds. Bromley, D. A., and Vogt, E. W., Univ. of Toronto Press, Toronto 1960
62. Pitthan, R., and Walcher, Th., Phys. Letters **36B**, 563 (1971) and Z. Naturforsch. 27a, 1683 (1972)
63. see, for example, the review by Satchler, G. R., Phys. Reports **14**, 97 (1974)
64. Polikanov, S. M., Druin, V. A., Karnaughov, V. A., Mikheev, V. L. Pleve, A. A., Skobelev, N. K., Subbotin, V. G., Ter-Akop'yan, G. M., and Fomichev, V. A., J. Exptl. Theoret. Phys. (USSR) 42, 1464 (1962)
65. Strutinsky, V. M., Nuclear Phys. **A95**, 420 (1967)
66. Balian, R., and Bloch, C. Ann. Phys. 69, 76 (1971)
67. Wheeler, J. A. in ***Atti del Convegno Mendeleeviano***, p. 189, ed. M. Verde, Accad. delle Scienze di Torino. Torino 1971; Swiatecki, W. J., private communication (1971)
68. Landau, L. D., J. Exptl. Theoret. Phys. (USSR) 30, 1058 (1956) and 32, 59 (1957)
69. Migdal, A. B., ***Theory of Finite Fermi Systems and Applications to Atomic Nuclei***, Wiley (Interscience) New York 1967
70. Arima, A., and Horie, H., Progr. Theoret. Phys. (Kyoto) 12, 623 (1954)
71. Blin-Styole, R. J., and Perks, M. A., Proc. Phys. Soc. (London) **67A**, 885 (1954)
72. Berman, B. L., Kelly, M. A., Bramblett, R. L., Caldwell, J. T. Davis, H. S., and Fultz, S. C., Phys. Rev. 185, 1576 (1969); Pitthan, R., and Walcher, Th. Z. Naturforsch. **27A**, 1683 (1972)
73. Kurath, D., Phys. Rev. 130, 1525 (1963)
74. Ungrin, J. Diamond, R. M., Tjøm, P. O., and Elbek, B., Mat. Fys. Medd. Dan. Vid. Selsk. 38, no. 8 (1971)
75. Alster, J., Phys. Rev. 141, 1138 (1966); Hafele, J. C., and Woods, R., Phys. Letters **23B** 579 (1966)
76. Mottelson, B. R. in ***Proc. Intern. Conf. on Nuclear Structure*** p. 87, ed. J. Sanada, Phys. Soc. of Japan (1968); Hamamoto, I., Nuclear Phys. **A126**, 545 (1969) **A141** 1, and **A155**, 362 (1970); Bès, D. R., and Broglia, R. A. Phys. Rev. 3C, 2349 and 2389 (1971); Broglia, R. A., Paar, V., and Bès, D. R., Phys. Letters **37B**, 159 (1971)
77. Bès, D. R., Dussel, G. G., Broglia, R. A., Liotta, R., and Mottelson, B. R., Phys. Letters **52B**, 253 (1974)



James Rinvater

## JAMES RAINWATER

I was born December 9, 1917 in a small town in Idaho (Council) where my parents had moved from California to operate a general store. My father, who had previously been a civil engineer, died in the great influenza epidemic of 1918. My mother then moved with me and her mother to Hanford, Calif. in the San Joaquin Valley of California, where she was re-married to George Fowler a few years later. In my schooling through high school, I excelled mainly in chemistry, physics and mathematics. Due mainly to my record on an open chemistry competition given by Cal Tech, I was admitted, graduating in 1939 as a physics major. Carl David Anderson was my physics group recitation instructor when he received his Nobel Prize and Milliken was the President of the Institute. I had a short biology course taught by Thomas Hunt Morgan. In 1939 I began graduate study in physics as a teaching assistant at Columbia University where I have remained. During the first two years, I had courses under I. I. Rabi, Enrico Fermi, Edward Teller and J. R. Dunning. Fermi was working on neutron moderator assemblies which led to the first working nuclear "pile" after his group was moved to Chicago. Dunning, Booth, Slack, and Von Grosse held the basic patent on the gaseous diffusion process for  $^{235}\text{U}$  enrichment and were working on its development. This evolved into the Oak Ridge enrichment plants and the present U.S. technology for  $^{235}\text{U}$  enrichment.

In March 1942, I married Emma Louise Smith. We have three sons, James, Robert and William who are all now adults. We also had a daughter, Elizabeth Ann, who died while young.

During W. W. II, I worked with W. W. Havens, Jr. and C. S. Wu under Dr. Dunning (Manhattan Project) mainly doing pulsed neutron spectroscopy using the small Columbia cyclotron. I received my Ph. D after my thesis was de-classified in 1946. I continued at Columbia, first as an instructor, reaching the rank of full professor in 1952. About 1946 funding was obtained from the Office of Naval Research to build a synchrocyclotron which became operational in early 1950. I was involved with the facility development from the beginning and my research has used that facility ever since. The research included neutron resonance spectroscopy, the angular distribution of pion elastic and inelastic scattering on nuclei with optical model fitting. Best known are the muonic-atom-x-ray studies starting with the pioneering 1953 paper with Val Fitch which first established the smaller proton charge radii of nuclei.

Starting in 1948, I taught an advanced nuclear physics graduate course. The Maria Mayer shell model suggestion in 1949 was a great triumph and

fitted my belief that a nuclear shell model should represent a proper approach to understanding nuclear structure. Combined with developments of Weizsäcker's semi-empirical explanation of nuclear binding, and the Bohr-Wheeler 1939 paper on nuclear fission, emphasizing distorted nuclear shapes, I was prepared to see an explanation of large nuclear quadrupole moments. The full concept came to me in late 1949 when attending a colloquium by Prof. C. H. Townes who described the experimental situation for nuclear quadrupole moments. It was a fortuitous situation made even more so by the fact that I was sharing an office with Aage Bohr that year. We had many discussions of the implications, subsequently very successfully exploited by Bohr, Mottelson, and others of the Copenhagen Institute.

Since I joined the Columbia Physics Dept., in 1939, it has been my privilege to have as teachers and/or colleagues many previous Nobel Laureates in Physics: E. Fermi, I. I. Rabi, H. Bethe (Visiting Prof.), P. Kusch, W. Lamb, C. H. Townes, T. D. Lee and L. Cooper in addition to R. A. Milliken, C. D. Anderson, and T. H. Morgan (Biology) while I was an undergraduate at Cal Tech.

Organization Membership, etc.

Fellow : American Physical Society, Institute of Electrical and Electronic Engineers, New York Academy of Sciences, American Association for the Advancement of Sciences.

Member: National Academy of Sciences, Optical Society of America, American Association of Physics Teachers

Recipient: Ernest Orlando Lawrence Award for Physics, 1963.

*Dr. Rainwater died in 1986.*

# BACKGROUND FOR THE SPHEROIDAL NUCLEAR MODEL PROPOSAL

Nobel Lecture, December 11, 1975

by JAMES RAINWATER

Columbia University, New York, N.Y., USA

The conceptual developments on which my award is based occurred to me about twenty-six years ago in late 1949. I shall attempt, as accurately as I can remember, to reconstruct how I viewed the situation of the nuclear shell model and non-spherical nuclear shape at that time.

In a sense the subject began in 1910 when Ernest Rutherford's a particle scattering experiments (1) showed that the nuclear size is  $\leq 10^{-12}$  cm radius, although the atom size is  $\sim 10^{-8}$  cm. This led to Niels Bohr's (2) 1913 theory of the hydrogen atom in terms of quantized electron orbits about the nucleus. This was extended by many workers, especially via the Wilson-Sommerfeld quantization rule that  $\int p_i dq_i = n_i \hbar$  for each degree of freedom, where  $q_i$  and  $p_i$  are the generalized coordinates and momenta of an electron in its orbit about the nucleus. The proposal in 1925 by Goudsmit and Uhlenbeck (3) of the concept of spin 1/2 for the electron and the statement by Pauli (4) of the exclusion principle for electrons, later generalized to all spin 1/2 particles, led to an understanding of the Periodic Table of the Elements, using the old quantum theory, in terms of filling electron shells.

The development of quantum mechanics in 1926 placed the subject on a proper foundation and led to an explosion of the development of atomic physics as is evident from a perusal of the 1935 treatise by E. U. Condon and G. H. Shortley, *The Theory of Atomic Spectra*, Cambridge University Press (1935 and 1951). In the case of the electron orbits or shells about the nucleus, the potential is dominated by the central coulomb attraction of the nucleus, thus permitting treatment of angular momentum as a good quantum number to a good approximation. The coulomb force law was completely known. For the nucleus, early attempts to treat it as composed of protons and electrons were unsatisfactory. When the neutron was discovered by Chadwick in 1932, the picture shifted to a nucleus composed of neutrons and protons bound by strong short range forces. Measurements of nuclear spins soon established that the neutron and proton should probably be taken to have spin 1/2 and to obey Dirac Theory and the Pauli exclusion principle, thus providing a basis for a nuclear shell model. My own detailed introduction to the subject was mainly provided by Bethe's massive review of Nuclear Physics (5) in the 1936, 1937 issues of *Reviews of Modern Physics*.

The subject of attempts at a nuclear shell model are reviewed by Bethe and Bather. (5) I was particularly familiar with the 1937 article by Feenberg and Phillips, (6) "On the Structure of Light Nuclei", where the Hartree method was used with a simplified assumed potential to investigate possible spin orbit Russell Saunders coupling states in filling the first  $I = 1$  shell between

$^4\text{He}$  and  $^{16}\text{O}$ , to explain the behavior of ground and excited nuclear states, etc. A model of particles in a spherical box has the first  $1\ s$  ( $l = 0$ ) state filled by 2 neutrons ( $N$ ) and 2 protons ( $Z$ ) at  $^4\text{He}$ . This nucleus is certainly exceptionally stable, having a binding energy of over 20 MeV for the last nucleon. The first  $p$  shell ( $I = 1$ ) then begins, which is closed at  $^{16}\text{O}$ . It is interesting that the mass  $A = 5$  system is unable to bind the last nucleon and appears as a resonance for neutron or proton scattering on helium. The third shell holds the second  $s$  and the first  $d$  ( $l = 2$ ) shell and is filled at  $^{40}\text{Ca}$  ( $Z = N = 20$ ) which is also unusually bound. It is the heaviest stable nucleus having  $N = Z$ . Beyond this the predicted shell closings disagreed with experiment. The basic force law between nucleons was poorly known.

Before 1940 it was known that the nuclear volume and total nuclear binding both increased roughly linearly with  $A$ , the number of nucleons. The range of the nuclear force between nucleons was known to be  $\approx 2 \times 10^{-18}$  cm and to be deep enough to give the single bound  $s$  ground state for the deuteron when  $n$  and  $p$  spins were parallel, but not when they are antiparallel. A major question involved the reason for "saturation of nuclear forces": i.e., why binding did not increase as  $A(A-1)$ , the number of possible pairings with a "collapsed" nucleus having radius  $\approx 10^{-18}$  cm. This was "answered" by Heisenberg, Wigner, Majorana and others in an ad-hoc fashion by assuming "exchange forces" which were attractive or repulsive depending on the wave function exchange properties. Only after 1950 did Jastrow introduce the concept of a short range repulsion which is now accepted as the reason.

In 1935, Weizsäcker introduced his semi-empirical binding energy formula (7) including volume, surface, isotope, coulomb, and "odd-even" or pairing terms to explain the general trend of nuclear binding. The surface term noted that surface nucleons were less bound, giving a decrease in binding proportional to  $A^{2/3}$  for the radius proportional to  $A^{1/3}$ . This gives less binding for light nuclei and partially explains why maximum stability occurs near  $^{56}\text{Fe}$ . The isotope term is easily understood on a shell model basis or using a Fermi-Thomas statistical model. The number of filled space states increases as  $(Z/2)$  or  $(N/2)$  for protons and neutrons. For a given  $A$ , minimum kinetic energy occurs for  $N = Z$ . For  $N > Z$ , one must change  $(N-Z)/2$  protons to neutrons of higher kinetic energy, with the average kinetic energy change per transferred nucleon proportional to  $(N-Z)$  for a total kinetic energy increase proportional to  $(N-Z)^2$ . This favors  $N = Z$  for stability. This is balanced by the coulomb repulsion energy of the protons which is proportional to  $Z(Z-1)/R$ . This favors having only neutrons. The stability balance for stable nuclei has an increasingly large fraction of the nucleons as neutrons as  $A$  becomes large. This term also gives reduced binding per nucleon beyond  $^{56}\text{Fe}$  and leads to instability against a decay beyond  $A \approx 208$  with not too long lifetimes for the  $^4\text{He}$  fragment to penetrate the coulomb barrier. It was observed that even  $N$ , even  $Z$  ( $e,e$ ) nuclei were unusually stable relative to odd, odd ( $0,0$ ) nuclei, such that after  $^{14}\text{N}$  the stable nuclei for even  $A$  all were ( $e,e$ ), often having two stable even  $Z$  values for each even  $A \geq 36$ . For odd  $A$ , there is almost always only one naturally occurring stable  $Z$  value,

with (e,o) and (o,e) equally favored. This extra binding,  $+\delta$  for ee, zero for  $A$  odd, and  $-\delta$  for (0,0) has  $\delta \sim 1$  to 3 MeV, decreasing as  $A$  increases approximately as  $12A^{-1/2}$  MeV. (See Bethe and Bacher (5), p. 104.) It is also observed that the ground states of even  $A$  nuclei have net spin zero, indicating a space pairing (potential energy) for strongest interaction to cancel the angular momentum contributions. Figure 1, from the Bohr-Mottelson text (8), plots the observed binding per nucleon for beta stable nuclei, vs  $A$ , with a best fit semi-empirical curve for comparison. The deviations of the experimental bindings from the smooth curve give hints of shell structure effects.

In the early 1930's, the energy dependence of the interaction cross section for reactions involving neutrons or protons incident on nuclei was treated by what is now referred to as an optical model approach. The incident nucleon-nucleus interaction was treated using a smoothed interaction potential for the

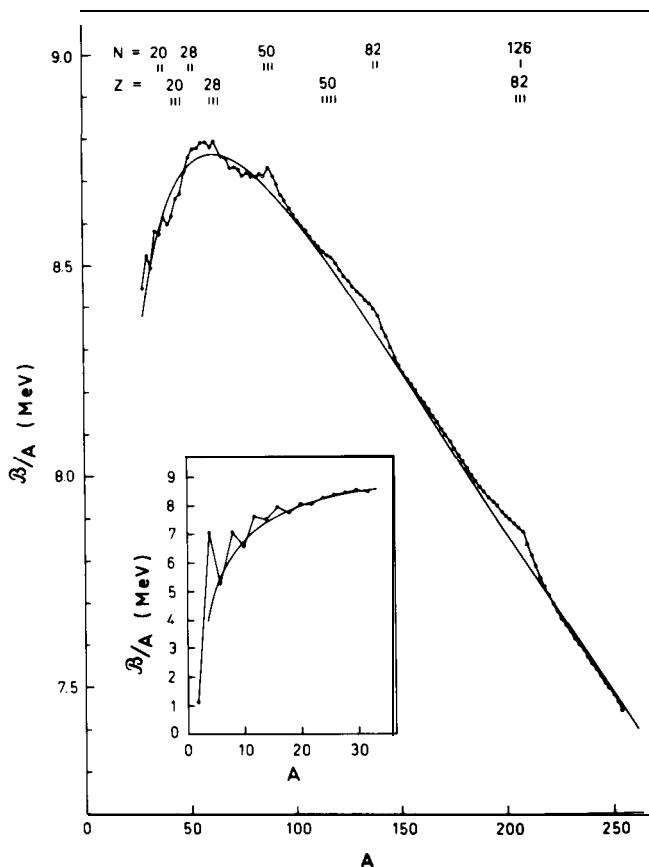


Fig. 1. The average binding energy per nucleon is plotted for nuclei stable against  $\beta$  decay. It is compared with the semi-empirical formula  $B/A = [15.56 - 17.23 A^{-1/3} - 23.28(N-Z)^2/A^2] \text{MeV} - 3\zeta^2 e^2/5R_c A$ , with  $R_c = 1.24A^{1/3} \text{ fm}$ . This figure is from Ref. 8, courtesy of W. A. Benjamin, Inc.

nucleon inside the nucleus. This model predicted "shape" resonances with huge resonance widths and spacings. Early experiments (5) using slow neutrons revealed cross section (compound nucleus) resonances for medium-heavy nuclei  $\sim 10$  to 100 eV apart, with  $< 1$  eV resonance widths. This led N. Bohr to suggest a liquid drop model (9) of the nucleus where the incoming nucleon, as for a molecule hitting a liquid drop, is absorbed near the surface and loses its identity. This is not necessarily incompatible with a shell model, since the shell model refers mainly to the lowest states of a set of fermions in the nuclear "container". However, when combined with the discouragingly poor fits with experiment of detailed shell model predictions (6), the situation  $\sim 1948$  was one of great discouragement concerning a shell model approach.

In the first part of 1949, three groups presented different "explanations" of nuclear shell structure (10) in the same issue of Physical Review. Of these, that of Maria Mayer became the now accepted model. A similar proposal by J. H. D. Jensen and colleagues at the same time led to the Nobel Award in Physics to Mayer and Jensen in 1963. From 1948 to  $\sim 1962$ , I taught a course in "Advanced Nuclear Physics" for graduate students at Columbia. I was also, as an experimental physicist, working on the completion of the Columbia University Nevis Synchrocyclotron which first became operational in March, 1950. During the 1949-50 academic year, I shared an office, Room 910 Pupin, with Aage Bohr who was visiting Columbia that year. I was particularly excited about the Mayer shell model which suddenly made understandable a vast amount of experimental data on spins, magnetic moments, isomeric states,  $\beta$  decay systematics, and the "magic numbers" at  $Z, N = 2, 8, 20$  (28), 50, 82, 126. I reviewed this material at a seminar at Columbia that year.

For over a year previously, I had felt that shell model aspects should have a large degree of validity for nuclei for the following reason. When one considers forming the nuclear wave functions, in  $3A$  dimensional coordinate space, for  $A$  nucleons in a spherical box the size of the nucleus, the shell model states result in lowest kinetic energy. The effective potential energy and the shell model kinetic energy (for  $r < R$ ) are both quite large compared with the net binding energy ( $\sim 8$  MeV) for the least bound nucleons. This is illustrated in Fig. 2 (from Ref. 8). The single particle state energies vs  $A$  have as the "valence" nucleon that with  $E$ , about  $-8$  MeV. If one attempts to use  $\psi$  functions wherein the spatial behavior for each nucleon is very different from that predicted by the shell model, the effect is equivalent to mixing in large amounts of higher energy states having compatible symmetry properties. This mixture of high curvature  $\psi$  states would greatly increase the  $\langle T \rangle$  for the least bound nucleons. I pictured the net  $\psi$  function *not* as a pure Hartree product of single particle  $\psi$  functions, but as being nearly so for the long wave length Fourier aspects of the functions. The short range nucleon-nucleon attractive force would lead to local distortions and clusterings in  $3 A$  dimensional space such as of deuterons and of  $\alpha$  particle structures, etc., but low energy studies would emphasize the long wavelength Fourier aspects which are suggested by the shell model. I was thus delighted by the success of the Mayer model. (I was

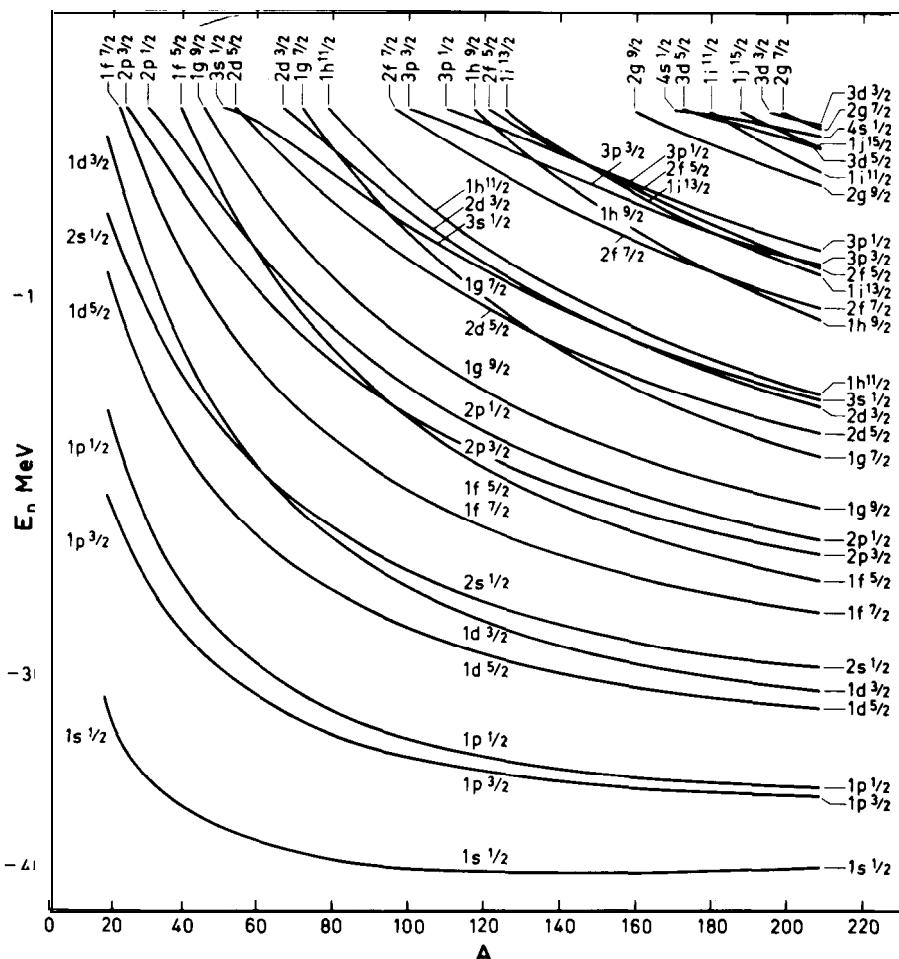


Fig. 2. Energies of neutron orbits using a model of C. J. Veje (from Ref. 8, Vol. I, p. 239). Courtesy of W. A. Benjamin, Inc. The least bound nucleons have energy \$\sim (-8 \text{ MeV})\$ which is small compared with their potential or kinetic energies inside the nucleus.

not then aware of Jensen's work.) The N. Bohr liquid drop model for nuclear reactions and fission did not seem to me to contradict the shell model since the concept of scattering is meaningless for a many fermion ground state, but not for an incident continuum state particle which is not inhibited by the Pauli principle from knocking bound nucleons to excited (unoccupied) states. The compound nucleus states emphasized by Bohr involved an eventual sharing of the excitation by many nucleons so \$\sim 10\$ eV level spacing for medium \$A\$ nuclei plus slow \$l=0\$ neutrons could result. Since \$\sim 1941\$, I had been using the small Columbia cyclotron to carry out slow neutron time of flight spectroscopy studies in collaboration with W. W. Havens, Jr., and C. S. Wu, under Professor J. R. Dunning. We were quite aware of the famous 1939 paper of N. Bohr and J. A. Wheeler on the theory of nuclear fission (11) which emphasized that excited nuclei need not be spherical.

In later 1949, Professor C. H. Townes gave a colloquium presenting the results of a review by Townes, Foley, and Low (12) of the currently available experimental data on nuclear electrical quadrupole moments. The figure which they presented, is shown in Fig. 3. The measured quadrupole moments are presented in the form  $Q/(1.5 \times 10^{13} A^{1/3} \text{ cm})^2$ . The trend shows a qualitative agreement with the Mayer-Jensen shell model, going to zero as one passes through closed neutron and proton shell numbers. For closed shell plus one extra high  $l$  proton, the value of  $Q$  is negative as expected for a proton in an equatorial orbit. As nucleons are removed from a high  $l$  closed shell, the value of  $Q$  becomes increasingly positive, reaching a maximum near where the  $l$  orbital is half filled, and subsequently decreasing. The problem expressed was that the value of  $Q/R^2$ , using  $R = 1.5 \times 10^{-13} A^{1/3} \text{ cm}$ , reaches 10 for  $^{176}\text{Lu}$  which is over 30 times what one might expect for spherical potential shell model wave functions coupled to give a 7 state ( $\mathcal{Z} = 71$ ,  $N = 105$ ,  $\tau = 4 \times 10^{10} \text{ yr}$ ). The rare earth nuclei particularly show much larger than expected  $Q$  values.

As Professor Townes was talking, what seemed like the obvious simple explanation suggested itself to me. Although the Mayer shell model used single particle wave functions based on a spherical potential, the Bohr-Wheeler fission paper showed that, if energetically favorable, the nucleus would distort to a spheroidal shape. For small values of the fractional difference  $\beta$  between the major and minor axes, for constant nuclear volume, the surface area term

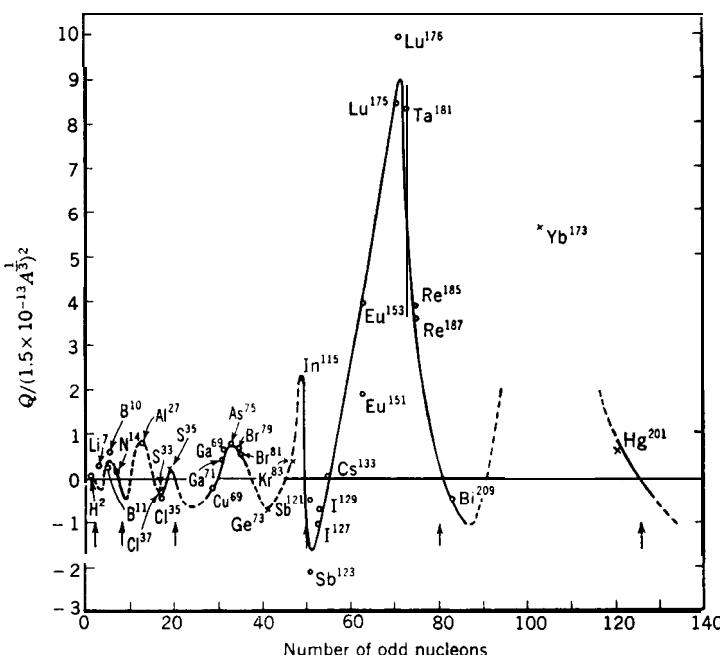


Fig. 3. The plot of  $Q/(1.5 A^{1/3} \times 10^{18} \text{ cm})^2$  for known nuclear quadrupole moments as presented by Townes, Foley, and Low (1949, Ref. 12).

increases as  $\beta^2$ , with the decrease in coulomb energy compensating in part (for high  $Z$ ). My picture assumed constant well depth, but with a distortion where  $R$  increased to  $(1 + 2\beta/3)R_0$  in the z direction and decreased to  $(1 - \beta/3)R_0$  in the x and y directions (or to  $R_0e^{2\beta/3}$  and  $R_0e^{-\beta/3}$ ). If one uses trial  $\psi$  functions which are identically distorted, the potential energy  $< V >$  is the same, but the kinetic energy  $< T_z > = (1 + 2\beta/3)^{-2} < T_{z0} >$  and  $< T_x >$  and  $< T_y >$  become  $(1 - \beta/3)^{-2}$  as large as before. For high  $|m|$  states, the orbits are nearly equatorial and  $< T >$  is nearly proportional to  $R_x^{-2}$  or  $R_y^{-2}$ , with  $< T_x >_0 \approx < T_y >_0 \gg < T_z >_0$ . This clearly favors  $\beta$  negative, or a bulge at the equator to disk (oblate) shape. Each 1 % increase in equator radius ( $R_x$  and  $R_y$ ) gives about 2 % decrease in  $< T >$ , or  $\delta T/T \approx +2\beta/3$ . For a closed shell,  $< T_x >_0 = < T_y >_0 = < T_z >_0$  averaged over all  $l_z (= m)$  for high  $l$ , so there is zero net linear term in the change in total kinetic energy with the distortion parameter  $\beta$ . For a high  $l$  closed shell minus equatorial (high  $|m|$ ) orbitals, the net nuclear angular momentum is the negative of the contribution of the missing nucleons (holes) and the contribution to the kinetic energy term linear in  $\beta$  is equal and opposite to that of the missing equatorial orbit nucleons. The important point is that this yields a term linear in  $\beta$  favoring  $|\beta| \neq 0$ , while the restoring terms are quadratic in  $\beta$ . The expected equilibrium  $\beta$  is thus  $\neq 0$  and is proportional to the coefficient of the linear term for not too large deviations of  $\beta$  from unity. This gives a prolate (cigar) shaped distortion.

The next step was to attempt a more quantitative evaluation of the  $\beta^2$  restoring term. For this, I found the 1939 paper (13, 7) by E. Feenberg useful.

He noted that the surface energy increased as  $E_s = E_{s0} [1 + \frac{8}{45} \beta^2 \dots]$  and

the coulomb energy decreased as  $E_c = E_{c0} / [1 - \frac{4}{45} \beta^2 \dots]$  which requires

$F = 2E_{s0}/E_{c0}$  ( $\approx 42.6 A/Z^2$ )  $> 1$  for a net positive restoring  $\beta^2$  term. This predicted zero net  $\beta^2$  restoring term for  $Z \sim 125$  for beta stable nuclei (no resistance to fission). The net term was  $\beta^2 [2.74 A^{2/3} - 0.054 Z^2 A^{-1/3}]$  MeV. Using this value gave  $^4Q/R^2 = -11$  for a single high  $l$  nucleon above closed shell for a fictitious case of  $A \sim 176$ . The picture, if anything, seemed capable of giving even larger  $Q/R^2$  values than were observed experimentally.

For a prolate spheroidal potential, with the distortion axis in the z direction, the  $\psi$  dependence of the single particle  $\psi$  for  $l_z = m$  is still  $e^{im\phi}$ . However 1, and  $l_y$  and  $l_z$  cannot be good quantum numbers. The core must somehow share the net angular momentum. This consideration helps when one considers the deviations of the observed magnetic moments from the Schmidt limits predicted by the simple shell model.

Aage Bohr pointed out to me at the time (14) that if the nucleus is a spheroid with an "intrinsic" quadrupole moment  $Q_0$  relative to its distortion axis, and total angular momentum is  $I$ , the maximum "observed"  $Q$  is reduced by a factor  $I(2I-1)/(I+1)(2I+3) = 1/10, 2/7, 5/12$ , and  $28/55$  for  $I = 1, 2, 3, 4$ . This emphasizes that  $Q = 0$  for  $I = 0$  or  $1/2$ , but  $Q_0$  may not be zero. Bohr,

Mottelson and colleagues (15) subsequently treated the situation for coulomb excitation cross sections for low lying rotational states. The excitation cross sections uniquely establish the intrinsic quadrupole moment  $Q_0$  for the ground states of distorted even-even nuclei as well as for odd  $A$  nuclei. Figure 4 was prepared by Professor Townes ~ 1957 for a review article on measured quadrupole moments.

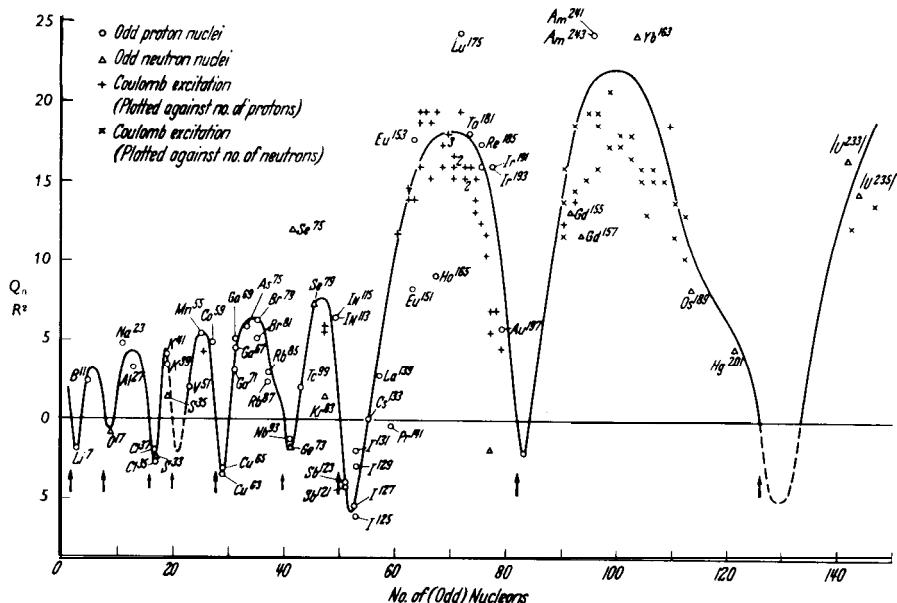


Fig. 4. A later plot of the intrinsic quadrupole moments,  $Q_0 / R^2$ , prepared by C. H. Townes (Ref. 16), using  $R = 1.2 A^{1/3} \times 10^{-13}$  cm. This figure supercedes Fig. 3. It emphasizes the large size of the quadrupole moments relative to values  $|Q_0/R^2| < 1$  expected for a spherical nucleus shell model.

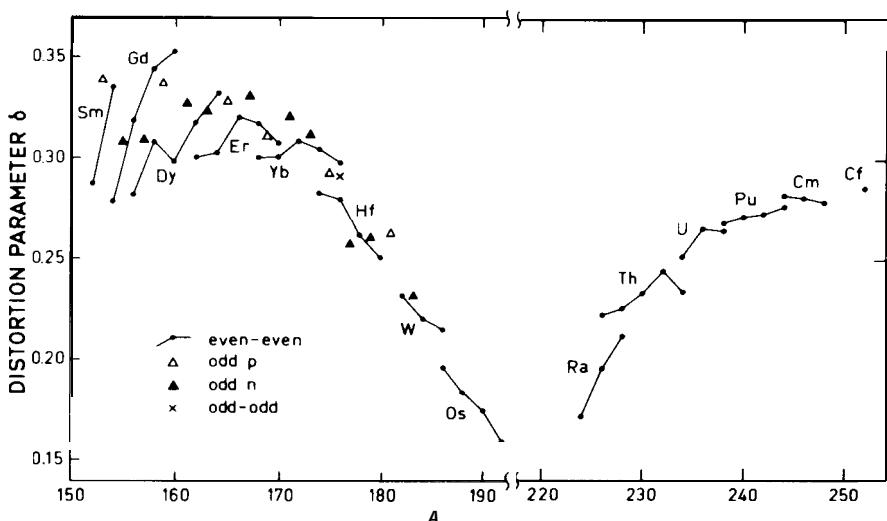


Fig. 5. A plot of the experimental distortion parameter  $\delta$  ( $\approx \beta$  of this paper) in the rare earth region and beyond mass ~ 220. From Ref. 8, Vol. II.

pole moments (16). The largest intrinsic quadrupole moments occur for the rare earth region before the double closed shell  $Z = 82$ ,  $N = 126$ , and beyond  $A \sim 230$  where even higher  $j$  single particle states are involved. Figure 5 shows a recent plot from the just released Vol. II of Bohr- Mottelson, Nuclear Structure (8). The distortion parameter  $\delta$  is nearly the same as the parameter  $\beta$  discussed above. It is seen, as was evident from Professor Townes' 1949 colloquium (12), that many nuclei deviate quite strongly from spherical shape so it does not make sense to use a spherical nuclear model in these regions of atomic size.

After Professor Townes' colloquium, Dr. Bohr and I had many discussions of my concept. He was particularly interested in the dynamical aspects. The distortion bulge could in principle vibrate or move around to give the effect of rotational levels. The first result was his January 1951 paper (17), "On the Quantization of Angular Momenta in Heavy Nuclei". The subsequent exploitation of the subject by Bohr, Mottelson and their colleagues is now history and the main reason for our presence here at this time.

I should mention that the program of evaluating the energies of single particle states in distorted nuclei was subsequently carried out in proper form by Mottelson and Nilsson and by Nilsson alone in the form of "Nilsson diagrams" such as in Fig. 6, which is for proton single particle states beyond

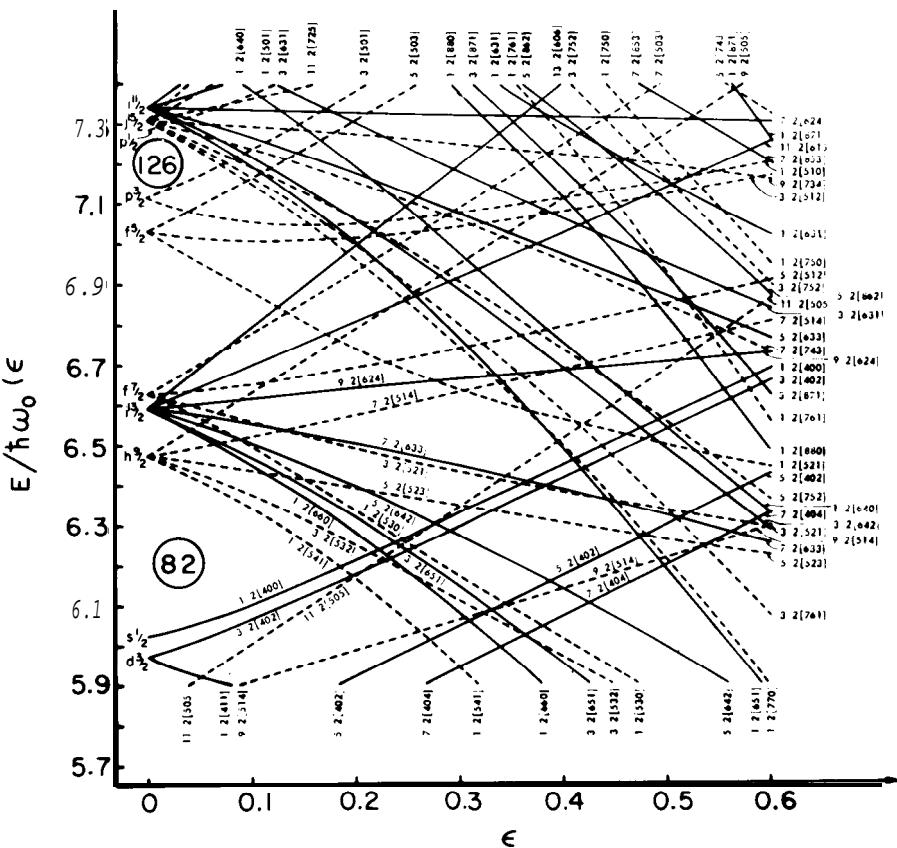


Fig. 6. Nilsson Diagram of single particle shell model proton states vs distortion for  $Z > 82$ .

$\zeta = 82$  vs the distortion. They have also made detailed comparisons, with experimental values of the predicted distortions, etc. with generally excellent results (8). It has also been established that some nuclei have appreciable octopole electric moments and distortions, a generalization of the concept.

One interesting feature of the distorted nucleus shell model is that as the distortion increases, the net energy may go through a minimum and then increase until the energy of an initially higher energy orbital, which decreases faster with deformation, crosses below the previous last filled orbital and subsequently becomes the defining least bound filled state. The net energy may then decrease and show a second minima, etc. vs distortion. This is shown in Fig. 7 which is Fig. 25 of Dr. S. A. Moszkowski's review article (18). This effect seems to be present in sub-threshold nuclear fission where the barrier shape has two minima as shown in Fig. 8 (from Ref. 8, Vol. II, p. 633). This was suggested by Dr. V. M. Strutinski (19) in 1967.

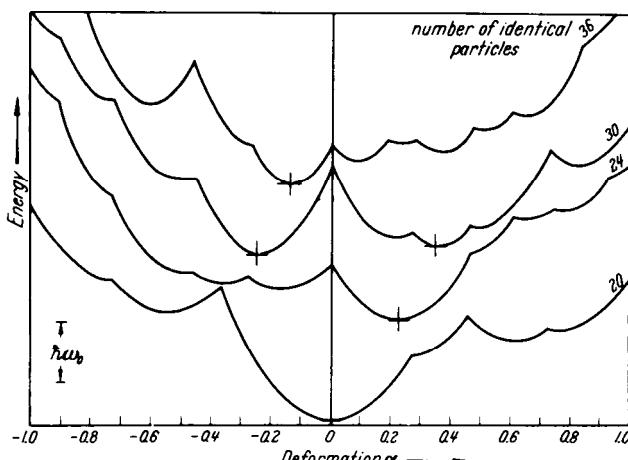


Fig. 7. Deformation potentials for various stages of shell filling-spheroidal harmonic oscillator binding potential. S. A. Moszkowski, Ref. 18, Fig. 25.

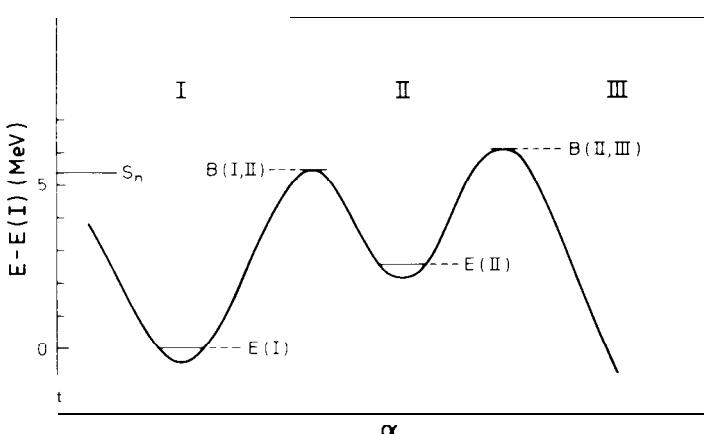


Fig. 8. Double hump energy vs distortion proposed by Strutinski to explain the observed features in sub-threshold nuclear fission. From Ref. 8, Vol. II.

There is one additional effect which I have not yet mentioned which favors spherical shape. If reference is made to the 1937 paper by Feenberg and Phillips (6) on the relative binding of different configurations having two or more  $1 = 1$  nucleons beyond the  ${}^4\text{He}$  core which are combined to form various total  $L$  and  $S$  ( $L$ - $S$  coupling) states for a short range attractive only force, it is seen that the overlap is sensitive to how this is done. As an example, for  $A = 6$ , the two p nucleon wave functions take on the form  $(x+iy)f(r)/\sqrt{2}, (x-iy)f(r)/\sqrt{2}$ , and  $zf(r)$ . The combination  $(x_1x_2+y_1y_2+z_1z_2)/\sqrt{3}$  for  $L = 0$  is more strongly bound than such choices as  $z_1z_2$ , or  $(x_1x_2+y_1y_2)/\sqrt{2}$  which are favored by a spheroidal potential but do not correspond to an eigenstate of  $\vec{L}^2$ . Such an effect may inhibit the distortion for small distortions until the gain from the distortion is more overwhelming relative to such symmetry effects on the interaction potential energy.

Since 1950, I have been mainly concerned with experimental physics research using the Nevis Synchrocyclotron. I have been an admiring spectator of the developments of the theory by the Copenhagen group. My main other (experimental) contribution was in the muonic atom x-ray studies started with Val Fitch (20) in 1953 where we first established the smaller charge radii for nuclei. When I made my proposal for use of a spheroidal nuclear model (14), it seemed to be an obvious answer which would immediately be simultaneously suggested by all theorists in the field. I do not understand why it was not. I was also surprised and dismayed to hear one or more respected theorists announce in every Nuclear Physics Conference which I attended through ~ 1955 some such comment as, "Although the Nuclear Shell Model seems empirically to work very well, there is at present no theoretical justification as to why it should apply." Fortunately, such opinions are no longer expressed.

Although my consideration of the "forcing term" for spheroidal nuclear distortion considered the dependence of the single particle kinetic energy on the distortion, I have never seen a description of my work elsewhere in those terms. A common equivalent phrasing is the "centrifugal force exerted on the barrier" by the orbit. Another method is to compute the increase in the potential energy interaction on distortion. This is equivalent, since for a single particle eigenstate, there is zero rate of change of energy with distortions of  $\psi$ . Thus  $\langle V \rangle$  and  $\langle T \rangle$  must give equal but opposite contributions to the term linear in  $\beta$ .

I wish to thank the Physical Review, W. A. Benjamin, Inc., and Springer-Verlag for permission to use the various figures.

## REFERENCES

1. Rutherford, E., Cambridge Phil. Soc. Proc. 15, 465 (1910).
2. Bohr, N., Phil. Mag. 26, 1 (1913).
3. Uhlenbeck and Goudsmit, Naturweis 13, 593 (1925), Nature 117, 264 (1926).
4. Pauli, W., Zeits für Phys. 31, 765 (1925).
5. Bethe, H. A. and Bather, R. F., Rev. Mod. Phys. 8, 82-229 (1936). Bethe, H. A., Rev. Mod. Phys. 9, 69-224 (1937); Livingston, M. S. and Bethe, H. A., Rev. Mod. Phys. 9, 245-390 (1937).
6. Feenberg, E. and Phillips, M. L., Phys. Rev. 51, 597 (1937).
7. Feenberg, of. E., "Semi-Emperical Theory of the Nuclear Energy Surface", Rev. Mod. Phys. 19, 239 (1947).
8. Bohr, A. and Mottelson, B., Nuclear Structure, Vol. I (1969) and II (1975). Benjamin, W. A., Inc.
9. Bohr, N., Nature 137, 344 (1936). Bohr and Kalckar, Mat. Fys. Medd. Dan. Vid. Selsk. 14, #10 (1937).
10. Maria Mayer, Phys. Rev. 75, 1969 (1949). Feenberg, E. and Hammack, K. C., Phys. Rev. 75, 1877 (1949). Nordheim, L., Phys. Rev. 75, 1894 (1949). Also, see Mayer, M. G., and Jensen, J. H. D., *Elementary Theory OF Nuclear Shell Structure* (1955), Wiley, New York, N.Y.
11. Bohr, N. and Wheeler, J. A., Phys. Rev. 56, 426 (1939).
12. Townes, C. H., Foley, H. and Low, W., Phys. Rev. 76, 1415 (1949).
13. Feenberg, E., Phys. Rev. 55, 504 (1939).
14. Rainwater, J., "Nuclear Energy Level Argument for a Spheroidal Nuclear Model", Phys. Rev. 79, 432 (1950).
15. Alder, K., Bohr, A., Huus, T., Mottelson, B., Winther, A., Rev. Mod. Phys. 28, 432 (1956).
16. Townes, C. H., "Determination of Nuclear Quadrupole Moments", Handbuch der Physick Vol. 38, (1959), edited by Flügge, S., Springer-Verlag, Berlin.
17. Bohr, A., Phys. Rev. 81, 134 (1951).
18. Moszkowski, S. A., "Models of Nuclear Structure", Handbuch der Physick Vol. 39, (1957), edited by Flügge, S., Springer-Verlag, Berlin.
19. Strutinski, V. M., Nucl. Phys. A 95, 420 (1967).
20. Fitch, V. and Rainwater, J., Phys. Rev. 92, 789 (1953).