Founded 1925

Incorporated
by Royal Charter 1961

To promote the advancement of radio, electronics and kindred subjects by the exchange of information in these branches of engineering

Volume 50 No. 11/12 November/December 1980

The Radio and Electronic Engineer

The Journal of the Institution of Electronic and Radio Engineers

Remote Sensing of the Environment

EVER since the early experiments by Marconi and others, the electromagnetic wave has been of paramount importance to those concerned with providing the very large and extensive radio communication services which are now regarded as an essential ingredient to modern societies. Research and technological development, which has taken place over some seven decades, has provided the world with an almost complete communication facility. In addition, more recent developments have allowed interplanetary communications which are very necessary to the current space programmes.

Whilst these advances in radio communication were initially stimulated by both commercial enterprise and demand in a climate of 'communication explosion', system feasibility and high reliability have become criteria which are of utmost importance in the design of the modern system. The achievement of these criteria has required the study of the very wide spectrum of factors which affect the propagation of the electromagnetic wave used as a 'carrier' and which is launched into the geophysical medium for onward transmission to a distant receiving station.

The many very extensive research programmes which have already been carried out in different parts of the world have yielded a great deal of information on some of the particular effects of a number of geophysical factors on the transient wave. For example, in radio systems which are predominantly concerned with ground-wave propagation, both the vegetation distribution and the general transmission path topography are known to have considerable effect on the systems performance. H.f. radio systems, which rely on both the ionosphere and the Earth's surface for long distance propagation, are now known to be subject to significant variations in performance as a result of sea-surface irregularities. Similarly, the many tropospheric radio systems which operate successfully throughout the world, rely entirely upon irregularities in the air mass structures of the lowest part of the Earth's atmosphere for transmissions well beyond the radio horizon.

With the acquisition of such information and the subsequent increase in our knowledge of the behaviour of the electromagnetic wave in the natural environment, there has been the realization that the radio wave could be regarded, in its own right, as a stratagem which had extremely high potential when applied to the study of a wide range of physical phenomena in both terrestrial and extra-terrestrial regions. Indeed, such a realization has, over the years, established the field of Radio Science with a significant sub-range of specialist disciplines, for example radio astronomy, radio geophysics, radio meteorology etc.

A very similar evolution has more recently extended both the frequency range and the character of the wave which may be used for exploratory purposes following the development of both the coherent monochromatic light source of high energy capability and the high power, highly directional, acoustic source. In addition, very sensitive detectors have also been developed in recent years which provide a very valuable complement to the wave sources now available.

Recent research, mainly in the field of communications, has also given rise to very significant developments in signal processing techniques which now allow large amounts of information to be processed rapidly and efficiently. Also, during the last three decades or so, research workers have acquired a much greater appreciation of the more subtle and powerful aspects of the works of Fourier. In pioneering the principle of the hologram Professor Dennis Gabor demonstrated that the correct combination of both amplitude and phase information on the components of a coherent wave front reflected or refracted by an inhomogeneity, gave a 'signature' that was unique to that region or object.

By combining all these facilities and techniques, that have been mainly developed for quite separate and different application, it can be readily appreciated that a very powerful system can be assembled for the purpose of exploring the environment in which we live. Moreover, by the

very nature of such a system, these studies may be carried out at a point very remote from the region or object of interest.

Apart from the obvious scientific interest in the physical properties of our planet and its environment, we have become very conscious, in recent years, of the need for more mineral resources. We have also become increasingly aware of the pollution of both the Earth and its atmosphere as a result of the vast array of industrial processes to which those same minerals are subjected. There is without doubt, therefore, an increasing and urgent demand for methods by which detailed and sensitive surveillances may be made to aid the search for mineral sources and control pollution. Such methods must, of course, be capable of identifying specific concentrations with accurate resolution in both time and space. These methods must also be capable, in the case of pollutants, of providing detailed information on the dynamic processes which determine the movements of concentrations.

In spite of the complexity of these requirements, every endeavour is now being made to develop methods that will provide the necessary information. It is not surprising therefore that the potential of electromagnetic or acoustic wave should be exploited as a remote sensor. Although a great deal of research is yet to be completed, a number of remote sensing systems have already been developed. Many of these systems, operating at wavelengths in the decametric to nanometric range, are already capable of identifying specific elements and their spatial distributions by either transform or scanning techniques. In some cases both techniques are used to advantage, particularly where the spatial distributions of pollutants are being mapped over a period of time. Currently, the range of elements which can be readily identified extends from the harmless but useful algae to the more toxic and harmful chemical concentrations in the atmospheric gas.

In general, the remote sensors that have been developed so far may be divided into two distinct categories—the active sensor and the passive sensor. The active system provides its own wave source to irradiate the region under observation whilst the passive system utilizes either the Sun or background sky as the reference wave source.

In the active system, the object or region of interest may be irradiated from a source of single wavelength or from a source of variable wavelength. The identification of the composition and geometry of the target may be achieved by use of matched or Kalman filtering techniques in association with the 'signatures' of both properties of the target which are represented in the reflected or scattered wave components. In the passive system, the spectral range of wavelengths is effectively fixed and spectroscopic techniques are the basis of the identification process.

When the identification process is completed in either system the information thus obtained is then available for further processing for presentation in the required form.

If the many facilities that are offered by the modern computer are considered, together with the variety of situations in which this type of sensor has already been used, it is clear that this particular branch of radio science has already reached the threshold of a major activity of the future. It is also clear that remote sensing techniques will develop very much further and will have applications of considerable importance in both the civil and military areas of the future.

W. G. BURROWS

Gordon Burrows, guest editor for this special issue, was formerly Senior Tutor to the Department of Electrical Engineering at Imperial College, London, and is now Head of the Department of Marine Electronics at Brunel Technical College, Bristol.

After some ten years in the radar and radio engineering industry he has for the last thirty years been engaged in the academic training of professional engineers in the Polytechnic and University arenas. For some twenty years he has also been engaged in personal research on wave propagation in inhomogeneous media with particular reference to tropospheric radio wave propagation.

In 1963 he pioneered a laboratory model technique to aid the study of tropospheric radio system performance in relation to the behaviour of the propagation medium. More recently his research activities have been extended to embrace the mechanisms of environmental pollution and submarine wave propagation.

Dr Burrows has presented papers at international meetings and has contributed to a number of NATO Advanced Study Institutes. He has also recently visited the Indian Institute of Technology in Delhi to assist in the development of a new research programme on tropospheric propagation. He has made contributions to the Institution's Journal and is currently a member of the Papers Committee.

Members' Appointments

CORPORATE MEMBERS

Sir Raymond Brown, OBE (Fellow 1974, Companion 1969) has recently had conferred on him the honorary degree of Doctor of Science of the University of Bath. Sir Raymond, co-founder of the Racal Electronics Group and present Chairman of Muirhead, was a member of the Institution's Council from 1976 to 1978.

M. James, Dip. El. (Fellow 1958, Member 1955, Graduate 1952) has been appointed managing director of the Vinten Group, based at Bury St. Edmunds. For the past seven years he has been director and general manager of Rank Taylor Hobson, Leicester. Mr James has served on the Institution's Council, on conference organizing committees and on the former Technical Committee from 1958 to 1968, for the last four years as its chairman; it was during this period that the present structure for Specialized Group Committees was drawn up. For several years he represented the Institution of the BSI Standards Committee. In 1961, while at De Havilland Propellors, he contributed a paper to the Journal on system engineering based on work on electronic equipment for aerospace vehicles, notably *Blue Streak*. He subsequently held appointments with the General Electric Company of New York, the Plessey Company and I.T.T.

Dr D. Makow, D.Sc. (Fellow 1965) has for the past thirty years been with the Division of Physics at the National Research Council of Canada and latterly his work has been concerned with studying the structure, optical properties and applications of liquid crystals. Recently he has combined his professional knowledge with his spare-time interest of painting in oils and acrylics and sculpturing in clay, bronze and aluminium to create paintings and sculptures using liquid crystal materials which change hue and shading as the angle of viewing or the temperature changes. A note about his paintings and sculptures appeared in *IEEE Spectrum* for April 1980. Dr Makow has occasionally contributed to the Institution's Journal over the years and with co-authors received the Lord Brabazon Premium in 1959 for a paper on 'A low-drain distress beacon for a crash position indicator'.

With the reorganization of Cable and Wireless into separate corporate groups, several senior members of the Institution within the Technology Directorate have new responsibilities. **J. Powell, TD, M.Sc.,** (President of the Institution) as Engineer-in-Chief is responsible for corporate technological policy and deputizes for the Director Technology. **W. B. Hopkins** (Fellow 1974, Member 1970, Associate 1969) becomes Group General Manager, Technology and is responsible for both administration and management of resources, while **S. G. Bird, B.Sc.** (Fellow 1975, Member 1967, Associate 1962) continues as Chief Engineer Submarine Systems and manages the cableship fleet as a 'revenue-earning unit'.

M. F. Jollyman (Member 1971, Graduate 1967) has been promoted to London Regional Engineering Manager of Chubb Alarms. Mr Jollyman has been with the company for the past three years, previously holding the post of Project Leader.

M. G. Kervell (Member 1966, Graduate 1969) has transferred within the Plessey Group and is now Engineering Manager for Plessey Radio Systems, a new company formed within the Electronic Systems Division and based at West Leigh, Hampshire.

D. A. Larder (Member 1963, Graduate 1961) who has been with International Aeradio since 1970, is now Manager of the IAL communications network project for the Saudi Arabia Directorate of Meteorology and Environmental Protection based in Jeddah. He previously held posts as Manager, Engineering Planning for Emirtel and from 1979 to 1980 as Libya Branch Manager for the company.

M. J. Moorhouse, P.Eng. (Member 1973, Graduate 1967) who has been with the Canadian Broadcasting Corporation since 1975, has been appointed to the position of Energy Conservation Co-ordinator in the office of the Divisional Engineer for the CBCs English Services Division. Mr Moorhouse was previously Assistant Regional Engineer in Alberta.

Sqdn Ldr M. A. Nixon, RAF (Member 1972, Graduate 1970) has taken up an appointment with Air Engineering 14b at the Ministry of Defence in London. He was previously Communications Security Officer for HQ AFCENT.

D. C. Patel, B.Sc., Ph.D. (Member 1973) has taken up an appointment as Product Manager for electronic modular converters with Gardners Transformers in Christchurch, Dorset. He was formerly a Senior Development Engineer with Vickers Medical engineering.

P. Scargill (Member 1965, Graduate 1962) has been appointed General Manager of Hilton Industries of Sarasota, Florida. Mr Scargill was previously with Union Carbide Corporation and held the posts of Product manager with the company's subsidiary in Geneva and Marketing Manager, Consumer Electronics at the plant in Simpsonville, South Carolina.

Sqdn Ldr I. Shephard, RAF (Ret.) (Member 1973) who was Senior Engineering Officer, 9 Squadron, RAF Waddington, has left the Royal Air Force and joined Marconi Radar Systems as a Project Engineer.

A. C. Wood (Member 1972, Graduate 1963) has returned to the United Kingdom after a year with the Civil Aviation Department in Zambia and has taken up a post as lecturer at IAL's Bailbrook College in Bath.

NON-CORPORATE MEMBERS

Lt Cdr C. I. F. Attanayake, SLN (Graduate 1971, Associate 1969) has recently been promoted in the appointment of Electrical and Electronic Engineer, Sri Lanka Navy HQ and Superintendent of the Lighthouse Service.

Flt Lt B. I. Grundon, RAF (Graduate 1967) has taken up the appointment of Engineering Services 1 at Support Command Signals HQ, RAF Benson. He was previously Officer Commanding Ground Radar Flight at RAF St Mawgan.

L. R. Kazipe (Graduate 1977, Student 1972) has returned to the University of Malawi to take up his post as Lecturer II after a two-year course in Technical Teaching at Huddersfield Polytechnic.

M. U. Zauro (Graduate 1977) has been appointed Director, Electricity Supply with the Ministry of Water and Electricity Supply, Sokoto State, Nigeria. He was formerly Deputy Chief Electrical Engineer of the State Workshop of the Ministry of Works in Sokoto.

A. S. S. Wong (Associate Member 1979) who was a Technical Officer with the Telecommunication Authority of Singapore, is now a Field Service Technician with the Victoria Cash Register Company in British Columbia.

Obituary

The Council has learned with regret the deaths of the following members

Clive Leonard Lacey (Member 1944) died on 21st May 1980, aged 63. At the time of his death he was Managing Director of Courtenay Photonics, Dorking, a company which he formed shortly after the war in which he had served as a Squadron Leader in the Technical Signals Branch of the Royal Air Force. For many years he was a keen and successful amateur motor racing driver.

Neil McIntyre (Member) died on the 10th November 1979 at Wilmslow, Cheshire, aged 61. He served in the Royal Signals from 1942 to 1965, holding appointments in the Far East and in the UK where he was with the School of Signals Catterick Camp. During this period he was a member of the North East Section Committee. At the time of his death Mr McIntyre was Regional Engineer with ICI's Central Management Services based at Wilmslow.

Charles Edward Sydney Ridgers (Fellow 1968, Member 1949, Associate 1945) died in March 1980 aged 60, leaving a widow, sons and a daughter. After a short period with electrical engineering companies in the thirties, Mr Ridgers joined the Royal Navy, serving as a Petty Officer until he was invalided out in September 1944. He worked subsequently for Romac Radio Corporation, Philco Radio and Television Laboratories, and Avimo Taunton where he was Electrical Engineer in charge. He then joined Standard Telephone and Cables to take charge of a department designing process equipment for valves and semiconductors, remaining with the company for five years until 1954 when he went to the Radar and Display Laboratories of Decca Radar. From 1960 to 1962 he worked on the development of flight simulators with Miles Electronics and for the next four years was Technical Director of Worthing Electronic Laboratories, dealing with logic systems and transistor inverters on which he contributed a paper to the Journal in 1965. In 1966 he went to Elliott Automation Space and Advanced Military Systems as a Principal Engineer and Consultant in Electronics. Four years later he fulfilled his ambition to form his own Consultancy and worked with two electronics companies to develop the original electronic petrol pump with digital read-out. In 1974 he joined EMI Electronics, at Feltham, where he worked in the Weapons and Systems Division. Charles Ridgers' interests included industrial archaeology and electronic musical instruments.

Peter Talbot Spencer (Member 1964, Graduate 1951) died on 20th June 1980, aged 56. He joined the BBC Engineering Department in 1941 and following 3 years in the RAF, returned to hold posts at various broadcast and television stations, more recently at North Hessary Tor.

Paul Harvard Taylor, D.F.H. (Member 1937) died recently aged 72. For most of his professional life he was concerned with audio frequency engineering and held positions of Technical Director and subsequently Chairman of Pamphonic Reproducers and Technical Director of W. Bryan Savage.

James Gladstone Stewart Wood, B.Sc. (Member 1964, Graduate 1955) died recently in Wellington, New Zealand, aged 52. He was for a number of years Telecommunications Engineer with the New Zealand Civil Aviation Administration and had more recently been on the staff of the Wellington Polytechnic.

In May 1978 a note was published on the career of **Raymond Gray Rippon Robinson** who had died a few months previously. At the time of his death he had been working on a thesis on ion bombardment for the degree of Doctor of Philosophy at the University of Surrey. We have now learned that the Senate of the University have been advised by the assessors that the notes he left justify the posthumous award of the degree to Mr Robinson.

Microprocessor-controlled Torque-speed Analyser

There is a need for an improved method of assessing quickly the torque-speed characteristics of a motor, or other prime mover, without connecting it to a dynamometer. Scientists and engineers at the GEC Hirst Research Centre, in collaboration with GEC Machines Limited, have solved this problem by using the acceleration method and exploiting the speed and accuracy of digital techniques and microcomputers. This approach forms the basis of the Torque-Speed Analyser (or T.S.A.) which measures the speed and acceleration of a machine during run-up to full speed. Knowing the inertia, it then calculates the torque-speed characteristic and, in one version, compares the measured data with a previously stored specification. A decision as to whether the machine has passed or failed a given test can then be displayed along with the detailed test results at the operator's discretion.

The machine is coupled, sometimes to added inertia, but always to a high resolution digital encoder whose output is fed to a binary counter. During the run-up to full speed, the contents of the binary counter are sampled by the microcomputer at regular intervals, and the data stored in random access memory. With this information the microprocessor calculates the speed, acceleration and, knowing inertia, the torque. The software can also multiply torque and speed to give the output power at various speeds.

In principle, the T.S.A. can be used to analyse the performance of electric motors, internal combustion engines, turbines and braking systems. When the T.S.A. is used with motors it complements, and is independent of, steady state testing methods, e.g. the dynamometer. The speed of the T.S.A. method allows many points on a torque-speed curve to be measured before thermal or other effects start to alter the results. For example, it is particularly useful for a.c. induction motors where excessive heating can occur during slow or repeated start-up. It is also of interest where transient torque, in the presence of rapidly changing demand, differs from the equilibrium torque-speed characteristic.

In production applications, the T.S.A. could be used to test many different designs of motor. At the start of the test the operator enters an identification code via a suitable terminal which immediately accesses the specification and value of inertia which has been stored earlier in programmable read only memory. After the run the measured torque-speed curve and power can then be compared automatically with the motor specification and a pass or fail decision given.

Some additional features can be provided which are particularly relevant to the testing of induction motors. In one example, the speed at which the centrifugal switch mechanism operates can be measured for capacitor start and resistance-split-phase induction motors. This is achieved by monitoring the change in motor current during run-up.

Membership Elections and Transfers

The Council confirms the elections and transfers of the following members whose candidatures have been recommended by the Membership Committee and announced in Membership Approval Lists Nos 274 and 275, published in issues of *The Electronics Engineer* dated 3rd July and 7th August 1980.

CORPORATE MEMBERS

Direct Election to Fellow

PERRY, Michael Aubrey.

Transfer from Member to Fellow

McMILLAN, Alan James.

Direct Election to Member

ALLEN, David Martin, ATTA, Nabiel Hussein Awni. BATE, Andrew Maxwell. BOURN, Stuart. DADI, Malkit Singh. FERNANDO, Thomas Michael. HILL, David Roy. HO, Ka-Leung. HOPE, George Edward. HUGHES, Kevin Arthur. KERNTHALER, John Ernest. PILGRIM, Michael George. TOMKINSON, Keith David. TUBBY, Donald. ULLAH, Irfan.

Transfer from Graduate to Member

ARIYARATNAM, Kandiah. DISSANAYAKE, Ratnasiri Bandara. GRANT, Dennis. HORNE, Robert John. MUAT, Alexander Terence. PIGGOTT, Stephen. THEIVASIHAMANY, Selliah.

Transfer from Associate Member to Member

YU, Kwok Chu Peter.

NON-CORPORATE MEMBERS

Transfer from Student to Graduate

BAILEY, Nicholas Raymond. CHEW, Teng Seng. LEUNG, Shing Tak Joseph. MARSHALL, Christopher John. SMITH, Leslie George. SMITH, Robert Arthur. WONG, Chun-Shing Jenson.

Transfer from Associate Member to Graduate

AKPABIO, George Akpabio.

Direct Election to Graduate

AJAYI, Omolayo. CHU, Thomas Chee Kheong. HERCMAN, Krzysztof Marek. LEIGHTON, Andrew Roger Charles. LUK, Pei Stanley. RUFF, Steven Charles. UDODONG, Sam Asukwo. WAKIM, Suheil Nasri N. WONG, Chong Keng.

Direct Election to Associate Member

AH-WANE, Evans Freddie. CHEW, Li Meng. CROWTHER, Ian Henry. D'SOUZA, Dennis. GRAYSTONE, Brian. HIGGINS, Graham Vincent. LIM, Hong Khong. PASKINS, Graham Roderick. WONG, Tew Huay.

Direct Election to Associate

AVERY, Edwin John. EVANS, Derek Keith.

Direct Election to Student

BAXTER, Keith Richard. CHAN, Sai Fan. CHAPMAN, Lynne. CHENG, Chun Keung. CHENG, Hok Ping. CHENG, Man-Chau. CHENG, Shu Chuen. CHEUNG, Chien Hung. CHEUNG, Fook Tsuen. CHEUNG, King Tong Glenn. CHEUNG, Moon Wah. CHOI, Kin Sun. CHOI, Hon Man. CHOW, Pak Hon. CHU, Ka Kin. FONG, Yin Ming. FROST, Graham. FUNG, Wai Shung. HENG, Siang Gek. HO, Chung Kwok. HO, Kim Hung John. IP, Wai-Yi. KUNG, Tat Wing. KWOK, Soo Chu Lora. LAM, Yee Loy. LAU, Shui Ying. LAW, Yu-Kwan. LEE, Kai Chiu. LEUNG, Chi Shing. LEUNG, Chun Wah. LI, Kwong. LIM, Pek Luang. LO, Chi-Hang. LO, Hoi Wah. LO, Wing Fai. LUI, Muk Wai. MAK, Ping Kwan. MAND, Ranjit Singh. MOK, Chung Kong. NAM, Wai Chung. OMOTOSHO, Lawrence. OWEN, Neil Huw Richard. PO, Siu Hing Cindy. POON, Ming Tak. SHAE, Yar Tai. SHAM, Hong Kwong Alfred. SHERRY, John Felix. SUNG, Chi Keung. TAI, Leung Chung. TANG, Catherine Sheng Lin. TANG, Kai Wing. TANG, Kin Cheung. TONG, Fat Cheung. TONG, Tang Soong. TSE, Man Wai. WONG, Man Yuen Daniel. WONG, Ngai Chung. WONG, Wai Chue. WONG, Yat Fung Joseph. WU, Pui Wah. YAN, Kwok Wah Paul.

Award of Royal Medal to Radio Scientist

One of the Queen's Gold Medals for 1980 is to be awarded by the Royal Society to Dr J. P. Wild, C.B.E., F.R.S., Chairman of the Commonwealth Scientific and Industrial Research Organization, Australia, in recognition of his conception of the basic principles of the *Interscan* aircraft instrument landing system and the guidance of its development to a successful conclusion.

The *Interscan* instrument landing system enables any aircraft to locate itself precisely at any moment in the zone covered by the system, which spreads to 40° either side of the runway, to an elevation of 20° above it and extending out to a distance of about 30 nautical miles. Using the system, an aircraft can make a fully instrumental approach and landing on any path, straight or curved, and at different elevation angles to suit the particular type of aircraft. The basic concepts of *Interscan*, its Australian name, were the work of Dr Wild who, with a small team of radio engineers, also achieved practical realization. The system was designed to meet the requirements of the International Civil Aviation Organization, which adopted it in 1978 after extensive trials of *Interscan* and alternative systems. During the next decade *Interscan* will replace the less advanced Instrument Landing System installed at the world's major airports.

Dr Wild, a graduate of Cambridge University served as a radar officer in the Royal Navy from 1943–47, following which he joined the staff of the Division of Radiophysics of the Commonwealth Scientific and Industrial Research Organization (CSIRO), Australia. He was appointed chairman of CSIRO in 1979. He was elected a Fellow of the Royal Society in 1970.

Three Royal Medals, known also as The Queen's Gold Medals, are awarded annually by the Sovereign upon the recommendation of the Council of the Royal Society. Two, initiated in 1825, are awarded for the most important contributions to the advancement of Natural Knowledge (one to each of the physical sciences and the biological sciences) and the other, initiated in 1965, is awarded for distinguished contributions in the applied sciences. The contributions must have been published originally in the Commonwealth within a period of not more than ten years, and of not less than one year of the date of the award.

The other recipients of the Royal Medals are Sir Denys Wilkinson, F.R.S., now the Vice-Chancellor of the University of Sussex, for his highly original research in nuclear physics and Professor Henry Harris, F.R.S., Regius Professor of Medicine in the University of Oxford, who has made numerous advances in the study of cell growth in malignant diseases.

Unconventional Cuts for Quartz Resonators

Work at Philips Research Laboratories and Cathodeon Crystals has shown that the stability of quartz resonators can be much improved by the use of plates cut from a quartz crystal at an orientation different from the one conventionally used.

The most important improvement obtained is the larger temperature of the new SC (strain compensated)-cut as compared with the AT-cut normally used, as is shown in Fig. 1. Other points of interest are the lower sensitivity to thermal and mechanical shock. Two unexpected bonuses were found in the further investigations of the properties of the SC-cut: the new devices aged much more slowly and showed much higher Q values. A comparison of SC-cut properties with those of the AT-cut is given in Table 1. The penalties to be paid for this improved performance are only minor: it is harder to cut the quartz at the correct angles, and marginally more complex circuits are needed to suppress an unwanted vibration mode at a frequency 11% above that of the temperature stable one.

Table 1

Comparison of numerical data pertinent to the new SC-cut frequency stabilizer quartz plates and the conventional AT plates

	AT	SC
Frequency change with temperature in the range 70 to 100°C	±2	±0·2 parts in 10^6
Sensitivity to bending forces equivalent to 100 times the plate weight	0·25	0·02 parts in 10^6
Sensitivity to 0·2°C temperature pulses	0·2	0·01 parts in 10^6
Ageing	0·005	0·002 parts in 10^6/day
Q	$0·2 \times 10^6$	$0·4 \times 10^6$

The results described here have been obtained in investigations partly funded by the British Ministry of Defence. Further information may be obtained from: Cathodeon Crystals Ltd., Linton, Cambridge CB1 6JU.

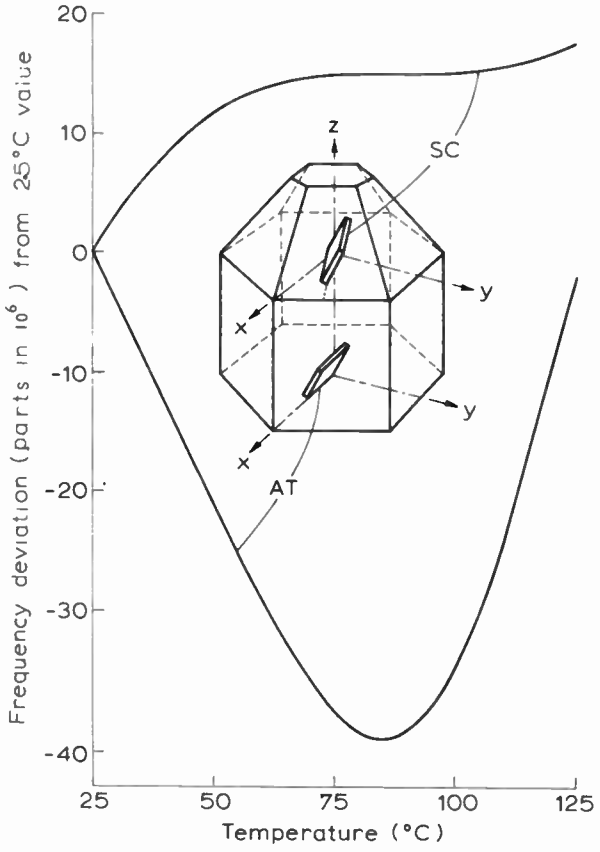


Fig. 1. The frequency deviation df in parts in 10^6 from the 25°C value as a function of temperature T for conventional AT-cut quartz crystals and the new SC device optimized to work in an oven at 85°C. For an SC device any temperature between 75 and 95°C will give a frequency temperature coefficient between ±0·02 parts in 10^6/degC.

Letter to the Editor

From: **A. Kumar, Ph.D.**

Comments on 'A high gain multimode dielectric-coated rectangular horn antenna'*

I read with interest the above paper the results of which appear to be contradictory. The following remarks are in order.

The authors have shown the radiation patterns of dielectric coated horn in Figs. 2(a), 2(b) and 5(a) in E and H-planes. The 3dB beamwidths at 8·36, 9·64 and 11·4 GHz have been calculated from these figures and are shown in Table 1 together with the values obtained from Fig. 5(b).

I was surprised to see the contradictory results, which have made me doubt the validity of the results obtained by the authors.

A. KUMAR

The Higher Institute of Electronics
Hal-Far, Malta.

Table 1 Comparison of results from Figs. 2(a), 2(b), 5(a) and 5(b) of the paper

Frequency, GHz	3-dB beamwidth, deg.	3-dB beamwidth, deg. (From Fig. 5(b))
8·36	21·0 (from Fig. 5(a))	40
9·64	18·0 (from Figs. 2(a), 2(b) and 5(a))	39
11·40	14·0 (Fig. 5(a))	37

*Nair, R. A., Kamal, A. K. and Gupta, S. C., 'A high-gain multimode dielectric-coated rectangular horn antenna', *The Radio and Electronic Engineer*, **48**, pp. 439–43, September 1978.

Engineering and Education

L. H. THORNLEY*

Over a year ago the monthly magazine of the Design Council, Engineering, *published a letter in reply to a mother who had asked why able scholars did not choose engineering as a career. At the end of 1980, considering this country's precarious industrial position, Mr Thornley's letter, reproduced here by permission of the Editor of* Engineering, *seems more pertinent than ever. The writer entered teaching over 25 years ago after a spell in industry and has taught technical studies (including mathematics) to O and A level first in a technical high school, then in a grammar school and now in a sixth form college.*

Social, educational and industrial attitudes towards an engineering or technological career for intelligent pupils have not effectively improved over the years. Society's attitude, often reflected through the various media, tends to ignore or at best play down the role of the engineer. Television glamorizes doctors, lawyers, police, vets, bankers–even motel keepers. But there is a notable lack of interest in our great technical achievements.

Education has weaned people away from the applied sciences. The grammar schools spent years copying the public schools, inculcating in students and parents alike the belief that a white collar and a pen was better in every way than a white coat and slide rule. In my very early youth a clerk was almost aristocratic compared with a skilled fitter and a doctor was a god to people who did not know what a chartered engineer was!

Now the comprehensives are at it. Taking the bright (and not so bright) and turning them into what the grammar schools mercifully only inflicted on the top 25%. Setting geography against technology (if there *is* technology!) and at the lower level throwing the odd period of craftwork with the Z stream to a staff of malcontents who are still turning out teapot stands and pokers.

At higher levels it is just as bad. 'Give us the 'A' levels we have accepted for the past 40 years. Do not bother us with these new-fangled subjects which show the lad might just have an interest in engineeering', moan the departments of engineering. And then they grumble about the low attainment of many of their students on entry. What do they expect, when youngsters have been force-fed on an academic and pure-science diet from ten onwards?

Intelligent students who come to me for 'Engineering Design' are dumbfounded when I talk about Mach 1 speeds on things that never leave the ground or even 'plastic' clothes. They have no idea about materials or sources or problems in technology.

Education, in its broadest sense, must take most of the blame for the decline in our engineering status. It has steadily indoctrinated our young with the idea that engineering is somewhat dirty, rather inferior and plainly plebian. It takes a determined youngster to overcome the cards stacked against him by teachers who have never felt the thrill of satisfaction that an engineering product can give to those involved in its creation, or have never worked alongside other adults in any kind of 'works' situation. No teacher in my opinion should be allowed to teach after a school–college–school situation. He or she has no first-hand knowledge, and little conception of the working world most of their charges will enter.

And what of industry itself? The skills of which we were justifiably proud, the quality which was once the keystone of our products and the reliability which was world renowned, all have been systematically demoted by mergers and takeovers. By groups of accountants who have brought anonymity to the workers and distance to the bosses. The modern industrial world demands profits and efficiency at all costs. This does not necessarily bring quality or reliability.

Industry's attitude to students is often cavalier. It imposes restrictions on age and training which are discouraging and archaic. Good, sound students of 18+ have a difficult time getting into engineering. They are too late for apprenticeships and too early for graduate training. And I must say that of all the visitors who talk to students at the College, undoubtedly the most boring are the engineers and technologists. That students do go on to study these subjects in spite of them is a tribute to their determination.

Of course I am being purposely provocative. It is not all as black as this but a lot of it is and a lot more is grey. In your magazine I am probably preaching to the converted, if not the committed, But if our skills are our lifeblood, then are we going to let our important engineering skills, developed over three centuries now waste away and let others take advantage of our deficiencies? If we do, we shall surely slip further behind and become more dependent on other, more forward-looking nations.

Re-education, perhaps, is one answer. That engineering is a service to people and without it our standard of living will decline; that reaching for a switch, turning a key or turning a tap requires skill, ingenuity and dedication from many people must never be despised or taken for granted; that the Earth is not boundless in its resources and if we are to survive upon it we must husband those resources with skill and care, these are important factors to be instilled into the very young alongside and equal to the other studies we have accepted as fit and proper for our society.

Technology and culture must grow together, neither superior to the other. Striking the right balance, is difficult but it has to be done. The answer in my opinion is to treat all subjects at school level as equal. We should introduce into our junior schools an appreciation of the service that technology brings to people and how it affects all our lives. Thus, one hopes, young people would learn not to take for granted achievements that have taken years of skill and development to come to fruition.

* *93 Kingsway, Scunthorpe, South Humberside DN15 7ER*

Laying Starts of World's Largest Optical Fibre Network

Laying has begun of the first cables for Britain's new optical fibre network intended to show massive savings in the cost of running and enlarging future telecommunications services.

The Post Office, in collaboration with British Industry, has planned a network that, when complete at the end of 1982 will use more than 3300 km (2200 miles) of fibre. The fibre is being made up into nearly 450 km (280 miles) of optical fibre cable which will be installed on 15 routes in England, Wales and Scotland.

The cables used in the new network will contain eight fibres, and can therefore be able to carry up to some 8000 calls—but their overall diameter is only some 10 mm. By comparison, the most commonly-used inter-city coaxial telephone cable—which carries broadly the same number of calls—is about 35 mm in diameter. More optical fibre cables can thus be installed in the Post Office's network of underground ducts and their size and flexibility make them much easier to handle than the stiff and cumbersome telephone cables.

The systems will transmit information using the internationally agreed standards of 8 Mbit/s, 34 Mbit/s and 140 Mbit/s, with respective capacities of 120, 480 and 1920 bothway telephone channels.

The 140 Mbit/s systems which will be used in the trunk network require dependent regenerators located at intervals of about 8 km compared to the 2 km spacing of repeaters used in 12 MHz analogue coaxial cable systems. Lasers are the light sources and the transmission equipment also includes a system for feeding power to the regenerators along the route.

The 34 Mbit/s systems will have many of the features of the 140 Mbit/s systems, and these are being designed to be specially suited to the needs of export markets.

Of the twenty-four 8 Mbit/s systems being ordered, two are for one route on the trunk network (Oxford–Banbury), while the remainder are for nine routes in the junction network, interconnecting local exchanges. Junction routes are relatively short, and usually in urban areas; some routes will not need regenerators, but when regenerators are required, they can usually be housed in Post Office buildings rather than in underground manholes. Regenerator spacings may be up to 12 km.

The fibres used in the cables will be mainly of graded index type. Low-loss fibre—less than 4 dB/km at the operating wavelength, which lies in the range of 820 to 900 nm—will be used for the long haul systems, to exploit the economic advantage obtained.

Transmitting devices are lasers—on the higher bit-rate systems—or light-emitting diodes, while the optical receivers are generally avalanche photodiodes.

New Optic Fibre Link

An optical communication cable combined with an overhead power conductor has been developed by BICC.

Known as Fibral, the optical fibre communication cable is placed in the centre of an overhead power line conductor, thus providing a telecommunications link whilst preserving the electrical and mechanical characteristics of the conductor.

Fibral is claimed to be competitive with existing methods and to offer virtually unlimited capacity which could be for general communication purposes not just the power authority. This advantage is combined with complete absence of interference from or by the system, maintains the control of the system within the power authority and avoids additional way-leave problems.

Fibral has been designed to withstand the rigours of conventional overhead line construction procedures and the vicissitudes of overhead conductor operation, including the various forms of vibration, and changes in length due to temperature, ice and wind loading, and creep. It can be used as either a shield wire or a phase conductor at the line design stage, or substituted for a shield or phase conductor on existing lines.

A 900-metre length of Fibral erected under normal overhead line conditions at the Central Electricity Research Laboratories, Leatherhead in July 1979 has since been successfully transmitting various types of data, including television signals. Two further types of Fibral are planned for erection this summer as additional conductors on an existing South-Eastern Electricity Board 11 kV line.

Industrial Robotics: A New Initiative

The Science Research Council has launched a major new initiative in industrial robotics research to be undertaken in selected universities and polytechnics. Drawing on at least £500 000 a year over the next five years, the programme aims to 'leap-frog' the present generation of robotics devices and provide the research needed to ensure that UK industry can take full advantage of the intelligent robot as it emerges in the mid-1980s. The programme will be co-ordinated by a small technical management team based at the Council's Rutherford and Appleton Laboratories.

UK industry has lagged seriously behind its overseas competitors in the take-up of present robotic techniques and existing research activity in the field is sparse. It is proposed that most of the research to be supported by SRC in the new initiative should at first take place in partnerships to be set up between individual university groups and firms using (or manufacturing) robots. The partnership idea will help to ensure that future research is firmly rooted in real problems encountered now when trying to introduce present generation robots into current manufacture. The experience SRC has already gained in operating the Teaching Company and other co-operative schemes is expected to be particularly valuable in creating such partnerships. The programme will be fully co-ordinated with Department of Industry initiatives in this area.

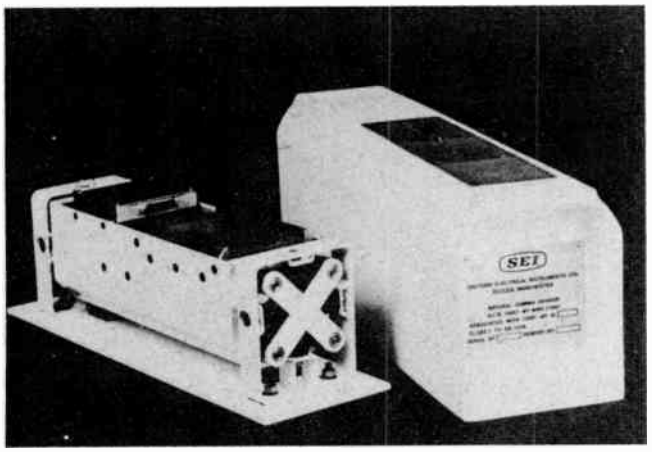
Future industrial robots, particularly for assembly tasks, will need to be much faster, cheaper, more reliable, and more accurate than current models. They must be more tolerant of imperfections in the parts being assembled, and in the way these are presented. To achieve such improvements SRC has identified a number of research topics as central to the programme. These include the development of fast and cheap tactile, visual and aural sensory devices; the development of modular robot construction; research on better, cheaper, lighter linkages and actuators, and the effect of wear on better, cheaper, lighter linkages and actuators, and the effect of wear on accuracy; the optimization of robot dynamics; research on supervisory functions such as safety; and the development of standards.

The present initiative arises from the recommendations of the 'Roberts' Report ('Proposed New Initiatives in Computing and Computer Applications' published by SRC in March 1979). This contained the recommendations of a Council Panel chaired by Mr D. H. Roberts (GEC) calling for a major programme of additional research and training to help the UK to gain maximum benefit from the advent of low-cost microprocessor systems.

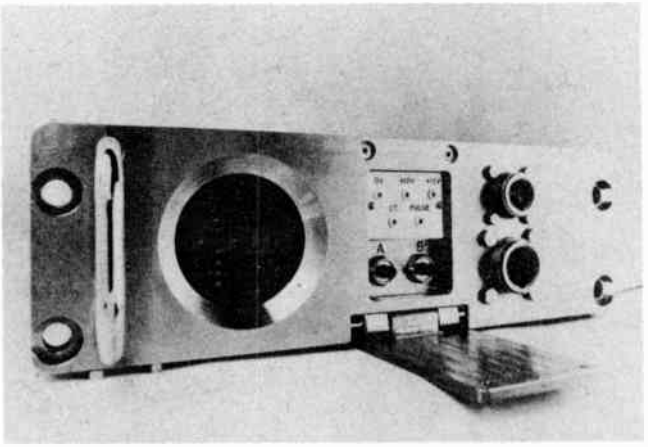
Natural Gamma Radiation Guidance for Coal-cutting Machines

A novel technique for guiding coal-cutting machinery, developed by Salford Electrical Instruments under contract from the National Coal Board, uses the natural gamma radiation present in certain rocks and shales as a means of determining the thickness of coal in a seam. The technique is incorporated in a new guidance system using a natural gamma probe which can be fitted directly onto a coal-cutting machine, a microprocessor-based indicator unit, a remote indicator and a power supply.

The new technique depends for its operation on the fact that many materials in their natural state exhibit gamma-emission effects somewhat more pronounced than their surrounding environment, and this effect is particularly noticeable with shales and rock above and below many coal seams, where the gamma activity can be up to ten times higher than that from the intervening coal.



The natural gamma sensor developed by SEI (under contract from the NCB) for guiding coal-cutting machines.



The microprocessor-controlled indicator unit used in the SEI natural-gamma system.

An important characteristic of this radiation is that it tends to be consistent for a given site—even along the entire length of a coalface—and the absorption of the radiation by different thicknesses of the coal bears a well-established relationship to the actual thickness. The levels of radiation are very low—activity is typically 20 picocuries per gram of material—and hence SEI have had to develop extremely sensitive equipment for monitoring and control purposes.

The radiation detector used in the system is the SEI Type 801 natural gamma probe, and would normally be mounted on the actual coal-cutting machine. It monitors the level of natural gamma flux penetrating the roof or floor of the coal seam, and provides both pulsed and analogue outputs proportional to the count rate. The probe can monitor coal thicknesses of up to 50 cm where the gamma activity of the overlying or underlying material is nominally 20 picocuries per gram in the energy band 30keV–2MeV.

The electronic characteristics of the SEI probe and associated indicator units allow machine speeds of up to 6m/min to be accommodated. The probe is of rugged construction to withstand a mining environment, and can withstand mechanical shocks of up to 50*g*.

Signals from the probe unit are coupled to a microprocessor-based indicator unit which converts radiation-level measurements into coal-thickness information; this is then displayed on a vertical column of light-emitting diodes, which also gives the machine operator a clear indication of the predetermined coal thickness to be left on the roof or the floor and the deviation in 2 cm steps from this nominal thickness.

The indicator unit includes self-checking and automatic setup and calibration facilities which enable it to accommodate varying activity levels for different coal faces, and help to minimize possible commissioning and operational errors. A data-transmission output is provided, and a remote slave indicator is also available which duplicates all the information on the main display and enables the machine to be guided from two locations.

The SEI equipment is ruggedly constructed for use underground or in other demanding environments, and has been designed to comply with UK Intrinsic Safety requirements (Group 1, methane-cadmium standard). SEI also supplies a battery-operated surveying system to determine whether the natural gamma probe can be usefully employed in a particular location, and 'in-seam' monitoring units for use where space is limited.

New Books Received

The following books which have been received recently have been placed in the Institution's Library and may be borrowed by members resident in the British Isles.

Energy Flow and Power Factor in Nonsinusoidal Circuits

W. SHEPHERD (*Professor of Electrical Power Applications, University of Bradford*) and P. ZAND (*United States Transmission Systems*). Cambridge University Press, 1979. 15·5 × 23·5 cm. 275 pages. £23.

CONTENTS: Basic circuit laws. Linear circuits with sinusoidal supply voltage. Power factor compensation in linear sinusoidal circuits. Nonlinear circuits with sinusoidal supply voltage. Chopper circuits with sinusoidal supply voltage. Linear circuits with nonsinusoidal supply voltage. Voltamperes in linear circuits with nonsinusoidal voltage. Nonlinear circuits with nonsinusoidal supply voltage. Compensation of nonlinear circuits with nonsinusoidal supply voltage. Parallel combination of a linear and a nonlinear load.

Intended for engineers dealing with power systems and applications.

Radio and Electronics for Technician Engineers

D. A. JACOBS (*Formerly Matthew Boulton Technical College*). McGraw-Hill, Maidenhead, 1979. 15 × 22·5 cm. 286 pages. £3.95.

CONTENTS: Circuit theorems. Power measurements and attenuators. Capacitor resistor networks. Negative feedback. Positive feedback—sinusoidal oscillators. Positive feedback—relaxation oscillators. Voltage stabilization. Miscellaneous circuits. Modulation. Demodulation. Field and line output circuits.

Aims to provide a broad basis of understanding of theory.

Elementary Semiconductor Physics

H. C. WRIGHT (*Allen Clark Research Centre*). Van Nostrand Reinhold, Wokingham, 1979. 21 × 30 cm. 72 pages. £3.75 paperback. £8.00 cloth.

CONTENTS: Electrical conduction. Basis of semiconductor behaviour. Carrier mobility and lifetime. Junctions and junction devices. Optical, thermal and magnetic effects. Unipolar devices. Microwave devices. Noise.

An introduction for the student.

Interlinking of Computer Networks

Edited by K. G. BEAUCHAMP (*Director of Computer Services, University of Lancaster*). D. Reidel Publishing Company, Dordrecht, 1979. 16 × 24·5 cm. 468 pages. $42.00.

CONTENTS: International networking. Network planning. Network design. Network application. Network management.

A collection of papers given at a NATO Advanced Study Institute held at Bonas, France, in 1978.

Electronic Designer's Handbook
(3rd Edition)

K. HEMINGWAY (*British Aerospace Dynamics Group*). Business Books London, 1979. 14 × 21 cm. 338 pages. £6.50.

CONTENTS: Basic elements of analogue circuit. Practical circuits. Useful techniques.

Concentrates on the design of transistor analogue circuits.

Radio and Electronic Laboratory Handbook
(9th Edition)

M. G. SCROGGIE and G. G. JOHNSTONE. Newnes-Butterworths, Sevenoaks, 1979. 13·5 × 21·5 cm. 587 pages. £17.95.

CONTENTS: The end and the means. Premises and layout. Fundamental principles of measurement. Sources of power and signals. Indicators. Standards. Composite apparatus. Choice and care of equipment. Measurement of circuit parameters. Signal measurements. Measurement of equipment characteristics. Dealing with results.

Extensively updated revision of a highly-regarded 'classic', first published in 1938.

The New Penguin Dictionary of Electronics

CAROL YOUNG. Penguin Books, London, 1979. 14 × 22·5 cm. 598 pages. £7.95.

The emphasis is on solid-state devices and circuits and their basic physics as well as fabrication.

Electronic Games, Design, Programming and Troubleshooting

WALTER H. BUCHSBAUM (*Department of Biomedical Engineering, Brookdale Hospital Medical Center, Brooklyn, N.Y.*) and ROBERT MAURO (*Professor of Electrical Engineering, Manhattan College, New York*). McGraw Hill, New York, 1979. 16 × 24 cm. 324 pages, £8.75.

CONTENTS: Electronic games fundamentals. TV picture parameters. Video effects. Sound effects for electronic games. Microprocessor fundamentals. Programming. Program storage techniques. Microprocessor applications to games. Electronic game parameters. Design examples. Typical electronic games. Troubleshooting techniques.

Probably the most thorough and authoritative work to date on a rapidly expanding microprocessor application.

Integrated Circuits: How to Make Them Work

R. H. WARRING. Lutterworth Press, Guildford, 1979. 14 × 22 cm. 127 pages. £3.95.

CONTENTS: Introduction to integrated circuits. General-purpose i.c.s. (arrays. Op-amps. Audio amplifiers. Heat sinks. Complete radio circuits. Multivibrators. Voltage regulators. Electric motor speed controllers. Filters. Introducing digital circuits. Miscellaneous circuits.

Over 80 circuits are dealt with in this guide for the amateur.

Television Broadcasting Tape Recording Systems

HAROLD E. ENNES. Howard W. Sams, Indianapolis, 1979. 14 × 22 cm. 510 pages. £6.45.

CONTENTS: Basic concepts in video tape systems. The rotating head as the basic time base. Basic requirements: what the system must do. The tape transport, air systems, and control circuitry. The video and fm system. The video processing system. The servo system. Tape dropout compensation. Advanced color-error correction systems. Splicing and editing. Tape system operations. Tape system maintenance. Tape cassette/cartridge systems.

Intended both for students and for practising engineers. Authoritatively backed by Ampex and RCA technical information.

Waveguide Tapers, Transitions and Couplers

F. SPORLEDER (*German Federal Post Office*) and H.-G. UNGER (*University of Brunswick*). IEE Electromagnetic Waves Series. Peter Peregrinus, Stevenage, 1979. 14 × 22 cm. 299 pages. £13.50 UK: £15.80 overseas.

CONTENTS: Coupled wave equations. Taper analysis with horn modes. Abrupt waveguide transitions. Transitions and couplers with one parasitic wave. Gradual waveguide transitions. Waveguide tapers with spurious mode degeneracies. Wave coupling with complete and nearly complete power. Optical waveguide tapers.

A comprehensive monograph in which the essential mathematics is supported by verbal descriptions.

Integration of Analogue Electronic Circuits

J. DAVIDSE (*Department of Electrical Engineering, Delft University of Technology*). Academic Press, London, 1979. 15·5 × 23·5 cm. 270 pages. £18.80.

CONTENTS: Active electronic components. A survey. Monolithic techniques. Some notes on BJT models. Integrable basic analogue circuits. Operational amplifiers. Translinear circuits and multipliers for analogue signals. Noise in analogue integrated circuits. Wideband amplifiers. Special applications of monolithic circuits.

Based on a university course but expanded with requirements of the design or application engineer in mind.

Audio IC Op-Amp Applications
(2nd Edition)

WALTER G. JUNG. Howard W. Sams, Indianapolis, 1978. 13·5 × 21 cm. 202 pages. £5.15.

CONTENTS: IC op-amp parameters important in audio applications. The basic op-amp configurations translated to audio applications. Practical audio circuits using i.c op-amps. Equalized amplifiers and active filters. Miscellaneous audio circuits.

Intended for the designer working on audio signal processing.

Management in the Professions—Guidelines to Improved Professional Performance

MICHAEL ALLSOP. Business Books, London, 1979. 14 × 22 cm. 187 pages. £12.

CONTENTS: Management and the professions. The objectives of professional practice. Planning and organization. Control and review. Personnel recruitment and selection. Personnel management. Communications. The professional as manager and executive. Financial management. Marketing and selling professional services.

Focuses attention on the benefits available to the professions from the adoption of management techniques and attitudes in applying professional knowledge.

Television Principles and Practice

J. S. ZARACH (*Senior Lecturer, North Staffordshire Polytechnic*) and NOEL M. MORRIS (*Principal Lecturer, North Staffordshire Polytechnic*). Macmillan Press, London, 1979. 15 × 23 cm. 288 pages. £12.50 hardcover; £5.95 paper.

CONTENTS: Principles of colour and colour perception. Transmission of monochrome and colour television information. The composite video signal. The essential features of a tv receiver. Tuners and i.f. amplifiers. Vision detectors and video amplifiers. Chrominance signal stages. TV sound. Synchronising pulse separator and field time base circuits. Line time base circuits. Picture tubes and associated circuits. Receiver setting-up procedure. Television aerials and systems. Receiver power supplies.

Emphasis is on PAL-D and it is intended for those taking television examinations for technicians.

Software Portability—An Advanced Course

Edited by P. J. BROWN (*Professor of Computer Science, University of Kent at Canterbury*). Cambridge University Press, 1979. 15 × 22 cm. 320 pages. £5.50.

CONTENTS: Basic concepts. Tools. Pragmatics. Legal aspects. Case studies. Research and the future.

A collection of 27 lectures given at the University of Kent at Canterbury in 1976 under the sponsorship of the SRC and EEC.

The Microprocessor and its Applications

Edited by Professor D. ASPINALL. Cambridge University Press, 1980. 15 × 23 cm. 383 pages. £6.95 paperback.

(See 'New Books Received', April 1980, p. 142, for details of contents of the original hard cover edition.)

Based on an advanced course in 1977 at the University College of Swansea sponsored by the SRC and EEC.

Thermistors

E. D. MACKLEN (*ITT Thermistor Product Division, Taunton*). Electrochemical Publications, Ayr, 1979. 15 × 21·5 cm. 232 pages.

CONTENTS: N.t.c. materials and their fundamental properties. Preparation of practical n.t.c. devices. Electrical characteristics of n.t.c. thermistors. N.t.c. thermistor applications involving temperature measurement, control and compensation. Applications of self-heated n.t.c. thermistors. Miscellaneous applications of n.t.c. thermistors. Properties and applications of indirectly-heated n.t.c. thermistors. High temperature and other unusual forms of n.t.c. thermistors. Elemental p.t.c. thermistors. Materials and fundamental properties of switching p.t.c. thermistors. Preparation and properties of p.t.c. switching thermistors. Applications of switching p.t.c. thermistors. Thermistor selection.

A comprehensive monograph for the design and development engineer.

A Guide to Amateur Radio
(17th Edition)

PAT HAWKER (*G3VA*). Newnes-Butterworths, Sevenoaks, 1979. 19 × 25 cm. 126 pages. £5.40.

CONTENTS: This is amateur radio. Getting started. Communication receivers. Transmitters. The antenna, amateur radio equipment. Workshop practice. The licence examinations. Operating an amateur station. The RSGB and the radio amateur. International amateur radio organizations. Fundamentals of electronics.

A useful introduction for the new radio amateur.

Digital Integrated Circuits and Computers

BARRY WOOLLARD (*Lecturer in Industrial Electronics, Instrumentation and Control Engineering, Walsall College of Technology*). McGraw-Hill, Maidenhead, 1979. 15 × 22 cm. 193 pages. £3.25.

CONTENTS: Integrated circuits. Logic, basic logic gates and numbering systems. Digital integrated circuits. Logic networks. Karnaugh maps. Bistable elements. Binary arithmetic processes. Registers, shift registers and counters. Industrial logic control systems. The digital computer. Computer memories. Concepts of programming.

Based on a short course of lectures, this book gives a useful and comprehensive introduction.

High Performance Loudspeakers

MARTIN COLLOMS (*Consultant*). Pentech Press, London. 1978. 14 × 22·2 cm. 246 pages. £8.95.

CONTENTS: Theoretical aspects of diaphragm radiators. The performance of practical diaphragms. Acoustic loading and low frequency system analysis. Moving-coil direct radiator drivers. Systems and crossovers. The enclosure. Loudspeaker assessment.

An up-to-date analysis and review of techniques with comprehensive references to research material.

Modern Power Transformer Practice

Edited by R. FEINBERG (*Consultant*). Macmillan Press, London, 1979. 16 × 24 cm. 350 pages. £20.

CONTENTS: Theory of transformer design and principles. The use of the automatic electronic digital computer. Transformer cores. Windings. On-load tap-changing equipment. Transformer processing and testing. Transformer noise. Distribution transformers. Power system transformers and inductors. Special transformers. Transformers in distribution systems.

An authoritative collection of articles on transformers for electric power supply.

Semiconductors and Electronic Devices

ADIR BAR-LEV (*Department of Electrical Engineering, Technicon-Israel Institute of Technology*). Prentice-Hall International, London, 1979. 15 × 23 cm. 347 pages. £11.95.

CONTENTS: Semiconductors and their preparation for engineering use. Crystal structure and valence model of a pure and doped semiconductor. Mobility and electrical conductivity. Excess carriers, lifetime, diffusion and transport phenomena. Measuring the electrical parameters of a semiconductor. Energy bands in solids. Homogeneous semiconductor in thermodynamic equilibrium. The pn junction diodes and their static characteristics. Semiconductors and diodes in opto-electronics. The dynamic behaviour of a diode under small and large signals. The field effect transistors. The amplifying stage: transistor small-signal low- and high-frequency models. Introduction to integrated circuits. Microwave semiconductor devices. Additional electronic devices, the thyristor.

A teaching text, aimed at the requirements of a first course in the physics of electronic devices.

The Forces of Nature

P. C. W. DAVIES (*Lecturer in Applied Mathematics at King's College London*). Cambridge University Press, 1979. 15 × 23·5 cm. 231 pages. £12 hardcover, £4.25 paperback.

CONTENTS: Forces and fields. The structure of matter. Breaking matter apart. The new forces. Law and order in the microworld. Unity or diversity?

Relates sub-atomic physics to relativity and quantum theory using only basic mathematics in aiming to appeal to a wide range of readers.

Radio Handbook
(21st Edition)

WILLIAM I. ORR. Howard W. Sams, Indianapolis, 1979. 16·5 × 24 cm. £12.70.

CONTENTS: Direct current circuits. Alternating current, impedance, and resonant circuits. Semiconductor devices. Vacuum-tube principles. Vacuum-tube amplifiers. Radio-frequency power amplifiers. Transmitter keying and control. Mobile and portable equipment. Receivers and exciters. H.f. and v.h.f. power-amplifier design construction. Power supply. Radiation and propagation. The transmission line. Antenna matching systems. H.f. general-purpose fixed directive rotary beam antennas. V.h.f. and u.h.f. antennas. Electronic test equipment. The oscilloscope. Construction practices. Electronic mathematics and calculations.

A long-established US work of reference for the radio amateur.

110 Semiconductor Projects for the Home Constructor

R. M. MARSTON. Newnes-Butterworths, Sevenoaks, 1978. 13 × 21 cm. 118 pages. £2.85.

CONTENTS: 30 silicon-planar transistor projects. 15 field-effect transistor projects. 20 unijunction transistor projects. 15 silicon controlled-rectifier projects. 30 COSMOS digital I.C. projects.

Op-Amps Their Principles and Applications
J. BRIAN DANCE. Newnes-Butterworths, Sevenoaks, 1978. 13·5 × 21·5 cm. 81 pages. £2.25.

CONTENTS: Some basic 741 circuits. Further 741 circuits. Various integrated circuit amplifiers. FET input devices. Audio power circuits. Low noise audio preamplifiers.

Basically intended for non-specialist engineers and home constructors.

Concise Dictionary of Physics and Related Subjects
(2nd Edition)
J. THEWLIS. Pergamon Press, Oxford, 1979. 17·5 × 25 cm. 366 pages. £20.00.

Covers all branches of physics and some related disciplines.

Fundamentals of Electric Circuits
DAVID A. BELL (*Lambton College of Applied Arts and Technology, Sarnia, Ontario*). Reston Publishing Company, Reston, Virginia, 1978. 18 × 24 cm. 645 pages. £13.10.

CONTENTS: The nature of electricity. The international systems of units (SIT). The electrical units. Measuring current, voltage and resistance. Conductors, insulators and resistors. Voltage cells and batteries. Series resistive circuits. Parallel resistive circuits. Series-parallel circuits. Network analysis techniques. Network theorems. Magnetism. Magnetic circuits. D.c. measuring instruments. Inductance. Capacitance. Inductance and capacitance in d.c. circuits. Alternating current and voltage. Phasors and complex numbers. Inductance and capacitance in an a.c. circuit. Series and parallel a.c. circuits. Power in a.c. circuits. A.c. network analysis. Resonance. Transformers. A.c. measuring instruments.

Intended for first-year electricity courses at a level equivalent roughly in the UK to a college of further education.

Computing Systems Reliability
Edited by T. ANDERSON and B. RANDELL (*University of Newcastle upon Tyne*). Cambridge University Press, 1979. 15·5 × 23·5 cm. 461 pages. £15.

CONTENTS: System reliability and structuring. System specification. Program validation. Programming languages. System fault tolerance. Hardware fault tolerance. Protection. Data processing integrity. Software reliability.

A collection of nine lectures given at an advanced course at the University of Newcastle upon Tyne sponsored by the EEC and SRC.

CB'ers Guide to Ham Radio
G. W. MCCARTHY. Van Nostrand-Reinhold, New York, 1979. 15·5 × 23 cm. 304 pages.

Active Filters for Communications and Instrumentation
P. BOWRON (*University of Bradford*) and F. W. STEPHENSON (*Virginia Polytechnic Institute and State University*). McGraw-Hill, Maidenhead, 1979. 15 × 22·5 cm. 279 pages. £6.50 softcover, £12.95 hardcover.

CONTENTS: The approximation problem. Passive networks. Selectivity, sensitivity and stability. Active elements. Controlled-source realizations. Realization using single operational amplifiers. Realization by analogue simulation. Practical considerations. A feature of the text is the reliance on design tables to provide comparative information on the various configurations.

Standard Frequency Transmissions

(Communication from the National Physical Laboratory)

Relative Phase Readings in Microseconds NPL—Station
(Readings at 1500 UT)

JULY 1980	MSF 60 kHz	GBR 16 kHz	Droitwich 200 kHz	AUGUST 1980	MSF 60 kHz	GBR 16 kHz	Droitwich 200 kHz
1	1·6	20·6	18·2	1	1·7	20·0	20·3
2	1·8	20·4	18·2	2	1·7	•	20·4
3	1·8	20·4	18·3	3	1·5	20·0	20·6
4	1·9	20·6	18·4	4	1·7	20·5	20·7
5	1·8	20·4	18·3	5	1·7	20·5	20·8
6	1·8	20·2	18·3	6	1·7	20·0	20·8
7	1·8	20·5	18·4	7	1·3	19·6	20·9
8	2·0	20·3	18·5	8	1·6	20·2	20·9
9	1·9	20·8	18·5	9	1·4	•	21·0
10	1·9	20·8	18·5	10	1·6	20·2	21·1
11	1·9	21·0	18·6	11	2·0	19·8	21·1
12	1·8	•	18·6	12	2·1	19·9	21·2
13	1·8	•	18·6	13	1·9	20·1	21·3
14	1·7	20·6	18·6	14	1·9	19·6	21·5
15	1·9	20·7	18·6	15	1·7	•	21·6
16	1·9	20·8	18·7	16	1·9	•	•
17	1·8	20·7	18·7	17	2·0	•	•
18	1·6	20·9	18·8	18	2·2	20·2	21·7
19	1·6	20·6	18·9	19	2·2	20·1	21·7
20	1·6	20·4	19·0	20	2·3	20·0	21·8
21	1·6	20·4	19·1	21	2·2	20·2	21·9
22	1·8	20·2	19·2	22	2·0	20·6	21·9
23	1·7	20·2	19·3	23	2·2	21·1	22·0
24	1·6	20·0	19·5	24	2·2	21·1	22·1
25	1·5	20·0	19·7	25	2·4	21·8	22·3
26	1·5	19·7	19·8	26	2·4	21·4	22·6
27	1·7	19·5	19·8	27	2·4	21·2	22·7
28	1·7	19·9	19·9	28	2·5	21·1	22·8
29	1·6	20·0	20·1	29	2·5	21·1	22·8
30	1·9	19·7	20·1	30	2·6	21·3	22·9
31	1·9	20·2	20·2	31	2·4	21·6	22·9

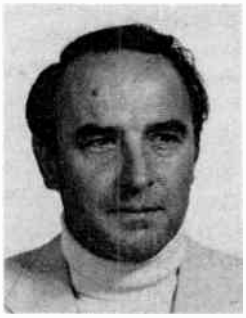
Notes: (a) Relative to UTC scale (UTC_{NPL}–Station) = +10 at 1500 UT, 1st January 1977.
(b) The convention followed is that a decrease in phase reading represents an increase in frequency.
(c) 1 μs represents a frequency change of 1 part in 10^{11} per day.

Contributors to this issue

Michael Lockwood obtained a first class B.Sc. honours degree in physics from the University of Exeter in 1975 and a Ph.D. from the same university for research on ionospheric radio propagation at the Sir Norman Lockyer Observatory in 1978. He has been studying the polar ionosphere at the Radio Research Centre at Auckland University before returning to England to join the Radio and Navigation Department of the Royal Aircraft Establishment, Farnborough. Dr Lockwood has now taken up an appointment at the Appleton Laboratory of the Science Research Council.

Bernard Mitchell received his B.Sc. degree in physics from the University of Exeter in 1967. He remained at the University to work for his doctorate which was awarded in 1971. Since then Dr Mitchell has been engaged on studies of the modulations of h.f. and v.h.f. radio signals caused by the ionosphere.

Karl Rothe qualified in physics and mathematics at the University of Bonn in 1970 and 1971 respectively, and for the next three years was a research assistant at the University of Cologne. In 1975 he was awarded a doctorate in physics from the University of Bonn and then moved to the University of Munich; after a year as a research assistant he was appointed to the academic staff. Dr Rothe's main research activities have been in the application of lasers to atmospheric studies and pollution monitoring.



Wolfgang Fell studied mechanical engineering at the Technical University of Braunschweig and graduated in 1965. In 1969 he joined the Development Division for Magnetic Tape Devices of AEG–Telefunken at Konstanz and a year later was granted his doctorate by the Technical University for a dissertation on the mechanical behaviour of magnetic tape. In 1975 Dr Fell joined Robert Bosch of Darmstadt, as Manager, Research Mechanics in the Television Equipment Division.

Professor Philip M'Pherson has been operating, designing, analysing and thinking about systems for 30 years. Qualifying in 1948 as a marine engineer officer in the Royal Navy, his involvement progressed through propulsion, weapons and inertial navigation systems. To these he added nuclear power systems in 1960 when he joined the Atomic Energy Authority and led the Dynamics Group at the Atomic Energy Establishment, Winfrith. Elected a Fellow of St John's College, Oxford, in 1965, he moved on after two years to the new City University in London becoming Professor and Head of the Department which is now named Systems Science. The Department vigorously pursues educational and research programmes in systems engineering and in the interdisciplinary aspects of systems science.

Torleiv Aa Orhaug graduated at the Chalmers University of Technology, Gothenburg, Sweden, by whom he was awarded a Ph.D. in 1965. His background has been in electron physics, electromagnetic wave propagation and information theory. He subsequently specialized in radio astronomy, working first at Chalmers University and afterwards at the National Radio Astronomy Observatory, Green Bank, West Virginia. In 1969 Dr Orhaug joined the Swedish National Defence Research Institute as a senior staff scientist and for the past ten years he has been head of the Image Processing Laboratory.

Ingvar Akersten is a senior scientist at the Swedish National Defence Research Institute where his main interests are in integrated information processing systems with particular reference to their application to decision support systems and remote sensing. Since graduating, Mr Åkersten has also worked on plasma physics under Professor Alfven and on space physics in the USA under Dr James Van Allen.

Jan Larsson is an associate scientist at the Royal Institute of Technology in Stockholm. After graduating he worked on the development of techniques for photogrammetry and cartography and earlier this year he was awarded his doctorate in photogrammetry for research work into computer image rectification of remotely sensed data.

Ramsay Shearman graduated at Imperial College, London, in 1945 where he held an Imperial College exhibition. He then joined the Admiralty Signal Establishment to work on h.f. radio-communications and in 1947 moved to the DSIR Radio Research Station as a Scientific Officer, where he worked on instrumentation and carried out research on radio propagation. Subsequently, he initiated research on h.f. radar backscatter techniques and their application to broadcasting, communications and geophysics. In 1960 as a Principal Scientific Officer he was attached to the Canadian *Alouette* Satellite team, assisting with the design and also initiating UK participation in the programme. In 1962 he joined the University of Birmingham, becoming Professor of Electronic and Electrical Engineering in 1965. There he initiated research on microwave antennas, propagation, mobile radio, radar and remote sensing. Professor Shearman is at present head of the Postgraduate School and Head of the Radio Research Group. His main research interest is in remote sensing of sea-state by radar.

Roy Cole graduated in physics at University College London in 1954 and was awarded the Ph.D. degree for his work on electron beam spectroscopy at Birkbeck College, London in 1969. In 1954 he joined EMI Electronics and was employed in the Microwave Division on the design and development of microwave ferrite devices and components and radar systems. He was appointed Lecturer in the Physics Department of the North London Polytechnic in 1962 and in 1965 moved to the Electronic and Electrical Engineering Department at University College London where he is currently a senior lecturer. Dr Cole's research interests are in the fields of millimetre wave propagation, where he has worked at 36, 55 and 110GHz, and atmospheric measurements, the latter giving rise to his work on quantitative sounding.

Tim Moulsley graduated in physics and electronics at Chelsea College, London, in 1976. He was awarded the Ph.D. degree in 1979 in the Department of Electronic and Electrical Engineering, University College London for his work on quantitative atmospheric acoustic sounding. He is currently a Research Fellow with the Department of Electrical and Electronic Engineering, Portsmouth Polytechnic, engaged in research on microwave propagation on Earth-space links. In the summer of 1980 he held a temporary appointment as a Research Associate with CIRES (Cooperative Institute for Research in Environmental Sciences), University of Colorado, to investigate low-level atmospheric acoustic sounding.

Bernard Crease graduated in physics at Durham University in 1973 and was awarded the M.Sc. degree in meteorology at Imperial College, London in 1974. He joined the Meteorological Office at Bracknell in 1974, being subsequently transferred to the Meteorological Research Unit, Cardington, where he was associated with a project to investigate the feasibility of using an acoustic sounder for quantitative boundary layer studies. He is at present on two-years leave of absence teaching physics at Harlington Upper School in Bedfordshire.

Jim Caughey graduated in physics in 1962 at the Queen's University of Belfast and was awarded the Ph.D. degree at the same university in 1970 for his work on the generation and structure of picosecond optical pulses. He joined the Meteorological Office in 1970 as a Scientific Officer and began research on the structure of turbulence in the atmospheric boundary layer at the Meteorological Research Unit, RAF Cardington; he was promoted to Principal Scientific Officer in charge of the Meteorological Research Unit in 1974. Dr Caughey was awarded the L. G. Groves Memorial prize by the Meteorological Office in 1977 for contributions to the understanding of boundary layer processes. In 1978 he was transferred to the Meteorological Office Headquarters at Bracknell and began work on the turbulent and microphysical structure of clouds.

Demosthenes Asimakopoulos graduated in electronic engineering from the College of Advanced Electronics in Athens in 1969. He was awarded the Diploma in Microwaves and Quantum Electronics of the Department of Electronic and Electrical Engineering, University College London in 1973 and the Ph.D. degree in the same department in 1976 for his work on quantitative atmospheric acoustic sounding, He held a Research Fellowship at University College in 1977 and was appointed a Lecturer in the Department of Meteorology, University of Athens in 1978. His research interests include the use of the atmospheric acoustic sounder in boundary layer studies and the design and development of meteorological instrumentation. In the summer of 1980 he held a temporary appointment as Research Associate with CIRES (Co-operative Institute for Research in Environmental Sciences), University of Colorado, Boulder, to investigate low-level atmospheric acoustic sounding. He is currently a Reader in Applied Meteorology at the University of Athens and an Associate Professor at the Greek Air Force College for Officers in Athens.

UDC 62 -
Indexing Terms: Systems science, Operations research, Production control, Terotechnology

Systems Engineering: an approach to whole-system design

P. K. M'PHERSON, S.M., M.A., C.Eng., F.I.Mech.E., F.I.Mar.E., F.Inst.M.C.*

SUMMARY

Since World War II the magnitude and complexity of technological enterprise has increased dramatically. So too have the costs of development, mistakes and failures. The old-style project engineer who could often be his own designer and project manager has been replaced by a team—the systems engineering staff—who integrate, co-ordinate and evaluate the realization of a project through all its evolutionary phases. In particular, systems engineering has produced an approach to design and project appraisal that integrates and balances the various technical, economic, reliability, safety, logistic, support criteria oriented towards future market and operational requirements evaluated over the whole-life of the system.

This introductory paper provides an entry to the systems concepts that underlie systems engineering, and then illuminates the practice through the perspectives of systems organization, design and planning.

* *Department of Systems Science, The City University, Northampton Square, London EC1V 0HB.*

TWO INTRODUCTORY QUESTIONS

Who holds an Engineering Plant Together and Keeps it Working?

The project engineer replies that it is the proper design of the functional parts, the structures in which they reside, the networks connecting the parts. The control and instrument engineer then claims his pre-eminence by saying that without proper instrumentation, information networks and control systems, the engineering plant—however well designed in itself—would be unable to operate and satisfy the required demand reliably and safely, let alone start-up or shut-down. The mention of reliability brings in the maintenance engineer who points out with considerable justification that without his team the plant would hardly hold together and would certainly stop working. Which cues in the logistics engineer with his argument that nothing could function or stay operational without the regular provision of feed-stock and energy, spare parts and provisions. By now the personnel manager can keep quiet no longer: it is not the machines but the men who matter—designers, engineers and operators. The management team then states its obvious claim to being the correct answer to the question, but the sales team cuts in—unless they correctly identify the market and then successfully sell the product everybody else starves. The reminder of the fundamental need to produce a marketable product that produces a profit infuriates the engineers. But there is a peacemaker present. The systems engineer calms everybody by saying that they—designer, engineer, operator, manager, salesman—are all important to the plant. The plant is a man–machine system requiring the proper co-ordination and optimization of their several interdependent contributions throughout the plant's life-cycle. It is his—the systems engineer's—privilege to help them by harmonizing their conflicting viewpoints at the design stage so that the plant not only satisfies the cost and performance requirements during its commissioning trials but is also *cost-effective with respect to the customer's objectives during the whole of its life.*

What is a Cost-effective Plant?

The squabble starts again. The project engineer stresses high operational efficiency for a given capital cost, the control engineer goes on about optimal load-following to minimize wastage, the maintenance engineer argues for high reliability supported by excellent maintenance to minimize downtime, the logistics engineer mentions the need for adequate provisioning to minimize the risk of stock-out. The personnel manager talks about training standards and on-the-job learning to maximize human potential, the managers underline the importance of good financial control and asset management to ensure a high rate-of-return on the capital invested, and the

The Radio and Electronic Engineer, Vol. 50, No. 11/12, pp. 545–558, November/December 1980 0033-7722/80/110545+14 £1.50/0

salesman points out the need for a flexible product and a sensitive pricing policy to meet changing market demands. They all glare at the systems engineer. Again he prefers no particular viewpoint. The true cost of a plant (he says) is its whole-life cost, and that includes the capital cost to produce a plant of a certain performance, all operating and maintenance costs from commissioning to retirement, the unreliability costs and penalties due to failures and spurious shut-downs, the marketing and field costs, the cost consequent on bad man–machine interfacing and management (maloperation, low job satisfaction, poor matching to operational context), the cost of insuring against potential risks and hazards. An effective plant (the systems engineer continues) has a high probability of being, for all foreseeable operational environments:

(i) operationally ready at any time;
(ii) able to complete any specific operational task;
(iii) able to meet the sales and performance requirements at any operational moment.

These three probabilities are called availability, dependability and capability; their product is effectiveness (as defined in terms of the customer's objectives). The optimal plant design is not necessarily the one with the best performance, or the higher reliability, or the lowest operating cost, or the largest safety margins, or the one that the manufacturer would most like to sell. The optimal plant design is that one which promises the best *overall trade-off between whole-life effectiveness and cost under the conditions of customer's objectives and style of operation*. The system engineer's role is to integrate the various design, engineering, operational and support tasks so that the cost-effectiveness concept can be achieved, monitored and maintained throughout the plant's life.

THE ORIGINS OF SYSTEMS ENGINEERING

An early (1950s) advertisement from RCA reads as follows:

> *The Systems Engineer at RCA*
>
> Systems Engineers: conduct studies to determine operational requirements . . .
> create and synthesize military equipment concepts . . .
> guide development of new integral elements . . .
> conduct evaluation programs to determine operational effectiveness.[1]

Twenty-five years later Ferranti advertised (1979) for a systems engineer using this job-description:

> We require engineers to work in a team which has responsibility for systems design work involving initial definition, interface specifications, cabling, alignment, setting-to-work, integration of hardware and software, trials and performance evaluation. The Weapons System Engineer is involved throughout the life of a project, from initial design to post acceptance.

These two descriptions of a system engineer's activities underline the electronic and military soil from which systems engineering grew. The need for something more than conventional compartmented engineering and management methods for pulling a project together grew urgently during the 1940s due to the advent of really large-scale communications and military systems: it was the Bell Telephone Laboratories who coined the term 'systems engineering' during the early 1940s.[2] In one of the first books on systems engineering (vintage 1962 and still one of the best) Hall[3] high-lighted the following evolutionary forces that led to the concept of systems engineering as a separate function in the organization of technological activity, namely complexity, expanding needs and environment, and shortage of technological skilled manpower. These evolutionary tendencies are amplified as follows:

Complexity

> Growth of large-scale systems with many interconnecting components both human and machine—from different branches and specializations of technology.
>
> The often wide geographical spread of such systems (e.g. interlinked manufacturing complexes, power distribution, transport, communications).
>
> The consequent complexity of design, planning and management problems posed by the integration, implementation and effective operation of such systems.

Expanding Needs and Environment

> Such large-scale systems have marked interactions with the environment into which they are inserted—the environment being the host society, economy and geographical territory. Indeed such systems are often used to 'tie together modern society'.
>
> The expanding needs of modern complex society within the usual economic constraints place a premium on defining long-term goals and strategies so that the 'best option' (from the rapidly developing technological cornucopia) could be selected for development. Large-scale high-technology is too big and expensive to experiment with.

Shortage of Technically Skilled Manpower

> The need to deploy optimally the limited R & D effort available raises the need to study carefully the research strategy first, leading to a deployment of effort from hardware development to policy analysis.
>
> The need to reduce manpower wastage on costly late abortions of development projects means that high quality scientific skills must be deflected to the design, scheduling and management of development programmes.
>
> The very difficult decisions required from the parent organizations produce the need for high quality and comprehensive analysis and evaluations to help senior management.

It is difficult to draw a strict boundary between the concept of systems engineering that emerges from Hall's 1962 viewpoint and the related concept of Systems Analysis that was developed as a problem-solving and decision-aiding methodology by the RAND Corporation (founded in 1946 by the USAF[4]). In fact systems analysis

Table 1: Development of systems methodologies

Initiating need	Methodologies		Emerging disciplines
	Phase I (1940–60)	Phase II (1960–75)	
Tactical and short-term resource management	Operations Research	Operations Research Operations Management Decision Analysis	Management Science
Strategical and long-term resource management	Systems Analysis	Applied Systems Analysis Regional Planning Planning, Programming Budgeting System	Industrial Engineering
Integrated design and management of complex man–machine systems	Systems Engineering	Logistics Engineering Systems Engineering Systems Design Systems Architecture	Systems Engineering
Design of radar and weapons systems	Servomechanism Theory	Control Engineering Control Theory Systems Theory	Systems Theory

(not to be confused with Business Systems Analysis or Computer Systems Analysis) has settled down as a comprehensive methodology for first surveying complex multi-objective, long time-horizon problems to which there are many alternative feasible strategical solutions, and then evaluating the alternatives to indicate the 'best' one. It is a kind of large-scale search and decision analysis and, as such, forms the initial phase of the systems engineering methodology—which goes on to evolve an optimal design for the whole system implied by that best alternative.

Both systems engineering and systems analysis are members of the brood of problem-solving methodologies that emerged from World War II under the collective heading of the Systems Approach. Table 1 summarizes the development of the initial systems methodologies of OR, Systems Analysis, Systems Engineering and Control Engineering.

The many elements in Phase II of the development of systems methodology have been regrouping, since about 1970, under the headings of new disciplines, e.g. Management Science, Industrial Engineering, Systems Engineering and Mathematical Systems Theory. Within each discipline the Phase II methodologies merge, adapt or develop according to the character and needs of the discipline's primary fields of interest. Another merger is also in progress as systems engineering is no longer limited strictly to the field of high-technology systems. It is absorbing the meaning of Applied Systems Analysis such that systems engineering is now often understood (particularly in the USA) to refer not only to the design and evaluation of cost-effective technological systems, but also to the resolution of conflicts that arise due to the impacts of technology on society and—thence—to the assessment and selection of systems that are compatible with their external social and natural environments as well as being well-designed within themselves.[5,6]

Systems Engineering as Good Engineering

Both systems engineering and industrial engineering are enthusiastic American 'think big' approaches that integrate the appraisal, planning, design, production and testing, of every aspect of a technological enterprise from conception to commission. The aim is not just to get the new system delivered and working to specification within the time and budget constraints of the schedule, nor is the accent of design on providing, within a price, a configuration that meets a stated performance specification. By being 'whole-system' and 'whole-life' oriented, the design includes reliability, logistics, human and environmental impact factors as well as marketing and operational requirements; the objective of the design process is to obtain trade-offs between all the conflicting factors such that future ownership costs over the planned life of the system are minimized. And ownership costs are not just capital and operational costs, they include all the support costs, the costs of not meeting customers' or users' objectives adequately, and the potential costs of undesirable side-effects via social and environmental impacts.

All this sounds like nothing more than good engineering applied at *project management* level.[7] Of

course, good systems engineering implies good chemical engineering, civil engineering, communications engineering, electrical engineering, naval architecture, power engineering, weapons engineering, etc., as appropriate (e.g. Refs. 8, 9, 10). But it must be admitted that the reverse is not always the case as good engineering is too often understood to imply only clever solutions to tricky design problems at the device or performance levels: reliability, logistics, human factors are something else. The conventional education and training of engineers tends to make them device and function oriented, which is then compounded by the professional designer's zeal to produce the best performing machine (or whatever) at the prescribed cost. The designer's performance horizon is often limited to the specification and the acceptance trials, whereas good systems engineering requires that all the factors in Table 2 be taken into account during the early appraisal and design stages.

These factors, taken together, contribute towards the whole-system concept of the system engineer. Such a system is composed of a closely-interwoven organization of personnel, equipment and operations; its design must be oriented towards the optimization of post-launch operations; the interdisciplinary nature of the exercise requires careful attention to the interfaces between different equipment, between different stages of development, between different subsystems within the overall concept and between the system, its future owner/user and its environment. Finally, the planning and management of the interrelated tasks within the overall project plan is a major task in itself. Such systems certainly warrant the epithet 'complex'. The factors that combine to produce complexity in contemporary organizations are legion. They may be quite unimportant in themselves, but, through the gearing up that often results from the interactive loops formed with other factors, they can lead to consequences quite out of proportion with the initial trigger. The list of factors in Table 3 does not claim to be comprehensive, but it does serve to show their range diversity, and potential interactions. All of these contribute to the meaning of the word 'complexity' in the context of systems. The engineering of such systems is considerably more than just 'good engineering': Jenkins[11] captures the overall problem as 'the science of designing complex systems in their totality'.

Table 2: Factors in the Whole-System concept

A MAN–MACHINE SYSTEM IMPLIES	
BEFORE LAUNCH	AFTER LAUNCH
Long-range planning	Marketing
Market research	Operations
Appraisal	Competition
R & D	Environmental impacts
Design	Waste products
Human factors	Monitoring
Specification	Information processing
Proving	Maintenance
Tooling	Logistics
Production	Facilities
Quality control	Risks
Testing	Insurance
Despatching	Debt servicing
Installation	Training
Acceptance	Recruiting
Commissioning	Posting
Documenting	Operations Management
Recruiting	Support Management
Training	Personnel Management
Planning	Financial Management
Financing	Modification
Project Management	Retirement
WHICH AGGREGATE TO THE WHOLE-SYSTEM AND WHOLE-LIFE COSTS	

Systems Engineering in the UK

Unlike OR, systems engineering in its large sense has not obtained very widespread recognition in the UK as a discipline in its own right outside the original stamping ground of military, naval and aerospace systems. A more specialized form of systems engineering flourishes for avionics, telecommunications and weapons systems where the underlying discipline may be said to be an amalgam of computers, control, electronic and information engineering. The large-scale high-technology industrial fields such as the chemical engineering, nuclear engineering and electrical power industries tend to interpret the whole-system concept as a project management problem. From there, spreading further through the technological spectrum, the concept is either degraded and confused with control engineering or disappears completely. But the need for systems engineering's whole-system whole-life concepts exists just the same, and it keeps on reappearing under various guises, e.g. Systems Assurance, Terotechnology. The latter name is the Department of Industry's 1970 invention being one of the multidisciplinary industrial technologies warranting special attention.† A look at the definition given to Terotechnology suggests that it is a simplified cross between Systems Engineering and Industrial Engineering. The accompanying stream of booklets confirms this correlation as—so far—there is no content that does not also belong to OR/SE/IE.[12,13] Terotechnology is a plain English non-mathematical version of SE/IE, and it may be that this is what is needed to get the concept across to the great majority of good but non-systems engineers and managers, for it has to be admitted that SE, in its full American flowering, is high-pressure, high-jargon and high-mathematical stuff (e.g. Refs. 14, 15, 16, 17).

† The other three are Tribology, Corrosion, Materials Handling.

FUNDAMENTAL CONCEPTS OF SYSTEMS ENGINEERING

Systems

The specialized concept of a system has developed over the last 30 years to deal with the problems that arise from the complex structure and behaviour of modern human activity and technological systems. In this sense a system is usually defined as a complex organization of men, equipment, resources and operations through whose integrated functions an operational need is satisfied. This definition barely conveys the intricacy of the system concept with its hierarchical levels and lateral functions that change and evolve in time—all nesting in enclosing system shells. Today 'system' has become a vogue word on the tip of lazy tongues when, e.g., 'network', 'organization', 'method', 'servo' could provide a more focused meaning. So it is necessary to know exactly what we mean when we talk about a 'system'.

Systems are observable but not directly: whether an aggregate of related objects is a system or merely a conglomerate depends on the results of tests to see if the

Table 3: Factors inducing problems of complexification

TECHNOLOGICAL
Modularization; miniaturization; interlinking instrumentation for monitoring, control, safety; automation; innovation bringing higher technology and greater specialization; equipment sophistication compensating for lower human skills.
Integrated high-performance, high-sophistication, minimum-maintenance packages; complicated matching and interfacing between specialized packages and processes; interdependence of multiple processes introducing involved dynamic behaviour.
Multiple fault patterns through complex linkages between units and processes; parasitic failure patterns induced through complex instrumentation; reliability design developed to a high level; higher diagnostic skills required.
Interdependence of performance, reliability, maintainability, logistics, information and human factors in integrated design; need for design for total integration of all functions.
HUMAN
Man–machine interfacing; matching of available skills to advancing technology; conflicts induced through automation; job and responsibility degradation resulting from automation and computer-centralized management.
Shorter half-life of given knowledge and skill levels; requirement for greater human understanding and flexibility; more training to higher skills.
Changing social expectations; higher standards of work environment and interest; resistance to innovation and change; trade union involvement in decision-making and labour law.
EMBEDDING
Need to match the process/system to its real operational requirement and environment, developing appropriate technological infrastructure.
Increasing need to bring social and environmental side-effects into major design arena: wastes, pollution, health and safety risks, potential hazards; impingement on ecological, aesthetic and cultural factors.
COSTS
Ownership costs to be analysed over whole-life and all aspects, e.g. design and acquisition, trials and launch, preparation and marketing, operations, maintenance and logistic support, modification cycles, insurance against shut-down, risks and loss, quasi-costs of potential environmental and social damage, monitoring and quality control, information network, management.
Complexity of multiple economic trade-offs between various heads (difficulty of quantifying environmental and ethical factors); costs to be discounted over long time-horizons (problems over validity of discounting non-commercial and risk factors).
COMMERCIAL
Designing for high survivability against competition and threats, flexibility for shifting patterns of resource use, trade, geopolitics and public taste.
Higher selectivity and expectation of customers demanding sophistication with reliability and low cost; need to match product design closely to customer's real needs, skills and objectives; customers requiring proof of whole-life cost-effectiveness, reliability and safety.
Emergence of technology transfer market with developing nations (sophisticated shoppers but low technological/industrial/managerial back-up); turnkey contracts for whole systems and industrial complexes involving design, installation, operations support, technician and management training.
MANAGEMENT PROBLEMS
Complexity of systems management, conflict between hierarchical and lateral functions, different systems-oriented outlook and management skills; training and motivation of personnel in complex and dynamically changing organizations; attention to departmental interfaces, project management, customer liaison.
Long lead time of innovation and development requires comprehensive long-range planning, technological and market forecasting; extensive survey and analysis to ensure that 'right' options are selected.
Difficulty of planning for overall design, acquisition, marketing and operation of highly integrated multi-function processes; difficulty of evaluation and selection with multiple alternatives and multiple objectives.
Computer-aided management and design techniques, computer-based data processing and information systems; more information processing power but often less organizational flexibility; conflict between centralized and decentralized management.

aggregate has certain properties of structure, behaviour, purpose and cognition:

- the structure must consist of many elements that are connected through many interrelationships—which implies the existence of loops and complexity;
- the organization's behaviour must be dynamic and of such a character that it is not easily inferred from the reaction to their inputs of the individual elements in the organization;
- the behaviour should be purposeful and directed towards the achievement of some goals.

Such are commonly accepted properties of 'systems', but *successful* systems, i.e. ones that survive, improve and achieve their objectives, require additional attributes:

- the structure is compartmented into operational, directional and supporting subsystems interlinked by an information subsystem;
- the structure as a whole must be at least reliable and resilient, while—for longer-term survival—the structure may also have to be adaptive or capable of reorganization;
- at this level a system would have to be intelligent and normative, i.e. it is aware of deficiencies within itself, and defines better structures or goals.

The Proto-System: A Fundamental System Structure

Study of biological and human systems suggests that a goal-oriented (purposeful) system that also survives for a useful life-time must have an essential structure similar to that of Fig. 1. It is dubbed the Proto-System as it is a template or prototype for all organizations that are to qualify as surviving purposeful systems.

The proto-system has four essential subsystems:

Operational System ($P_1, P_2, \ldots, P_j, \ldots$; IP; IC; OC)

The functional processes that produce the useful output ($P_1, P_2, \ldots, P_j, \ldots$). This output is often partially consumed within the system (i.e. intermediate products) before the final output is discharged to the operational arena. The Information Processes IP, Information Control IC and Operations Control OC are integral parts of the Operational System.

Management System (PD, CC; IC)

To direct and co-ordinate the functional processes (CC—co-ordinational control) according to the received policy/strategy whereby the system is to achieve its objectives (PD—policy development). The management system needs to be linked to the information system (IC—information control).

Information System (IP, IC)

The central core of information processes that collects information about the state of the functional processes and of the environment for processing and transmission to the appropriate level of the management and control hierarchy.

Now any operational system that supplies an action, product or service (a) consumes resources and energy that will have to be replaced, and (b) will deteriorate due to wear, tear or failure. No system on earth can escape the second law of thermodynamics. Consequently a *surviving* system needs another subsystem to counteract resource consumption and performance degradation:

Support System (L, M; IP, IC; SC)

L is the logistic subsystem processing resources from the environment, and M is the maintenance process through which all operational units have to cycle periodically for repair, maintenance and restoration (to aid clarity Fig. 1 shows resource flow and maintenance cycles for the functional and information processes P2, P1 only. All other processes, including management, should also be connected to the support system).

The basic structure of the proto-system displays all the descriptive systemic properties (elements, links, loops, levels, clusters, boundaries, hierarchy, functions, transmissions and dependencies) which have now been organized to provide a model of any operational and surviving system. The model will be very familiar to systems engineers and biodynamicists. It is an open system with five input/output pathways across the system boundary:

- Resource inputs and outputs.
- Wastes and undesirable outputs discharging to the environment.
- Information inputs and transmissions.

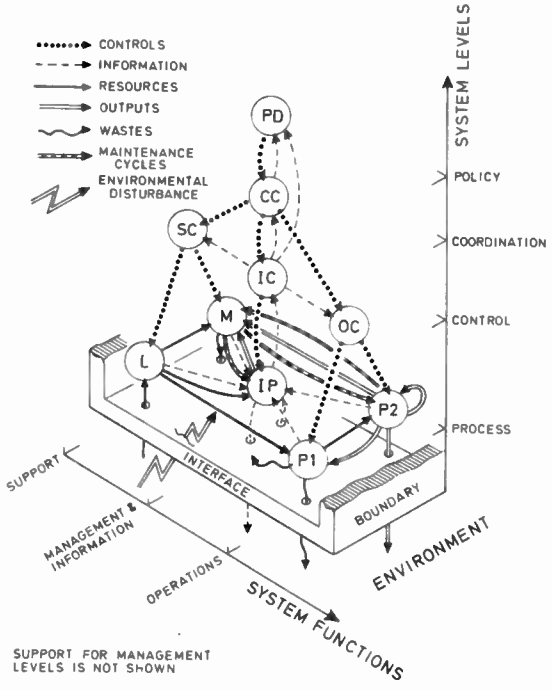


Fig. 1. Structure for the Proto-System.

- Perceptions, being subjective information inputs to the human managers and operators that may influence the way that they react to the processed information that they receive.
- Environmental buffetting, e.g., noise in the information inputs, environmental changes to which the system is sensitive, disruptive forces that damage or destroy the system structure.

As a model it is an outline representation of, for example:

- a living biological system from body function up to brain;
- a technological system with operational and support echelons co-ordinated and directed by a central management system;
- an economic system with resource, service and productive sectors, each with its sectoral control under central (government) direction.

If all this seems rather large scale, take the model down to the scale of a single humble servo. Now P_1, P_2 represent the process under control and (IP, IC, OC) is the feedback measurement and control chain; L provides the power supplies and spare parts, M is the maintenance facility, while SC is the spare parts inventory and maintenance schedule. CC is the integrated design process for the concept PD.

The proto-system really is a general model for purposeful surviving systems of all kinds. It tells the systems' investigator what to look for in detail once the systemic scanning filters have recognized the possibility of a system organization, and it gives the systems' designer a general plan for assembling successful systems. In particular, the system—whether biological, technological or social—will break down if the system structure is of finite robustness and the support system is inadequate and unable to meet the operational stress. Surviving systems must be robust and reliable systems.

The Whole-System Concept

For systems engineering, the 'system' is a complex organization of men, equipment, resources and operations through whose integrated functions an operational need is satisfied. The main components of the systems engineering concept, with their interrelations, are shown in Fig. 2. The basic building block is the man–machine system with its four Ms: men, money, materials and machines.

Man–Machine System

While it is easy to discern that technological aggregates, e.g. a ship, a chemical plant, a telephone network, are man–machine systems, it is often difficult to remember that quite small technological units are also man–machine systems. For example, an accelerometer is clearly a rather small piece of equipment, but it needs resources for its construction and operation, its *sole* purpose is to provide a useful function in aid of some operational activity, and it needs people to design it, manufacture it, sell it, buy it, operate it and maintain it. The mere existence of the accelerometer entails resources, operations and people.

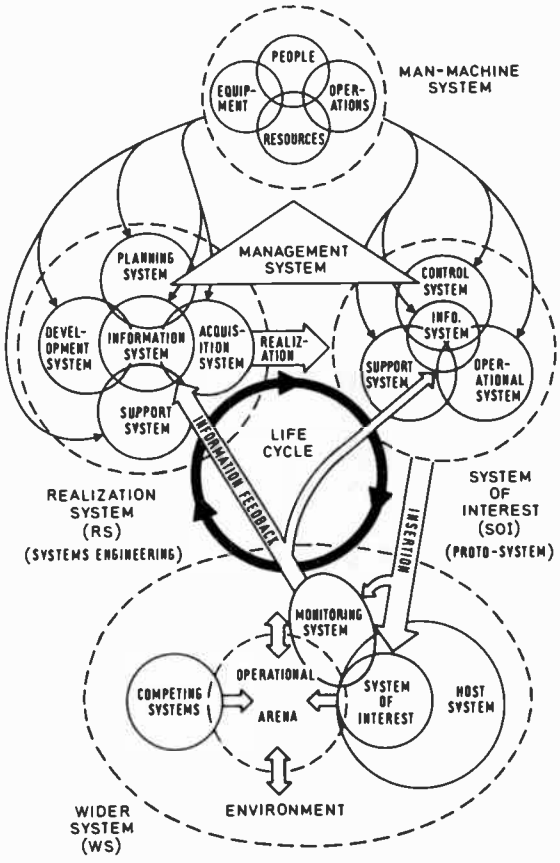


Fig. 2. The Whole System concept of systems engineering.

The man–machine system building block provides the elements from which the major systems in the concept are structured. These are in turn:

System-of-Interest (SOI) (or Proto-System (PS))

The system-of-interest is the system that is to be engineered (or modified) whose operations satisfy the owner's purpose. It could be that accelerometer again, or the equipment system in which it is housed, or the plant, or the parent organization depending on the level of scrutiny. All purposeful systems require four essential service systems if they are to operate correctly, fulfil their purpose and survive: they are:

(1) the Operational System—to provide the useful output;
(2) the Control System—to regulate the operational system and the other service systems;
(3) the Information System—to provide feedback to the control system from the service systems, and to monitor the state of the environment to the SOI;
(4) the Support System—to supply and maintain the other service systems.

Realization System (RS) (or Systems Engineering (SE))

If the SOI is a new system it has to be brought into existence, i.e. REALIZED from an initial concept which is Systems Engineering. The structure and methodology of SE will be studied in more detail later: for the moment the RS is

compartmented into:

(1) the Planning System—to survey the problem, produce possible solutions, analyse and evaluate the alternatives; and develop the project plan;
(2) the Development System—to design, develop and engineer the selected SOI concept;
(3) the Acquisition System–to manufacture, construct, test, commission and launch the SOI, thus converting the SOI from a design concept into an operational system inserted into the real world;
(4) the Information System } to provide these necessary
(5) the Support System } functions in the RS.

If the SOI is an existing system the RS may be considered to be the instrument whereby operational experience is analysed so that modifications or improvements can be incorporated in the SOI (thus renewing it).

Management System (MS)

The management system represents the policy making, co-ordinating, administrative, marketing and financial aspects of the organization that requires the SOI. It may, in fact, be two distinctly different management systems: that of the producing organization and that of the owning (i.e. operating) organization. The interface between these two management systems represents producer–customer liaison.

The SOI after commissioning and acceptance trials is 'launched' into its operational role, the commissioning–acceptance–launching sequence being referred to as INSERTION into the:

Wider System (WS)

which is compartmented into:

(1) The Host System—the parent system in which the SOI nests, or the parent organization that owns the SOI, or the geographical region (town, ocean, nation, etc.) in which the SOI is placed.
(2) The Operational Arena—being that part of the environment that is directly affected by the SOI because of its existence, its resource requirements and its output (operations).
(3) The Competing Systems—there may be other systems competing for the resources/market/territory/power.
(4) The Environment—the economic, social and natural environment that is external to the operational area, but is affected in one way or another, directly or indirectly, by the SOI through its operations.
(5) The Monitoring System—the information retrieval and feedback system. It is composed of two parts:
    (i) that specifically designed and inserted to provide information feedback to the management, designers and operational controllers;
    (ii) the general information and data gathering instruments (e.g. journals, economic forecasts, press, espionage, grapevine, etc.) that contribute to the available information (or misinformation) on the Wider System.

The aggregate of all the systems and concepts in Fig. 2 provides the Whole-System concept for systems engineering. The Whole-System is built up from:

(i) the SOI nesting in the various system layers of the Wider System;
(ii) the Realization process to provide the SOI in the first place;
(iii) the Management of the realization and insertion processes;
(iv) the planned Monitoring of operations to provide information for improvement of the SOI-type.

The systems engineering team is both the realizer (creator) of the SOI and the manager of that realization. Which is a flowery way of saying that SE provides an integrated framework for capital project appraisal and selection, whole-system realization with respect to whole-life criteria, and whole-project management. Systems engineering is a style of engineering.

System Life-Cycle

If one looks at Fig. 2 again a cyclic pattern may be detected by paying less attention to the content of the internal systems RS, SOI and WS and focusing more on the sequence of events implied in the Figure. Systems exist in time. For any one system there is a point of time prior to which its particular concept did not exist, and there is another point of time beyond which it will not exist because it will have been scrapped, destroyed or evolved into another form. This is a birth–life–death–rebirth process that may be called the System Life-Cycle.

There is a self-evident and familiar sequence of phases during a project's development and operations. These are listed as the Phases in the life-cycle programme listed in Table 4. The phases may be grouped into the major Modes of a system's life, and each phase is itself a group of Activities. Table 4 is fairly standard with respect to the many other existing definitions of a life-cycle programme (e.g. Refs. 3, 11, 14, 15, 16, 17). The fundamental sequence of the life-cycle is shown in Fig. 3(a) in which the phases are arranged to form a loop. But the life-cycle is itself an amalgam of several cyclic processes:

- Design cycles spanning the conceptual and design phases.
- Development cycles spanning the production and design phases.
- Support cycles spanning the operational and support phases.
- Information cycles from the operational and support phases back to design causing improvement in equipment performance or operational procedures.
- Renewal cycles from the operational and support phases back to design in which the information feedback is now used to modify the existing system or to initiate the design of a successor system.

Some idea of the interlinking between those cycles is given in Fig. 3(b), in which the cycles become spirals stretched out in a time axis. The same time axis can also

Table 4: System life-cycle programme

THE SYSTEM LIFE CYCLE implies		
Modes	Phases	Activities
REALIZATION	Conceptual and preliminary design	System survey and analysis System synthesis Initial design studies Project plan Evaluation and sanction
	Design and development	Detailed design and development Equipment spec. System spec. Logistic spec.
ACQUISITION	Production	Acquire facilities and resources Manufacturing engineering Production and quality control Assemble and test Packaging Recruit and train personnel
	Insertion	Advance parties Distribution, site preparation Construction and trials Commission and hand-over
OPERATION		Activate and provision support
	Market and deployment	Deployment of operational units and personnel Marketing and sales
	Operational	Operate system Monitor performance
	Support	Maintain, repair, overhaul Provisioning, personnel support
RENOVATION		Evaluate operational information
	Renewal	Modify Design successor system
RETIREMENT	Phase out	Place in reserve, sell, scrap

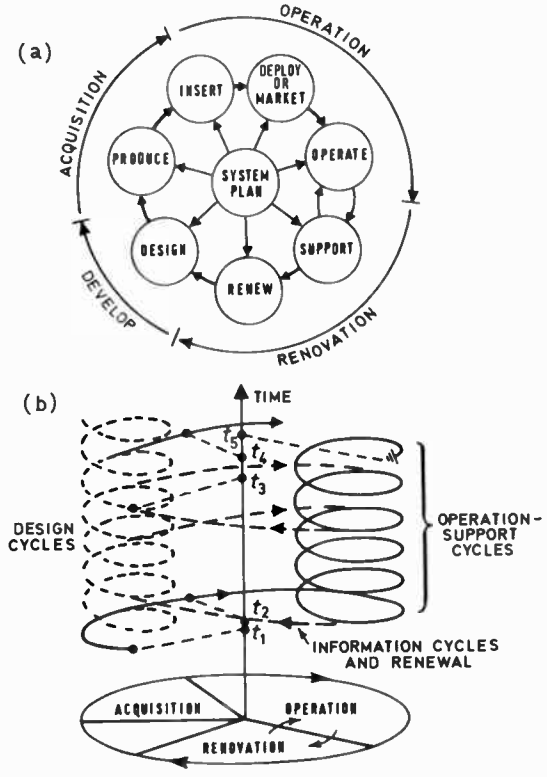


Fig. 3. Systems life cycle.

be thought of as an axis of technological advance because the system learns from the information feedback and incorporates the new knowledge to provide better performance. One complete life-cycle is shown in Fig. 3(b) starting at t_1. At t_2 the system is inserted and becomes operational, it then goes into a sequence of support cycles during which time it is twice advanced via information/renewal cycles. At t_5 it is phased out. The successor system starts its life at t_3, becoming operational at t_4. Part of the practice of systems engineering is to design an information system that captures, channels and utilizes operational data and experience so that these improving cycles can be exploited in the interest of better design and better operation.

The life-cycle is an important concept in systems engineering as it orients the design of the SOI towards effective operations throughout its useful life rather than just satisfying its performance specification at the start of its operational life. The extent of the useful life has to be defined at the design stage from predictions of the future operational effectiveness and costs when the system is in service—*which may extend to a time-horizon anything from 5 to 30 years ahead.* Thus systems engineering leads to the consideration of Whole-Life Costs as the basis on which to value the economic aspects of the SOI, rather than the initial cost to produce. Whole-Life costs include all possible costs entailed throughout the useful life of the SOI, including support, modification and retirement costs. The Whole-Life orientation in systems engineering should make the systems engineer the customer's friend as it is the customer who operates the system and who wants cost-effectiveness defined with respect to his—the customer's—objectives.

THREE PERSPECTIVES OF SYSTEMS ENGINEERING

One could say that the essence of systems engineering is in its emphasis on the whole-system and life-cycle concepts. These two concepts provide, in combination, a basis for the systems engineer to build a framework with which to tackle the problems of planning, designing and managing the whole-system realization problem. Some perspectives are presented in the next three Figures to indicate the scope and depth of systems engineering as process and discipline.

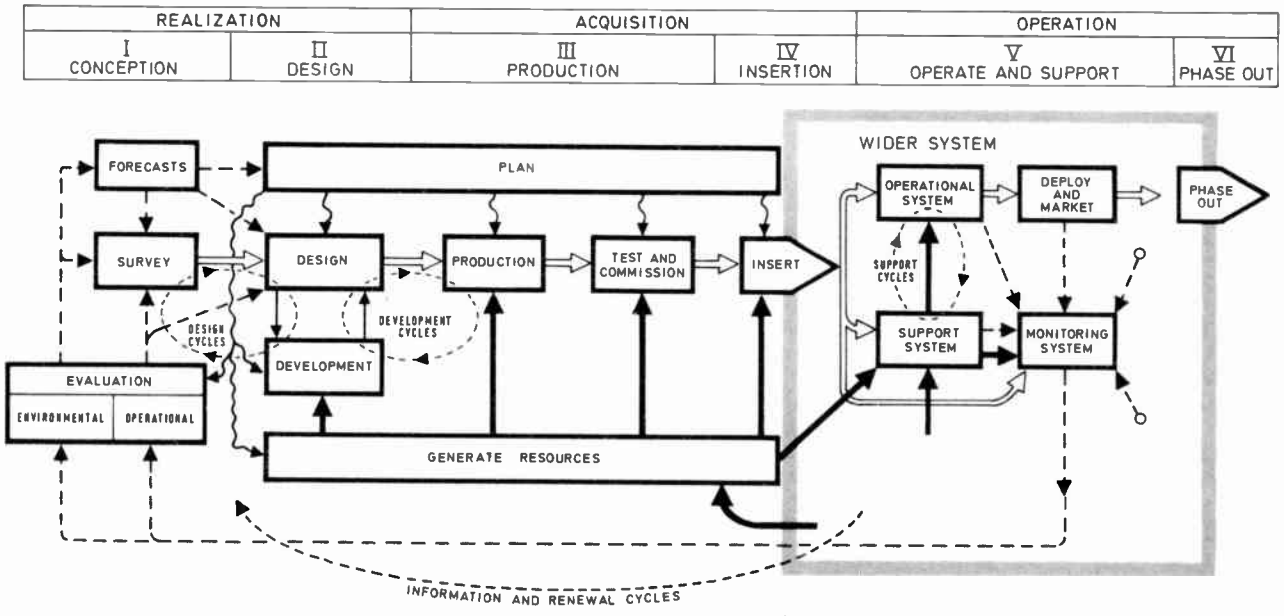


Fig. 4. Organizational perspective of systems engineering.

- *Organization Perspective*

Systems engineering is not just a large-scale planning process—it is an organic process that has to be conceived, organized and managed as a whole. Figure 4 brings the whole-system and life-cycle concepts (Figs. 2 and 3) together and rearranges them in the form of a process with resource inputs, operations, controls (the plan) and information feedbacks. The cycles indicate the main loop and feedback clusters in the process. Viewed in this light, systems engineering becomes a dynamic multivariable, multiloop system which will have an interesting dynamic behaviour that is likely to require careful tuning, co-ordinating and control if it is not to display signs of instability.[18,19] Projects can go wrong, go unstable or fail just as well as any operational process if they are not designed properly. Hence the tendency in systems engineering to see project management more as 'the design of design', with design, or any operational process for that matter, as a feedback system rather than an open-loop sequence.[20]

- *System Design Perspective*

No matter how well the project is planned and organized, the effort will be largely wasted if the operational system that results is the wrong system. A system can be wrong in many ways even though it may be excellently designed in itself. For example:

(1) Not meeting the customer's real objectives or the actual operational environment:

This often occurs because the objectives and operating conditions were poorly articulated in the first place. In these days of increasing complexity the customer may not be technically competent to specify exactly what he needs. In which case the onus is on the systems engineer to put himself genuinely in the customer's shoes and define the customer's problem for him.

(2) Not matched to the operational environment:

Operational environment different from or more extreme than that used for the design specification.

Too sophisticated for the user's managerial or operational skills.

Creates additional problems (interference, unreliability, muddle) due to unforeseen interactions with other systems.

(3) Badly timed:

Inserted just before a new technology will make it obsolescent.

Inserted into a competitive field before the parent organization has the technical strength to deal with the competition.

Misjudging the economic or market climate.

(4) Not matched to the social or natural environments:

Generates antagonism due to inattention to such matters as waste products, pollution, health risks, location, landscaping.

Causes cultural or political strain through the injection of a high rate of technological change in a society (e.g. a developing country) without adequate thought, or preparation, or matching development in the local educational system and technological infrastructures.

And, of course, the system may be wrong because it is badly designed with respect to any or many of the following criteria:

capability, availability, reliability, maintainability, supportability, survivability, marketability, operatability, ergonomic factors, safety, internal or external compatibility (i.e. poor interfacing), trade-off balance, life-cycle costs. No doubt many more could be listed.

The whole-system life-cycle approach of systems engineering endeavours to guard against such errors by

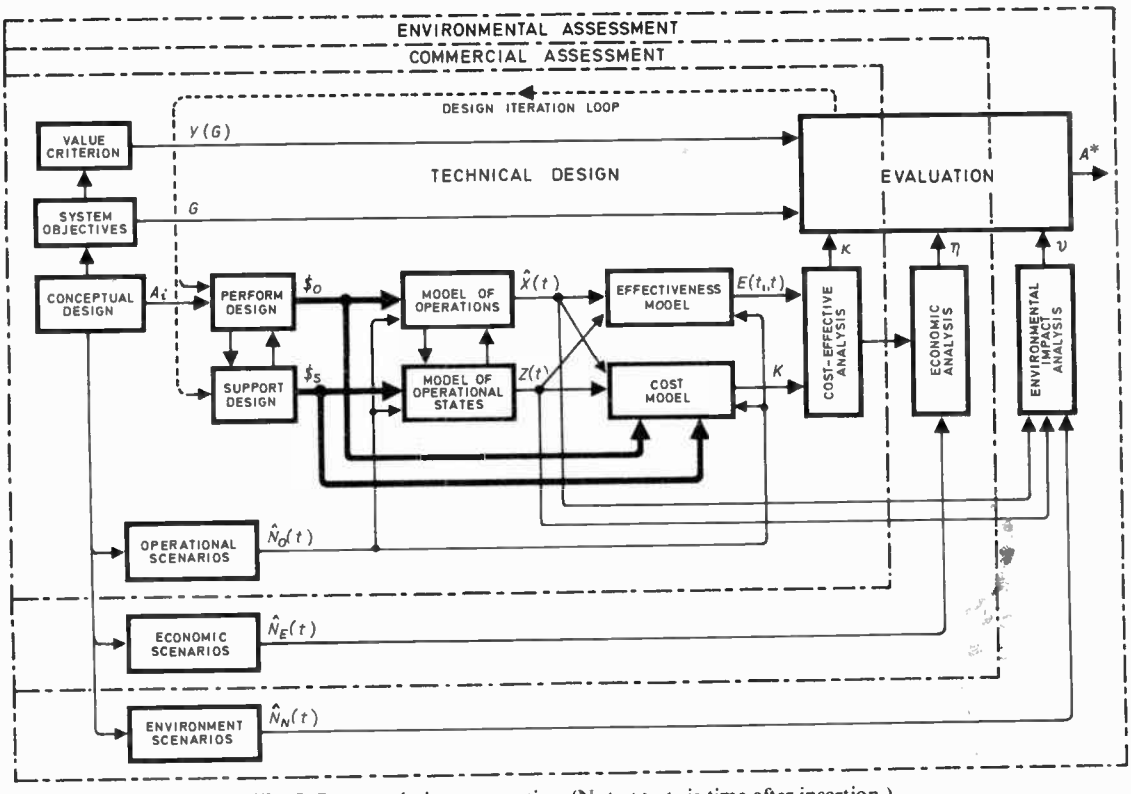


Fig. 5. Systems design perspective. (Note $t > t_1$ is time after insertion.)

bringing as many considerations and criteria as necessary into the design process. The result is Whole-System Design as illustrated in Fig. 5: it is a sketch only, the system design process is studied more deeply in a later paper. The object of the exercise is to develop a system concept A_i to its optimal configuration A*, or to select the optimal system A* from a set of alternative concepts {A}.

The word 'optimal' has four layers of meaning:

- Life-cycle optimality implying that the system is evaluated over its entire life, which may mean that the designer will have to indulge in the uncertain business of long range forecasting.
- Technical optimality in which the attributes κ refer only to technological performance and non-commercial costs, i.e. realization, operational and support costs. Typically, cost-effectiveness analysis deals with the search for technical optimality.
- Economic optimality where the attributes η might be return-on-investments, break-even time, market penetration.
- Environmental optimality. Here the attributes v are derived from assessment of the impacts of the system on the social and natural environment.

The attributes are the consequences (results, output) of three associated design processes that involve a cyclic or interactive search of the optimal system configuration. The processes—Technical Design, Commercial Assessment, Environmental Assessment—are self-evident from the figure. Each process requires an appropriate set of future scenarios with which to test the system's projected performance. System performance itself is analysed with the aid of three mathematical models:

- Model of operation to examine what performance results from a design specification in the context of prescribed operational scenarios, e.g. the accelerometer in different working environments, the ship on different routes under varying weather conditions, the telecommunications system in different locations and atmospheric conditions.
- Model of operational states to examine to what extent the support system (maintenance and logistic designs) maintains the operational system in a satisfactory operational state given the prevailing operational scenario and the operational history.
- Cost model to track and aggregate the likely future realization, operational and support costs, discounting them back to a reference start time.

The predictions from the models depend, of course, on the particular performance and support design configurations $\$_0$, $\$_s$ of the concept A_i that is being studied. The design loop is closed by adjusting the alterables in the configurations so that the evaluation of the projected system indicates increasing value.

Systems design pays great attention to the evaluation process as, no matter how clever the technical aspects of the design or how realistic the scenarios, what is

evaluated as a 'good' design will depend entirely on the initial statement of system objectives and the value criterion in use. The value criterion is an extension of the objectives. Assuming that the objectives are all oriented in a maximizing sense the value is increased the more the system meets the objectives. Thus if the objectives as stated are not comprehensive or are not matched to the real operational needs, the result of the optimization process will be a 'wrong' system as discussed earlier. More often than not there is a considerable conflict between the objectives. It is regrettable, but cheap technology is not usually efficient, reliable, safe or clean; economies of scale are enjoyed only from large-scale specialized operations; largeness of scale entails problems of concentration, e.g., complexity of construction and operation, intensity of capital investment, increased financial or operational risk. Consequently optimization in a system design sense really means harmonization of multiple conflicting objectives.

System Planning Perspective

The full implication of integrated system design is illustrated by Fig. 6. Here the SOI is the end result of a concerted design process covering six sectors: the operational plant and its associated equipment, the maintenance and logistics systems to support operations, the manpower system providing the required management and operational skills, and the management system for operational, support and financial control. All combine to produce a cost-effective system in the sense that the combined objectives of the company(s) producing the SOI and the customer using the SOI (or its products) are satisfied. This perspective is clearly the realm of project management. The project plan must schedule, co-ordinate and integrate the tasks in each sector of the SOI's development so that it may become an operational reality to the customer's satisfaction. For a major project, the project manager may have to organize and co-ordinate the effort of hundreds of individual decision-makers and designers in the parent organization and in the subcontractors. The project effort can entail thousands of man-years' work with an overall expenditure running into many millions of pounds spread over a space of several years. During all this time the project team has to be steered along the

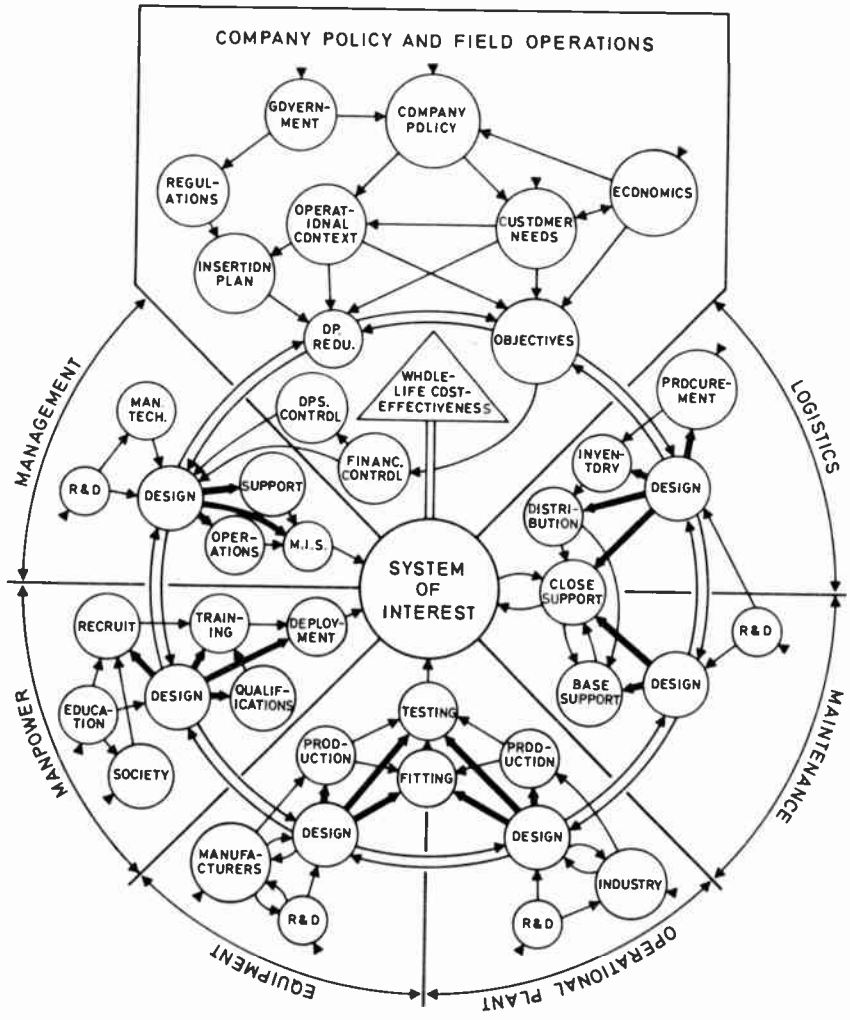


Fig. 6. Integrated system design.

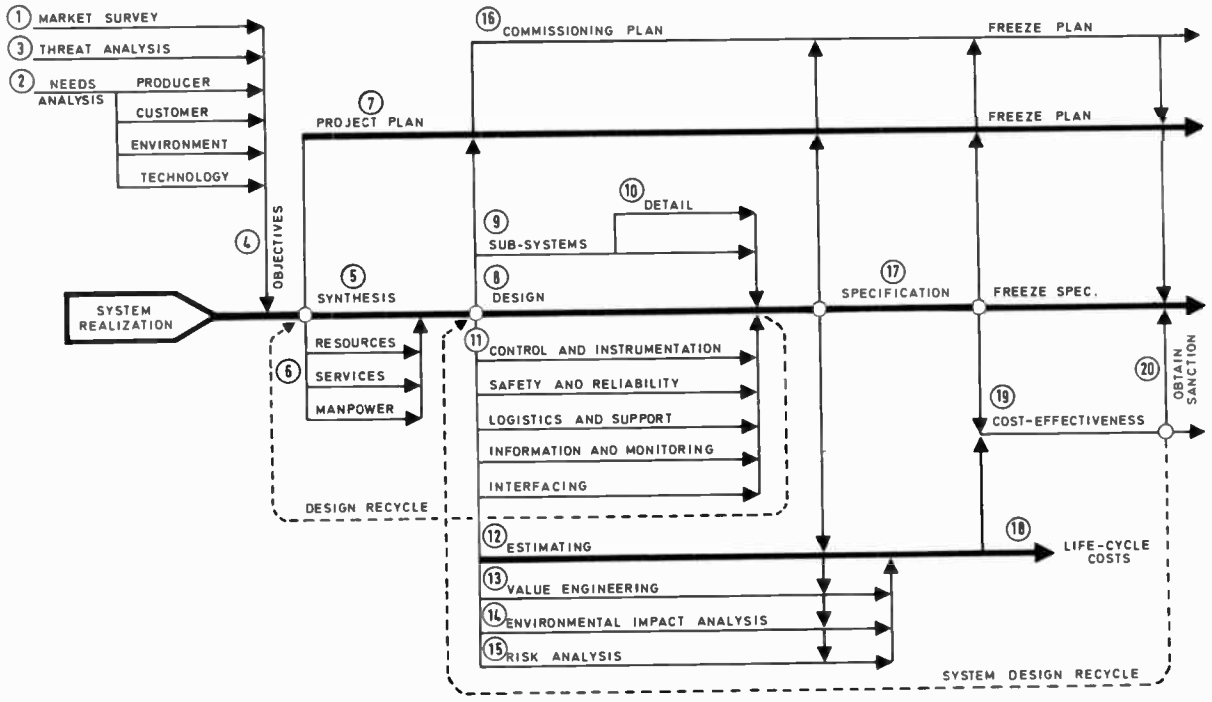


Fig. 7. Outline of system realization prior to sanction.

chosen route, unforeseen crises dealt with, and close liaison maintained with the expectant customer. Careful plans must be prepared for the handover, commissioning and insertion of the system into the customer's operational environment. A representation of a planning network for a project is given in Fig. 7: it shows the major problem/study tasks that any project manager is likely to encounter (in approximate sequence) as he drives the project through the preliminary conceptual phase that leads up to sanction for the development of a specific system.

The accent has been placed on the size of the management problem generated by large capital projects. But the same range of problems face the project managers or system designers responsible for smaller activities right down to subsystem level. For example, the design of the accelerometer subsystem for a stable inertial platform should pass through the route of the twenty study tasks outlined in Fig. 7, just as the project for developing the parent guidance system (in which the accelerometer is a small component) has to follow the same course. Indeed every design engineer ought to use Fig. 7 as a 'check list' and think in systems engineering terms.

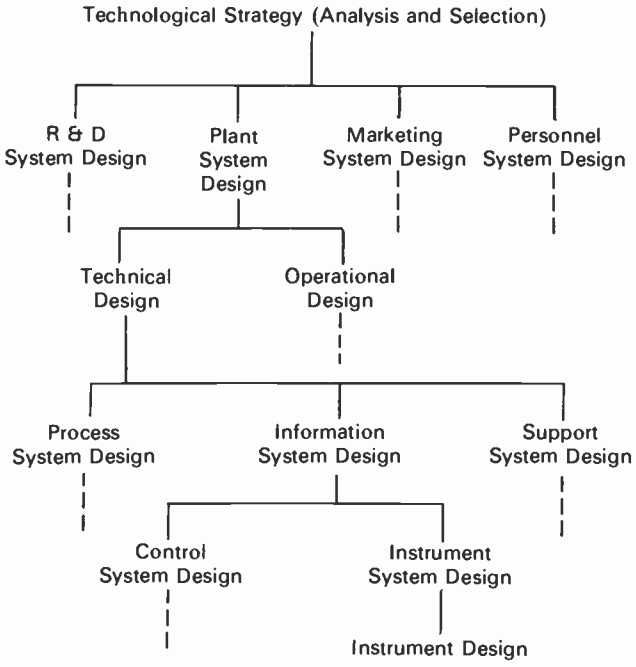
OVERVIEW

The broad concepts of systems engineering with their attention to whole-system, whole-life have been stressed so that the perspective of system design can be placed in context. All else follows from the realization phase of the systems engineering process (Table 4, Fig. 4) as it is during this phase that the problem is studied, alternative solutions to the problem considered and evaluated, leading to the selection of a particular alternative to go forward for detailed design, development and acquisition. In a very real sense the system is designed during the realization phase, as the system specification that emerges at the end of the phase has been obtained by developing an advance design of the system-as-a-whole which can be evaluated for whole-life performance and costs with the aid of mathematical models. So by system design is meant the overall process necessary to produce a system specification that has been selected on the basis of its predicted whole-life performance, and this implies that the subsidiary processes of performance design, reliability, maintenance and logistic designs have been carefully integrated within the overall design concept.

The systems design approach is not limited to the conceptualization of vast technological or industrial complexes. It represents good design practice that can be applied at many levels of largeness and complexity to any man–machine organization that qualifies as a system. There is a hierarchy of system design levels ranging from the analysis and selection of an overall technological strategy down to the design of finite technical systems. Some members of such a hierarchy are shown in Table 5.

A 'good' system is not one that has been delivered on time at the quoted price and has passed its acceptance trials: that implies only that the project management has been good, i.e. the performance specification and the

Table 5: Hierarchy of systems design levels



production schedule have been met, although that is often difficult enough. The potential goodness of the system is only realized after it begins to operate and is observed by experience to satisfy the following operational criteria:

- operational matching (timeliness, captures the desired operational niches, matched to the owner/operator needs and skills, and marketability);
- whole-life cost-effectiveness (with respect to the owner and the future operational contexts);
- resilience (can meet potential threats, competition, breakdowns and survive through flexibility and adaptability);
- environment and safety (undesirable outputs and risks kept to a minimum);
- embedding (tuned to the technical, social and environmental contexts of the host and wider systems into which the operating system is launched).

The mistakes and faults in a system (apart from running-in defects) emerge with time after the point of sale and after it is operating. Thus the design of good systems means attention to the proper interfacing, integration and co-ordination of all the technical, marketing, operational and social functions of the system so that an appropriate balance of advantage and disadvantage over the many conflicting objectives involved may be realized over the system's life-time.

The accent is on design for wholeness: whole system, whole life. This adds many more dimensions to the already difficult problem of design for performance. Systems engineering has developed to meet these enlargements of the designer's task. The next paper in this series will outline some of the main theories and methodologies used by the systems engineer within an integrated framework for whole-system design. The methods will not improve on good engineering, but they do mean that the designer and project manager have a set of matching 'programmes' with which to tackle the problems of integration and balanced completeness that large scale complexity brings to good engineering.

References

1 Warfield, J. N., 'Systems Engineering', NOrd 7958–307 Ordnance Research Laboratory, Pennsylvania State University, 1955.
2 Schlager, K. J., 'Systems engineering—key to modern development', *IRE Trans. on Engineering Management*, **EM-3**, pp. 64–6, 1956.
3 Hall, A. D., 'A Methodology for Systems Engineering' (Van Nostrand, Princeton, 1962).
4 McKean, R. N., 'Efficiency in Government Through Systems Analysis' (Wiley, New York, 1958).
5 Warfield, J. N., 'Societal Systems: Planning, Policy and Complexity' (Wiley, New York, 1976).
6 Sage, A. P., 'Methodology for Large Scale Systems' (McGraw-Hill, New York, 1977).
7 Lewin, D., 'The relevance of science to engineering—a reappraisal', *The Radio and Electronic Engineer*, **49**, no. 3, pp. 119–24, March 1979.
8 Williams, T. J., 'Systems Engineering for the Process Industries' (McGraw-Hill, New York, 1961).
9 Sargent, R. W. H., 'Forecasts and trends in systems engineering', *The Chemical Engineer*, no. 262, p. 226, June 1972.
10 Rawson, K. J., 'Towards economic warship acquisition and ownership', *Trans. R. Inst. Nav. Archit.*, **115**, pp. 43–56, 1973.
11 Jenkins, G. M., 'The systems approach', *J. Systems Engineering*, **1**, no. 1, pp. 3–49, 1969.
12 Terotechnology publications from the National Terotechnology Centre, Cleeve Road, Leatherhead, Surrey.
13 M'Pherson, P. K. and Charlwood, F. J., 'A Teaching Framework for Systems Engineering and Terotechnology: Ten years experience at B.Sc. and M.Sc. levels'. DSS/PKM/FJC/184 Department of Systems Science, The City University, London, March 1979.
14 Shinners, S. M., 'Techniques of Systems Engineering' (McGraw-Hill, New York, 1967).
15 Kline, M. B. and Lifson, M. W., 'Systems engineering', in J. M. English (ed), 'Cost-Effectiveness' (Wiley, New York, 1968).
16 Blanchard, B. S. 'Logistics Engineering and Management' (Prentice-Hall, Englewood Cliffs, New Jersey, 1971).
17 Coutinho, J. de S., 'Advanced Systems Development Management' (Wiley, New York, 1977).
18 Shinners, S. M., *op. cit.*, pp. 87–95.
19 Wilcox, R. B., 'Analysis and synthesis of dynamic performance of industrial organizations—the application of feedback control techniques to organizational systems', *IRE Trans. on Automatic Control*, AC-3, no. 2, pp. 55–67, March 1962.
20 Rogers, J. R., 'Towards a philosophy of systems design', in S. Lees (ed.), 'Air, Space and Instruments' (McGraw-Hill, New York, 1963).

Manuscript first received by the Institution on 13th February 1980 and in final form on 29th July 1980
(Paper No. 1959/ACS 81)

UDC 621.391.812.62.029.55
Indexing Terms: Radio transmission, propagation effects, ionosphere

Oblique h.f. radiowave propagation in the main trough region of the ionosphere

M. LOCKWOOD, B.Sc., Ph.D.*
and
V. B. MITCHELL, B.Sc., Ph.D.†

SUMMARY
The propagation of 7·335 MHz, c.w. signals over a 5212 km sub-auroral, west-east path is studied. Measurements and semi-empirical predictions are made of the amplitude distributions and Doppler shifts of the received signals. The observed amplitude distribution is fitted with one produced by a numerical fading model, yielding the power losses suffered by the signals during propagation via the predominating modes. The signals are found to suffer exceptionally low losses at certain local times under geomagnetically quiet conditions. The mid-latitude trough in the F2 peak ionization density is predicted by a statistical model to be at the latitudes of this path at these times and at low K_p values. A sharp cut-off in low-power losses at a mean K_p of 2·75 strongly implicates the trough in the propagation of these signals. The Doppler shifts observed at these times cannot be explained by a simple ray-tracing model. It is shown however, that a simple extension of this model to allow for the trough can reproduce the form of the observed diurnal variation.

* *Formerly at the University of Exeter; now with the Appleton Laboratory, Ditton Park, Slough, Berkshire SL3 9JX*

† *Department of Physics, University of Exeter, Stocker Road, Exeter, Devon EX4 4QL.*

1 Introduction

The 'mid-latitude' or 'main' trough in nocturnal F-layer ionization densities was first observed in topside sounder data from the *Alouette 1* sounder[1]. The trough has since been found to be a regular feature of the topside ionospheres of both hemispheres (unlike the shorter-lived polar troughs which are also observed), having a 95% occurrence frequency at local midnight.[2] The trough is seen as a sharp decrease in the ionization density which is of the order of 10 degrees of latitude wide, and which lies along a geomagnetic latitude which varies approximately linearly with the K_p index value[3]. The F-layer peak is typically reduced by a factor of about four at the trough centre.[4] The trough moves equatorward and becomes much narrower at higher K_p values, giving it much steeper walls. The morphology of the trough under quiet conditions has been reviewed by Mendillo and Chacko.[5]

The trough is also observed in total electron content measurements[6] hence it is a depletion, and not a redistribution, of plasma. Very few direct observations of the trough below the F2 peak have been possible. The trough cannot be seen by ground-based ionosondes, due to the effects of off-vertical reflections, except at times of very low geomagnetic activity when the trough is at its widest.[7] The bottomside trough has been observed by various other direct methods including: ion trap experiments;[8] incoherent scatter radar[9] and an ionosonde with direction-finding capabilities.[10] Its presence, and in particular that of its steep poleward wall, have been inferred from various 'oblique echoes' on vertical ionograms by Stanley.[11] Some of the relatively scarce observations of the trough by ground-based and airborne h.f. experiments have been reviewed by Hartmann.[12]

The signals studied in the work described here are propagated along a path that lies approximately geomagnetically west to east at the latitudes at which the trough is observed under quiet magnetic conditions.

2 Amplitude Distributions of the Received Signals

2.1 Observed Distributions

The c.w. signals are transmitted from CHU, Ottawa (geographic co-ordinates: 45°18′N, 75°45′W) at a frequency of 7·335 MHz, with a power of 10 kW and by a vertical monopole antenna. They were received on a Beverage antenna at the Norman Lockyer Observatory, Sidmouth (geographic co-ordinates: 50°41′N, 3°13′W), a great circle distance of 5212 km from the transmitter. A receiver bandwidth of 1 kHz was used and the measured input signal amplitude calibrated frequently against known signal levels to allow for the receiver a.g.c. The amplitude was digitized and then recorded by a mini-computer data-logger. The receiver, data-logger and calibration signal synthesizer were all housed in separate buildings, about 100 m apart to prevent any mutual

0033-7722/80/110559+08 $1.50/0

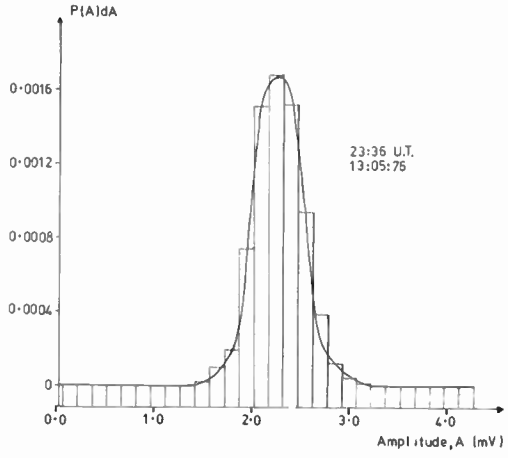


Fig. 1. Histogram of observed probability, $P(A)\mathrm{d}A$ (averaged over 0·15 mV ranges for a dA of 0·001 mV) as a function of amplitude A. The continuous curve is the model distribution for a single coherent signal of amplitude 2·2 mV (equivalent to a total loss of 113 dB) and random phase scattered signals.

interference problems. The amplitude was sampled at regular intervals from a clock based on the 60kHz transmission from Rugby. The recordings consisted of 512 samples taken first at 500 samples per second and then a further 512 samples taken at a sampling rate of 4 per second. The recordings thus lasted just over 2 minutes and were made at 36 minutes past each hour. The numbers of the 'slow' (4 per second) samples falling in certain amplitude intervals during a recording were counted (after a correction to allow for the time-constants of the receiver a.g.c.) and gave amplitude distributions of which examples are shown in the histograms of Figs. 1 and 2.

From the continuous recordings which were made in a three-month period (April, May and June 1976), 79% had to be discarded as the signals were below the noise level. This was usually due to large non-deviative absorption when all or part of the path was under daytime conditions. A further 12% of the recordings could not be analysed as they yielded amplitude distributions which were broad and complex. This was caused by severe multi-path conditions at the time of the recording, or by large variations in the received power spectrum during the recording. Therefore only 9% of all the recordings could be analysed, a total of 77.

2.2 Model Distributions

The amplitude and phase of the resultant of two interfering signals of different amplitudes and Doppler shifts are given by equations derived by Toman.[13] To this resultant a third signal can be added at that time, t, using the same equations. In this way n specular, coherent signals and a Rayleigh-distributed vector (the scattered signal component) can be added together at the time t. (The resultant of n Rayleigh-distributed vectors is itself a Rayleigh-distributed vector.[14]) This is repeated at different values of t until the resultant phase and amplitude variation in a whole number of beat periods is obtained. The amplitude distribution in a whole number of beat periods can thus be modelled. Adding various r.m.s. scattered components to a single coherent signal yields the well-known Rice–Nakagami distribution,[15, 16], although fluctuations of the coherent component amplitude cause deviations from this.[17] The distribution for the interference of two coherent signals (of amplitudes W_1 and W_2) with various r.m.s. scattered components (W_s) are shown in Fig. 3. It can be seen that here the presence of a second coherent signal gives a second peak in the distribution which can be resolved only if the secondary signal is larger than about three times the r.m.s. scattered signal. Distributions produced by this numerical fading model were fitted to the observed distributions using a chi-squared fit.

Examples of best-fit distributions are shown by the continuous curves in Figs. 1 and 2.

3 Power Losses

The maximum power losses of the resolvable signals, L_{TO}, were evaluated from the signal amplitudes used in the production of the best-fit distribution by using the minima of the measured feeder losses. Figure 4 shows n, the number of observations of L_{TO} in 1dB ranges. It can be seen that in 33 of the usable 77 recordings in this period L_{TO} was less than L_{min}, the minimum loss for a hypothetical link using the same apparatus but over a free-space, straight-line path of length equal to that of the Ottawa–Sidmouth great circle arc. Figure 5 shows an identical histogram for measurements taken with the same apparatus during this period when monitoring signals received from Tirane, Albania.

For this second path, which lies nearly north-south, the L_{TO} value was less than L'_{min} in only four out of a total of 71 usable recordings. Thus the anomalously low losses observed on the Ottawa–Sidmouth path cannot be due to measurement or calibration errors at the receiver as these would also cause anomalously low losses to be observed on the Tirane–Sidmouth path.

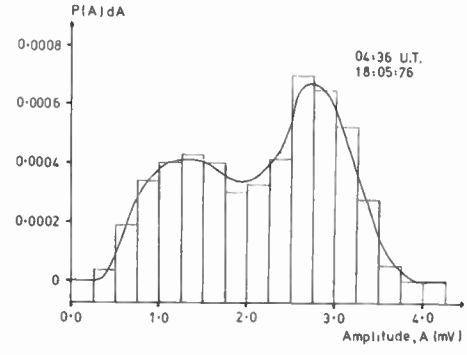


Fig. 2. Histogram of observed probability, $P(A)\mathrm{d}A$ (averaged over 0·25 mV ranges for a dA of 0·001 mV) as a function of amplitude A. The continuous curve is the model prediction for two coherent signals of amplitudes 2·0 and 1·1 mV (equivalent to total path losses of 114 dB and 123 dB respectively).

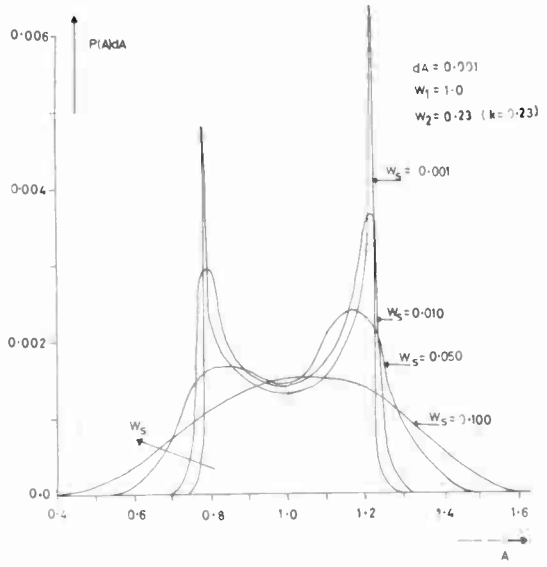


Fig. 3. Probability, $P(A)dA$ (for a dA of 0.001 mV), as a function of amplitude A, for two coherent signals of amplitude 1·0 mV and 0·23 mV and various r.m.s. scattered components, W_s.

L_{min} is a minimum possible loss for the path as it is the sum of the minimum of the spatial attenuation, the minimum of the feeder losses and the maxima of the two antenna gains. Hence there are a large number of other factors not included in L_{min} which cause the loss for any one mode of propagation to be considerably greater than L_{min}, for example: absorption losses; polarization losses; reflection losses; additional spatial attenuation (as the ray path length of any ionospheric mode must be greater than the great circle arc length). Semi-empirical predictions of the power loss for the possible modes of propagation were made using a method based on that of Barghausen *et al.*[18] Such predictions are based on monthly means of ionospheric parameters and it is well known that they can differ greatly from the actual losses at any one time. This is due to: approximations used in making the prediction; the large spreads in the values of the parameters used (due to the variability of F-layer propagation), and factors which are not allowed for in the prediction at all. The first two of these possibilities were eliminated from the analysis by evaluating a minimum loss for each propagation mode, L_M. This was done by using a minimum of each loss term in the expression for the total loss (evaluated by using the relevant decile values), and by omitting some other loss terms completely (e.g. deviative absorption and polarization losses). The only gain factors included were focusing due to a spherical ionosphere and maxima of the antenna gains. It was thus possible to get estimates of the 'additional' gain of the observed signals over the minimum predicted loss.

As the predicted loss was minimised this gives a minimum of the additional gain, G_{min}, given by

$$G_{min} = L_M - L_{TO}. \quad (1)$$

Figures 6 and 7 show the predicted field strengths (for isotropic antennae) and elevation angles at the receiver for the modes predicted to be present for median, upper and lower decile nocturnal ionospheres for May 1976. The sum of the transmitter and receiver antenna gains, $(G_T + G_R)$, is shown in Fig. 8 as a function of elevation angles (at both antennae), ϕ, for propagation in the great circle plane. (These curves only strictly apply to propagation via a spherically stratified ionosphere.) These antenna gains were modelled theoretically using a range of ground conditions applicable to the sites in question[19] and then checked by direct calibration over the low elevation angle range used in this experiment.

Figures 6, 7 and 8 show that the predicted dominant propagation mode is the 2F2 mode during these nights, as it always suffers the lowest transmission loss and has the highest total antenna gain. High and low angle

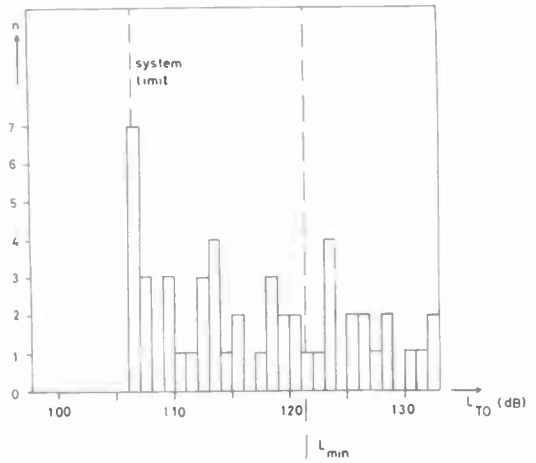


Fig. 4. Histogram of n, the number of observations for which the observed loss on the Ottawa–Sidmouth path is between L ($L_{TO}+1$) dB in April, May and June 1976. L_{min} is the minimum loss of hypothetical link using the same apparatus over a straight-line path of length 5212 km.

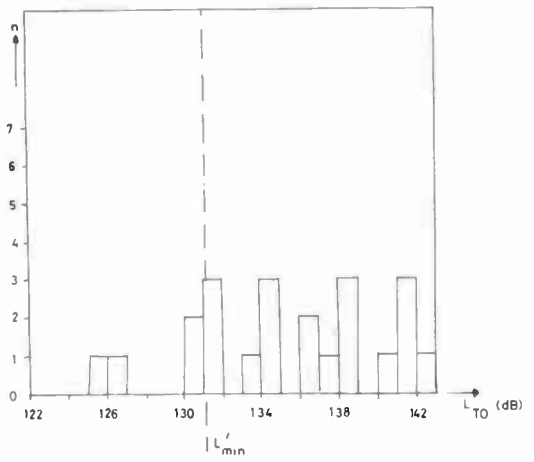


Fig. 5. Histogram of n, the number of observations for which the observed loss on the Tirane–Sidmouth path is between L_{TO} and ($L_{TO}+1$) dB in April, May and June 1976. L'_{min} is the minimum loss of a hypothetical link using the same apparatus over a straight-line path of length 2048 km (the Tirane–Sidmouth great circle distance).

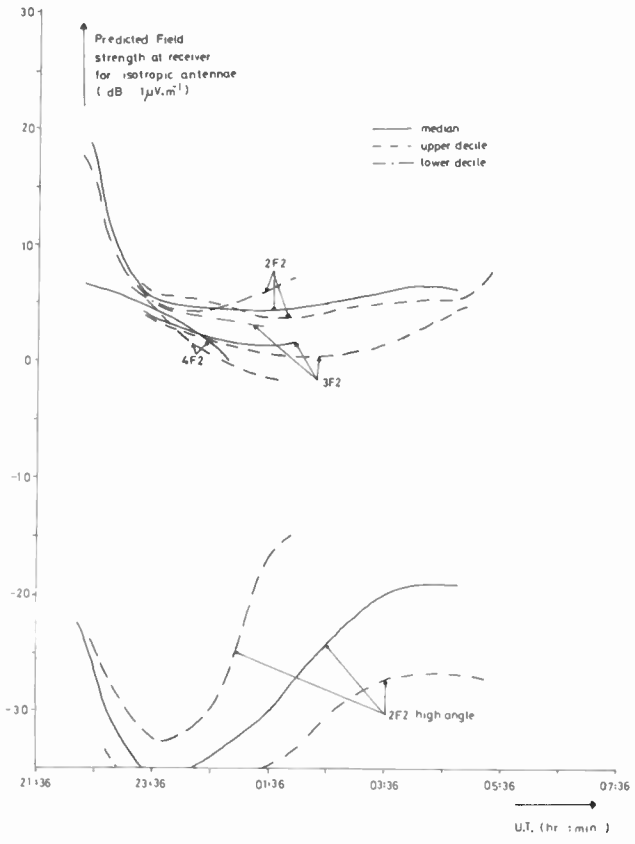


Fig. 6. Field strength at the receiver (for isotropic antennae) for the propagation modes predicted for median, upper and lower decile nocturnal ionospheres in May, 1976.

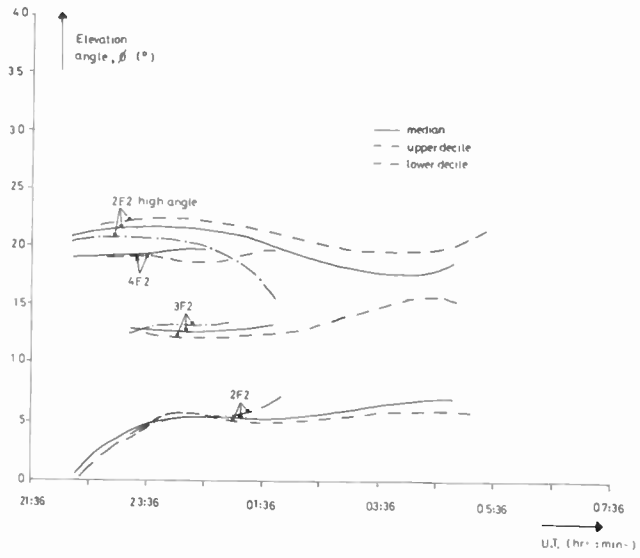


Fig. 7. Elevation angles at the receiver for modes shown in Fig. 6.

modes were predicted but only the simplest of mixed modes (1F, 1E and 1E, 1F).

The variations in the calculated values of G_{min} are shown in Fig. 9 for five consecutive days. It can be seen that in this time positive values of G_{min} were often observed during the night. Large positive values were very rarely observed in isolation, rather they were usually part of a regular variation of the type shown by days 1, 2 and 5 of Fig. 9. On days 3 and 4 no G_{min} values in excess of -40 dB were found. In April, May and June 1976 (the period used to compile Fig. 4) values of G_{min} up to $+35$ dB were observed. In 27 of the usable 77 recordings in this period G_{min} was greater than $+10$ dB and in 43 it was positive.

4 Discussion of the Amplitude Results

In a few recordings in the 3-month period the amplitude distribution of the signal received over the west-east path was of a form that could be fitted with a model prediction and that yielded a positive value of G_{min}. Such examples account for only 5% of the total number of recordings made during this period, but this is over half of the 9% of usable recordings. 40% of the recordings taken when the path was totally in darkness were found to be usable. For the north-south path, by way of comparison, only 3% of the usable recordings gave positive G_{min} values and none gave it to be larger than $+7$ dB (compared with 43% of usable recordings on the west-east path for which G_{min} exceeds this value).

By using the definition of G_{min} given in equation (1) there are only a few possible causes of any positive G_{min} values. Firstly, the propagation may be due to a mode deemed impossible by the prediction scheme. The most important candidate for this is the high angle 1F2 mode, which the prediction scheme does not consider over such long paths, but which has been observed to propagate over a similar path by Warren and Hagg and by Kift.[20] Muldrew and Maliphant[21] showed by ray tracing that such propagation was possible and observed one-hop

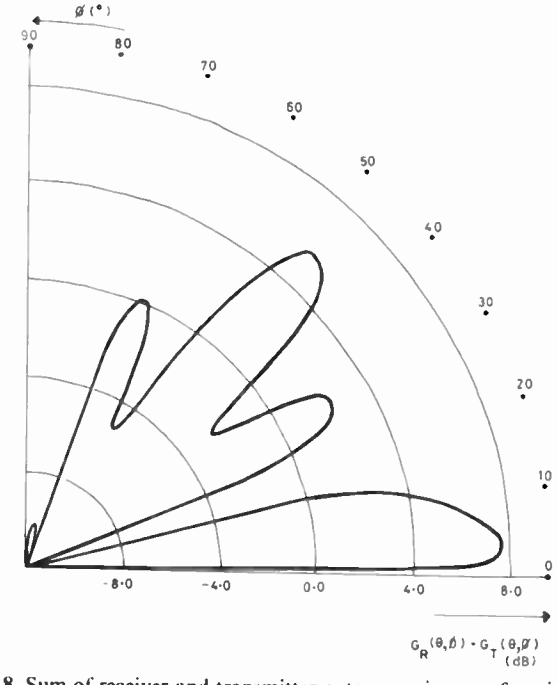


Fig. 8. Sum of receiver and transmitter antenna gains as a function of elevation angle, ϕ, modelled for great circle propagation and equal elevation angles at receiver and transmitter.

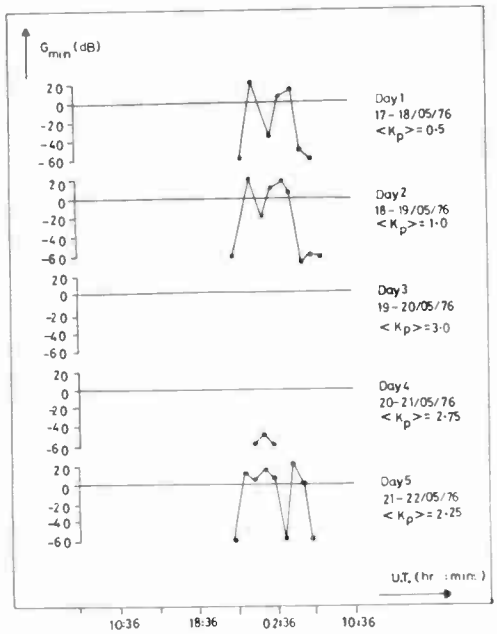


Fig. 9. The variation of the additional ionospheric gain of the predominating mode, G_{min}, during five consecutive days.

propagation to be present on average 35% of the time two-hop propagation was present on an Ottawa–Hague path. In the summer this ratio rose to 80–90%. However they only observed the one-hop mode at frequencies above 4 or 5 MHz below the m.u.f.; (Warren and Hagg's measurement was also at 27–33 MHz). The absolute lower frequency limit of the high angle one-hop mode was shown to be the layer critical frequency for certain simple models. Thus the frequency of these signals is lower than those for which Pedersen one-hop has been observed. Such a mode would also have a lower antenna gain than the 2F2 mode (Fig. 8) and suffer larger deviative absorption and divergence losses (although the latter can be reduced by an increase in the layer thickness[22]).

Another possibility arises from the 'maximum range' focusing effect on long hops caused by a continuous increase in the ionization between the E and F-layers. This effect is usually not present at night due to the deeper valley between the layers.[23] Positive G_{min} values may also be due to modes which have not been resolved by the amplitude distribution. This would be the case for signals which had Doppler shifts differing by less than the inverse of the recordings duration (i.e. less than about 0·01 Hz). However large numbers of such signals, all in phase at the time of recordings, are needed to explain the magnitude of the G_{min} values, a situation which is statistically improbable.

Other factors which could contribute to positive G_{min} values are focusing due to tilts of the iso-ionic contours of the ionosphere, and tilt-supported 'chordal' ('supertrapezoidal') propagation modes as are observed in trans-equatorial propagation.[24] Such modes involve two reflections from an ionospheric layer without an intermediate reflection by the ground or by the topside of a lower layer.

Hence it appears possible that tilts may play a large role in the propagation of these signals. Large tilts of the F-layer are expected near the latitudes of this path at the walls of the mid-latitude trough. In order to evaluate their possible role in the propagation of these signals results were used concerning the mean morphology of the trough. Various such statistical models of the trough (as a function of local time, season, K_p index and sunspot number) have been compiled from the existing data.[25–27] It must be remembered that there are several pitfalls in the use of such statistical morphological results[4] and the regression equations cannot be used to predict the actual boundary locations of the trough at any one instant.[28] Hence their use here can only indicate that any one hypothesis is consistent with the mean statistical position of the trough.

Consider two points, A and B, situated at an altitude of 200 km, one-quarter and three-quarters of the way along the Ottawa–Sidmouth arc. The reflection points of the 2F2 mode will in general be removed from A and B, but they can be regarded as sufficiently accurate indicators considering the uncertainties in the predicted trough wall positions. The trough depletion is not very large during the day as then photoionization dominates over the loss mechanisms which are responsible for the trough's formation. Hence at dusk the trough 'opens' with increasing local time and it 'closes' at dawn giving sunset and sunrise 'end walls' to the trough as viewed in a LT—invariant latitude plot. The mean times that the trough begins and ceases opening and closing at the points A and B can be calculated for any one night from the solar zenith angles and the K_p index value using the regression equations given by Feinblum.[26] 80% of the observed positive values of G_{min} were at times when the trough was fully open at both A and B, 89% were at times when the trough was at least partially opened at both A and B and all were within an hour of these times. The times between which the trough opened at A and closed at B varied with the K_p index in exactly the same way as the times that positive G_{min} values began and ceased respectively. The two maxima in G_{min} were also at times when A and B were nearest to the end walls. This is qualitatively consistent with the idea of a 2F2 mode focused by the trough because at these times focusing would occur in the elevational plane as well as the azimuthal plane.[7] To test if this hypothesis can explain the magnitudes of the G_{min} values quantitatively requires a three-dimensional ray tracing study of propagation along the length of the trough.

Figure 10 shows the invariant latitudes of the equatorward and poleward edges of the trough walls as a function of local time for thee K_p values. These are given by the regression equation given by Halcrow and

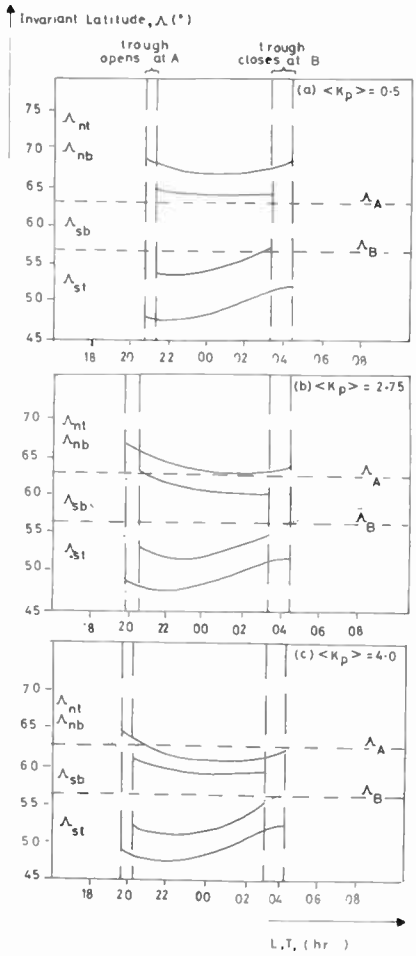


Fig. 10. The invariant latitudes of the tops and bases of the north and south troughwalls (Λ_{nt}, Λ_{nb}, Λ_{st} and Λ_{sb} respectively) of the northern hemisphere trough for 17–18/05/76 and for mean K_p values of (a) 0·5, (b) 2·75 and (c) 4·0.

Nisbet[4] from 8 years' *Alouette I* and *II* data. Also shown are the invariant latitudes of the points A and B, Λ_A and Λ_B. It can be seen that under quiet magnetic conditions both points are within the trough walls at the times at which the trough would be open at them. As K_p increases the trough moves equatorward, especially its poleward wall, making it narrower and equatorward of the point A, B remains within the trough wall at all K_p at these times. Hence at low K_p all the path is at the latitudes of the trough. In days 3 and 4 of Fig. 9 no large values of G_{min} were observed. On these two days the mean K_p is significantly higher. Figure 11 shows the observed values of G_{min} in April, May and June 1976 as a function of the mean K_p on the day of recording. The daily mean of the K_p value (from 12 UT to 12 UT) was used because the plasmapause, and hence the trough, positions are determined by the K_p history as well as its current value.[28] It can be seen that G_{min} is more consistently higher at the lowest K_p values, 80% at mean K_p less than 1·0 giving positive G_{min} values; this falls to 60% for mean K_p in the range 1·0 to 3·0 and for K_p greater than 3·0 it is zero. No positive G_{min} values are observed at all when the mean K_p exceeds 2·75, there being a sharp cut-off in large G_{min} about this value. In Fig. 10(b) it can be seen that at this K_p the point A lies within the poleward trough wall for most of the night. At lower K_p, A is inside the trough and at higher K_p it is outside it. This, again, qualitatively suggests the trough is focusing a mode. For example in a two-hop mode the first hop would be focused at low K_p but defocused at higher values, giving a cut-off in G_{min} at 2·75 as observed. It is unlikely that the 42% of recordings at mean K_p below 2·75 which yield negative G_{min} are due to the absence of a bottomside trough as the topside trough is present on roughly 95% of all nights.[2] These values are more likely to be due to the losses minimized in calculating G_{min} (e.g. non-zero polarization losses, auroral absorption, etc).

5 Doppler Shifts

The local oscillators used at the receiver and transmitter have a frequency stability in excess of 1 part in 10^9. This is the principal limitation on the accuracy of the Doppler shift measurement and corresponds to an error of less than $\pm 0{\cdot}01$ Hz on these signals.

The measurement of the Doppler shift of oblique h.f., c.w. waves was first done by Essen in 1935.[29] He viewed the ionospheric shift as an error in standard frequency transmissions, rather than as a source of information on the ionosphere. The technique developed into the latter as the accuracy of local oscillators increased and data storage techniques improved (e.g. Watts and Davies[30]). Full analysis of Doppler shifts is complicated by various effects, even at vertical incidence.[31] The simplest model

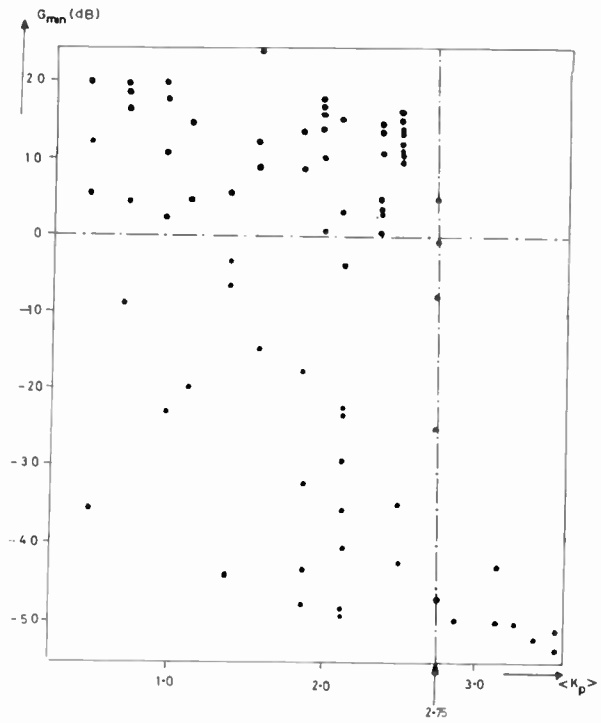


Fig. 11. Observed values of G_{min} in April, May and June, 1976, as a function of the mean K_p value.

for very oblique signals is a mirror-like reflecting ionosphere moving vertically. This model gives a shift which is proportional to the signal frequency, the layer velocity and the cosine of the angle of incidence of the rays onto this mirror-like layer. Hence shifts are smaller for longer paths.

The received signals were mixed down with signals derived from the local oscillator and their phase compared with that of another signal derived from the local oscillator. This phase comparison was done digitally by timing the delay between the zero crossovers of the two signals. The rate of change of this phase difference yields the Doppler shift of the composite sky wave. This equals the Doppler shift of the dominant mode, Δf, when averaged over a whole number of beat periods.[13] The mean shift in a 2-minute period was recorded hourly in the same way as was the amplitude, immediately prior to the amplitude recording.

An example of the observed Doppler shift variation is shown by the points in Fig. 12. Shifts could only be observed when G_{min} was large (approximately greater than -10 dB for midnight at the receiver) as the Doppler measuring system was designed to be triggered only by signals above a certain level. The Doppler variations always showed a regular diurnal variation of the kind seen in Fig. 12. This has features similar to the observed G_{min} variations in that it showed two maxima during a night with an intermediate minimum. The second peak is surprisingly large, being of amplitude between 0·2 and 0·5 Hz and occurring between 04 and 06 UT.

These Doppler shift results cannot be explained easily without involving the mid-latitude trough. The curves in Fig. 12 are produced by two-dimensional ray tracing through a parabolic F-layer of constant semi-thickness. The layer peak density and height was taken as a function of local time, for the date of the shown recordings, from the world maps used in the propagation loss prediction. This crude model does not explain either the form or the magnitudes of the observed Doppler shift variations (the continuous curve). If however, a trough in the peak density and height are introduced (of the model form used by Halcrow and Nisbet[4]) a variation like the one shown by the dashed curve is produced. In this example the peak density is reduced by a factor of four inside the trough relative to its value outside (after Halcrow and Nisbet[4]), the peak height is increased by 40 km and the semi-thickness is kept constant. It can be seen that this curve is considerably closer to the observed values. It is interesting to note that in the ray-tracing prediction the predominating mode was initially a two-hop mode, the angle of incidence at the ground reflection then became larger until that reflection did not occur at all. After this the mode was thus a 'chordal' type mode. The amount by which the ray 'missed' the Earth's surface between the two-layer

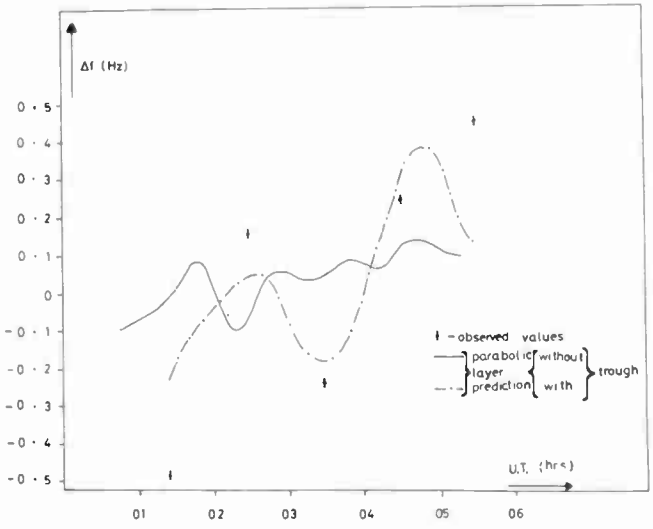


Fig. 12. Observed Doppler shift variation or 05–06/06/77 (mean $K_p = 1{\cdot}6$). The continuous line is a ray tracing model prediction with no allowance for the mid-latitude trough. In the dashed curve a trough model has been included in the prediction.

reflections firstly increased but then began to decrease until the conventional two-hop mode, with a ground reflection, was restored.

6 Conclusions

The presence of anomalously low power losses on a path which lies along the length of the mid-latitude trough and their absence from a path which lies nearly perpendicular to it, suggests it may be affecting the propagation of radio waves over the former path. The trough is strongly implicated by the cut-off in low power losses at the mean K_p when the path would, on average, no longer be entirely within the trough. The times of observations of low losses vary with K_p in agreement with the times that the trough opens and closes in the regions of ionospheric reflection of a 2F mode. Further evidence comes from the Doppler shift measurements which reveal exceptionally large, regular shifts at these times. The shifts can be modelled easily only if a crude model of the trough is introduced. This suggests the trough is supporting a low-loss 2F2 mode or a 'chordal' hop type mode at these times.

7 Acknowledgments

The authors are grateful to the staff of the Radio and Navigation Department, Royal Aircraft Establishment, Farnborough, for their help with the computer predictions. Thanks are also due to the staff of the Radio Research Centre, Auckland University, for their help with the ray-tracing modelling of the Doppler shifts. One of us (M. Lockwood) is indebted to the Science Research Council for a grant under the 'Co-operative Awards in Science and Engineering' scheme.

8 References

1 Muldrew, D. B., 'F region ionisation troughs deduced from Alouette data', *J. Geophys. Res.*, **70**, pp. 2635–50, 1965.

2 Ahmed, M., Sagalyn, R. C., Wildman, P. J. L. and Burke, W. J., 'Topside ionospheric trough morphology: Occurrence frequency and diurnal, seasonal and altitude variations,' *J. Geophys. Res.*, **84**, pp. 489–98, 1979.

3 Köhnlein, W. and Raitt, W. J., 'Position of the mid-latitude trough in the topside ionosphere as deduced from ESRO 4 observations,' *Planet. Space Sci.*, **25**, pp. 600–2, 1977.

4 Halcrow, B. W. and Nisbet, J. S., 'A model of F2 peak electron densities in the main trough region of the ionsphere', *Radio Sci.*, **12**, pp. 815–20, 1977.

5 Mendillo, M. and Chacko, C. C., 'The base level ionospheric trough,' *J. Geophys. Res.*, **82**, pp. 5129–37, 1977.

6 Mendillo, M. and Klobuchar, J. A., 'Investigations of the ionospheric F region using multistation electron content observations,' *J. Geophys. Res.*, **80**, pp. 643–50, 1975.

7 Helms, W. J. and Thompson, A. D., 'Ray tracing simulation of ionisation trough effects upon radio waves,' *Radio Sci.*, **8**, pp. 1125–32, 1973.

8 Sharp, G. W., 'Mid-latitude trough in the night ionosphere,' *J. Geophys. Res.*, **71**, pp. 1345–56, 1966.

9 Taylor, G. N., 'Structure at the poleward edge of the mid-latitude F region trough,' *J. Atmos. Terr. Phys.*, **35**, pp. 647–56, 1972.

10 Bowman, G. C., 'Ionisation troughs below the F2 layer maximum,' *Planet. Space Sci.*, **17**, pp. 777–96, 1969.

11 Stanley, G. M., 'Ground based studies of the F region in the vicinity of the mid-latitude trough,' *J. Geophys. Res.*, **71**, pp. 5067–75, 1966.

12 Hartmann, G. K., 'H.f. and u.h.f. propagation studies of the mid-latitude ionosphere,' *Ann. Geophys.*, **31**, pp. 39–51, 1975.

13 Toman, K., 'Ionospheric phase—and group—path,' *J. Atmos. Terr. Phys.*, **29**, pp. 1019–24, 1967.

14 Norton, K. A., Völger, L. E., Mansfield, W. F., and Short, P. J., 'The probability distribution of the amplitude of a constant vector plus a Rayleigh distributed vector,' *Proc. I.R.E.* **43**, pp. 1354–61, 1955.

15 McNicol, R. W. E., 'The fading of radio waves of medium and high frequencies', *Proc. Instn Elect Engrs*, **96**, pt. 3, pp. 512–24, 1949.

16 Röttger, J., 'Influence of spread F on h.f. radio signals,' AGARD Conference Proceedings No. 173, Paper 26, 1975.

17 Vinogradova, M. B. and Gusev, V. D., 'Amplitude distribution of a scattered field with allowance for the focusing effect of large scale ionospheric inhomogenities,' *Geomag. and Aeron.*, **16**, No. 5, pp. 459–62 1977.

18 Barghausen, A. F., Finney, J. W., Proctor, L. L. and Schultz, L. D., 'Predicting long-term operational parameters of high-frequency sky-wave telecommunication systems,' ESSA Tech. Rept. WRL 1110–ITS 78, USA Govt. Printing Office, Washington, 1969.

19 Travers, D. N., Martin, P. E. and Sherill, W. M., 'Use of Beverage antenna in wide aperture h.f. direction finding,' Rep. Southwest Research Inst., San Antonio, 1964.

20 Kift, F., 'Single hop propagation of radio waves to a distance of 5300 km,' *Nature*, **181**, pp. 1459–60, 24th May 1958.

21 Muldrew, D. B. and Maliphant, R. G., 'Long distance one-hop ionospheric radio wave propagation,' *J. Geophys Res.*, **67**, pp. 1805–15, 1962.

22 Sazhin, V. I. and Tinin, M. V., 'Long distance propagation by means of the Pedersen ray,' *Geomag. and Aeron.*, **15**, pp. 564–5, 1975.

23 Croft, T. A., 'H.F. radio focusing caused by the electron distribution between ionospheric layers,' *J. Geophys. Res.*, **72**, pp. 2343–55, 1967.

24 Yeh, K. C. and Villard, O. G., 'A new type of fading observable on high frequency radio transmissions propagated over paths crossing the magnetic equator,' *Proc. IRE*, **46**, pp. 1968–70, 1958.

25 Halcrow, B. W., 'F2 peak electron densities in the main trough region of the ionosphere,' Tech. Rept. PSU–ILR–IR–55, Ionospheric Research Laboratories, Pennsylvania State University, 1976.

26 Feinblum, D. A., 'Hilion—a model of the high latitude ionospheric F2 layer,' Tech. Rept, US Army Contract No. DAHC–60–71–C–0005, Bell Labs., New Jersey, 1973.

27 Pike, C. P., 'An analytic model of the main F-layer trough,' AFGL–TR–76–0098, Air Force Geophysics Labs., Cambridge, Mass., 1976.

28 Mendillo, M., Chacko, C. C., Lynch, F. and Wildman, P. L. J. 'Attempts to predict trough/plasmapause boundaries in real time,' AFGL–TR–78–0080, p. 105, Air Force Geophysics Labs., Cambridge, Mass., 1978.

29 Essen, L., 'International frequency comparisons by means of standard radio emissions,' *Proc. Roy. Soc.*, **149**, A, pp. 506–10, 1935.

30 Watts, J. and Davies, K., 'Rapid frequency analysis of fading radio signals,' *J. Geophys. Res.*, **65**, pp. 2295–301, 1960.

31 Robinson, L. and Dyson, P. L., 'Effects of ionospheric irregularities on radio waves—I. Phase path changes,' *J. Atmos. Terr. Phys.*, **37**, pp. 1459–67, 1975.

Manuscript first received by the Institution on 23rd March 1978, in revised form on 19th September 1979, and in final form on 15th May 1980. [Paper No. 1960/Comm 258)

UDC 551.510.4: 621.375.826

Indexing Terms: Atmospheric composition, Pollution, Lasers, chemical, Lasers, gas

Monitoring of various atmospheric constituents using a c.w. chemical hydrogen/deuterium laser and a pulsed carbon dioxide laser

K. W. ROTHE, Ph.D.*

SUMMARY

Applications of a c.w. chemical hydrogen/deuterium fluoride laser and a pulsed TEA CO_2 laser for monitoring of atmospheric trace constituents and process control are reported. The c.w. HF/DF laser has been used for field measurements of the HF concentration in the exhaust of an aluminium plant. The laser has also been operated as a DF laser for measurements of the HDO concentration in natural water samples. In addition, using a pulsed TEA CO_2 laser in connection with the differential absorption technique, range-resolved ambient air measurements of the water vapour distribution in the vicinity of a cooling tower as well as measurements of the ethylene concentration over the area of a refinery have been performed. The results have been further evaluated using simple and reasonable propagation models. From this both the diffusion parameters and the total mass emission of a source could be obtained.

* *Physics Department, University of Munich, Am Coulombwall 1, D-8046 Garching, Federal Republic of Germany.*

1 Introduction

The necessity for effective monitoring and control of atmospheric constituents—especially pollutants—has become obvious in recent years. The well-established and commonly-used chemical methods suffer from an essential drawback, as they only can perform point monitoring measurements. The various laser techniques, however, allow the mean concentration value to be determined not only over a certain path-length but also over a certain area or volume; therefore, they may, in many applications, be superior to chemical methods. In addition, high-power pulsed lasers can be used in connection with the differential absorption technique to obtain a complete range-resolved map of the distribution of a species over larger urban or industrial areas. Furthermore, it is of interest to compare observed expansions of gaseous components with theoretical models. From this both the diffusion parameters as well as the total mass emission of a source can be evaluated, which—obviously—is one of the key quantities for determining the burden on the environment.

Range-resolved pictures, for example of a pollutant concentration, also enable one to detect and localize emission sources in large plants; therefore, this technique is of some interest to industry too, as leakages can be detected in complex production processes, which may help to save energy and raw materials.

Different techniques for measurements of atmospheric trace constituents have often been studied and compared in the past; in most cases absorption measurements turned out to be the most straightforward and sensitive ones.[1,2,3] The three most frequently used absorption methods—each of them with its own and distinct class of application—are the following:

(a) Long Path Absorption Technique.
For this purpose both pulsed and continuous-wave (c.w.) lasers can be used, the latter generally being preferred due to the more sensitive detection techniques (phase-sensitive detection, heterodyne detection). This method gives the mean concentration of a species over the pathlength, which may vary from several metres up to several kilometres (for details see Refs. 1, 2).

(b) Differential Absorption Technique.
In this case pulsed lasers are used. Their light is back-scattered by small particles in the atmosphere (Mie-scattering) and recorded as a function of time. Such a radar-like method (it is called lidar, since the radio waves are replaced by the laser light) gives the concentration of a species as a function of distance and can thus be used for a detailed mapping of a constituent over an area of several kilometres in diameter (for details see Ref. 3).

(c) Opto-Acoustic Detection Technique.
In this case the laser beam is transmitted through a

0033–7722/80/110567+08 $1.50/0

cell which is filled with the air sample to be investigated. The absorbed energy gives rise to a small increase of temperature and consequently of pressure inside the cell. Therefore, if the laser beam is periodically modulated, this pressure variation can be detected by a microphone. Such an opto-acoustic cell (often called spectrophone) offers an extremely sensitive method for the detection of atmospheric trace constituents at very low concentration levels (for details see Ref. 4).

Considering the different field measurements which have been performed in recent years, it turns out that apart from NO_2 (Refs. 5, 6, 7) and SO_2 (Refs. 8, 9) nearly all investigations have been performed in the infra-red spectral region, where almost all constituents show their individual absorption bands. With the present state of laser technology, essentially only the semiconductor diode laser[10,11] and the fixed-frequency infra-red laser[3,12] are usable for practical field applications. All other available infra-red lasers are still lacking in reliability and ease of handling. This is made plain by the fact that in spite of numerous laboratory investigations with different laser systems only a limited number of field measurements have been performed.

In the following Sections some measurements performed with fixed frequency infra-red lasers will be reported. These lasers are easy to handle and can be operated at rather high intensities. There are no problems associated with frequency control as the laser lines are fixed, but on the other hand one can only make use of accidental coincidence between the laser lines and pressure-broadened absorption lines of atmospheric constituents. The multitude of coincidences found, however, indicates that this is not a very severe limitation, provided that the absorption coefficients and the influence of self and foreign-gas broadening have been carefully evaluated.[2,3]

2 Long Path Absorption Measurements of Hydrogen Fluoride Emission

The control of the hydrogen fluoride (HF) emission, e.g. from an aluminium plant, is a serious problem. The sensitivity of chemical methods for the detection of HF is rather poor, so that even for HF concentrations at the maximum-allowed emission level long sampling times up to one hour are necessary. Consequently, short-term emission peaks cannot be detected and a real-time measurement is not possible. Another problem associated with the chemical methods is the distinction between gaseous HF on one side and HF absorbed on dust and aerosols on the other side. This problem does not exist for optical absorption measurements, which only detect the more dangerous gaseous HF.

The most sensitive optical method for detection of HF is by way of resonance absorption of the HF laser radiation itself. The active medium is produced from a chemical reaction in excited vibrational states. Thus there are transitions to the ground state, which of course coincide with the HF absorption lines in an optimum way. This advantage is somewhat reduced by the fact that the emission of the HF laser lies within a region of rather strong atmospheric attenuation due to water vapour and carbon dioxide absorption. Therefore, these additional absorptions have to be taken into account for the evaluation of HF concentrations. Details of the set-up as well as of the evaluation scheme are given in Ref. 13.

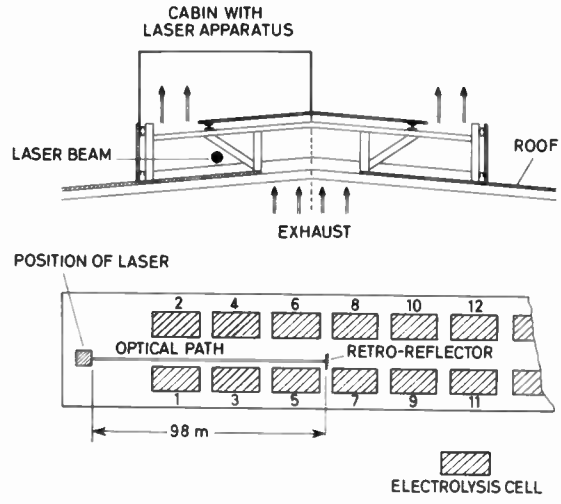


Fig. 1. Location of the absorption path under the roof top of the aluminium plant above the electrolysis cells.

For the field measurements of the HF emission the laser system had been installed under the roof top of an aluminium plant close to the exhaust of a row of electrolysis cells. Figure 1 shows the location of the optical absorption path with respect to the electrolysis cells and the exhaust. A typical result of the measurements is shown in Fig. 2 in which the HF concentration is plotted as a function of time. The short-term emission peak is caused by some activities at the electrolysis cells underneath (e.g. anode effect or crust breaking). This fast response of the laser method to the rapidly-changing conditions suggests a potential

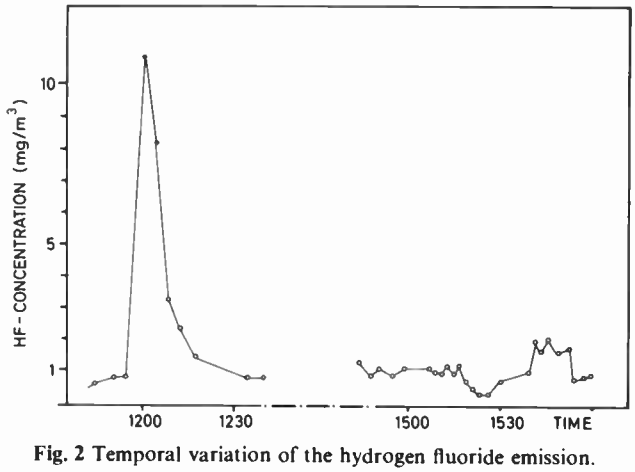


Fig. 2 Temporal variation of the hydrogen fluoride emission.

application of the laser detection technique for the control of bulk chemical processes.

At the end of the HF field experiments simultaneous measurements have been performed to compare the laser method with two nearly-identical chemical methods. For this purpose an absorption cell was used to ensure that all methods were probing the same object. The exhaust air was circulated through the cell by fans at either side of it. To establish a sufficiently high concentration level for the chemical methods a flexible pipe attached to the cell delivered exhaust air directly from different points in the vicinity of the electrolysis cells. As the sampling time of the chemical methods was between 15 and 30 minutes, only the mean value of the individual laser measurements over the corresponding time interval could be used for a comparison. The results clearly indicate a significant disagreement between the different methods up to a factor of 5. The averaged concentrations determined by the laser absorption technique fell within the two independent chemical measurements. The HF concentrations obtained by the 'Landesanstalt für Immissions- und Bodennutzungsschutz Essen' exceeded the laser results by a constant factor of 1·6. The reason for this, especially for the discrepancies of the two chemical measurements, is currently subject to speculations and can only be clarified by further investigations and comparisons.

3 Opto-Acoustic Detection of the HDO Content in Water Samples

For many hydrological investigations a precise knowledge of the abundance values of water isotopes is very helpful. Due to slight differences in the specific gravity and the vapour-pressure, e.g. of H_2O versus HDO, there are small deviations from the normal abundance, from which some information can be deduced, e.g. regarding origin and age of water samples as well as evaporation rates of lakes. The normal abundance of HDO is only 0·03% and the deviations are of the order of 1% of this value.[14] Therefore, very sensitive techniques are necessary to give the required accuracy, which up to now could only be achieved by mass-spectrometers. In this Section a straightforward and reliable laser method is described which is believed to give at least the same accuracy as mass-spectrometers without being as complicated.

Due to the low HDO concentrations the measurements should be performed in a spectral region with negligible absorption by the major atmospheric constituents as H_2O and CO_2. Therefore, the DF laser was used, the lines of which lie within the 3·5 μm–4·0 μm region of high atmospheric transmission. The laser was essentially the same as for the HF measurements discussed in the previous Section, only two modifications being made: for DF laser operation the reagent gas H_2 was exchanged for D_2 and an intra-cavity grating was used to tune over the individual laser lines. In this way the number of lines available could be increased and a significant reduction of the noise was achieved in the measurements.

As the HDO concentration to be measured is rather small the sensitive opto-acoustic detection technique was used for the measurements. The opto-acoustic cell was made of stainless steel with the microphone placed in the middle of it. At both ends, close to the cell windows, there were reservoirs which could be filled with the water samples. The microphone signal was preamplified, filtered, and synchronously detected by a lock-in amplifier. As the absorption inside the cell was very small, the intensity of the laser beam could be measured immediately behind the spectrophone, thus making the set-up as well as its alignment very simple. Testing and calibration of the apparatus was performed by filling the cell with a small amount of CO_2 and measuring the spectrophone response due to the known absorption of the 2P4 HF laser line.[13]

The HDO measurements were performed in two steps. The first part aimed to give a survey on the HDO absorption at the different laser lines in order to identify those lines which are best suited for a sensitive HDO detection. In the second step these lines were examined closer with regard to the minimum detectable HDO concentration. For this purpose a deuterium-enriched water sample was stepwise diluted over several decades.

For reasons of higher accuracy the absorption coefficients were investigated at increased HDO concentrations. For this purpose small amounts of pure D_2O were added to distilled water samples. Unfortunately, the DF laser radiation is absorbed not only by lines of the HDO molecule; in addition, there is also some line absorption by D_2O as well as line and continuum absorption by H_2O (Refs. 15, 16). While the H_2O absorption could be neglected for the investigations, the influence of the D_2O absorption had to be taken into account. This was done by recording the spectrophone response for two differently deuterium enriched water samples. From this the HDO absorption and the D_2O absorption could be evaluated separately.

In order to provide a better comparison of the relative HDO sensitivity, Fig. 3 shows in its lower part the DF laser spectrum and in its upper part the signal obtained from the opto-acoustic cell filled with a deuterium-enriched water sample. The D_2O contribution is less than 10% for all laser lines. From this the 2P2 DF laser line can easily be identified to be the best-suited line for HDO measurements. To the knowledge of the author this is the highest detection sensitivity for HDO obtained so far with a DF laser. This might be due to the fact that the 2P2 line will in general only be emitted from a grating tuned c.w. DF laser.

The second part of the measurements aimed to give information on the sensitivity of the HDO detection. For

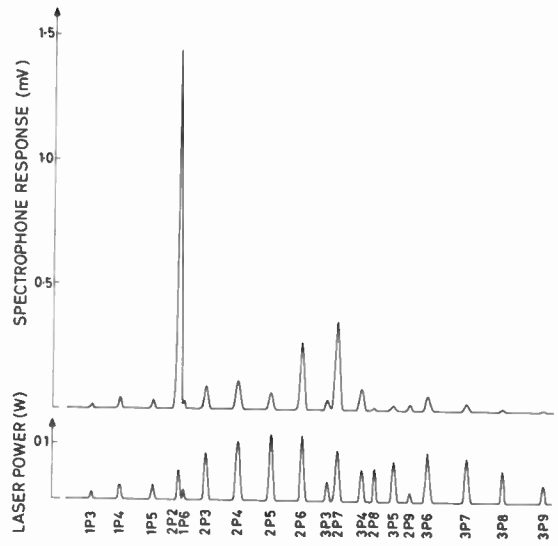


Fig. 3 Opto-acoustic response due to HDO absorption of the DF laser lines.

this purpose a deuterium-enriched water sample with a HDO concentration of 7·9% was diluted stepwise up to the natural concentration of 0·03%. These measurements were performed only at those DF laser lines which gave rise to a relatively large HDO response signal. For a comparison the measured spectrophone signals were normalized for a laser power of 1 W and a total water vapour pressure of 28·1 mbar. Figure 4 shows the normalized spectrophone signals as function of the HDO concentration. The arrow indicates the natural HDO concentration of 0·03%. A double logarithmic scale was chosen because of the large concentration variations over more than two decades. As can be seen from the figure there is no significant interference due to other substances. The measured signals have been extended by the dotted lines to a signal-to-noise ratio of one. The noise level for all lines was found to be independent of the laser power and predominantly due to electromagnetic noise produced from the microwave discharge of the chemical DF laser.

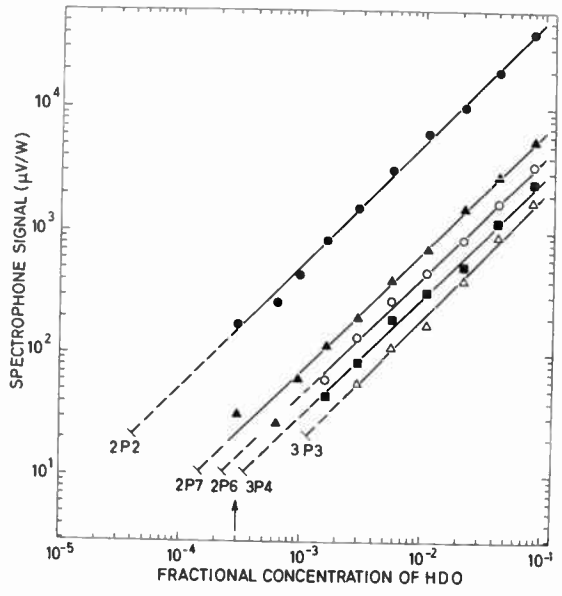


Fig. 4. HDO absorption signals for five DF laser lines as a function of the HDO concentration.

Figure 4 indicates that already with the present set-up, natural HDO concentrations can be measured up to an accuracy of 10%. More effective shieldings of the microwave discharge should improve the sensitivity to 1% or even better.

4 Range-Resolved Measurements of Atmospheric Constituents with a Pulsed CO_2 Laser by the Differential Absorption Technique

In addition to measurements of spatially-averaged concentration or *in situ* measurements as discussed in the last two Sections, there is also a need for detailed mapping of different constituents over large areas. With such a range-resolved picture of the concentration distribution it is possible to detect and localize emission sources over large urban or industrial areas. In addition, the observed expansion of a species can be compared with theoretical diffusion models. In this way long range prediction of the propagation, e.g. of a pollutant, becomes possible.

Measurements like this require a lidar system with a high-power pulsed laser where Mie-scattering is used as a spatially distributed reflector. The differential absorption technique has been described in detail in the literature[1,2,3,5] and its applicability has been proven for several species.[3,5,6,7,9,12] In this Section a mobile lidar system is described which has been used to measure the three-dimensional distribution of water vapour in the vicinity of a cooling tower as well as the ethylene distribution and propagation over a refinery.

The lidar system is equipped with a pulsed TEA CO_2 laser with a beam diameter of 4 cm, a pulse energy of 1·3 J, and a repetition rate of up to 70 Hz. The individual laser lines are selected by a grating and can be monitored by means of a spectrum analyser. Before being emitted into the atmosphere the beam is expanded by a factor of 8. In this way not only is eye safety improved, but also the divergence of the laser beam is reduced to about 0·4 mrad which is significantly smaller than the field of view of the receiving optics. The latter consists of a spherical mirror of 60 cm diameter and 240 cm focal length, which focuses the backscattered light onto a fast infra-red detector. The detector signal is amplified, stored in a transient recorder as a function of time and then transferred into a computer for further data evaluations. The computer also performs wavelength control of the laser as well as the mechanical steering of the whole lidar system which can be positioned hydraulically both in horizontal and vertical direction.

4.1 Field Measurements of Water Vapour Distributions

The first field measurements programme with the lidar set aimed to demonstrate the suitability of the system for

remote mapping of atmospheric constituents. For this purpose measurements of the atmospheric water vapour concentration in the vicinity of the cooling tower of a gas combustion power plant were performed as part of a concerted measurement programme near Meppen in Northern Germany. The R(20) and R(18) lines of the 00°1–10°0 transition of the CO_2 laser (with absorption coefficients of $8{\cdot}65 \times 10^{-4}$ (atm cm)$^{-1}$ and $0{\cdot}935 \times 10^{-4}$ (atm cm)$^{-1}$ respectively[17]) were used to study the expansion of the water vapour out of the visible plume as function of different wind directions.

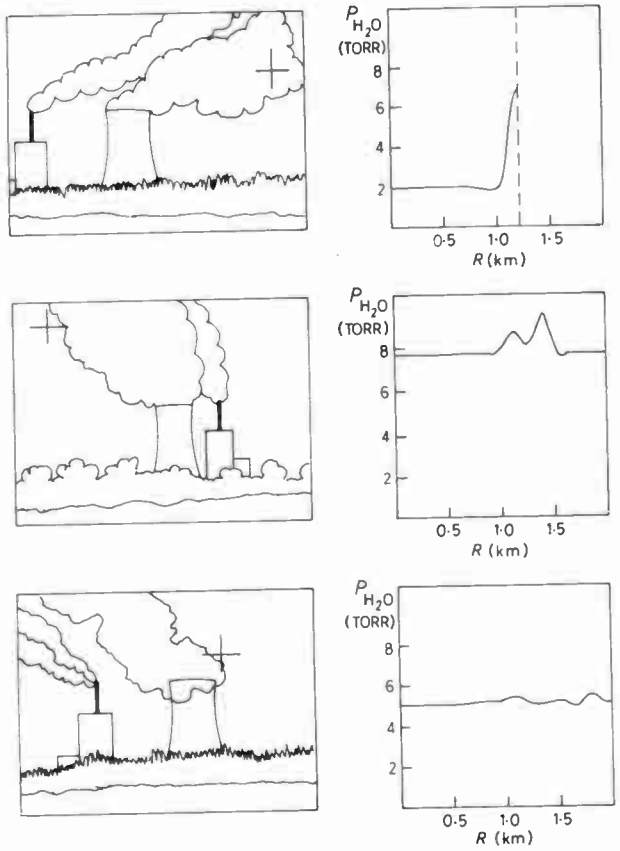


Fig. 5. Different types of water vapour measurements.

Some typical examples for the different types of measurements are demonstrated in Fig. 5. On the left-hand side of the Figure the cooling tower and the visible plume consisting of condensed water vapour is shown as seen from the laser. The cross marks the direction of the laser beam. On the right-hand side the corresponding water vapour concentration is plotted as a function of distance. The first type of investigation comprises all measurements in which the laser was directed onto the plume itself. In this case a sharp increase of the concentration can be observed, starting about 150 m before the beginning of the visible plume. There was no possibility of extending the measurements into the condensed phase, as all light is reflected back from the very beginning of the plume. This distance is indicated by the dotted line in the upper right part of Fig. 5. In the second type of investigation the laser beam was directed at the visible plume on its lee-side (centre part of Fig. 5). In this case a distinct increase of the water vapour concentration could be observed even several hundred metres away from the plume, indicating the strong influence of the wind. This behaviour became obvious from the third type of measurement with the laser beam directed at the windward side of the plume (lower part of Fig. 5). In this case no significant increase of the water vapour concentration was observed even in measurements performed very close to the plume.

A three-dimensional profile of the water vapour concentration was obtained by scanning the lidar system horizontally at different altitudes. Figure 6 shows the result of one horizontal scan at a fixed altitude (200 m). The area of the power plant, the location of the cooling tower and lines of equal water vapour concentrations are shown. The numbers give the concentration values in torr. The dotted line (near 12 torr) marks the part of the plume where the water vapour was condensed to droplets. The concentration over the power plant area itself results from the plume of a smaller smoke stack (see also the inset). These measurements can give detailed information on the distribution of water vapour outside the visible plume and therefore allow the influence of the wind and other meteorological parameters to be studied.

4.2 Field Measurements of Ethylene Distributions

The second field measurement programme was aimed at demonstrating the feasibility of the lidar system and the

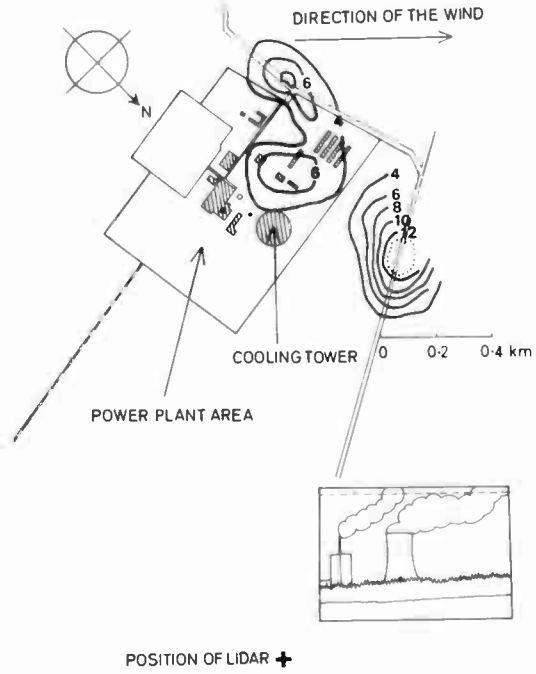


Fig. 6. Water vapour pressure in the vicinity of a cooling tower. In the lower right-hand corner the cooling tower is shown as it is seen from the position of the lidar set-up. The dashed line shows the altitude of observation.

differential absorption technique for detection of trace constituents in the atmosphere at the parts per billion level, which might help to find leakages in larger industrial areas. Therefore, measurements of the ethylene distribution over a refinery were performed.

Like most hydrocarbons ethylene can be measured near 3 μm, but this is a region of strong interference not only with other hydrocarbons but also with water vapour and carbon dioxide. In addition, ethylene represents only a very small fraction of the total amount of hydrocarbons (typically less than 1%). Therefore, to avoid interferences with other species the measurements were performed in another spectral region, namely with the P (14) line of the 00°1–10°0 vibrational transition of the CO_2 laser at 10·5 μm. For this line an absorption coefficient of 36·1 $(\text{atm cm})^{-1}$ for ethylene has been determined.[18] Also in this case some care has to be taken in choosing the appropriate non-absorbing laser line for the reference signal in order to minimize the water vapour interference. Calculations of atmospheric transmission in the spectral region of interest showed that the P (16) line is not the most suitable candidate for a reference line as it is influenced by a wing of a water vapour line nearby.[15] Therefore, if the P (16) line is used, water vapour absorption has to be taken into account for the evaluation of the ethylene concentration.[19] In order to avoid this complication the P (14) and P (18) laser lines were used which should show negligible interference. First topographical measurements confirmed this behaviour and gave a mean ethylene concentration of about 25 parts in 10^9 over a path-length of 700 m.

Range resolved measurements of the ethylene distribution were performed using Mie-scattering and the differential absorption technique. Typical background concentrations of 10–20 parts in 10^9 have been found as well as steep rises to values of 100–200 parts in 10^9 on the lee-side of the distillation towers of the refinery, identifying them as the main source of the ethylene emission. A typical example of the measurements is shown in Fig. 7. In the left part of the Figure lines of equal ethylene concentrations are plotted over the area of the refinery. Starting with 20 parts in 10^9, the concentration increases by 40 parts in 10^9 from line to line. This figure represents the mean ethylene concentration over a period of two hours and approx. 100,000 laser pulses. In the right part of Fig. 7 the same measurement is shown again using another kind of data presentation. From the evaluated lines of equal concentrations a computerized three-dimensional picture of the ethylene distribution has been constructed which gives less quantitative information but a more vivid illustration of the concentration profile.

4.3 Applicability of Lidar Measurements for Propagation Studies

One of the key factors for the air quality is undoubtedly the pollution concentration. The reason for the concentration of a pollutant in many cases (e.g. if it is caused by industrial processes) is based on the fact that there is a strong remote source from which the pollutant is emitted. It then spreads out into the atmosphere and is transported to urban areas. Consequently, the concentration measured there is essentially determined by the total mass emission from the source per unit time and the process of transportation and rarefaction in the atmosphere. Further evaluation of the lidar data by means of propagation methods can give some significant information on both aspects.

The diffusion of a plume from its source is mainly caused by the wind and—perpendicular to it—also by atmospheric turbulences. Compared to these factors the molecular diffusion can be neglected. Assuming constant wind velocity in the x-direction, the dispersion of a plume from a single source can approximately be

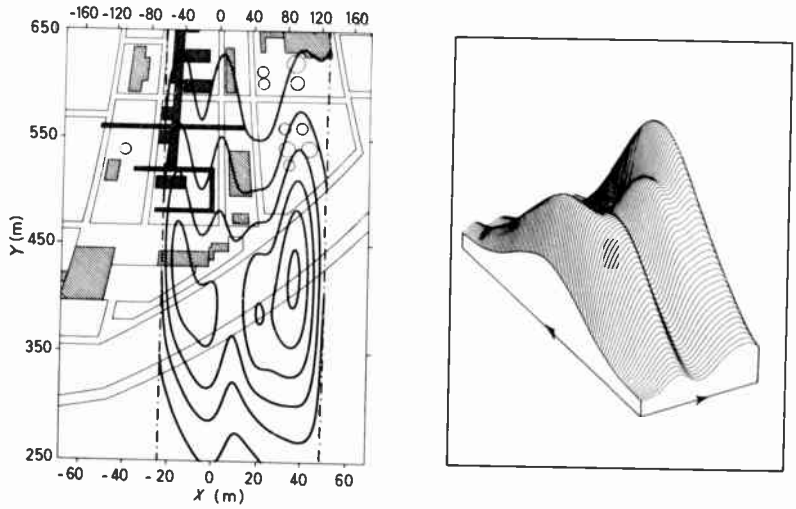


Fig. 7. Ethylene distribution over the area of a refinery (left) and its graphical illustration (right).

described by:[20]

$$c(x, y, z) = A(x) \exp\left[-(a|y|^r + b|z|^s)\right] \quad (1)$$

$c(x, y, z)$ gives the spatial concentration of the pollutant, the exponential term describing the diffusion of the plume in the (y, z) plane perpendicular to the direction of the wind; the parameters a, b, r, and s—which describe the width and the shape of the plume—must be determined experimentally. (For simplicity the exponentials r and s are assumed to be 2 in many cases. With this assumption one gets a simple Gaussian profile of the concentration distribution.)

Conventional methods to measure the above parameters are difficult to perform as the dispersion of the plume must not be disturbed by the measurement itself. This problem does not exist for lidar measurements which are therefore well suited for propagation studies. As can be seen from eqn. (1), the measurements should be performed in a vertical plane—perpendicular to the direction of the wind. The evaluated concentration profile from such a vertical scan is shown in the upper part of Fig. 8. This Figure gives the ethylene concentration in a plane at the periphery of the refinery

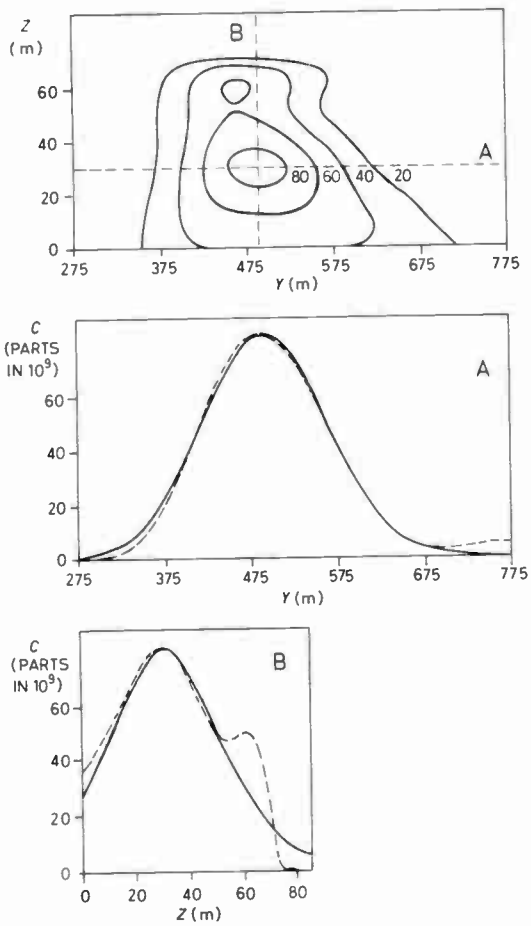


Fig. 8. Measurement of the vertical ethylene distribution at the periphery of a refinery and the derived and calculated diffusion profiles (for details see text).

(the measurements were performed at the end of the ethylene field measurement programme, described in Section 4.2). For further evaluation both the horizontal (section A) and the vertical (section B) profile through the centre of mass of this distribution was determined and compared with the theoretical profile (eqn. (1)). The results are shown in the middle and lower part of Fig. 8. The dashed lines give the measured concentration profile and the solid lines represent a least-square fit of the theoretical curves to these data points. The best approximation was found to be:

$$c(x, y, z) = A(x) \exp\left(-6 \times 10^{-5} |y|^{2.0} - 4 \times 10^{-3} |z|^{1.7}\right) \quad (2)$$

This result confirms that the often-made assumption $r = s = 2$ might be a fairly good approximation, at least for horizontal sections. The deviation of the measured vertical concentration profile from the theoretical curve (see lower part of Fig. 8) is caused by a local maximum in the concentration distribution, which can also be identified from the lines of equal ethylene concentration (upper part of Fig. 8). This indicates the existence of a secondary source for ethylene, which was not taken into account in the theoretical consideration.

In addition to the description of the diffusion of a plume the total mass emission of a source can also be determined from the lidar measurements. As all the pollution is transported through the plane of observation, the mass emission Q is simply given by:[21]

$$Q = \iint u(y, z) c(y, z)\, dy\, dz \quad (3)$$

u is the wind velocity in the (y, z) plane. If the emitted pollutant is subject to chemical reactions in the atmosphere or is partially absorbed at the ground, eqn (3) only gives a lower limit of the mass emission and must be adequately corrected. With a wind velocity of $u = 3$ m/s a total mass emission of about 250 kg ethylene per day was calculated from the measurements.

5 Conclusion

Summarizing the present situation, it can be stated that the application of laser methods is leading to considerable progress in the field of monitoring atmospheric constituents. For many applications laser techniques proved to be superior to conventional methods with respect to sensitivity and speed of response. The use of fixed frequency infra-red lasers for monitoring purposes has been demonstrated successfully. Laser measurements will not only help to monitor trace constituents and identify leakages, but will also enable propagation studies to be made which are very useful for the prediction of the long-range burden on the environment. Further measurements have to be performed in comparison to conventional methods in order to establish the laser technique as a very helpful instrument in sensitive, remote, and on-line monitoring of atmospheric trace constituents.

6 Acknowledgments

The financial support of the Bundesministerium für Forschung und Technologie, the Minister für Arbeit, Gesundheit und Soziales des Landes Nordrhein-Westfalen, and the Electricité de France is gratefully acknowledged. The investigations discussed in this paper were performed together with H. Walther, J. Wanner, A. Tönnissen, W. Baumer and M. Trautmann. I greatly appreciate their collaboration.

7 References

1 Byer, R. L., 'Review—remote air pollution measurement', *Opt. Quantum Electronics*, **7**, pp. 147–77, 1975.
2 Hinkley, E. D., 'Laser Monitoring of the Atmosphere' (Springer Verlag, Berlin, 1976).
3 Rothe, K. W. and Walther, H., 'Remote sensing using tunable lasers', in Mooradian, A., Jaeger, T. and Stokseth, P. (Ed.), 'Tunable Lasers and Applications', Proceedings of the Leon Conference, Norway, 1976 (Springer Verlag, Berlin, (1976).
4 Pao, Y. H., 'Optoacoustic Spectroscopy and Detection' (Academic Press, New York, San Francisco, London, 1977).
5 Rothe, K. W., Brinkmann, U. and Walther, H., 'Applications of tunable dye lasers to air pollution detection: measurements of atmospheric NO_2 concentrations by differential absorption', *Appl. Phys.*, **3**, pp. 115–9, 1974.
6 Grant, W. B., Hake, R. D., Jr., Liston, E. M., Robins, R. C. and Proctor, B. K., Jr., 'Calibrated remote measurement using the differential-absorption backscatter technique', *Appl. Phys. Lett.*, **24**, no. 11, pp. 550–2, 1974.
7 Rothe, K. W., Brinkmann, U. and Walther, H., 'Remote measurement of NO_2 emission from a chemical factory by the differential absorption technique', *Appl. Phys.*, **4**, pp. 181–2, 1974.
8 Kuhl, J. and Spitschan, H., 'A frequency-doubled dye laser with a servo-tuned crystal', *Optics Communications*, **13**, pp. 6–12, 1975.
9 Grant, W. B. and Hake, R. D., Jr., 'Calibrated remote measurements of SO_2 and O_3 using atmospheric backscatter', *J. Appl. Phys.*, **46**, no. 7, pp. 3019–23, 1975.
10 Hinkley, E. D., 'Laser spectroscopic instrumentation and techniques: long-path monitoring by resonance absorption', *Opt. Quantum Electronics*, **8**, no. 2, pp. 155–67, 1976.
11 Reid, J., Shewchun, J., Garside, B. K. and Ballik, E. A., 'High sensitivity pollution detection employing tunable diode lasers', *Appl. Optics*, **17**, no. 2, pp. 300–7, 1978.
12 Murray, E. R., 'Remote measurements of gases using discretely tunable infrared lasers', *Opt. Engng*, **16**, no. 3, pp. 284–90, 1977.
13 Tönnissen, A., Wanner, J., Rothe, K. W. and Walther, H., 'Application of a c.w. chemical laser for remote pollution monitoring and process control: Measurement of the hydrogen fluoride concentration in an aluminium plant', *Appl. Phys.*, **18**, pp. 297–304, 1979.
14 Moser, H. and Stichler, W., 'Die Verwendung des Deuterium- und Sauerstoff-18-Gehalts bei hydrologischen Untersuchungen', *Geologica Bavarica*, **64**, pp. 7–35, 1971.
15 McClatchey, R. A., Benedict, W. S., Clough, S. A., Burch, D. E., Calfee, R. E., Fox, K., Rothman, L. S. and Garing, J. S., 'AFCRL atmospheric absorption line parameter compilation', AFCRL-TR-73-0096 (1973). Evaluations were performed using the most recent data tape.
16 White, K. O., Watkins, W. R., Bruce, C. W., Meredith, R. E. and Smith, F. G., 'Water vapor continuum absorption in the 3·5–4·0 μm region', *Appl. Optics*, **17**, no. 17, pp. 2711–20, 1978.
17 Shumate, M. S., Menzies, R. T., Margolis, J. S. and Rosengreen, L. G., 'Water vapor absorption of carbon dioxide laser radiation', *Appl. Optics*, **15**, no. 10, pp. 2480–8, 1976.
18 Boscher, J., 'Entwicklung eines flugzeuggetragenen Messsystems zur Erkundung der Erdoberfläche und Atmosphäre: Analytische Untersuchungen', BMFT Contract 01 TL 026-AK/RT/WRT 2074, 1978.
19 Murray, E. R. and van der Laan, J. E., 'Remote measurement of ethylene using a CO_2 differential-absorption lidar', *Appl. Optics*, **17**, no. 5, pp. 814–7, 1978.
20 Pasquill, F., 'Atmospheric Diffusion' (Van Nostrand, New York, 1974.
21 Sutton, O. G., 'A theory of eddy diffusion in the atmosphere', *Proc. R. Soc.*, Series A, **135**, pp. 143–65, 1932.

Manuscript received by the Institution on 4th February 1980
(Paper No. 1961/MI 15)

UDC 621.394.142.6: 681.3: 778.38
Indexing Terms: Computer applications, Imaging, Aerial photography, Satellite photography, Photogrammetry.

Computer processing of multi-layer imagery data

T. ORHAUG, Ph.D.*
S. I. ÅKERSTEN, B.Sc.*
and
J. LARSSON, Ph.D.†

SUMMARY

The technique and methodology of computer-aided image processing has become a valuable tool for handling, processing and analysing multi-spectral imagery data acquired from sensors in aircraft and satellites. This paper describes some of the background and the rationale for performing computer processing and also discusses types and kinds of operations which come into use. The main purpose of the paper is to give a description of a software system for performing various tasks in computer-aided image processing. One interesting part of this system is a supervised classification program system for pixelwise (or objectwise) multi-layer classification. The classification algorithm performs class assignment in a linearly transformed space. It is a classifier whose c.p.u. time is approximately linearly dependent upon the number of layers (e.g. spectral features) compared with an approximate quadratic dependence for a classical implementation of the maximum-likelihood Bayes classifier for normally-distributed features.

Examples of pilot and application studies of various problems in environmental monitoring and land use are also given. Forest monitoring is considered to be an important application in Sweden. The paper also shows the importance of incorporating non-spectral (topographic) data for increased classification accuracy of detecting clear-cut areas.

* *National Defence Research Institute, Box 1165, S-581 11 Linköping, Sweden.*
† *Department of Photogrammetry, Royal Institute of Technology, S-100 44 Stockholm, Sweden.*

1 Introduction

Computer-aided techniques have become important tools for handling, processing and analysis of imagery data. This special type of computer application may be considered from many different viewpoints: pattern recognition, signal theory, applied mathematics, scene analysis, computer vision to name just a few. The methods developed are now being applied to various fields and for many different types of imagery data with different success. The hopes one may have had a decade ago of 'automatic' picture interpretation have not been so easy to achieve and to a large degree the methods being applied and the solutions found are problem-dependent. One area in which the computer has been relatively successfully introduced and applied is the area of handling, processing and analysing multi-spectral imagery in remote sensing. The reasons for this are interesting to note and they are mainly twofold. First, the multi-spectral scanner (m.s.s.) imagery data acquired by sensors on aircraft or satellite platforms are often digitized in the process of recording. Computer systems are therefore necessary for image handling and image presentation. Since the data are in digital form, they are also adapted for computer handling and computer analysis. Secondly, the multi-spectral recording of image data makes it possible to utilize the spectral characteristics or the spectral signatures of objects and ground properties for analysis purposes. To use the spectral features is far more direct and simple than using geometrical and spatial properties as are used in vision for object detection and identification. The application of pattern recognition methodology is thus more simple and uncomplicated. The spectral features are more well-defined and also limited in number (defined by the number of spectral channels in the sensor) in comparison with the spatial features. It is therefore not surprising that particularly in applications fields utilizing *Landsat* and similar multi-spectral data, computer processing methods have become very significant tools indeed. This field is probably at present one of the fields where image computer processing methodology has advanced furthest in the application phase and consequently a number of commercial systems are being offered for such applications.

The present paper aims at reviewing some of the image processing problems encountered in remote sensing, particularly when dealing with m.s.s. data. The problem of pixelwise classification is dealt with in some detail. In this context we describe a very time-efficient classification algorithm and also a flexible program system for performing classification. Furthermore, examples of applying supervised and unsupervised classification are presented as well as an example of the value of incorporating auxiliary data with the spectral data by the addition of terrain topographic data for the problem of detecting clear-cut areas in forestry regions.

The Radio and Electronic Engineer, Vol. 50, No. 11/12, pp. 575–584, November/December 1980 0033–7722/80/110575+10 $1.50/0

2 Image Processing

Image processing has evolved into a very large and diversified field and a complete review is therefore beyond the scope of this paper. A number of textbooks reviewing the field or addressing specific aspects are available.[1-4] We shall therefore limit ourselves to some few general remarks. The structuring of the field of computer image processing can be achieved in many different ways. On a very gross level we may consider the purpose of the operation we apply to an image in the way it effects the output from the processor:

image-to-image transformation for visual presentation,
image-to-image transformation for further computer processing,
image-to-decision transformation (per pixel or per object)
image-to-description transformation (cf. scene analysis, image understanding or computer vision)

Another way of structuring is to consider the purpose of the operation applied to the image:

image coding (image → image)
image enhancement (image → image)
image restoration (image → image)
image classification (image → image
or image → description)
image analysis (image → description)

A more detailed discussion of the various processing schemes and operators for implementing the various schemes has been presented elsewhere.[5]

3 Processing of Multi-layer Imagery Data

We shall limit our discussion to multi-spectral images acquired with scanning spectrometers of the type used in the *Landsat* and *Tiros* series of satellites or in aircraft. An exemplifying structure of the various procedures and steps which might be incorporated in a remote sensing image processing and information integration system is shown in Fig. 1. A number of problems indicated in Fig. 1 are of importance for various parts of the information handling and processing system. There is a tendency to locate more and more of the processing functions with the data acquisition part of the system (i.e. preprocessing using sensor and platform data (cf. the *Landsat* data acquisition and distribution system)). For more local studies using aircraft m.s.s., the user is often required to do most of the preprocessing himself.

3.1 Software System for Handling and Processing M.S.S. Data

A cooperative research and application-oriented project was conducted between 1972–1978 in order to investigate the potential application of multi-spectral scanner images for Swedish conditions and the project started using *Landsat 1* digital data. In the later part of the project, analogue and digital VHRR data from NOAA satellites and analogue m.s.s. data from a 10-channel airborne m.s. scanner were utilized. The aim of the project was to introduce an awareness of the m.s.s. images as an important data source and also to introduce the importance of the digital computer for handling and processing m.s.s. data. The ways of achieving these aims have been:

(1) to develop software routines for transforming analogue m.s.s. data to digital data having special and suitable format needed for the subsequent software system for m.s.s. processing;
(2) to develop a batch-oriented general-purpose software system for handling, processing and

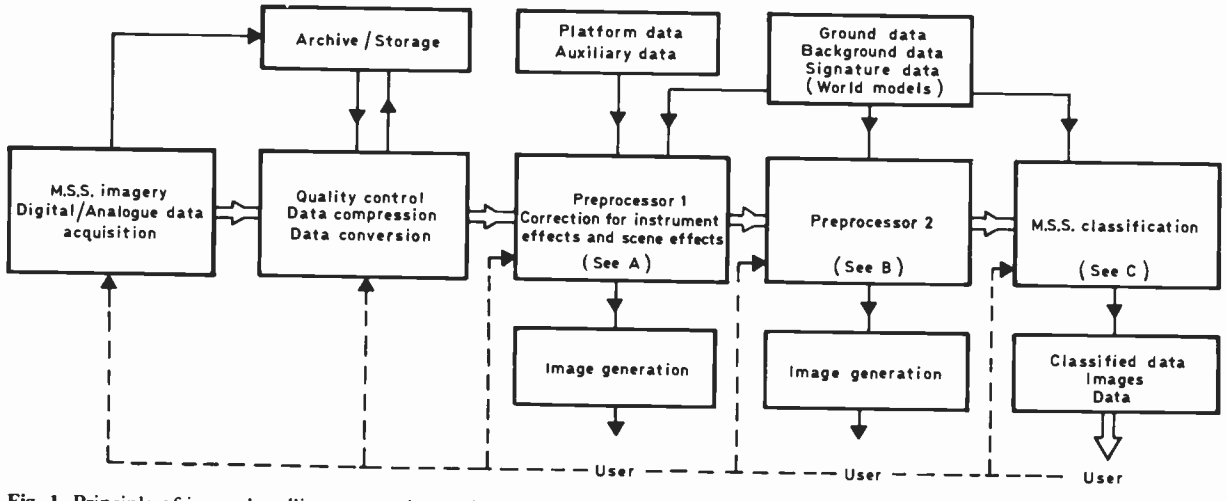


Fig. 1. Principle of image handling, processing and analysis of imagery multi-spectral data in remote sensing.

A: Sensor corrections
Normalization
Atmospheric corrections
Geometric corrections

B: Contrast stretching
Density slicing
Colour composites
Image smoothing
Image sharpening
Geometric registration
Geometric transform

C: Pixelwise classification
Image segmentation → classification
(using spectral and other data)

analysis of m.s.s. data on a general-purpose computer (IBB 370–165);
(3) to develop routines for geometric correction of *Landsat* and airborne m.s.s. data;
(4) to convert the main part of this system to remote terminal dialogue systems (IBB 370–165 (using a time-sharing terminal system) and DEC 10);
(5) to develop handling and processing routines on a dedicated computer (PDP 11/34-system with special purpose hardware for image I/O and image display);
(6) to apply the systems above to well-defined application-oriented studies in various fields of interest (forestry, environmental monitoring, vegetation studies, land-use studies etc.).

The software systems indicated above have been documented in various reports.[6–8]

3.2 Multi-spectral Classification

The use of computer analysis in remote sensing is very much based upon general methods of pattern recognition (see Refs 4, 5). We shall limit our introductory discussion to the case when spectral features are used although the methods utilized are not themselves subject to this restriction. Among the methods which come into use we may distinguish between the two main cases of classification, supervised and unsupervised. In the first case, regions in the image are identified according to the class they belong to (e.g. known vegetation). Characteristic data from these regions are then calculated and used as training areas to train the classifier. The detailed properties needed for the classification algorithms are thus determined from the training samples. In other words, the multi-spectral space is divided into decision regions and the boundaries of these regions are determined from the training phase (cf. Fig. 2).

The supervised methods in turn may be divided into the two groups of parametric and non-parametric methods. In the parametric methods one assumes that the functional form of the data cluster for each class is given as a probability distribution. In the non-parametric case, on the other hand, the functional form of the discriminant (decision) function is given. The training phase then amounts for the former case to the determination of the parameters of the probability distribution and as a consequence thereof, the shape and the position of the discriminant function, and for the latter case, the position of the discriminant function. A more detailed discussion of the various methods in pattern recognition can be found in the literature.[4,5]

Two different types of parametric methods come into frequent use: piecewise constant probability distribution (density) leading to hyperbox classification which normally is implemented with table look-up programs and normal (Gaussian) multivariate probability density leading to either second-order or to linear discriminant functions.

3.3 A Hybrid Algorithm for Supervised Classification

The classification algorithm implemented and being extensively used in various application studies can be labelled a hybrid classifier and it contains the following features:
(1) a simple threshold (a one-dimensional box classifier) for the purpose of discriminating water/non-water or similar uniquely defined classes,
(2) a multi-dimensional rectangular box classifier,
(3) a canonical classifier assuming Gaussian statistics and ordered features,
(4) a combination of (2) and (3) above,
(5) a local averaging 'filter' generalizing the classification result.

For a schematic illustration of the total classifier, reference is made to Fig. 3.

The hybrid classifier is a flexible classification system. It uses spectral (or other type) data for classifying

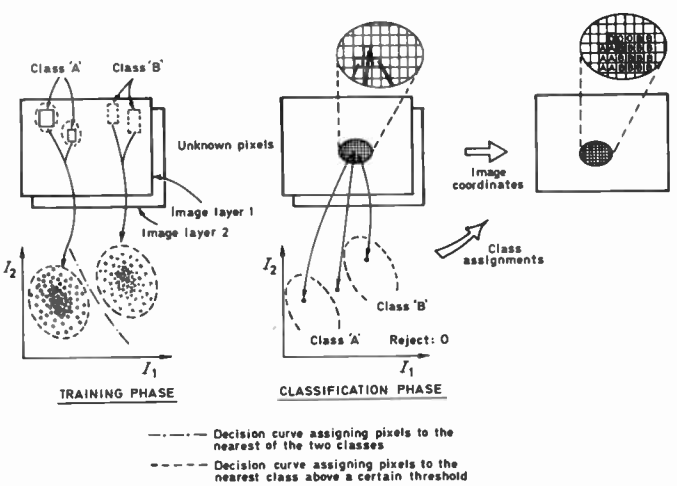


Fig. 2. Depicting the training and classification phases of pixelwise multi-spectral analysis.

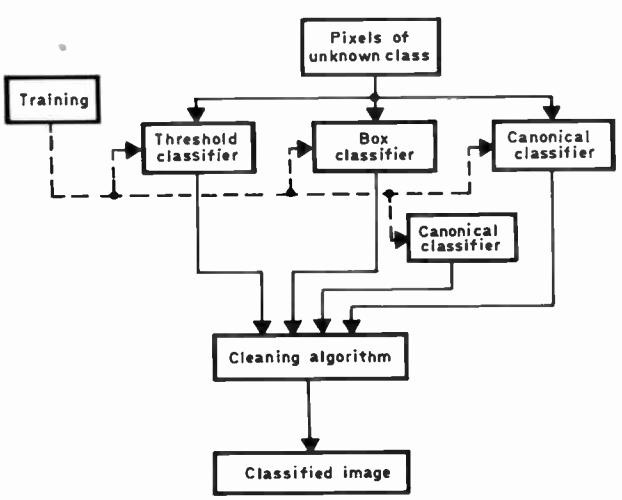


Fig. 3. Schematic flow diagram illustrating the classification program structure.

pictorial data in an efficient way. If one spectral band is sufficient for decision or class-assignment it uses one band only, and if data are 'well-behaved' so that rectangular cluster boundaries are sufficient for describing the data clusters, it reaches a decision using rapid table-look-up algorithms. Only for ambiguous cases above does it use full maximum-likelihood algorithms assuming Gaussian data statistics. But even for this more elaborate algorithm a decision is made in a very rapid way compared with 'ordinary' classifiers since the decision is made in a linearly transformed space where the transformed spectral features or values are ordered with respect to their efficiency to reach a decision. The theory of this classifier was published by Dye and Chen.[9] A detailed description of the method is outside the scope of this paper; for further details both of the mathematics and of the algorithmic implementation, reference is made to recent publications.[7,8]

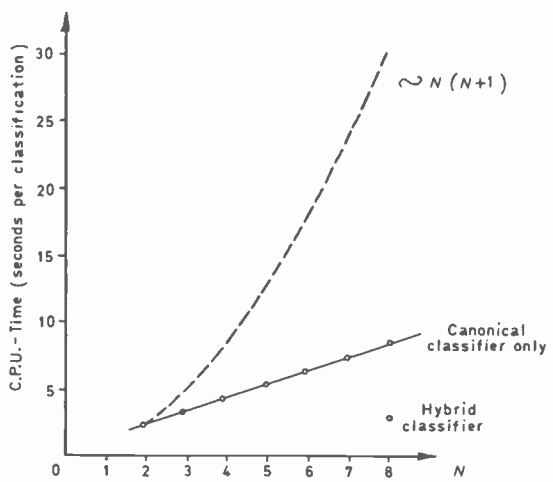


Fig. 4. Result from computer experiment showing c.p.u. time for classification using different numbers of channels for classification (canonical and hybrid classifiers compared with 'classical' implementation).

An interesting and important feature of the canonical classifier will, however, be mentioned here. The 'ordinary' implementation of the normal statistics classifier leads to a computational time (c.p.u.-time) proportional to $N(N+1)K$ where N is the number of channels ('features') and K is the number of classes used in the classification procedure. For $N \gg 1$ this leads to c.p.u.-time proportional to $N*N$, or a drastic increase in classification time for increasing the number of channels used. One main advantage of this canonical classifier is that the corresponding dependence of c.p.u.-time is linear, or proportional to NK. This is illustrated in Fig. 4 where c.p.u.-times for actual classification of a 256*256 subimage from a 10-channel m.s.s.-scene acquired by an airborne sensor is shown. Classification was carried out using various numbers of channels with the same training data. (This computer experiment is described in some detail elsewhere.[7]) In the same diagram is also illustrated the extra gain obtained with the hybrid classifier (compared with the canonical alone). If the data had contained pixels for which thresholding could be applied (as often is the case for 'normal' scenes), an even more appreciable gain would have been obtained.

3.4 Unsupervised Classification

In addition to the supervised classification methods described above, two various kinds of unsupervised methods have also been implemented and applied to m.s.s. imagery data. The first method is the 'classical' clustering method (hierarchical clustering) in a version described by Swain.[4] In this method, data are assigned to a particular (unlabelled) class or cluster if the data points in the N-dimensional representation are 'close' enough. The number of clusters wanted are given beforehand and also a termination rule.

The other method for unsupervised classification is a 'nonclassical' method by which data in a high-dimensional representation (N-space) can be transformed or mapped into a representation of lower dimension (M-space) by preserving data 'tightness' or clustering tendency. The advantage of this method is that no information (numbers of clusters, etc.) need to be introduced. Another advantage is that the computational time is independent of the ratio N/M. The method ('nonlinear mapping') has been described and tested on m.s.s.-data.[10]

4 Examples from Application Studies

As described earlier, the software system has been utilized in various pilot and application projects in Sweden. These projects cover the following areas of interest:

detection of clear-cut areas in forests,
identification of tree stands in forests,
study of water vegetation,
land use mapping,
water quality,
water ecology.

4.1 Local Filtering

We exemplify the use of a local, two-step filtering algorithm with the use of symbol-coded, pictorial illustration. The purpose of the local filter is to remove the 'graininess' which almost always results from pixelwise classifications. In its first step a 'majority vote' decision is made in a local neighbourhood, say 3×3, around the pixel in question. Should a user-selected number of pixels, say 5 or more, inside this window belong to one given class, the centre pixel is assigned to that class regardless of its original class-assignment. In its second step possible isolated pixels are assigned to that class which is most frequently occurring within the local, say 3×3 neighbourhood. Figure 6 shows the actual class assignment for a subscene from an application project

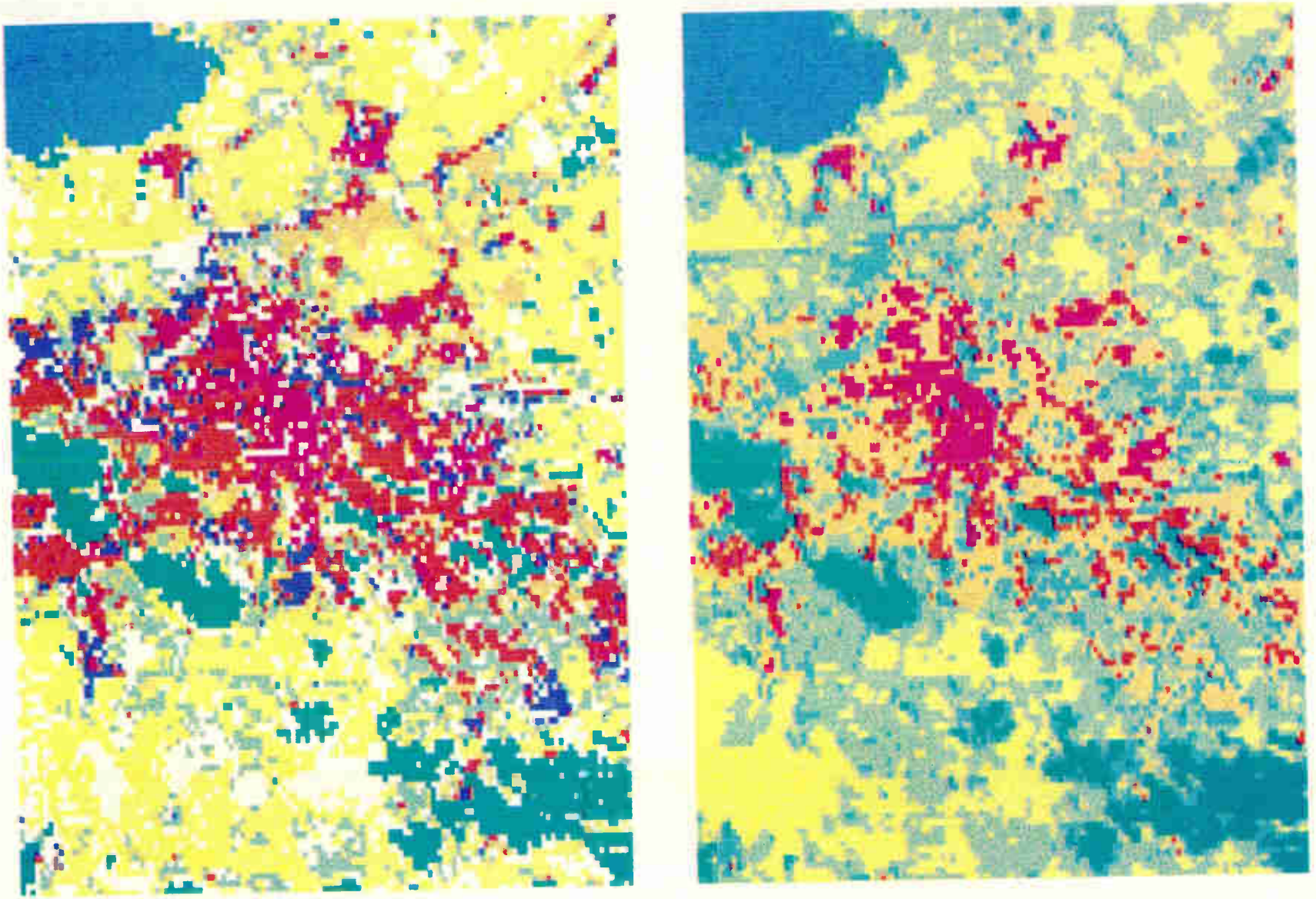


Fig. 5. Result of supervised (left) and unsupervised classification (right) applied to the same four-channel Landsat data. For details, see text. The colour representation used are: blue—water, dark green—forest, yellow—agriculture, reddish—urban.

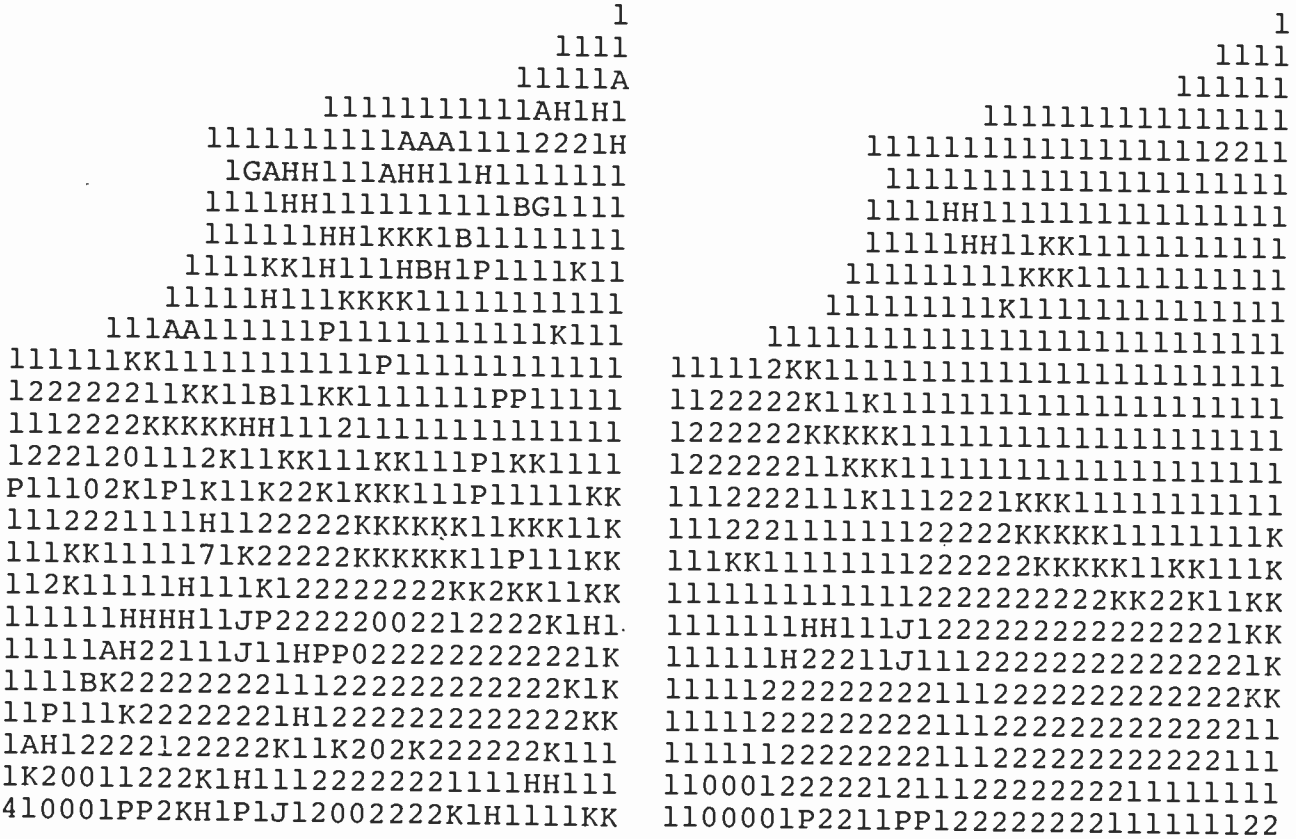


Fig. 6. Showing the effect of the cleaning algorithm by which isolated pixels will have class-assignment changed. The left-hand picture shows the pixelwise classification result in a symbol-coded form and the picture window to the right the area after a (two-step) local filtering.

before (left) and after (right) applying this CLEANUP filtering algorithm.

4.2 Supervised and Unsupervised Classification

Another example is selected to show a comparison of supervised and unsupervised classification using the same data. The example chosen is from land use studies using four-channel *Landsat* data (scene ref. 2162-09292, 75-07-03 (209, 19)) for investigating the Linköping area.[11] The supervised classification was done by training on 12 different areas and in Fig. 5. they have been grouped into 9 classes in displaying the classification results. The colour code used for image display is also indicated. The same image data were used as input to the algorithm for unsupervised classification into 14 clusters. When clustering was completed, 7 clusters were identified by comparing with the results from the supervised classification, and where possible, colours similar to those previously used were also applied to this latter result.

The data presented in Fig. 6 have been geometrically corrected for Earth-rotation. The pictorial result of Fig. 5 has been produced by displaying the digital data on a television-monitor using a refresh memory while the pictorial colour printouts of Fig. 7. and 8 have been made on ordinary white paper using a colour ink jet plotter.[12]

The two classification results presented in Fig. 5 show approximately the same geometrical pattern. The results from the unsupervised classification scheme consequently generate a similar result as that from the supervised classification scheme.

4.3 Supervised Classification using Spectral and Non-spectral Data

One of the first, and also one of the main studies carried out in the project described above was directed towards the investigation of the m.s.s.-technique applied to forestry problems. In the first study[11,13] it was found that in the supervised classification process regeneration areas with a preponderance of young deciduous weeds had not been separated from areas dominated by young coniferous stands. Also, the apportionment between regeneration areas with and without predominance of coniferous saplings was not entirely correct. The latter class contains regeneration areas where the contribution of reflected light from deciduous shrubs overpowers that from coniferous plants as a result of their relative frequency of occurrence. Also, it was found that the discrimination between regeneration areas and areas of bog was difficult due to the similarity between their spectral signatures. It has been shown that discrimination of vegetation may be improved by carrying out multi-temporal classification, e.g. using two scenes from different times during the vegetation period.[14] Another way has been suggested and applied by Larsson[15] utilizing the fact that the topographic

terrain variation of clear-cut areas often tends to be larger than that for bog areas. In this experiment, aircraft scanner data, rectified to register to the National Land grid, were used. This allowed the combined use of image and terrain height data. The idea here was to distinguish between water infested areas (bog) and other open areas (especially clear-cut forest areas), by using the fact that bogs tend to occur more frequently in flat terrain compared to clear-cut forest areas.

To be able to use this in an automated classification procedure, the terrain heights were transformed into terrain roughness values. These were computed for each pixel as the average absolute value of the difference in terrain height in the nearest four pixels surrounding the actual one. These values were then included in the original multi-dimensional image as an extra image layer and thus used in the subsequent supervised classification. In order to do this inclusion of an extra topographic layer in the m.s.s. data, the multi-spectral data have to be geometrically corrected to perfect registration with the coordinate system of the height information (see Larsson[15]). The results from this classification, compared to the classification without terrain information included, can be studied in Figs. 7 and 8. Table 1 is a comparison with ground truth data obtained from the economic map and the forest map. All numbers are in percentage of the area of the true class. An increase in classification accuracy can here be observed.

Table 1.

Improvement in classification when using digital terrain model. The image is classified into classes as follows: B: bog, CC: clear cut, F: forest, O: open areas, W: water and R: rejects. This classification is compared with the areas of the true classes, obtained from the economic and forest maps for the following classes: forest, clear cut and bog, and the numbers in the table are in percentage of the true areas

	Class	B	CC	F	O	W	R
Without digital terrain model	B	48	11	29	10	0	1
	CC	40	29	22	7	1	1
	F	5	1	92	2	0	0
With digital terrain model	B	59	8	19	13	0	0
	CC	23	48	16	12	0	1
	F	5	3	89	3	0	0
Improvement	B	+11	−3	−10	+3	0	−1
	CC	−17	+19	−6	+5	−1	0
	F	0	+2	−3	+1	0	0

However, there also exist other kinds of bog, not necessarily located on flat surfaces, where this simple measure of terrain variation will be misleading. A better approach would be to use terrain heights to construct an image layer of terrain types, such as convex, concave, slope of different degree and heading, and use this layer in the classification or in a following interpretation procedure.

The investigation discussed above shows, however, that remote sensing can not exclusively rely on spectral signatures. Other properties of the objects such as geometric data must be considered. When such data can be incorporated, a significant improvement in classification accuracy may result.

5 Summary of Experience using M.S.S. Data

Based upon several years of experience of using m.s.s. data we may summarize the experience as follows:[14]

1. The availability and overall quality of the digital m.s.s. data from the *Landsat* system has been of significant importance for the introduction of computer techniques to the user community.
2. For Swedish natural resource management inventory problems, forest inventory and land use mapping seem to be promising applications for *Landsat* data. Classification accuracy should, however, be improved, either by applying multi-data analysis, improved resolution and/or systematic use of complementary databases.
3. Operational uses of *Landsat* data in forest inventory and land use mapping may be of significance in Sweden first when the improved resolution of *Landsat D* will become available.
4. The airborne m.s.s. data have higher cost figures, more severe geometrical and radiometrical distortions than *Landsat* data.
5. As a result of higher distortions, classification accuracy for the airborne data decreases drastically towards the edges due to the scan-angle effect.
6. In the central region the accuracy of the classified airborne m.s.s. data is significantly higher than for *Landsat* data due to the better resolution.
7. The airborne m.s.s. data collection system will probably not be important for operational work, due to costs and distortions. As a research tool the technique is, however, useful.

6 Acknowledgments

In the work reported here a number of people have been active in various phases. The cooperation with Dr. L. Wastenson from the University of Stockholm has been most important through many phases of the project, particularly in the applications parts. Erik Carlsson and Sven Seeman have assisted both in the programming work and in providing illustrative examples of supervised and unsupervised classification of imagery data. We also acknowledge financial support given by the Swedish Board for Space Activities for various parts of the project.

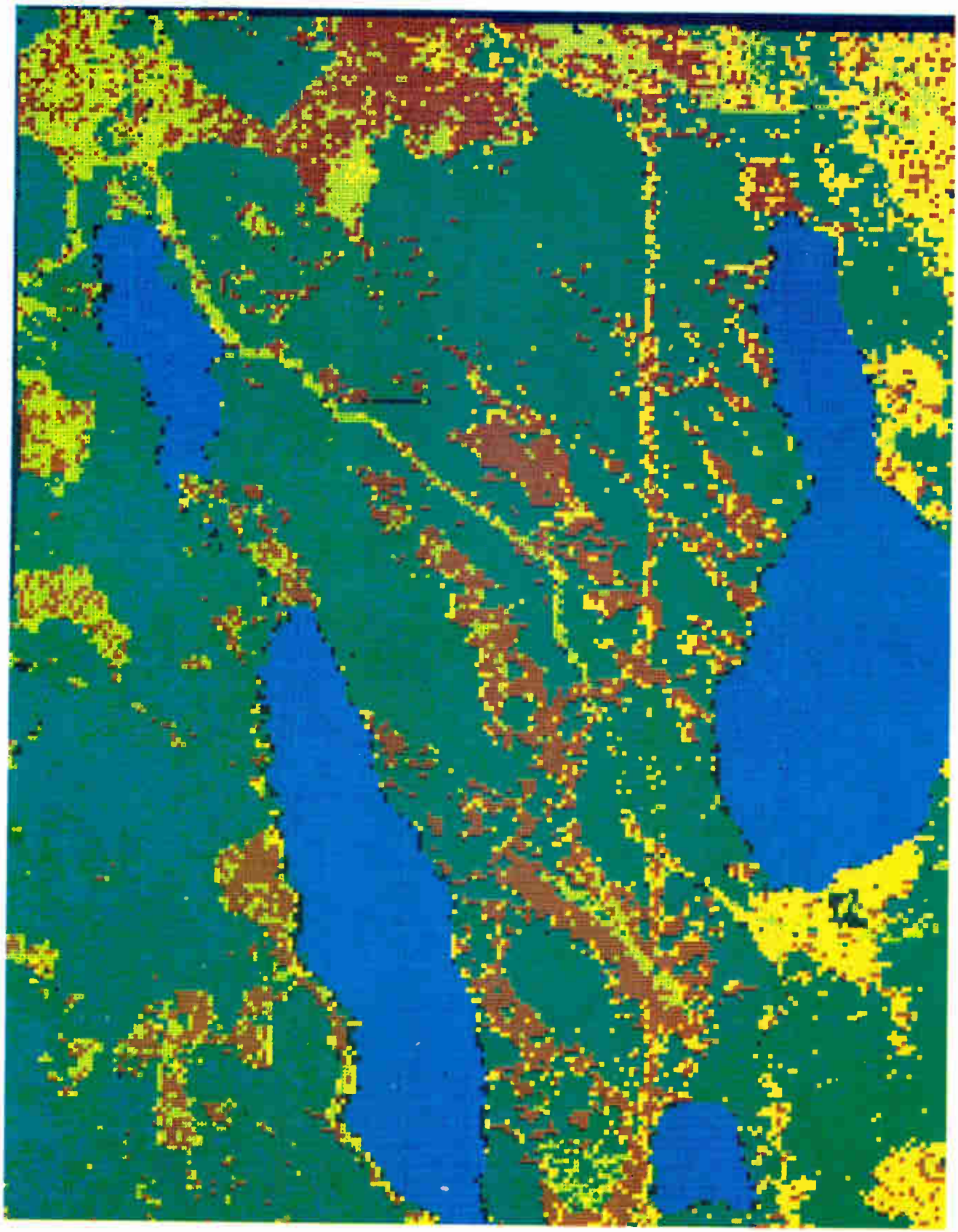


Fig. 7. The original pixelwise classification of ten channels aircraft m.s.s. data, using eight out of the ten available spectral channels. The pixel size on ground is 15 metres. The separate classes were assigned the following colours:

light green—clear cut, dark green—forest, brown—bog, yellow—agriculture, blue—water, black—rejects.

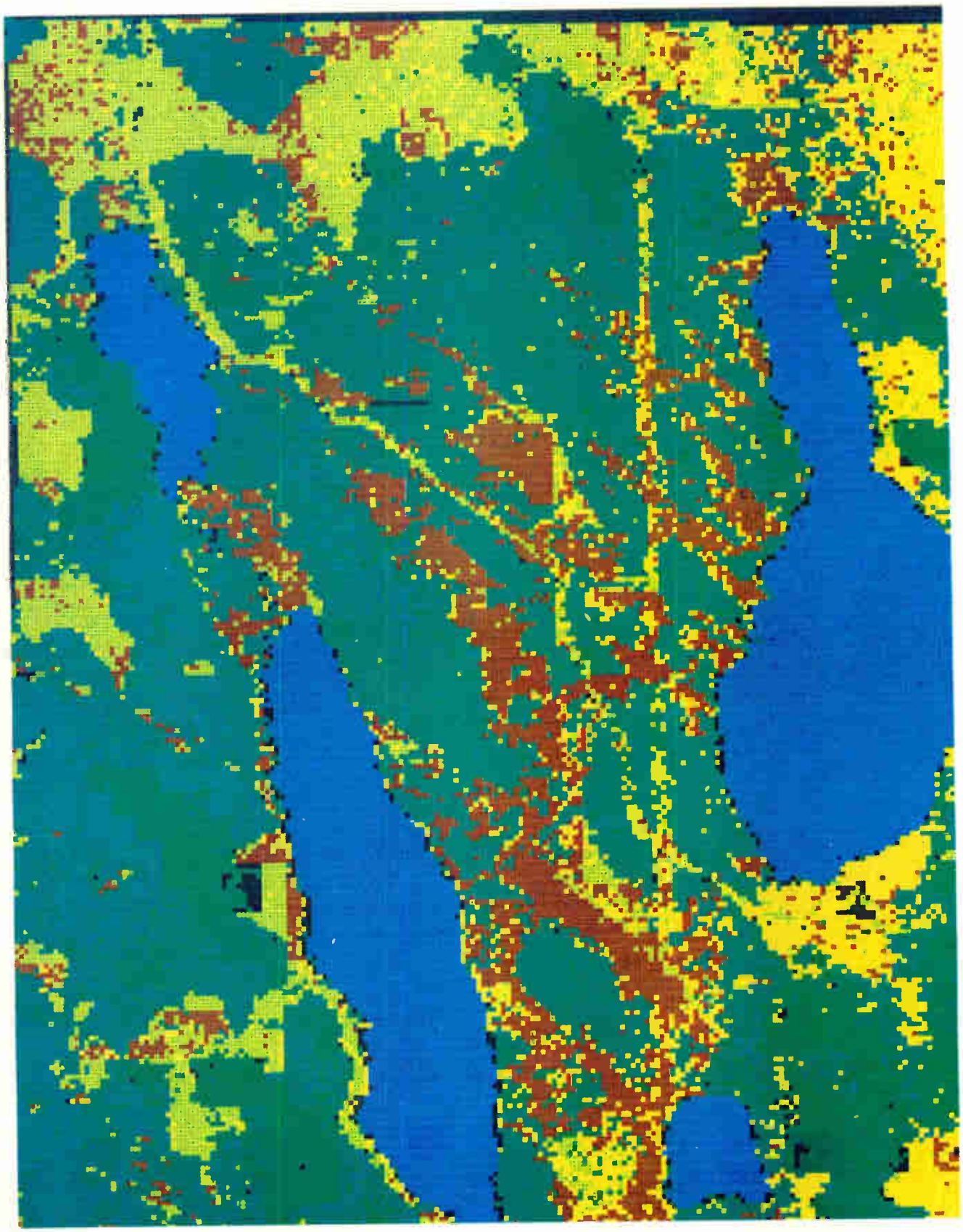


Fig. 8. Classification result using both eight-channel spectral data and terrain information. Note especially the clear-cut area in the north-west (*Paper No. 1962/Comp. 197*).

7 References

1 Rosenfeld, A. and Kak, A. C., 'Picture Processing' (Academic Press, New York, 1976).

2 Duda, R. O. and Hart, P. E., 'Pattern Classification and Scene Analysis' (Wiley, New York, 1973).

3 Pratt, W. K., 'Digital Image Processing' (Wiley, New York, 1978).

4 Swain, P. H. and Davies, S. M. (Editors), 'Remote Sensing, The Quantitative Approach' (McGraw-Hill, New York, 1978).

5 Orhaug, T., 'Pattern recognition—with special emphasis on image processing', in Lund, T., (Editor), 'Surveillance of Environmental Pollution and Resources by Electromagnetic Waves' (Reidel, Dordrecht, 1978).

6 Åkersten, S. I. and Gustafsson, L-E., 'Program Library for Handling and Processing of Remotely Sensed Multispectral Data'. FOA report C 30146-El, September 1978.

7 Åkersten, S. I., Gustafsson, L-E. and Orhaug, T., 'MSS-75: Data Handling and Data Processing'. FOA report A 30019-El, August 1978.

8 Orhaug, T. and Åkersten, S. I., 'A program system for efficient multispectral classification', Proc. of the First Scandinavian Conference on Image Analysis, Linköping, Sweden, January 1980.

9 Dye, R. H. and Chen, C. S., 'Divergence analysis of Bendix feature extraction and classification system', Symp. on Machine Processing of Remotely Sensed Data, June 3–5, 1975 Purdue University, West Lafayette, Indiana, USA.

10 Orhaug, T. and Severinsson-Eklundh, K., 'Clustering by nonlinear mapping', Proc. 3rd Int. Joint Conf. on Pattern Recognition, November 8–11, 1976 Corona, Calif., pp. 386–393. (IEEE Cat. No. 76 CM1140-3C; FOA reprints B 30018-El, 1976/77:8).

11 Wastenson, L., Orhaug, T. and Åkersten, S. I. 'Swedish experiences in forest inventory and land use mapping by automatic classification of digital MSS data from *Landsat* and aircraft'. Proc. Int. Symp. on Remote Sensing for Observation and Inventory of Earth Resources and the Endangered Environment, July 2–8, 1978, Freiburg, Fed. Rep. of Germany.

12 Hertz, C. H. and Orhaug, T., 'The ink jet plotter: a computer peripheral for producing hard copy color imagery', *Comp. Graph. Image Proc.*, **5**, pp. 1–62, 1976. (FOA Reprints 1976/77:3).

13 Orhaug, T., Wastenson, L. and Åkersten, S. I., 'Forest Inventory Using Landsat Digital Imagery'. FOA Report A 30008-El, October 1976).

14 Haberäcker, P., 'Zur Klassifizierung Multispektrale Bilddaten aus der Fernerkundung' (Technische Universität, Berlin, 1978).

15 You, K. C. and Fu, K. S., 'An approach to the design of a linear binary tree classifier, Proc. Symp. Machine Processing of Remotely Sensed Data, Purdue University, 1976, (IEEE Cat. No. 76CH 1103.1 MPRSD)

16 Larsson, J., 'Rectification of Digital Images for Remote Sensing Analysis', Fotagrammetrika Middelaman 2:43, Dept of Photogrammetry, Royal Institute of Technology, Stockholm, 1980.

Manuscript first received on 22nd February 1980 and in final form on 8th May 1980. (Paper No. 1962/Comp. 197).

UDC 551.501.796: 551.513
Indexing Terms: Atmospheric movements, Atmospheric structure, Sounding apparatus, Acoustic generators

Some aspects of the construction and use of atmospheric acoustic sounders

R. S. COLE, Ph.D.*,
D. N. ASIMAKOPOULOS, Ph.D.†,
T. J. MOULSLEY, Ph.D.‡
and
S. J. CAUGHEY, Ph.D.§ and
B. A. CREASE, M.Sc.§

SUMMARY

Acoustic sounder design and performance parameters are described. Various atmospheric features as recorded by the acoustic sounder, including thermal plumes, inversions and waves, are presented and their interpretation given. The use of the acoustic sounder to monitor the height of low level layer cloud and the depth of radiation fog is discussed as well as the detection and tracking of plumes of methane gas.

* *Department of Electronic and Electrical Engineering University College London, Gower Street, London WC1E 7JE.*
† *Formerly at University College London; now with the Department of Meteorology, University of Athens.*
‡ *Formerly at University College London; now with the Department of Electrical and Electronic Engineering, Portsmouth Polytechnic.*
§ *Meteorological Office, London Road, Bracknell, Berkshire RG12 2SZ.*

1 Introduction

The development of acoustic sounding of the lower atmosphere proceeded rapidly after the early work of McAllister[1] in 1969. The principle of the technique of so-called monostatic acoustic sounding is straightforward; a burst of sound is projected into the atmosphere in a well-collimated beam using a suitable acoustic antenna. Antennas commonly used include an array of loudspeakers,[2] a single loudspeaker at the focus of a reflecting dish or sometimes a simple horn loudspeaker. The antenna is switched to the listening mode after transmission when the loudspeakers have ceased ringing significantly. Received echoes are amplified and displayed, for an immediate visual record, in height–time form on a facsimile chart recorder, the intensity of the record giving an indication of the strength of the received echo. Since the same antenna is used for both transmission and reception only the backscattered sound is detected. The transmitted frequency is normally chosen as a compromise between the effect of the atmospheric absorption of sound, which increases at higher frequencies and ambient noise, which has a greater effect at lower frequencies.

Facsimile records presented in this work were produced by different sounders operating at four frequencies between 820 Hz and 6·5 kHz.[3,4] Transmitted acoustic powers from these systems range between 30 W at 820 Hz and 1·5 W at 6·5 kHz with pulse lengths varying between 5 ms and 74 ms corresponding to height resolutions from about 1 m to 12 m.

Acoustic sounding can offer a method of remotely estimating important turbulence quantities, such as the structure parameters for wind (C_V^2) and temperature (C_T^2) at any chosen height,[5] since the eddies producing the scattered sound are of a length scale found within the inertial sub-range of atmospheric turbulence.[6] Additionally the echoes provide information on the wind structure in the boundary layer through the magnitude of the Doppler shift. The accompanying paper* deals exclusively with the use of the acoustic sounder to provide quantitative estimates of these parameters. The object of this paper is to show that useful information can be obtained from the visual examination of monostatic sounder facsimile charts. In general terms the monostatic sounder will record any atmospheric structure having an associated fine scale turbulence which generates significant refractive index fluctuations, predominantly due to temperature fluctuations at acoustic wavelengths, on a scale approximately equal to that of the transmitted sound half wavelength, i.e. between about 2·5 and 20 cm for the range of sounders used in this work. Thus it will readily indicate the

* Caughey, S. J., Crease, B. A., Cole, R. S., Asimakopoulos, D. N. and Moulsley, T. J., 'Quantitative interpretation of acoustic echoes from the planetary boundary layer,' *The Radio and Electronic Engineer*, **50**, no. 11/12, pp. 598–610, November/December 1980.

0033–7722/80/110585+13 $1.50/0

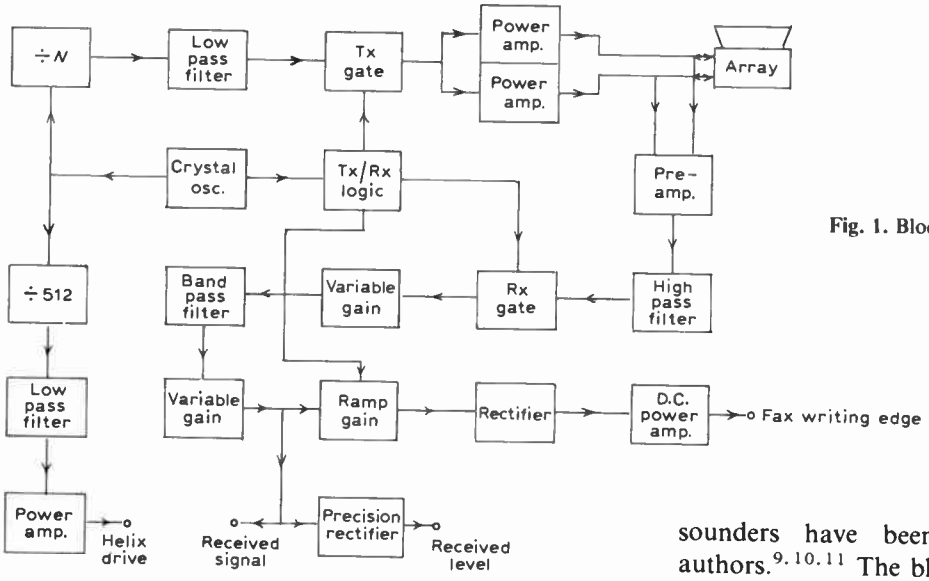


Fig. 1. Block diagram of an acoustic sounder system.

presence of ground-based thermal activity, synoptic and nocturnal inversions, breaking waves and the turbulent thermal structure associated with fog. Examples of these features are discussed in the following Sections.

Certain factors must be borne in mind when one attempts to relate the structure indicated on a monostatic sounder facsimile chart to the turbulent activity in the atmospheric boundary layer. The chart illustrates only that part of the atmosphere being advected above the sounder antenna, hence the sounder provides only a 'height–time' section through any structure and so inferences about the three-dimensional shape cannot be attempted nor can the Lagrangian evolution of any features of interest be usefully discussed. The correspondence between the chart record and the actual structure will depend on the gain and the dynamic range of the system. For example, the location of the edges of a thermal plume may be ill-defined. Recent experimental work has shown that the intensity of small-scale temperature fluctuations falls off rapidly with height.[7] This means, for example, that a monostatic sounder, typical of those in general use, will significantly underestimate the upper limit of convection in deep boundary layers. Indeed an examination of the charts reveals an apparent 'fading' of turbulent thermal activity with height.

In describing the application of acoustic sensing to boundary layer meteorology in this and the accompanying paper, we have drawn upon field studies conducted jointly by the Department of Electronic and Electrical Engineering, University College London and the Meteorological Research Unit, Cardington. Further more general information on the progress in atmospheric acoustics in the last decade may be found in the detailed review paper given by Brown and Hall.[8]

2 Acoustic Sounder Design

The considerations involved in the design of acoustic sounders have been discussed by a number of authors.[9,10,11] The block diagram of a typical sounder system[12] is shown in Fig. 1. The transmitted frequency is derived from frequency division of the output of a 32·768 kHz crystal oscillator. The same oscillator is also used to generate the timing signals for the control of the sounder transmission and reception, including the pulse length and pulse repetition frequency. The transmit gate is used to generate a tone burst of the required duration which is amplified and applied to the antenna which produces a narrow beam of high acoustic intensity.

In the case of the monostatic sounder a common transmit/receive antenna is used and a passive diode gate is employed to protect the sensitive receiver circuits from the high power transmit signal. The receiver also incorporates an active gate which opens after the 'dead time' resulting from the antenna ringing after transmission. The transmit and receive gates use c-m.o.s. analogue switches which have low distortion and high off-state rejection. The transmit-receive logic also uses c-m.o.s.

The received signal is amplified and filtered to remove out-of-band noise components and is detected using a precision rectifier having a dynamic range of about 50dB. The output of the rectifier can be recorded for quantitative analysis and can also be used to provide a visual display, in this case on a facsimile recorder. A time varying gain is used to compensate for the loss of amplitude resulting from spherical divergence of the acoustic echoes as they are received from longer ranges. The facsimile recorder helix drive is also derived from the crystal oscillator.

In order to produce detectable echo signals a high acoustic transmit power is desirable. This can be achieved with high electrical power to the antenna, implying high power handling capability and high antenna efficiency. For a high receive power it is necessary to use an antenna with a large collecting area and to operate at frequencies at which the atmospheric absorption is small, acoustic absorption increasing

rapidly with frequency. However, if the antenna is too large the beamwidth becomes too narrow resulting in significant sound energy being lost due to refraction by the wind.

It is usually possible to construct amplifiers with noise levels well below the environmental noise contribution, e.g. wind, traffic, aircraft, etc. However, in very quiet environments, the thermal noise in the receiver transducer may impose the limit on sensitivity. The intrusion of environmental noise can be minimized by good antenna design, e.g. low sidelobes and screening. Often a low-noise pre-amplifier is used in close proximity to the antenna to eliminate the problems caused by electrical pick-up when long cables (50–200 m) are used to connect the antenna to the receiver. The pre-amplifier incorporates a step-up transformer to match the antenna impedance of a few ohms to the input impedance of the receiver, a few kilo-ohms.

In the design of the receiver electronics the most important requirement is to separate the wanted narrowband acoustic signal from strong, usually broadband, noise signals and thus a narrow band filter is necessary. However, a filter bandwidth of at least a few per cent of the carrier frequency is usually required because of wind induced Doppler shifts and turbulent spectral broadening.

Most conventional acoustic sounders use a carrier frequency in the region of 1 to 2 kHz with maximum useful ranges from a few hundred metres to a few kilometers. Operation at higher frequencies, greater than about 4 kHz, has a number of advantages.[4] For example, external noise sources, such as traffic, often have little acoustic output at these frequencies which allows operation in a noisy town environment with a minimum of acoustic screening. Also the antenna is physically small which, besides the obvious advantage of being easily portable, means it has a short reverberation time, i.e. a short 'dead time' after transmission which allows sounding down to a few metres from the ground. However, these advantages are at the expense of reduced range because of the greater atmospheric acoustic attenuation at these frequencies.

3 The Acoustic Sounder Antenna

The main limitations in the performace of an acoustic sounder often depend to a great extent on the characteristics of the antenna, since present-day electronic techniques are capable of achieving almost ideal performance in a sounding system. However in typical sounders[13,14] there remains considerable scope for improvement in antenna design.[8]

The basic requirements for an effective monostatic antenna are:

(1) Narrow main beam to enable probing of a well defined atmospheric volume.
(2) Low sidelobe levels to minimize reflections from nearby solid objects and to reduce reception from external noise sources.
(3) High transmitting efficiency to allow high acoustic outputs using acceptable electrical input powers.
(4) High receiving efficiency (in many cases this is equal to the transmitting efficiency) to avoid degradation of the received signal to noise ratio by receiver noise.

There are three types of antenna in common use at frequencies in the range 1–5 kHz:

(1) Conical horn antenna
This consists of one or more transducers acoustically coupled to a large impedance matching horn. The aperture diameter is usually 1–2 m to give a beam width of 5–10° and this kind of antenna is reported to have an efficiency of up to 50%.[15] Very low sidelobe levels can be obtained with proper screening, e.g. at 90° off-axis, 50dB below axial level is possible.[16]

(2) Single horn and reflector
In this arrangement a transducer is mounted at or near the focus of a parabolic reflector. Acoustically damped microwave-type antenna reflectors of 1–2 m diameter are often used.[15] The sidelobe levels of a horn and reflector system may be determined by the radiation from the transducer itself and hence extra care must be exercised in constructing acoustic screening.

(3) Array antennas
Here the antenna consists of a number of transducers usually in the form of a uniform planar array. The array characteristics depend on both the array geometry and transducer properties and wide variations are possible in both of these areas.[4,17] A typical antenna consisting of 36 elements in an array 1 m × 1 m is described by Asimakopoulos and Cole.[13] The polar characteristics of an array can easily be adjusted by changing the array element separation. It is also possible to phase the drive to each element and use amplitude shading to control the polar diagram.

In quantitative sounding it is essential to know the polar diagram and the efficiency of the antenna. For Doppler wind measuring systems it is desirable that the antenna characteristics are known in order that measurement errors may be estimated.[18]

The estimation of antenna efficiency depends on an accurate knowledge of the polar diagram in the region of the main beam and this can be obtained by measurement with a microphone. The major difficulties encountered are in ensuring that the microphone is in the far-field region and at the same time that its position is known with sufficient accuracy. The necessary separation between antenna and microphone is often several metres.

The measurements can be made by suspending the microphone between masts[19] or from a tethered balloon and tracking with theodolites.[20] In order to determine the receiving sensitivity the suspended microphone is replaced with a calibrated sound source. Full details of the measurements on the arrays used in the quantitative work described in the accompanying paper can be found in Moulsley.[12]

4 Thermal Plumes

Thermal plumes are features present in convective conditions and have been observed by *in-situ* sensors for some time.[21] They originate in the superadiabatic region near the ground and normally extend to the top of the convective boundary layer.[22] Direct observations show that they are of great importance in the diurnal development of the boundary layer over land and are responsible for carrying aloft significant heat and moisture from the surface. Of particular importance in, for example, pollution studies is the time of onset of convection and the rate of change of the depth of the convective layer. Similarly in the evenings when convection dies down, turbulent mixing reduces dramatically and the ability of the atmosphere to disperse pollutants is much affected. Considerations such as these illustrate that a system which can remotely record the state of convection over long periods of time is likely to be of considerable usefulness.

The first atmospheric acoustic sounders showed strong echo regions in the lowest few hundred metres which were associated with thermal plumes.[1] Since the scattering cross-section involves the intensity of small-scale temperature fluctuations which fall off rapidly with height, most sounders do not show the full height of a plume. Normally, however, the upper limit of convection is marked by an inversion (see Sect. 5) and this can be taken as the depth of convection. Sounders can give a good indication of the strength of turbulence within a plume, its size in the horizontal plane as well as its speed and direction of movement, although rather complicated multi-antenna sounding systems are required to obtain all of this information.[23]

A typical example of thermal plume structures as recorded by an acoustic sounder is shown in Fig. 2 (a). This was recorded on a clear day in light wind conditions over a dry short grass surface at Cardington in the summer of 1978 using a monostatic sounder operating at 2048 Hz. The plume structure seen by the sounder extends to a height of about 200 metres while the apparent plume diameter contracts or sometimes remains approximately constant as a function of height. This facsimile record relates well to the schematic representation of thermal plumes in the boundary layer based on the direct measurements of Kaimal *et al.*[7] The determination of the plume diameter from the facsimile recording of a single sounder is not straightforward since

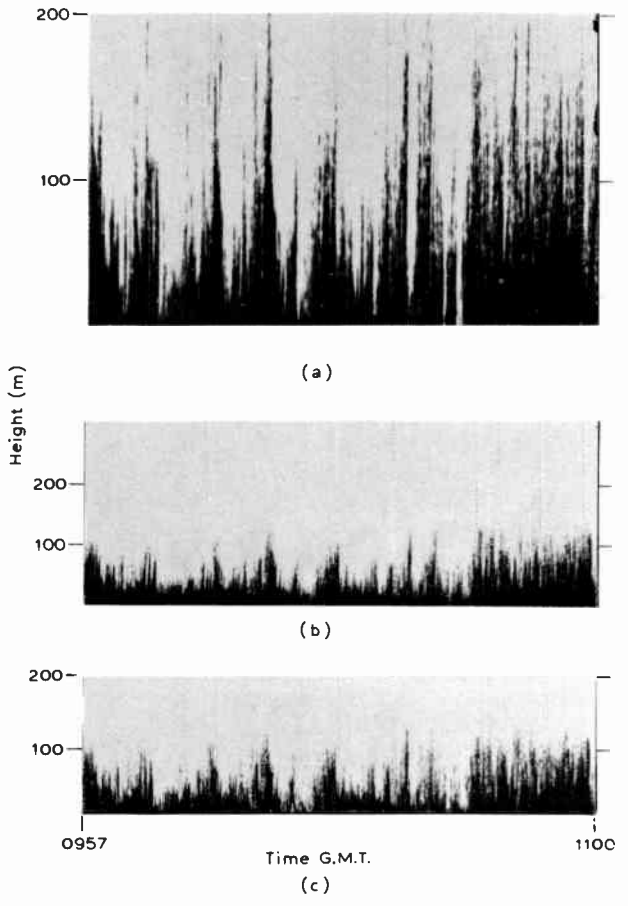


Fig. 2. Acoustic sounder facsimile records of thermal plumes for the period 0957–1100 G.M.T. on 18 August 1978, obtained using three sounders operating at frequencies (a) 2048 Hz, (b) 4·7 kHz, (c) 6·5 kHz.

it depends not only on the wind speed but also the direction of the plume motion relative to the sounder. However, plume diameters ranging from about 50 metres up to many hundreds of metres are observed. Within the plume boundaries, localized regions of greater temperature fluctuations sometimes occur which indicates a distinct plume substructure. Assemblies of plumes may be several kilometres in extent and take 10–15 minutes to pass through the sounder beam.

An important parameter in an acoustic sounding system is the transmit frequency. This determines not only the atmospheric absorption of the acoustic signal and hence the range of the system but is also related to the scale of the atmospheric turbulence which the sounder will see. Within the range of frequencies commonly used in acoustic systems this scale is only marginally different. An example of acoustic sounder facsimile recordings of thermal plumes obtained simultaneously using three sounding frequencies is shown in Fig. 2. Figures 2 (a), 2 (b) and 2 (c) were obtained from closely spaced sounders operating at 2·05 kHz, 4·7 kHz and 6·5 kHz respectively. These recordings clearly demonstrate that the apparent maximum height of the plumes depends on the individual

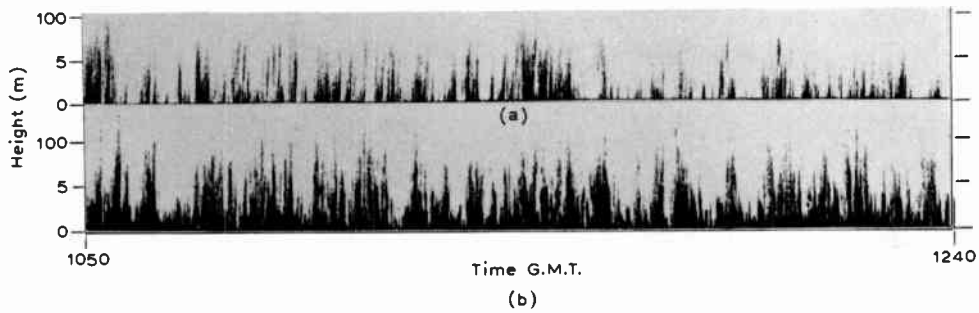


Fig. 3. Facsimile records for 18 September 1978 between 1050 and 1240 G.M.T. The upper trace is that of the 6·5 kHz horizontal sounder pointing downwind at an elevation angle of 5 deg. The lower trace is from the vertically pointing 4·7 kHz sounder.

characteristics of the sounders, i.e. frequency, power transmitted, antenna efficiencies and receiver gains etc. Clearly the lower frequency sounder shows that the plumes extend well beyond the limits shown by the high frequency sounders. The three records, although showing different plume heights, do nevertheless agree remarkably well as far as the relative position and basic shape of the plumes are concerned.

Normally acoustic sounders are used in the monostatic or bistatic (i.e. two separated antennas are used and off-axis sound is detected) configurations with the antennas being directed vertically or at some fairly large angle to the horizontal. This is necessary since spurious echoes from nearby objects would be received and mask the atmospheric returns. However the recently developed mini-sounders operating at higher frequencies[4] and using small antenna arrays are suitable for mounting on a tower and can be directed at angles close to the horizontal, so providing the opportunity to study the plumes in two different planes, i.e. horizontally and vertically.[24]

Typical examples of simultaneous horizontal and vertical sounding are shown in Figs 3 and 4. In both cases the horizontal record is obtained by a 6·5 kHz sounder mounted on the top of a 40 metre high tower and the vertical record from a 4·7 kHz sounder located some 75 metres from the tower. Figure 3 (a) shows the acoustic record obtained when the horizontally directed sounder was pointing downwind and set at an elevation angle of 5°. Figure 4 (a) is for the horizontal sounder pointing across wind and again set at an elevation angle of 5°. In each case the lower half of the Figure shows the conventional vertical monostatic record. For the downwind case the plume echo regions exhibit a small but persistent lean. This is presumed to be due to regions of high reflectivity moving away from the sounder between successive pulses. From the angle of lean and the facsimile chart speed it is possible to derive the component of plume velocity along the direction of sounding. When the direction of sounding is across wind the plume echo regions are not tilted since the component of motion in the direction of sounding is near zero.

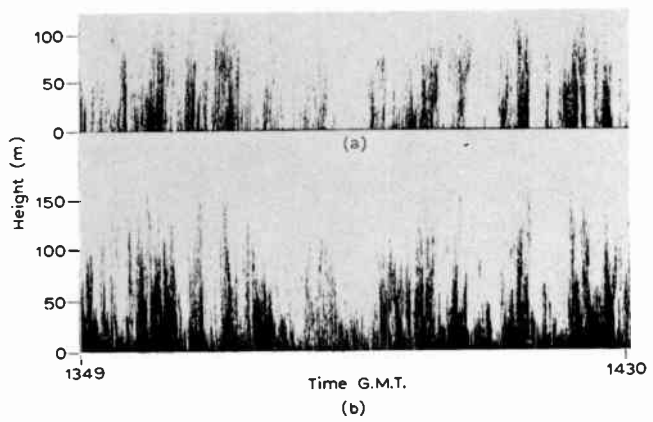


Fig. 4. Facsimile records for 18 September 1978 between 1349 and 1430 G.M.T. The upper trace is that of the 6·5 kHz horizontal sounder pointing crosswind at an elevation angle of 5 deg. The lower trace is due to the vertically pointing 4·7 kHz sounder.

A close examination of Fig. 4 (a) shows some detached echo regions at 1352 G.M.T. at 100 metre range and at 1415 G.M.T. at 70 metres range. These are echoes from plumes which pass through the sounder beam maintaining a roughly constant range. The number of detached plumes which will be observed depends on the plume sizes. The fact that the horizontal and vertical sounder records show good correlation implies that the plumes are much greater in size than the antenna separation (75 metres). From the known wind speed and the time duration of the plumes, obtained from the facsimile records, the plumes were estimated to be about 600 m wide at their base. This is much larger than the 100 m effective range of the horizontal sounder hence, on average, only about one-seventh of the plumes will be expected to appear detached. It is also worth noting that the horizontal records do not exhibit the continuously high echo levels at short ranges which are characteristic of a vertical record, and are associated with intense convective activity near the ground. This allows the short reverberation time of the high-frequency sounder antenna to be clearly distinguished and confirms the ability of these systems to record reliable information from very short ranges.

5 Inversions

One of the most useful applications of acoustic sounding may well prove to be the monitoring of inversions within

the lowest few kilometres of the atmosphere. These are regions in which the temperature increases with height contrary to the normal decrease typical of the troposphere. They are generally associated with anticyclones (regions of subsiding air) and may extend over many hundreds or thousands of kilometres in the horizontal plane whilst their vertical scale is usually of order several hundred metres, although sharp layers may exist in which the temperature can increase by as much as 5 K in less than one metre.[25,26]

Inversions are of great meteorological significance since their presence may indicate preferred regions for strong wind shear which can be of importance to aviation. They can also sometimes limit the growth of the convective boundary layer and hence influence such parameters as maximum daytime temperature, visibility etc. In some cases the height of the lowest inversion may be taken as the depth of the convective boundary layer which is an important scaling parameter in the description of turbulence in the lower atmosphere and is relevant to the dispersal of pollutants. Further, they can effectively impose a barrier to the transfer of heat, moisture and pollutants from the surface to the free atmosphere and thus may critically affect the formation of cloud and the air quality of the near surface layer. Inversions also frequently contain strong gradients of humidity and can therefore seriously affect radio wave propagation.[27] Clearly the ability to monitor the height and characteristics of inversions remotely with surface based instruments would be of great value and in this Section we describe the information that may be gained on these features from monostatic acoustic sounding.

An acoustic record taken on 28 October 1975 between 1335–1445 G.M.T. when the convective boundary layer was capped by a strong (~10 K) inversion with a base at about 200 m is shown in Fig. 5.[28] The interaction of turbulent convective plumes with the region of strong temperature gradient generates an intense echo layer. This interface represents the top of convection on this day and therefore may be taken as the boundary layer depth.[7] The echo layer is clearly distorted by a series of perturbations of similar vertical scale but markedly differing along wind scale. Larger-scale modulations, with a period of some 30 minutes, may originate from mesoscale variations of the inversion base associated with slightly deeper convection. The smaller-scale disturbances with durations from a few tens of seconds to several minutes have sharply defined boundaries and characteristics similar to those described by Rayment and Readings.[25] These authors observed 'hummocks' some 50 m in depth and 500 m in length propagating with the mean wind speed within the mixed layer. In accordance with their findings (i.e. that these were induced by convective activity) the acoustic record indicates some correlation between thermal activity and upward movement of the echo layer. Below the inversion the convective plumes have durations compatible with these disturbances and those that extend towards the 200 m level can often be associated with 'hummocks' of greater vertical displacement. It is of interest to note that within the larger plumes regions of enhanced return are evident which may imply a plume 'substructure' within which more intense temperature fluctuations occur.

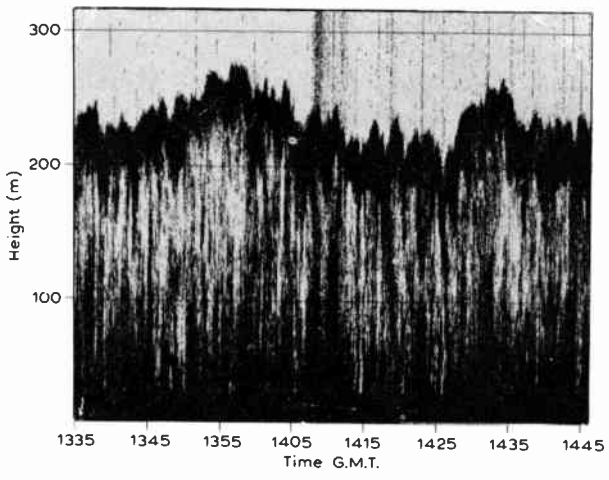


Fig. 5. Sounder facsimile chart for the period 1335–1445 G.M.T. on 28 October 1975. The intense oscillatory echo at ~200 m height is associated with the base of a strong (10K) subsidence inversion.

There is also evidence throughout this record of thermal activity originating from the inversion extending well down into the boundary layer. It is attractive to associate this with entrainment, a process in which air is transferred across the capping inversion from the free atmosphere and mixed into the boundary layer. Much work has been carried out to investigate the mechanisms by which entrainment may occur[29,30], but the resultant effect is that entrained air is carried well down into the boundary layer by the return flow between convective plumes.[7] The acoustic record provides direct qualitative confirmation of this process, e.g. the entrained feature at about 1426 G.M.T. extends down to about 100 m. Entrainment usually results in a lifting of the inversion base and hence an increase in the boundary layer depth. However in the presence of strong subsidence the two may be in approximate balance in which case no net increase in depth occurs.

Clearly the sounder is capable of yielding very useful information on the presence and nature of temperature inversions and of confirming the occurrence of entrainment. As well as these important aspects the sounder can be used to monitor the break-up of inversions.[31,32] This process can occur when the boundary layer continues to warm through daytime surface heating in which case the inversion temperature may be reached or exceeded. Convection then breaks through the inversion layer and continues upward to some higher-level inversion or more generally until its buoyancy is lost. An example of the disruption of an

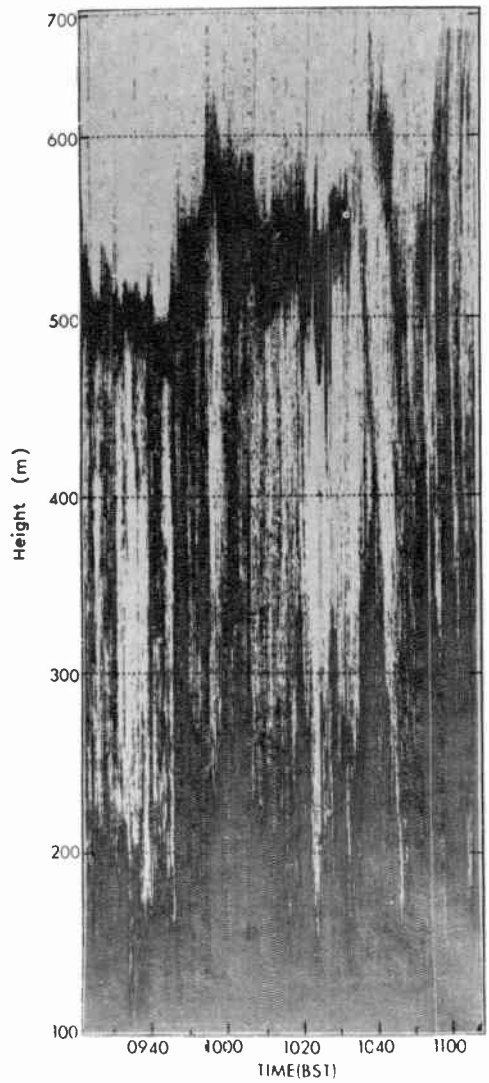


Fig. 6. Acoustic sounder record from a convective boundary layer with an inversion present at ~0940 B.S.T. The inversion subsequently rises and is eroded away by intense convection.

inversion is shown in Fig. 6. In this case intense convective motions extending through the full depth of the boundary layer, i.e. at 0940 B.S.T. to about 500 m, erode and lift a temperature inversion. By 1040 B.S.T. the inversion has reached about 600 m and the intense layer echo now appears broken and ragged.

Acoustic sounding has considerable potential for monitoring the evolution and character of synoptic scale and local temperature inversions in the lower atmosphere. This capability will doubtless prove of benefit to meterological and air pollution scientists and a specific example is discussed in the next Section.

6 Diurnal Variation of the Convective Boundary Layer

The depth of the surface based convective boundary layer (Z_i) is an important parameter for studies of atmospheric diffusion. A great deal of recently published work presents models which predict Z_i and the structure of the convectively mixed layer.[33,34] However, little experimental data to verify this work is available.[7,35] In general it is necessary to obtain information from heights greater than a kilometre in order to determine the depth and investigate the structure of the higher regions of the convective boundary layer[36,37] and this implies the use of a very-high-power sounder.[38] In some circumstances a lower-power sounder can be used on any available high ground. Recent work in Greece used a typical commercial acoustic sounder, operating at 1600 Hz with an effective range of 1000 m, on the top of a 1000 m high mountain overlooking Athens. A typical example of a facsimile record obtained is shown in Fig. 7. This shows the diurnal cycle of the convective boundary layer on a warm clear sunny day. Some intermittent cirrus and altocumulus clouds present around noon did not significantly affect the heating at the ground. As can be seen from the facsimile record, there is an elevated inversion layer in the morning (0400 G.M.T.) at about 250 metres which shows evidence of wave motions and dips below the mountain top at about 0600 G.M.T. This inversion is considered to have been formed by the previous day's convection. From about 0730 G.M.T. the onset of convection becomes apparent, increasing in intensity fairly regularly to produce strong thermal plumes which, by about noon, reach a height of 500 metres above the mountain top. Marking the upper limit of convection is the inversion layer which is highly distorted and shows evidence of active entrainment (see Sect. 5). It is worth noting that the maximum depth of the convective layer (approaching 2000 metres) occurs some time after the maximum intensity of convection which is reached at about 1400 G.M.T. As the strength of convection decreases during the afternoon the inversion height begins to decrease.

The midday radiosonde profile from a nearby meteorological station is shown in Fig. 8. Because of the light winds the balloon, by the time it reaches 1900

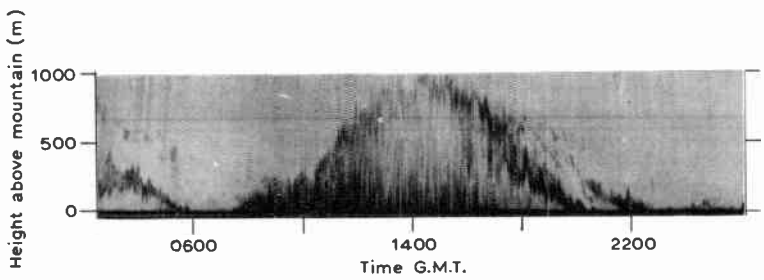


Fig. 7. Acoustic sounder record obtained from a 1000 m high mountain near Athens, Greece, on 15 April 1979.

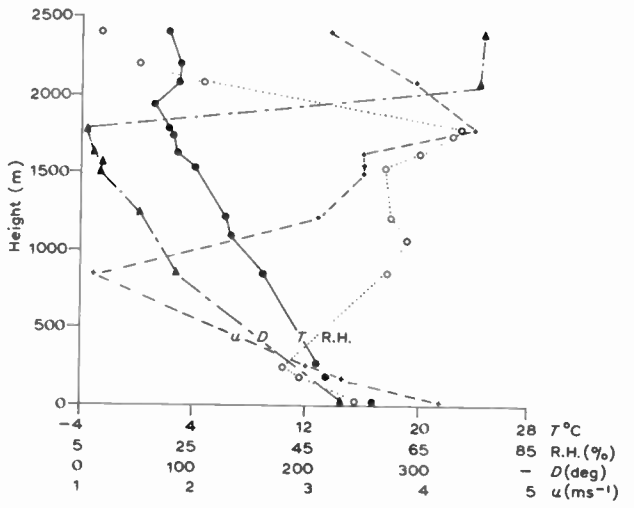


Fig. 8. Profiles of temperature ●——●, relative humidity ○— — —○, horizontal windspeed +— — —+, wind direction ▲— — —▲ obtained by a radiosonde at Athens, Greece, at about 1200 G.M.T. on 15 April 1979.

metres, is within 1·5 km from its point at launch, i.e. about 7–8 km S.W. of the sounder site. It is interesting to note that the mountain has no influence on the inversion height. Several other recordings have shown the same behaviour under similar meteorological conditions.[39]

It is clear that a sufficiently powerful acoustic system should be capable of monitoring the onset and cessation of convection as well as the diurnal variation in the mixed layer depth. This information would be invaluable in, for example, estimating the diffusion of pollutants in the lower atmosphere.

7 Stable Conditions and Waves

With a stable stratification the atmosphere can support wave motion and the interpretation of atmospheric data may well be complicated by the coexistence of waves and turbulence. This topic is particularly relevant to studies of the nocturnal boundary layer and in order to understand the structure of the turbulence field in these conditions it is necessary to eliminate the wave contributions from the data.[40,41] Acoustic sounding offers an excellent method for remotely assessing the significance of wave motion in particular situations.[42] Furthermore the waves in many instances may become unstable and form patches of turbulence which may be the major contributors to turbulent transfer on these occasions. This is particularly true of areas which are nearly continuously stable such as the polar regions in winter.

The most common form of wave-like activity in the boundary layer is undoubtedly Kelvin–Helmholtz (K–H) instability. This occurs in layers where the static stability is much larger than average, i.e. where there are large gradients of temperature and humidity. They are produced predominantly by tilting of these very stable layers.[43,44] Following initiation the billows develop and 'roll up', then secondary instabilities form and the initially regular array of billows may collapse to form irregular 'turbulent' motions containing a cascade of smaller eddies. This mixture of billows and turbulence is easily detected by the acoustic sounder since it generates significant temperature fluctuations on a scale of one-half the acoustic wavelength.

A second type of wave motion which is common in the lower atmosphere is the so-called gravity wave.[45,46] These waves may be triggered by disturbances within the stable environment. An important difference between gravity and K–H waves is that the former propagate through the air with significant phase velocity.[47] Gravity waves may also become unstable and generate local patches of turbulence[48] and hence it is difficult to distinguish between the two by simple inspection of an acoustic record. It is, however, possible to deduce wave speeds and directions from a network of sounders and in this way the generating mechanisms could be distinguished.[42]

Although turbulence levels are generally much lower in stable conditions the small vertical fluxes may produce large net transfers when maintained for long periods of time. Acoustic sounding provides a method for immediately assessing the state of turbulent motion in the stable boundary layer.

To illustrate the potential of sounding to this field we provide examples which show the boundary layer in various states. Figure 9 presents an example of well developed K–H breakdown in a near surface layer. The full line on the figure indicates the mean temperature profile at the time and this confirms the presence of a stable layer, extending from about 60 m to 300 m, across which the temperature increases by approximately 2 K. The wavelike 'herringbone' pattern is characteristic of

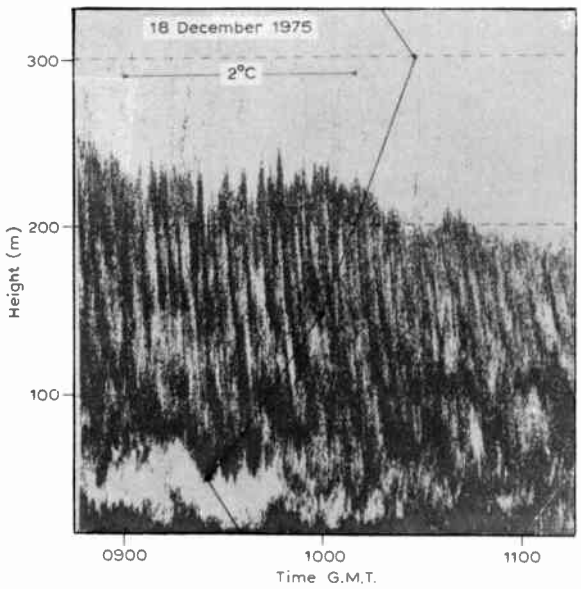


Fig. 9. Kelvin–Helmholtz waves in a surface based stable layer on 18 December 1975. The solid line indicates the temperature profile at this time.

K–H breakdown and has been observed in many acoustic and radar studies of the atmosphere.[49,50]

The acoustic chart for the period 1435–1600 G.M.T. on 14 November 1975 is shown in Fig. 10. Here the boundary layer above about 30 m was stable and between 1455 and 1555 G.M.T. a series of oscillating patterns appear at around the 100 m level. It is considered that these are generated by breaking gravity waves[45,46], although this interpretation must be open to some doubt since no measurement of the wave phase velocity is available. A noteworthy feature of this record is the relatively high resolution of this facsimile record. The sounder in this instance was operated with a 5 ms pulse duration enabling structures about 2 m deep to be resolved, hence these wavelike structures must contain fluctuations occurring within a depth of only a few metres.

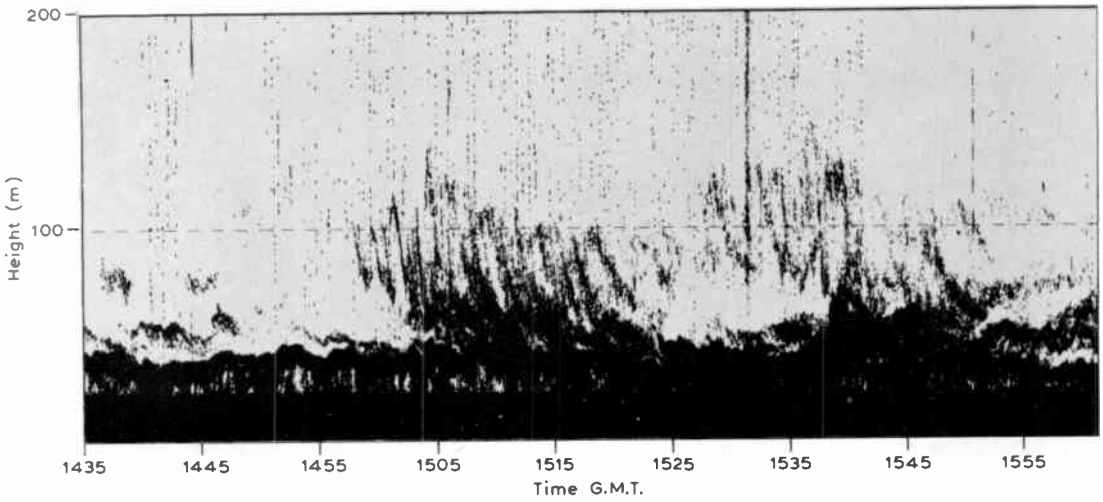


Fig. 10. Although no unambiguous interpretation is possible it is felt that the oscillatory patterns in this record taken on 14 November 1975 are due to breaking gravity waves.

Finally, Fig. 11 illustrates a quite different pattern of behaviour in the nocturnal boundary layer. Here the acoustic returns are generated by a high level of mechanical mixing acting on the slightly stable temperature profile, the darker regions therefore indicate regions of stronger mixing. This acoustic record would be typical of those for conditions with strong winds and low nocturnal cooling rates.

8 Acoustic Sounding of Clouds

A potentially very important application of acoustic sounding is to monitor the height of low level layer cloud or fog top. The latter topic is discussed in Section 9 and involves an essentially low level application since fog depth is rarely more than 300–400 m. On the other hand stratocumulus cloud forms much higher in the atmosphere and it is usually accompanied by strong gradients of temperature and humidity at its upper boundary. These changes can occur in remarkably short distances, e.g. in some recent studies Caughey *et al.*[26] found the mean temperature to increase abruptly, by 5 K in less than 1 m, at the top of nocturnal stratocumulus. Since the cloud layer is turbulent, intense small-scale variations in temperature are generated at cloud top and hence generate acoustic returns. Acoustic scattering has also been observed to occur at the boundaries of cumulus clouds. Shaw[51] convincingly demonstrated that convective clouds within the range of his sounder (~1500 m) were detected. Similar-shaped echo regions were observed by Gaynor[52] and these correlated well with low-level cumulus cloud. Normally direct observations of cloud base are available and hence sounding can offer a simple and effective method of monitoring cloud depth.

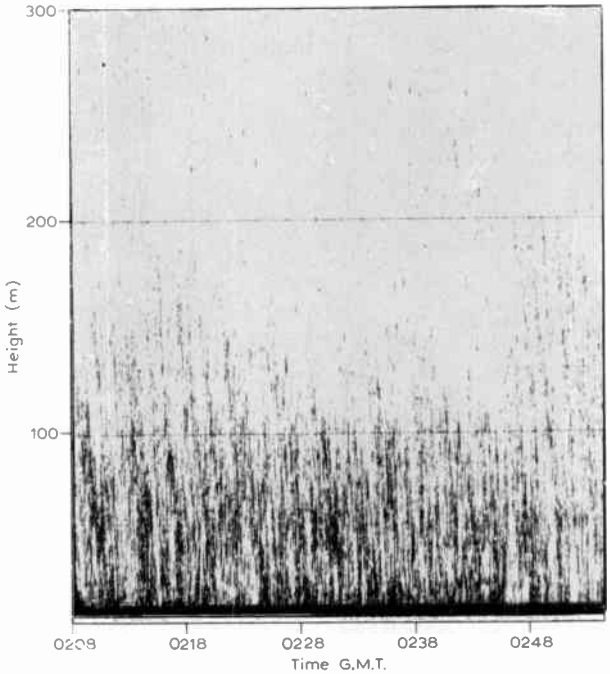


Fig. 11. Echo regions produced by the action of strong wind shear on a weakly stable temperature profile recorded on 7 November 1975.

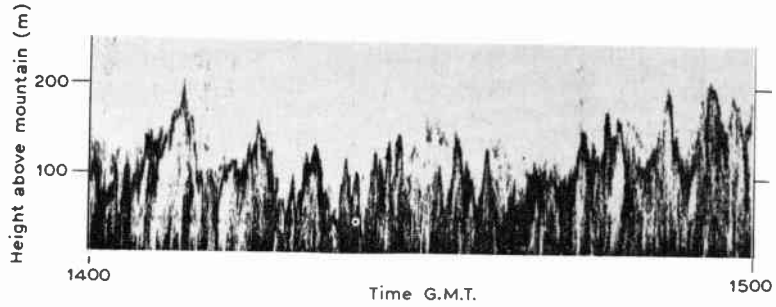


Fig. 12. This acoustic chart was recorded on the summit of Great Dun Fell, Cumbria (847 m) on 15 May 1979. The thin layer echo is from an inversion associated with the top of a cap cloud blowing over the hill. It is probable that entrainment was occurring since the layer echo appears highly contorted and broken.

The long range to the cloud top clearly limits the resolution with which the motion occuring may be resolved. This is because in order to achieve a reasonable echo level a very long acoustic pulse must be used ($\sim \frac{1}{2}$ s) and hence the vertical resolution is rather poor (~ 84 m). In order to study the motion occurring at cloud top in greater detail an acoustic sounder was positioned at the summit of Great Dun Fell (847 m), Cumbria, England. This means of course that the range to cloud top is greatly reduced and hence a short pulse may be used. Figure 12 shows a section of record obtained between 1400 and 1500 G.M.T. on 15 May 1979 using a 4·7 kHz sounder. In this the highly contorted and variable cloud top is clearly resolved and there is some evidence for entrainment effects in which air is drawn across the inversion and mixed with the cloudy air below.

9 Acoustic Sounding of Radiation Fog

A possibly important practical application of acoustic sounding is to assist in the measurement of the depth of radiation fog and the prediction of the time of its clearance. Present methods for estimating fog depth involve the measurement of temperature and humidity profiles using the Cardington Balthum[53] or meteorological radiosonde balloons. Profiles of potential temperature usually show a strong inversion associated with the fog top which is formed by the lifting of the surface-based nocturnal inversion following fog formation and continued radiation cooling from the fog top. With the ground radiatively-shielded a small superadiabatic lapse rate can develop[54] which may result in the establishment of a weakly convective regime within the fog, thereby transferring heat flux upwards from the ground and enhancing vertical mixing.

Echo patterns obtained in radiation fog have typically demonstrated a strong echo layer overlying a region usually exhibiting evidence of weak convective activity.[55] It has been suggested that the echo layer is generated by the interaction of the convective cells with sharp temperature gradients found in the vicinity of the fog top and also by fluctuations in temperature resulting from unstable internal waves forming in the inversion layer at the fog top. Caughey *et al*,[56] using a balloon-borne droplet spectrometer to directly monitor the height of the fog top, established a clear correlation between the lower boundary of the echo layer and the fog top for fog depths between 60 and 240 m.

An illustration of the potential of the acoustic sounder for monitoring the evolution of radiation fog is given by the following example. Figure 13 (a) shows the pattern of cooling in the lowest 16 m of the boundary layer on 19 December 1977 at Cardington. The cooling rate is strongest between 1700 and 1800 G.M.T. at all levels.

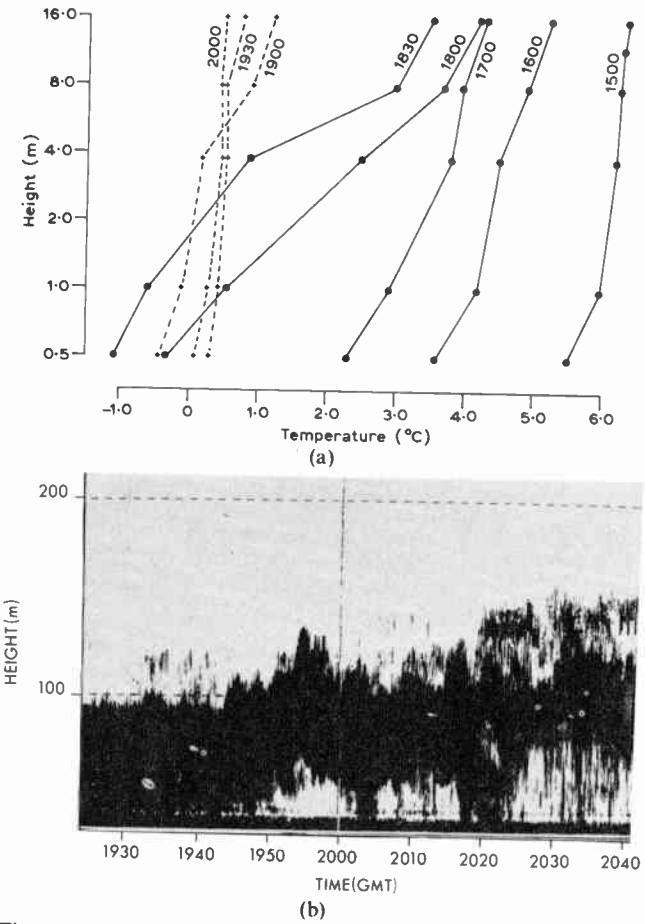


Fig. 13. (a) Temperature profiles up to a height of 16 m for the time period 1500 to 2000 G.M.T. on 19 December 1977. Dotted lines are profiles after the formation of fog. (b) Facsimile record for the period 1930 to 2040 G.M.T. on 19 December 1977. Fog is present underneath the lower edge of the echo layer.

After fog forms the surface based inversion shows signs of dissipation, i.e. by about 1900 G.M.T. as the fog begins to deepen. This is shown clearly on the facsimile chart (Fig. 13 (b)) as the deep echo layer lifts well clear from the surface at about 1940 G.M.T. when the temperature profiles show the top of the fog to be above 16 m. The fog continued to deepen and by 2200 G.M.T. the sky was

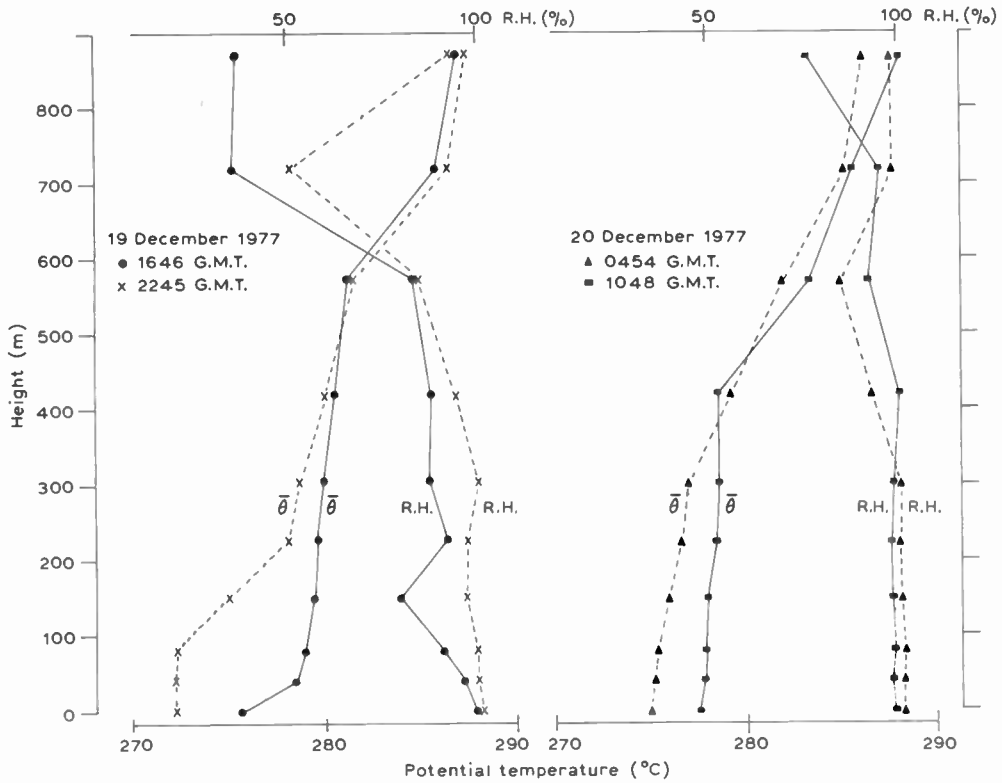


Fig. 14. Cardington Balthum profiles of potential temperature ($\bar{\theta}$) and relative humidity (RH) taken at 1646 and 2245 G.M.T. on 19 December 1977, and at 0454 and 1048 G.M.T. on 20 December 1977.

obscured. The Balthum profiles shown in Fig. 14 indicate an increase in the fog depth from approximately 130 metres to about 420 metres in the period from 2245 to 0454 G.M.T. on the next morning (20th) as indicated by the change in height of the temperature inversion. The facsimile chart shown in Fig. 15 confirms this rise over the time period 0010 to 0200 G.M.T. and shows the undulating inversion layer moving upwards at a rate of approximately 50 metres per hour. By 0500 G.M.T. the mean potential temperature of the fog layer had increased significantly. Visibility gradually improved as the fog lifted from the surface into stratus, being deep enough to allow the formation of drizzle. The surface wind speed increased slightly and by 1048 G.M.T. the visibility had increased to 3400 metres.

The ability of the acoustic sounder to monitor the development of radiation fog could be of great value to the meteorological forecaster on an airfield since it could remove some of the uncertainty associated with present methods of estimating fog depth. With better information on the fog depth it would be possible to prepare more accurate estimates of fog clearance time and to keep them under review by continuous acoustic sounding.

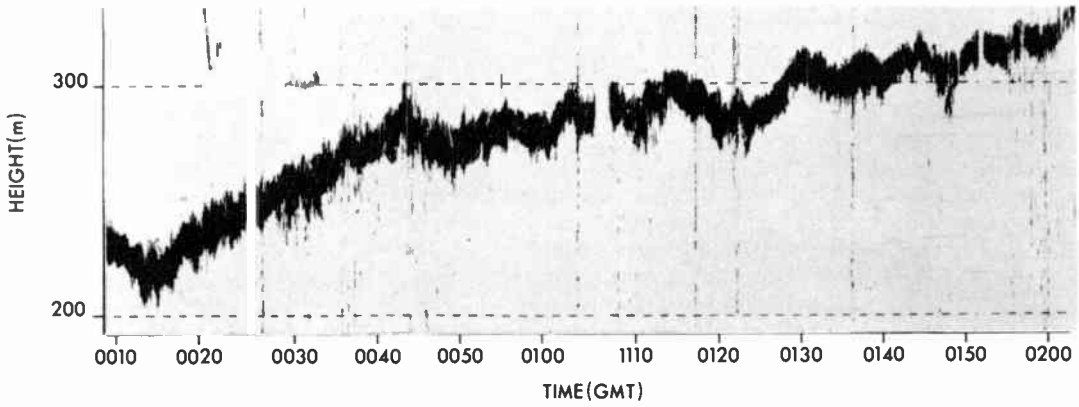


Fig. 15. Acoustic sounder facsimile record taken between 0010 and 0200 G.M.T. on 20 December 1977. Fog is present under the lower edge of the echo layer.

10 The Tracking of Methane Gas Plumes

As mentioned in the foregoing Sections, there is considerable interest in the use of the acoustic sounder in pollution studies. Usually the sounder is used to detect conditions in which pollutants are likely to be concentrated in the lowest levels of the atmosphere rather than dispersed, which might occur, for example, when a strong inversion is present. The acoustic sounder has, however, recently been successfully used to directly monitor a plume of methane gas.[57] It is occasionally necessary for the gas industry to vent a volume of methane gas to empty a plant for essential maintenance or in a possible emergency. Any information which can be obtained on the size and formation of the plume of released gas is of obvious interest to the gas and chemical engineering industries from the point of view of both air pollution and safety.

It could be anticipated that the difference in acoustic phase velocity between methane and air, 430 ms^{-1} and 341 ms^{-1} respectively at 20°C, would produce an acoustic scattering cross-section large enough to be detectable using an acoustic sounder. Figure 16 shows the facsimile recording obtained, using a 4·7 kHz monostatic sounder, of a plume of methane gas. The echoes from the plume can be clearly distinguished from those of the local thermal plumes and start a few seconds after zero time marked on the time axis, which corresponds to the time the vent commenced. The plume was observed some 75 m away from the 20 m high venting stack and as can be seen the gas ranges in height between about 20 m and 40 m. The recording also shows that the vertical motion of the gas in the plume appears to be strongly affected by the thermal convection from the ground.

This work illustrates the potential usefulness of acoustic sounder to track emissions from a chimney directly provided there is a difference in acoustic refractive index between the atmosphere and the emission or that the emission is at a different temperature to the surrounding atmosphere and produces temperature fluctuations which can be detected by the sounder.

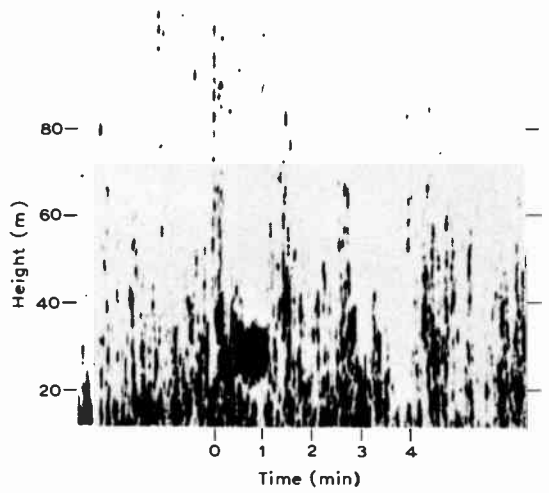


Fig. 16. Sounder facsimile chart of a methane gas plume which appears as the intense region shortly after zero time and occurs between 20 and 40 m height.

11 Conclusion

The acoustic sounder is being increasingly used as a continuous monitor of the lower atmosphere since it gives a real-time record of atmospheric features which cannot be readily identified by *in-situ* sensors without the expense and complication of tethered balloons, radiosondes or meteorological towers. For example, it will show the formation, presence and height of an inversion which can provide an indication of the likelihood of the deposition of pollutants onto the ground. It is proving a useful instrument in studies of the boundary layer and is adding to our understanding of the meteorological processes occurring in the lower atmosphere. In particular, it is a valuable aid in the interpretation of atmospheric *in-situ* instrument studies over a range of heights. The acoustic sounder promises to have an important application in monitoring the height of low level cloud and the depth of radiation fog. The ability to record the development of radiation fog could be of great value to the meteorologist since it could remove some of the uncertainty associated with present methods of estimating fog depth.

12 Acknowledgments

Thanks are due to the Science Research Council which provided the financial support for the design and construction of three of the sounders used in this work in a grant to University College London. Thanks are also due to the staff of the Meteorological Research Unit, Cardington, Bedfordshire, for assistance in the field work. One of the authors, T. J. Moulsley, is grateful for a Science Research Council CASE award in conjunction with the Meteorological Office.

13 References

1 McAllister, L. G., Pollard, J., Mahoney, A. and Shaw, P., 'Acoustic sounding—a new approach to the study of atmospheric structure', *Proc. IEEE*, **57**, pp. 579–87, 1969.

2 Moulsley, T. J., Asimakopoulos, D. N., Cole, R. S. and Crease, B. A., 'Design of arrays for acoustic sounder antennas', *J. Phys. E. (Sci. Instrum.)* **11**, 657–62, 1978.

3 Asimakopoulos, D. N. and Cole, R. S., 'An acoustic sounder for the remote probing of the lower atmosphere', *J. Phys. E. (Sci. Instrum.)* **10**, pp. 47–50, 1977.

4 Moulsley, T. J. and Cole, R. S., 'High frequency atmospheric acoustic sounders', *Atmospheric Environment*, **13**, pp. 347–50, 1979.

5 Tatarski, V. I., 'Wave Propagation in a Turbulent Medium' (Dover, New York, 1961).

6 Lumley, J. L. and Panofsky, H. A., 'The Structure of Atmospheric Turbulence' (Interscience, New York, 1964).

7 Kaimal, J. C., Wyngaard, J. C., Hangen, D. A., Cote, O. R., Izumi, Y., Caughey, S. J. and Readings, C. J., 'Turbulence structure in the convective boundary layer', *J. Atmos. Sci*, **11**, pp. 2152–69, 1976

8 Brown, E. H. and Hall, F. F., 'Advances in atmospheric acoustics', *Rev. Geophys Space Phys.*, **16**, pp. 47–110, 1978.

9 Little, C. G., 'Acoustic methods for the remote probing of the lower atmosphere', *Proc. IEEE*, **57**, pp. 571–8, 1969.

10 Wescott, J. W., Simmons, W. R. and Little C. G., 'Acoustic echo sounding measurements of temperature and wind fluctuations', Tech. Memo ERLTM-WPL5, Environ. Sci. Services Admin., Boulder, Colorado, 1970.

11 Simmons, W. R., Wescott, J. W. and Hall, F. F., 'Acoustic echo sounding as related to air pollution in urban environments' Tech. Rep. ERL 216-WPL 17, 77 pp. NOAA, Boulder, Colorado, 1971.

12 Moulsley, T. J., 'Acoustic sounding of the atmosphere', PhD Thesis, University of London, 1979.

13 Asimakopoulos, D. N. and Cole, R. S., 'An acoustic sounder for the remote probing of the lower atmosphere' *J. Sci. Instrum.*, **9**, pp. 47–50, 1977.

14 Owens, E. J., 'NOAA Mk VII acoustic echo sounder', Tech. Memo. ERL WPL-12, 71 pp, NOAA, Boulder, Colarado, 1975.

15 Hall, F. F. and Westcott, J. W., 'Acoustic antenna for atmospheric echo sounding' *J. Acoust. Soc. Am.*, **56**, pp. 1376–82, 1974.

16 Adekola, S. A., 'Concerning the influence of echo carrier frequencies and antenna dimensions on the performance of echosonde (acoustic radar) antennas', *J. Acoust. Soc. Am.*, **62**, pp. 524–42, 1977.

17 Bourne, I. A. and Keenan, T. D., 'High power acoustic radar', *Nature*, **251**, pp. 206–8, 1974.

18 Spizzichino, A., 'Discussion of the operating conditions of a Doppler radar', *J. Geophys. Res.*, **79**, pp. 5585–91, 1974.

19 Moulsley, T. J., Asimakopoulos, D. N., Cole, R. S. and Crease, B. A., 'Design of arrays for acoustic sounder antennas', *J. Phys. E. (Sci. Instrum.)*, **11**, pp. 657–62, 1978.

20 Thomson, D. W., Coulter, R. L. and Warhaft, Z., 'Simultaneous measurements of turbulence in the lower atmosphere using radar and aircraft', *J. Appl. Met.*, **17**, pp. 723–34, 1978.

21 Kaimal, J. C. and Businger, J. A., 'Case studies of a convective plume and dust devil', *J. Appl. Met.*, **9**, pp. 612–20, 1970.

22 Warner, J. and Telford, J. W., 'Convection below cloud base', *J. Atmos. Sci.*, **24**, pp. 374–381, 1967.

23 Hall, F. F., Edinger, J. G. and Neff, W. D., 'Convective plumes in the planetary boundary layer investigated with an acoustic echo sounder', *J. Appl. Met.*, **14**, pp. 513–23, 1975.

24 Moulsley, T. J., Cole, R. S., Asimakopoulos, D. N. and Caughey, S. J., 'Simultaneous horizontal and vertical acoustic sounding of the atmospheric boundary layers', *Bound.-Layer Met.*, 16, pp. 223–30, 1979.

25 Rayment, R. and Readings, C. J., 'A case study of the structure and energetics of an inversion', *Q. J. R. Met. Soc.*, **100**, pp. 221–33, 1974.

26 Caughey, S. J., Crease, B. A. and Roach, W. T., 'A field study of nocturnal stratocumulus: II Turbulence structure and entrainment', Submitted to *Q. J. R. Met. Soc.*, 1980.

27 Ruthroff, C. L., 'Multipath fading on line-of-sight microwave systems as a fraction of path length and frequency', *Bell Syst. Tech. J.*, **50**, pp. 2375–98, 1971.

28 Crease, B. A., Caughey, S. J. and Tribble, D. T., 'Information on the thermal structure of the atmospheric boundary layer for acoustic sounding', *Met. Mag.*, **106**, pp. 42–52, 1977.

29 Readings, C. J., Golton, E. and Browning, K. A., 'Fine scale structure and mixing within an inversion', *Bound.-Layer Met.*, **4**, pp. 275–87, 1973.

30 Palmer, S. G., Caughey, S. J. and Whyte, K. W., 'An observational study of entraining convection using balloon-borne turbulence probes and high power Doppler radar', *Bound.-Layer Met.*, **16**, pp. 261–78, 1979.

31 Cronenwett, W. T., Walker, G. B. and Inman, R. L., 'Acoustic sounding of meteorological phenomena in the planetary boundary layer', *J. Appl. Met.*, **11**, pp. 1351–8, 1972.

32 Wyckoff, R. J., Beran, D. W. and Hall, F. F., 'A comparison of the low level radiosonde and the acoustic echo sounder for monitoring atmospheric stability', *J. Appl. Met.*, **12**, pp. 1196–1204, 1973.

33 Tennekes, H., 'A model for the dynamics of the inversion above a convective boundary layer', *J. Atmos. Sci.*, **30**, pp. 558–67, 1973.

34 Zilitinkevich, S. S., 'Resistance laws and prediction equation for the depth of the planetary layer', *J. Atmos. Sci.*, **32**, pp. 741–52, 1975.

35 Chorley, L. G., Caughey, S. J. and Readings, C. J., 'The development of the atmospheric boundary layer: 3 case studies', *Met. Mag.*, **104**, pp. 349–60, 1975.

36 Konrad, T. G., 'The dynamics of the convective process in clear air as seen by radar', *J. Atmos. Sci.*, **27**, pp. 1138–47, 1970.

37 Rowland, J. R., 'Intensive probing of a clear air convective field by radar and instrumental aircraft', *J. Appl. Met.*, **12**, pp. 149–55, 1973.

38 Bourne, I. A. and Keenan, T. D., 'High power acoustic radar', *Nature*, **251**, p. 206, 1974.

39 Asimakopoulos, D. N., Deligiorgi, D. G. and Lalas, D. P., 'Acoustic sounder observations of the atmospheric boundary layer from the top of a steep mountain', to be published 1980.

40 Kaimal, J. C., 'Turbulence spectra, length scales and structure parameters in the stable boundary layer', *Bound.-Layer Met*, **4**, pp. 289–309, 1973.

41 Caughey, S. J., Wyngaard, J. C. and Kaimal, J. C., 'Turbulence structure in the evolving stable boundary layer', *J. Atmos. Sci.*, **36**, p. 1041, 1979.

42 Eymard, L., 'Ondes de gravité dans la couche limite planétaire étude experimentale par sondage acoustique', Technical Note No. 54, Centre National d'etudes des Telecommunications, France, 1978.

43 Woods, J. D., 'On Richardson's number as a criterion for laminar turbulence—laminar transition in the ocean and atmosphere', *Radio Sci.*, **4**, pp. 1289–98, 1969.

44 Scorer, R. S., 'Billow mechanics', *Radio Sci.*, **4**, pp. 1299–1308, 1969.

45 Caughey, S. J. and Readings, C. J., 'An observation of waves and turbulence in the earth's boundary layer', *Bound.-Layer Met.*, **9**, pp. 279–96, 1975.

46 Hooke, W. H., Hall, F. F. and Gossard, E. E., 'The observed generation of an atmospheric gravity wave by shear instability in the mean flow of the planetary boundary layer', *Bound.-LayerMet.*, **5**, pp. 29–41, 1973.

47 Gossard, E. E., Richter, J. H. and Atlas, D,, 'Internal waves in the atmosphere from high resolution radar measurements', *J. Geophys. Res.*, **75**, pp. 903–13, 1970.

48 Axford, D. N., 'An observation of gravity waves in shear flow in the lower stratosphere', *Q. J. R. Met. Soc.*, **96**, pp. 273–86, 1970.

49 Emmanuel, C. B., Bean, B. R., McAllister, L. G. and Pollard, J. R., 'Observations of Helmholtz waves in the lower atmosphere with an acoustic sounder', *J. Atmos. Sci.*, **29**, pp. 886–92, 1972.

50 Emmanuel, C. B., 'Richardson number profiles through shear instability wave regions in the lower planetary boundary layer', *Bound.-Layer Met*, **5**, pp. 19–27, 1973.

51 Shaw, N. A., 'Acoustic sounding of the atmosphere', Ph.D. dissertation, Roy. Aust. Air Force Acad., University of Melbourne, Melbourne, Australia, 1971.

52 Gaynor, J. E., Mandics, P. A., Wahr, A. B. and Hall, F. F., 'Studies of the tropical marine boundary layer using acoustic backscattering during GATE', in Reprints of the 17th Radar Meteorology Conference, pp. 303–6, 1976. American Meteorological Society, Boston, Mass.

53 Painter, H. E., 'The tethered radiosonde', *Met. Mag.*, **99**, pp. 93–98, 1970.

54 Roach, W. T., Brown, R., Caughey, S. J., Garland, J. A. and Readings, C. J., 'The physics of radiation fog: a field study', *Q. J. R. Met. Soc.*, **102**, pp. 313–33, 1976.

55 Crease, B. A., Caughey, S. J. and Tribble, D. T., 'Information on the thermal structure of the atmospheric boundary layer from acoustic sounding', *Met. Mag.*, **106**, pp. 42–52, 1977.

56 Caughey, S. J., Dare, W. M. and Crease, B. A., 'Acoustic sounding of radiation fog', *Met. Mag.*, **107**, pp. 103–13, 1978.

57 Cole, R. S. and Moulsley, T. J., 'The use of the acoustic sounder to track methane gas plumes', *Atmos. Env.*, **13**, pp. 1437 41, 1979.

Manuscript first received by the Institution on 19th March 1980 and in final form on 8th August 1980.
(Paper No. 1963/AMMS 103)

UDC 551.501.796:551.513

Indexing Terms: Atmospheric movements, Atmospheric structure, Sounding apparatus

Quantitative interpretation of acoustic echoes from the planetary boundary layer

S. J. CAUGHEY, Ph.D.*

B. A. CREASE, M.Sc.,*

R. S. COLE, Ph.D.,†

D. N. ASIMAKOPOULOS, Ph.D.‡

and

T. J. MOULSLEY, Ph.D.§

SUMMARY

Recent advances in the quantitative application of acoustic sounders to studies of the atmospheric boundary layer are described. Precise measurement of the echo strength is shown to provide a means of estimating the structure parameters for wind velocity and temperature. Extraction of the Doppler shift provides a method for obtaining information on the mean airflow and small-scale fluctuations in wind velocity

* *Meteorological Office, London Road, Bracknell, Berkshire RG12 2SZ.*

† *Department of Electronic and Electrical Engineering, University College London, Gower Street, London WC1E 7JE.*

‡ *Formerly at University College London; now with the Department of Meteorology, University of Athens.*

§ *Formerly at University College London; now with the Department of Electrical and Electronic Engineering, Portsmouth Polytechnic.*

1 Introduction

The last few years have seen much experimental effort devoted to the quantitative evaluation of acoustic sounding as a technique for remote probing of the lower atmosphere. Visual examination of acoustic facsimile records has provided a valuable insight into the occurrence and nature of common boundary layer features; thermal plumes, nocturnal inversions, breaking waves, etc.[1,2,3] However, of potentially greater importance are the amplitude and frequency spectra of the echoes which contain quantitative information on the structure of temperature and wind fluctuations in the boundary layer and, through the Doppler shift, on the airflow.

Quantitative acoustic data may find many useful applications in boundary layer studies. For example, Doppler sensing could be of importance in assessing the occurrence of low-level wind shears. Similarly, knowledge of the profiles of vertical velocity variance and turbulence dissipation rate could be useful to experiments concerned with the diffusion of pollutants in the atmosphere. The successful application of acoustic techniques will depend upon the development of sophisticated data processing systems capable of near real-time quality control, data reduction, display and analysis. Furthermore, long-term studies with *in-situ* measurements will be required to ascertain the performance limitations of the acoustic system under different atmospheric and environmental conditions. This paper reviews the present information available on the performance characteristics of some acoustic sensors, derived from short period (i.e. a few hours) comparison studies.

In comparing acoustically measured values of atmospheric parameters with the same parameters measured using direct instrumentation it is worth recalling some basic differences between these quantities:

(a) The acoustic sounder provides an estimate over a volume, ΔV, determined by the geometry of the system, the characteristics of the antenna and the acoustic pulse length, τ.

(b) The wind direction is important since features sampled by the *in-situ* instrumentation may not pass through the acoustic scattering volume. The significance of this effect can be reduced by comparing time-averaged statistics.

These considerations indicate that some variations and differences between acoustically and directly measured quantities must always be expected.

2 Theoretical Background

2.1 Acoustic Scattering

Acoustic sounding relies on the scattering of sound waves from inhomogeneities in refractive index on a scale of about half the acoustic wavelength, i.e. about 0·1 m for sounders in common use. At this scale the

The Radio and Electronic Engineer, Vol. 50, No. 11/12, pp. 598–610, November/December 1980 0033–7722/80/110598+13 $1.50/0

inhomogeneities correspond to the refractive index and wind velocity variations associated with turbulence. Under most conditions over land, variations in acoustic refractive index are determined mainly by temperature fluctuations, but in the marine boundary layer the contribution from humidity fluctuations may also be significant.[4]

The theory of scattering of sound from turbulence has been well developed for the case of locally isotropic and homogeneous turbulence.[5] The turbulence frequency spectrum is assumed to have a region with a $-\frac{5}{3}$ slope, known as the inertial subrange, which is bounded by a minimum wavenumber κ_0 (maximum wavelength $L_0 = 2\pi/\kappa_0$) and a maximum wavenumber k_0 (minimum wavelength $l_0 = 2\pi/k_0$). Wavenumbers below κ_0 correspond to the large input scales at which energy is supplied to the turbulent system from, for example, convection from solar heating of the ground. The large (anisotropic) eddies are unstable and break up, transferring energy to successively smaller scales down the inertial subrange in which region the turbulence is considered to be homogeneous and isotropic, to wavenumbers greater than k_0, the 'dissipation range', where viscous effects cause a rapid loss of turbulent energy and a spectral slope much steeper than $-5/3$. The outer scale of turbulence in the lower atmosphere is typically tens or hundreds of metres while the inner scale is the order of millimetres. Acoustic sounder wavelengths can therefore be expected to lie within the inertial subrange.

Monin derived an expression for the acoustic scattering cross-section of temperature and velocity spectra.[6] Assuming that the turbulence is homogeneous and isotropic on the scales probed by the acoustic waves the scattering cross-section can be written as

$$\sigma(\theta) = 0{\cdot}03k^{\frac{1}{3}}\cos^2\theta\left[\frac{C_V^2}{c^2}\cos^2\frac{\theta}{2} + 0{\cdot}136\frac{C_T^2}{\bar{T}^2}\right] \times \left[\sin\frac{\theta}{2}\right]^{-\frac{11}{3}} \quad (1)$$

where $\sigma(\theta)$ is the acoustic power scattered per unit incident flux per unit solid angle at a scattering angle θ, $k(=2\pi/\lambda)$ is the acoustic wavenumber, c is the velocity of sound and λ the wavelength, $\bar{T}$ is the temperature in the scattering volume. C_V^2 and C_T^2 are the velocity and temperature structure parameters defined by

$$C_V^2 = \overline{[V(x)-V(x+r)]^2}/r^{\frac{2}{3}} \quad (2)$$

and

$$C_T^2 = \overline{[T(x)-T(x+r)]^2}/r^{\frac{2}{3}}. \quad (3)$$

Here $V(x)$ and $V(x+r)$ are the wind speeds at points x and $x+r$, $T(x)$ and $T(x+r)$ are the temperatures at points x and $x+r$.

The most significant properties of the acoustic scattering cross-section indicated by equation (1) are:

(a) scattering is a weak function of wavelength and mainly in the forward direction ($\theta < \pi/2$);
(b) there is no scattering from either temperature or velocity fluctuations at an angle of $\pi/2$;
(c) only temperature fluctuations contribute to backscatter ($\theta = \pi$).

The scattering cross-section appears in the radar equation for a pulsed monostatic sounder (i.e. co-located transmitter and receiver, see Fig. 1(a)).

$$P_r = P_t\frac{\sigma(\pi)c\tau\eta AL}{2R^2} \quad (4)$$

where P_r is the received power, P_t is the radiated acoustic power, τ is the transmitted pulse duration, η is the receiver efficiency, A is the receiver area, L is the round-trip attenuation and R is the range to the scattering volume. From a measurement of P_r the scattering cross-section, $\sigma(\pi)$, and thus C_T^2 can be determined provided all the other quantities in equation (4) are known. The values of P_t, τ, η and A are all determined by the sounder system and can be measured, whilst the range, R, is found from the time interval between transmission and reception. The attenuation, L, can be attributed to the following:

(a) classical absorption,

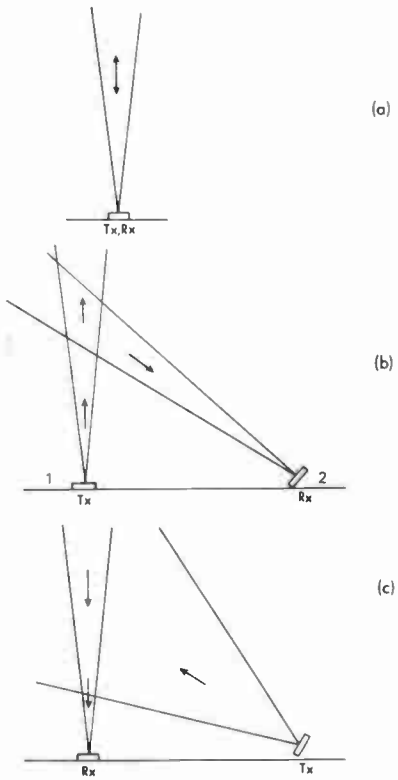


Fig. 1. Common acoustic sounding arrangements.
(a) Monostatic, single antenna used for transmission and reception.
(b) Narrow beam bistatic system for measurements at a single height.
(c) Broad beam transmitter probed by a narrow beam receiver used for obtaining profile information.
Tx—transmitter; Rx—receiver.

(b) molecular absorption,
(c) excess attenuation.

The first two mechanisms are well understood and accurate semi-empirical methods exist for the calculation of these absorption coefficients as functions of frequency, temperature and humidity.[7,8] Excess attenuation is the term applied to any other attenuation mechanism. Significant excess attenuation has been indicated by acoustic sounders[9] but these observations have not been entirely accounted for by theoretical models involving turbulent beam broadening[10] or refraction.[11]

In the case of a bistatic sounder (see Fig. 1(b) and (c)) the scattering angle is less than 180° and so the system is sensitive to scattering from both temperature and velocity fluctuations. If both antennas have narrow ($< \sim 10°$) beams (Fig. 1(b)) a modified form of the monostatic radar equation can be used to calculate the received power.[12] However, with broad beams (e.g. Fig. 1(c)) this equation is not sufficiently accurate for many applications and an integral formulation of the radar equation is necessary.[13]

2.2 Doppler Sounding

In general the scattered sound received by an acoustic sounder has a Doppler shift dependent on the wind velocity in the scattering volume. In the case of a monostatic sounder the frequency shift, Δf, is given by

$$\Delta f = -\frac{2f_0 V_r}{c} \tag{5}$$

where f_0 is the transmitted frequency, V_r is the radial velocity and c is the velocity of sound. Thus the total wind vector can be determined from three orthogonal Doppler shift measurements.

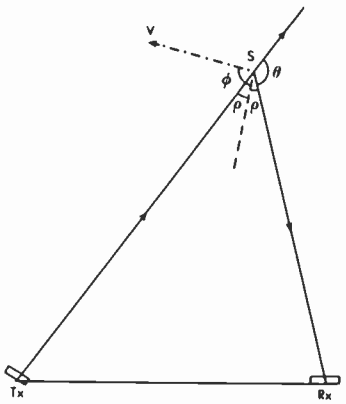


Fig. 2. Bistatic geometry for the Doppler shift analysis. The angle ϕ is between the wind direction (V) and the bisector of the angle Tx S Rx. The scattering angle is θ.

In a bistatic system the Doppler shift is proportional to the wind component along the bisector of the angle between transmitter and receiver, i.e.

$$\Delta f = \frac{2f_0 V}{c} \cos\phi \sin\left(\frac{\theta}{2}\right) \tag{6}$$

where V is the wind speed and ϕ and θ are defined in Fig. 2.

2.3 Determination of Structure Parameters

As discussed earlier, a simplifying assumption is made that the turbulence field, at the scale probed by the acoustic waves, is isotropic and homogeneous (for a discussion of supporting experimental evidence see Caughey *et al.*[14]). The acoustic scattering cross-section can then be expressed in terms of structure parameters, which can be measured with appropriate sensors. In the studies discussed here C_T^2 was measured by two fast response temperature sensors spaced by $\sim 0{\cdot}34$ m.

An alternative method for obtaining estimates of the structure parameters comes from application of the Kolmogorov expression for the power spectral densities of vertical velocity ($S_W(n)$) and temperature ($S_T(n)$) in the inertial subrange, i.e.

$$\begin{aligned} C_V^2 &= 10{\cdot}2\left(\frac{n}{\bar{u}}\right)^{\frac{2}{3}} nS_W(n) \\ C_T^2 &= 13{\cdot}6\left(\frac{n}{\bar{u}}\right)^{\frac{2}{3}} nS_T(n) \end{aligned} \tag{7}$$

where n is the frequency in hertz. (See, e.g. Caughey *et al.*[14] and Asimakopoulos *et al.*[15]) The velocity structure parameter is also related to the dissipation rate of turbulent kinetic energy, ε, through the expression,

$$C_V^2 = 2\varepsilon^{\frac{2}{3}}. \tag{8}$$

In the comparison experiments to be discussed these expressions have been used to compute estimates of the structure parameters.

3 Direct Measurements

For the proper evaluation of acoustic sensors it is clearly necessary to make direct measurements of the boundary layer flow as close as possible to the scattering volume. Instrumentation supported by towers can be used for this purpose, although the distance between the direct sensor and the sampling volume is then fairly large ($\sim \frac{1}{2}$ km), since the tower may generate spurious acoustic echoes.[16] In the studies described here balloon-borne instrumentation was employed.[17,18] This has the advantage that the balloon can be positioned between sidelobes of the acoustic antenna and hence avoid spurious reflections. The instrumentation attached to the balloon cable has been described fully elsewhere[18,19] and is capable of measuring the high frequency fluctuations of the longitudinal (u), lateral (v) and vertical (w) components of air motion, the temperature (T) and

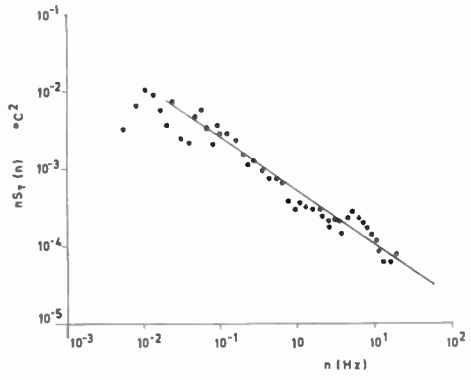


Fig. 3. The one-dimensional temperature spectrum, $nS_T(n)$, plotted versus frequency (n), for measurements made over a period of six minutes during convective conditions. The solid line has a slope of $-\frac{2}{3}$.

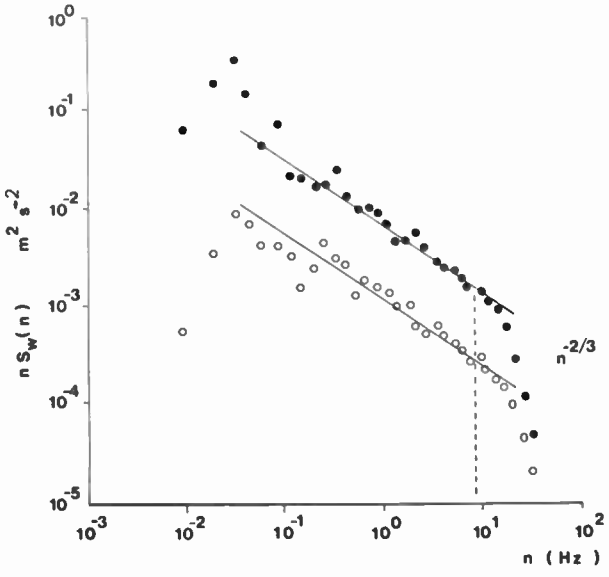


Fig. 4. Two typical vertical velocity spectra, $nS_w(n)$, plotted versus frequency (n) which exhibit a well-defined $-\frac{2}{3}$ slope in the inertial subrange, but significantly different spectral densities.

differential temperature (ΔT) across a distance of 0·34 m in the vertical. These basic quantities can be used to obtain estimates of the structure parameters (for wind and temperature) C_V^2 and C_T^2 and other statistics of the airflow for comparison with acoustically-derived data.

The assumption discussed in Section 2 concerning the nature of the turbulence field is generally supported by the direct measurements of the power spectra. Figure 3 shows an example of a temperature spectrum obtained in convective conditions which shows a well-defined slope at high frequency of approximately $-\frac{2}{3}$. Similarly, Fig. 4 indicates two vertical velocity spectra, again from convective conditions, which show a roll-off at high frequency of $-\frac{2}{3}$ but have markedly different spectral levels. In stable conditions the spectral behaviour is much less regular and in many instances departures from a $-\frac{2}{3}$ slope can occur.[20]

4 Measurement of Structure Parameters and Turbulence Dissipation Rate

4.1 Monostatic Sounding

Early work sought to demonstrate a correlation between the intensity of differential temperature fluctuations ($\overline{\Delta T^2}$), measured across a distance of 0·12 m, i.e. an acoustic half wavelength, and the intensity of back-scattered sound. A good example was obtained on 20 November 1974 (see Asimakopoulos *et al.*,[15] and Fig. 5). On this occasion the tethered balloon instrumentation was raised slowly through a surface based nocturnal inversion, eventually emerging from the region of strong echo. The correspondence between a marked reduction in ΔT fluctuations and cessation of the echo is striking. Note also that the upper limit of the echo layer corresponds well with the top of the nocturnal inversion and hence the sounder could be used to determine the depth of the stable boundary layer, h, a parameter which is basic to the structure of turbulence in these conditions.[21]

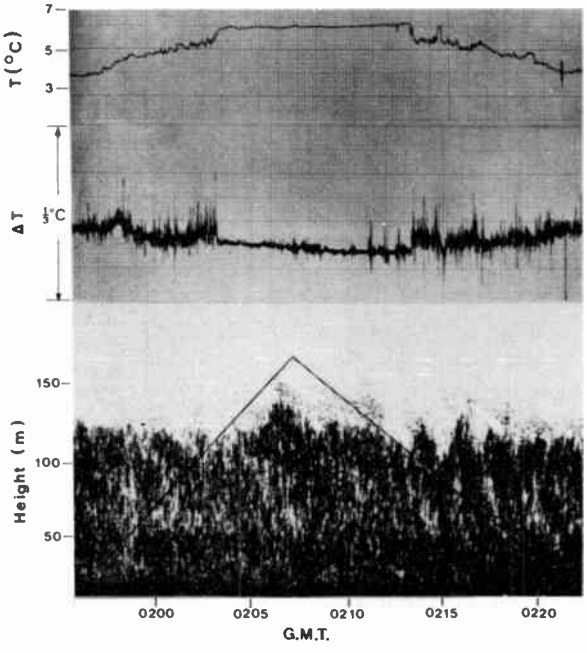


Fig. 5. Simultaneous recordings of the temperature (T) and differential temperature (ΔT) with the acoustic sounder facsimile record. The solid line represents the track of the instrument package as the balloon was raised and lowered. The wind speed increased from ~1 ms^{-1} at 40 m to ~4 ms^{-1} at 165 m.

More detailed studies required the accurate calibration of the acoustic system in the field and also the verification that the turbulence field was such that the simplified version of the scattering cross-section involving only C_T^2 and C_V^2 was justified.[22,14] It is simplest first to consider a simple vertically-directed monostatic sounder detecting backscattered sound and hence the contribution to the scattering cross-section from wind fluctuations is zero and the precise measurement of echo strength may be used to estimate C_T^2. Measurements were

made in two distinct atmospheric conditions:

(1) Stable conditions at night with moderate wind speeds.
(2) Strongly convective conditions with well-defined thermal plumes.

$$C_T^2 \text{ (probe)} = 1{\cdot}9 \times 10^{-3} \text{ degC m}^{-\frac{1}{3}}$$

The open circles represent estimates of C_T^2 from one-dimensional temperature spectra computed over a 26 s period. These are in good agreement with the other direct estimate of C_T^2 from the ΔT measurements.

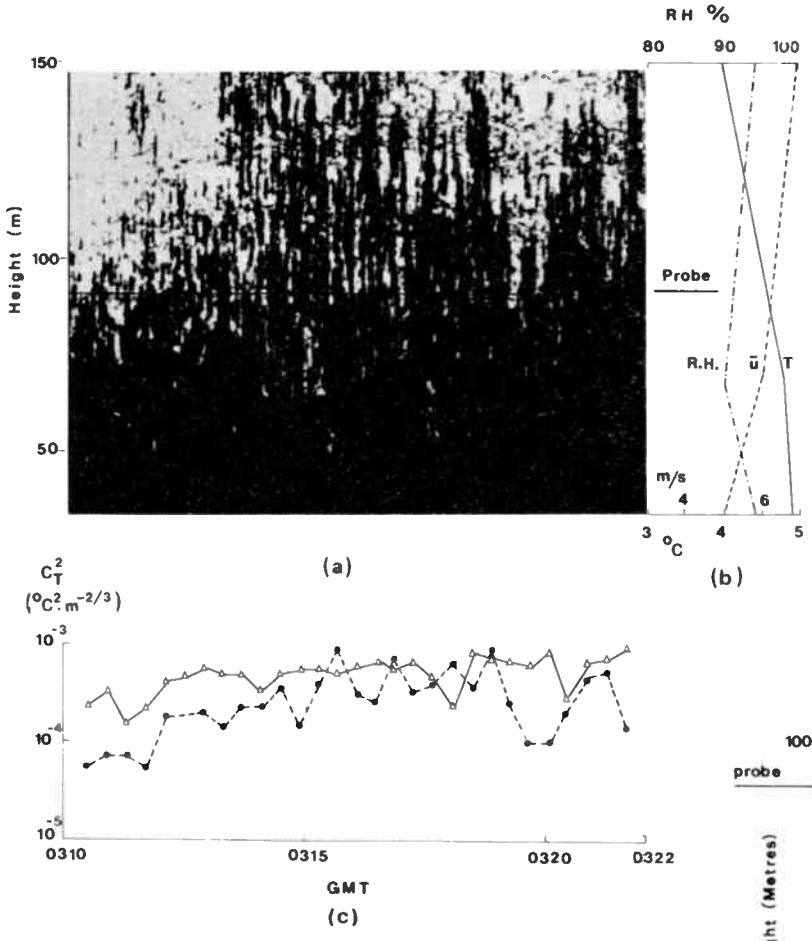


Fig. 6. (a) Facsimile recording of a nocturnal inversion on 17 January 1975. (b) Profiles of temperature (T), wind velocity ($\bar{u}$) and relative humidity (RH). (c) Comparison of measurements of C_T^2 from the acoustic sounder △ and turbulence probe ●.

Shown in Fig. 6 is the facsimile recording for the run carried out between 0310 and 0322 on 17 January 1975. This shows an echo pattern somewhat characteristic of stable conditions increasing in strength across the period of the run. The comparison between C_T^2 estimates is given in Fig. 6(c). Both show fairly good agreement in both magnitude and trend. The average values compare well being

$$C_T^2 \text{ (sounder)} = 5{\cdot}8 \times 10^{-4} \text{ degC}^2 \text{ m}^{-\frac{1}{3}}$$

and $$C_T^2 \text{ (probe)} = 4{\cdot}1 \times 10^{-4} \text{ degC}^2 \text{ m}^{-\frac{1}{3}}.$$

Results obtained in convective conditions on 4 June 1975 are shown in Fig. 7. The facsimile record indicates thermal plumes extending upwards to a height of 100–200 m. The probe was located at a height of 91 m. Again the sounder C_T^2 values show the same variations as the direct estimates and the overall averages are in excellent agreement:

$$C_T^2 \text{ (sounder)} = 1{\cdot}4 \times 10^{-3} \text{ degC}^2 \text{ m}^{-\frac{1}{3}},$$

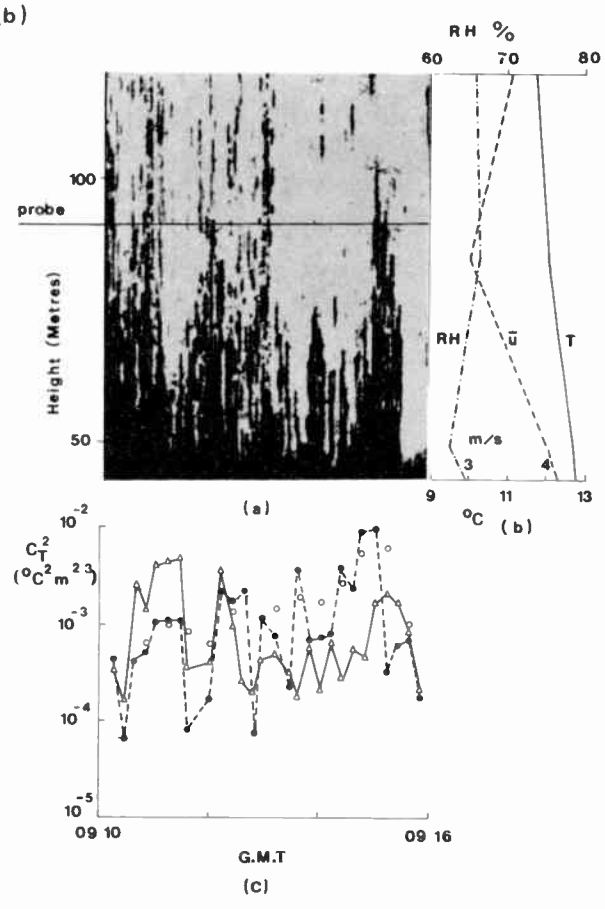


Fig. 7. (a) Facsimile recording of convective conditions on 4 June 1975. (b) Profiles of temperature (T), wind velocity ($\bar{u}$) and relative humidity (RH). (c) Comparison of measurements of C_T^2 from the acoustic sounder △, temperature spectrum ○, and differential temperature measurements ●.

4.2 Bistatic Sounding

More recent studies investigated the feasibility of using off-axis scattered sound, i.e. a bistatic sounder configuration, to estimate C_V^2 and hence the rate of dissipation of turbulent kinetic energy, ε. These experiments have been described in detail by Caughey *et al.*[14] A conventional bistatic configuration was used in this work, see Fig. 1(b). The vertically directed array 1 was used in both transmit and receive modes while array 2 was used only to receive. For comparison purposes the sounder ε estimates were averaged over 51 pulses (i.e. 102 s). Turbulence probe signals were sampled at 80 Hz and, for correspondence with the sounder data, vertical velocity power spectra were computed over 8192 data points, i.e. ~102 s. To sampled at 80 Hz and, for correspondence with the taken at 8 Hz and the mean wind speed across each interval was computed.

The facsimile records of the monostatic and bistatic returns for the period between 1505 and 1545 G.M.T. on 27 October 1975 are shown in Figs. 8(a) and (b) respectively. Tethered balloon profiles of wind speed, temperature and relative humidity across the height range of interest are given in Fig. 8(c). The monostatic echoes indicate only weak and quite variable thermal activity in the lowest 100 m of the boundary layer. At the turbulence probe height (~130 m) echoes were at or below noise level throughout the period. This constrasts strongly with the less variable and more intense bistatic echoes, although these do decrease to a fairly low level by the end of the period. It is worth noting that in the middle of the record when the monostatic echoes fall to the noise level at heights well below 50 m, the bistatic echoes show no appreciable variation. The comparison of the sounder and turbulence probe estimates of the dissipation rate, ε, is given in Fig. 9. It is apparent that the two estimates show the same general trend and decrease from $\sim 3 \times 10^{-3}\ m^2\ s^{-3}$ at the start of the afternoon to $\sim 10^{-5}\ m^2\ s^{-3}$ by the end. This progression fits in well with that expected as atmospheric conditions change from convective through neutral to stable during the afternoon. Overall mean values compare fairly well, bearing in mind the substantial overall change that occurs, being $9{\cdot}9 \times 10^{-4}$ and $3{\cdot}5 \times 10^{-4}\ m^2\ s^{-3}$ from sounder and probe respectively.

In stable conditions buoyancy effects may be significant and local generation of turbulence can occur through breaking wave phenomena.[23,16] This means

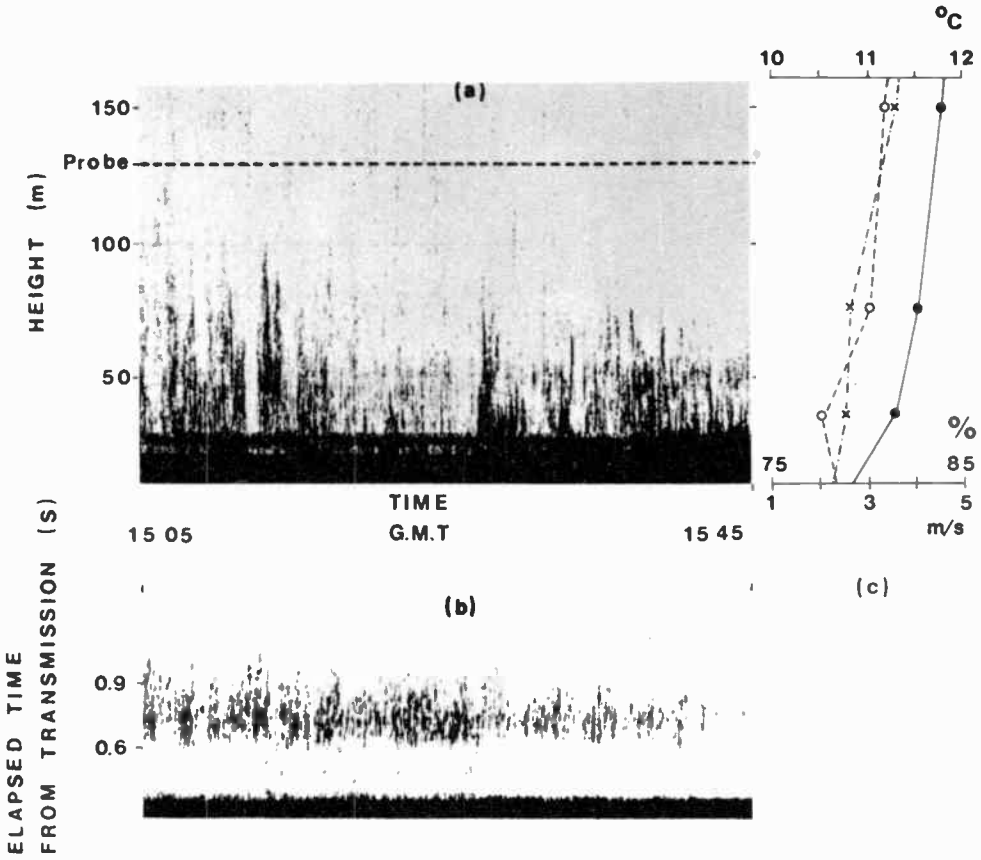


Fig. 8. (a) Monostatic acoustic echoes for the interval 1505–1545 G.M.T. on 27 October 1975. The turbulence probe height is denoted by the dashed line. (b) Bistatic acoustic echoes for the same time interval. The ordinate is elapsed time from transmission, i.e. most of the scattered sound is received at array 2 between 0·6 and 0·9 s after the transmission by array 1. (c) Profiles of wind speed ●, potential temperature × and humidity ○ across the height range of interest.

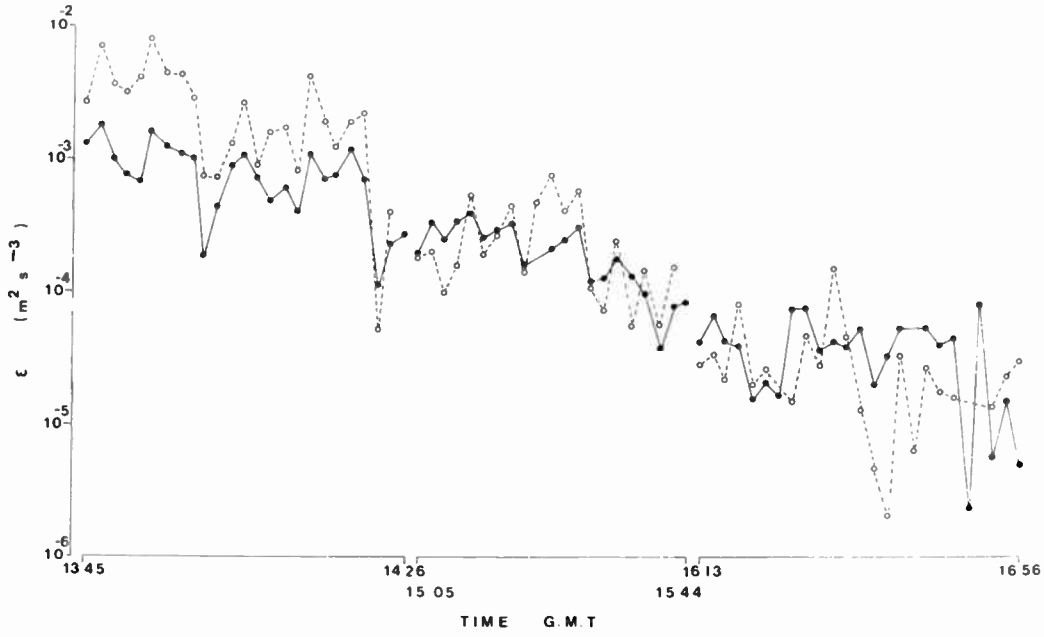


Fig. 9. Time histories of the dissipation rate (ε) estimates from acoustic sounder ○ and turbulence probe ● on the afternoon of 27 October 1975.

that the Kolmogorov inertial subrange expressions for the spectral densities $S_W(n)$ and $S_T(n)$ may not always be valid, particularly in strongly stable conditions when the turbulence becomes weak and intermittent.[30] These considerations indicate that the quantities C_V^2 and C_T^2 are no longer strictly meaningful and that the validity of the scattering cross-section given by equation (1) is in some doubt.

Runs in stable conditions were carried out between 0208 and 0607 G.M.T. on 7 November 1975. Tethered balloon profiles of wind speed and temperature indicated a strongly stable region (Richardson number $\sim 0{\cdot}18$) extending from the surface to ~ 150 m. Bistatic echoes were intense and indicated a significant level of turbulent mixing within the nocturnal inversion. Both the direct and acoustic estimates of ε were large ($\sim 10^{-2}\,m^2\,s^{-3}$). However, there was much greater variability in the acoustic data than from the probe. The scattering cross-section employed here does not include the effects of inhomogeneity and anisotropy, or the effects of refraction by background wind and temperature gradients, so marked discrepancies in the experimental comparison for stable conditions might be expected and this has proved to be the case.

4.3 Height Profiles of C_T^2 and C_V^2

More sophisticated bistatic systems may be used to obtain profiles of C_V^2 (Ref. 24). A rectangular array of loudspeakers is used to generate a wide beam of sound which is probed with a narrow beam square array (see Fig. 1(c)). A second square array operated as a monostatic sounder may be used to obtain a profile of C_T^2. In these experiments two turbulence probes were used to provide data for comparison purposes. Figures 10 and 11 show the comparisons of C_T^2 and C_V^2 data obtained in convective conditions at 46 m and 137 m. The agreement is seen to be excellent and it is noteworthy that a strong lapse in C_T^2 is indicated, whereas C_V^2 decreases only marginally with height. As noted in Section 4.2, stable conditions impose more difficult operating conditions for the sounder and once again a deterioration in the level of agreement is observed.

Profiles of C_T^2 derived from these studies are given in Fig. 12. Those numbered 1–3 correspond to convective conditions and show a rapid decrease in C_T^2 with height—approximating a $-\frac{4}{3}$ decrease as would be expected in free convection conditions.[25] The remainder are from various stable boundary layers and exhibit a much greater degree of scatter and only a slight decrease. Figure 13 presents the corresponding results for C_V^2. Again there is good agreement between direct and acoustic estimates in convective conditions and the data show only a fairly slow decrease with height.[25,26] The profiles in stable conditions are rather more variable and in some cases indicate a slight increase with height. This change in behaviour is also reflected in the turbulence probe data. The agreement obtained during stable stratification therefore depends upon the characteristics and structure of turbulence in the particular situations encountered.

5 Doppler Acoustic Sounding

The comparison studies described here were carried out over a two-year period (1975–1977) and include measurements taken in convective and stable conditions. The sounders employed were as previously described;

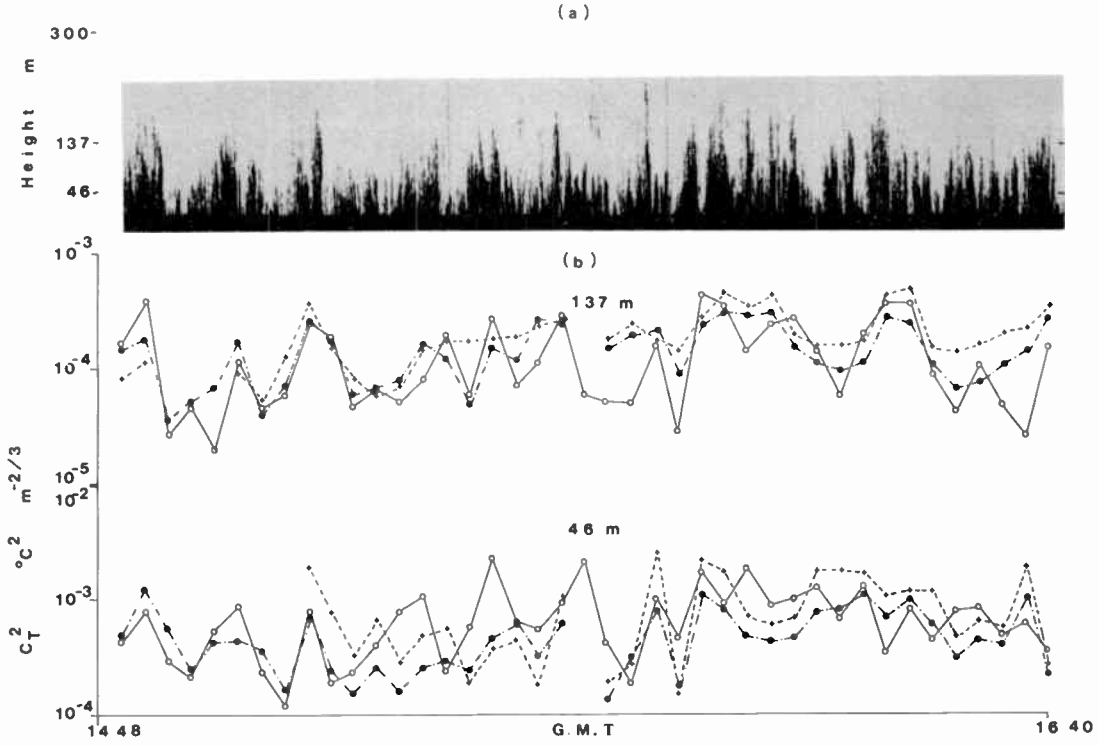


Fig. 10. (a) Monostatic echo pattern recorded in convective conditions between 1448 and 1640 G.M.T. on 21 June 1977. (b) Comparison between sounder estimates of C_T^2 ○, direct spectral estimates + and, from the differential temperature measurements, high-pass filtered at ~1 Hz ● at heights between 46 m and 137 m.

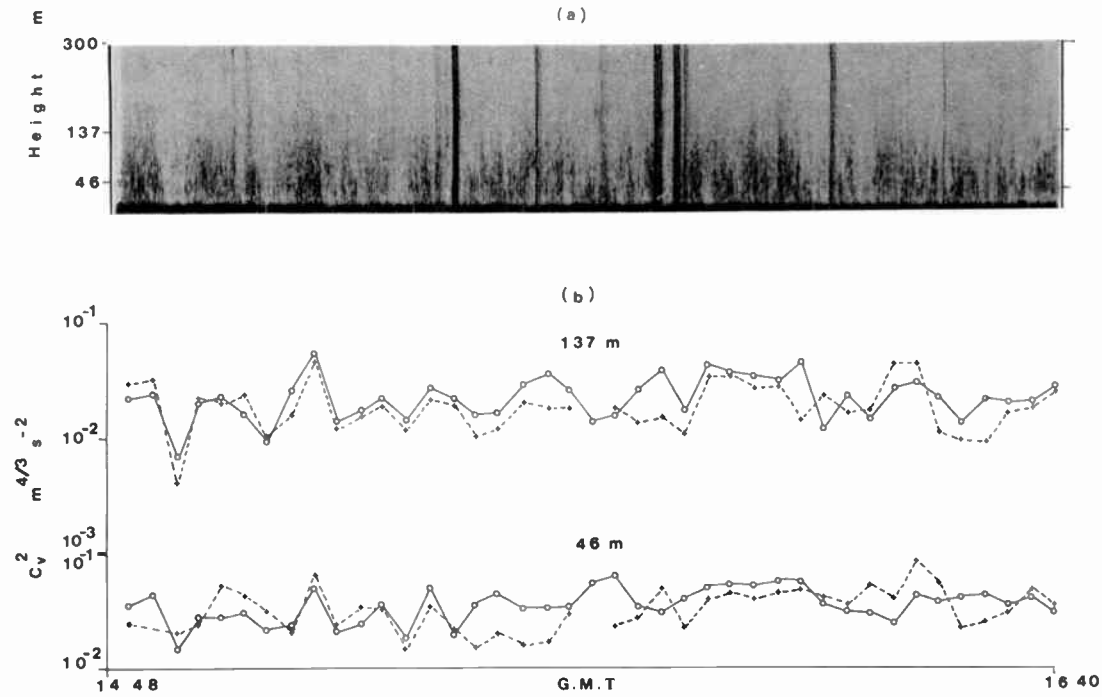


Fig. 11. (a) Bistatic echoes between 1448 and 1640 G.M.T. on 21 June 1977. (b) Comparison between sounder estimates of C_V^2 ○ and from the turbulence probe vertical power spectra + at heights of 46 m and 137 m.

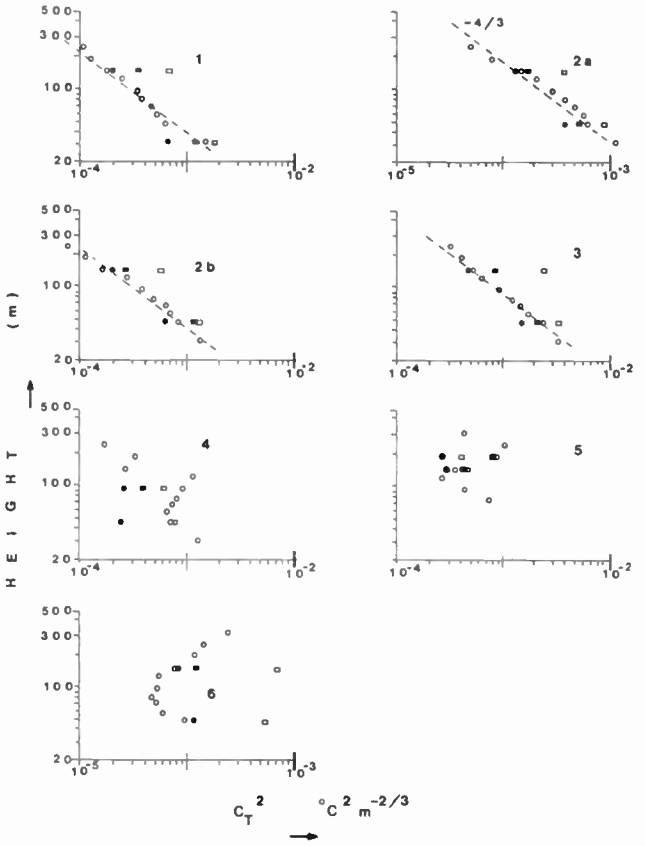


Fig. 12. Vertical profiles of the temperature structure parameter, C_T^2, for a series of runs in convective conditions (1, 2a, 2b and 3) and in stable conditions (4, 5 and 6). Sounder estimates ○. Turbulence probe differential temperature estimates □. Turbulence probe differential temperature estimates filtered at 1 Hz ●. One-dimensional temperature spectrum estimates ■.

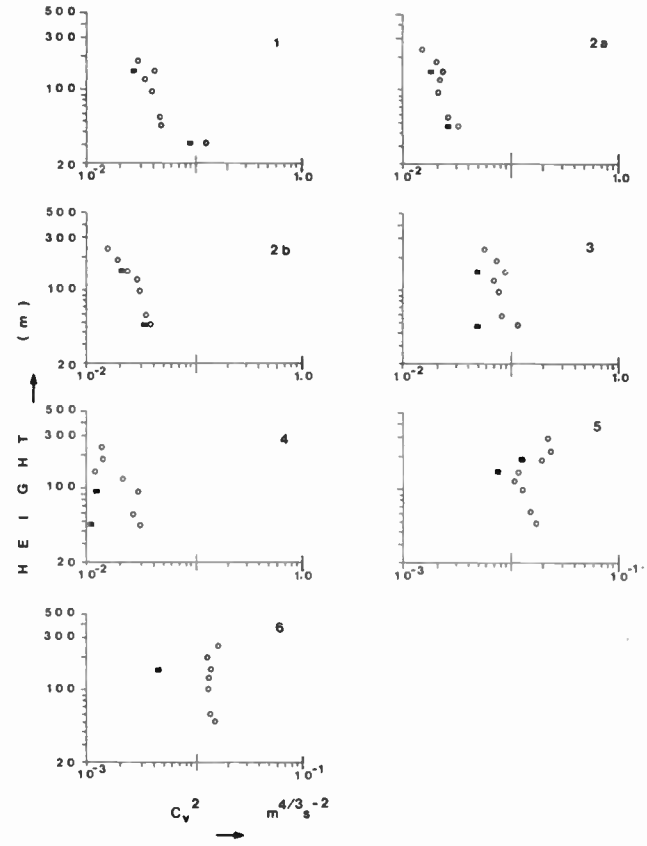


Fig. 13. Vertical profiles of the wind velocity structure parameter, C_V^2, for a series of runs in convective conditions (1, 2a, 2b and 3) and in stable conditions (4, 5 and 6). Sounder estimates ○. Turbulence probe vertical velocity spectral estimates ■.

however, the method used to extract the Doppler shift varied to match the experimental needs. In the early work[27] a simple monostatic sounder was used alongside the turbulence probe to compare estimates of vertical velocity. During these studies the Doppler shift was determined using a Hewlett–Packard spectrum analyser.[28] Figure 14(a) shows the facsimile recording for one of the runs under convective conditions on 18 June 1975. During this run the turbulence probe, providing the direct estimates, was located at a height of ~70 m and some 80 m away from the sounder. Figure 14(b) compares vertical velocities from the sounder and turbulence probe, the former being averaged over 20 sounder pulses and the latter over a corresponding time interval. It is worth noting that reliable estimates of the Doppler shift could be estimated for the weak echo regions between the periods of thermal activity. The results indicate a good correspondence between acoustically and directly estimated vertical velocities.

An example of comparisons obtained during stable conditions on 26 June 1975 is shown in Fig. 15. Even for these small values the correspondence between data from the two techniques is very good. Correlation coefficients between the acoustic and direct estimates for the runs in stable conditions fell in the range 0·7 to 0·9.

These early results were encouraging since they demonstrated for the first time that reasonable estimates of the vertical wind speed could be obtained from the echoes of a monostatic acoustic sounder, under both convective and stable stratifications. However, the range of scales over which useful information could be obtained remained uncertain. A further series of experiments was therefore conducted[30] with some minor modifications to the sounder, e.g. the pulse repetition rate was increased to 1 Hz to improve the spatial resolution. The power spectra of the vertical velocity time series from sounder and turbulence probe at a height of 56 m are shown in Fig. 16. The two sets of measurements compare well and indicate that the first half of the run (solid points) was much more turbulent than the second (open circles). The upper frequency limit for the sounder spectral estimates is set by the need to average over a fairly large number of pulses to obtain a reliable estimate for the vertical velocity since single pulse estimates are highly inaccurate. In the present case an average over 25 pulses was used corresponding to a theoretical instrument error of 0·25 ms^{-1} (Ref. 29).

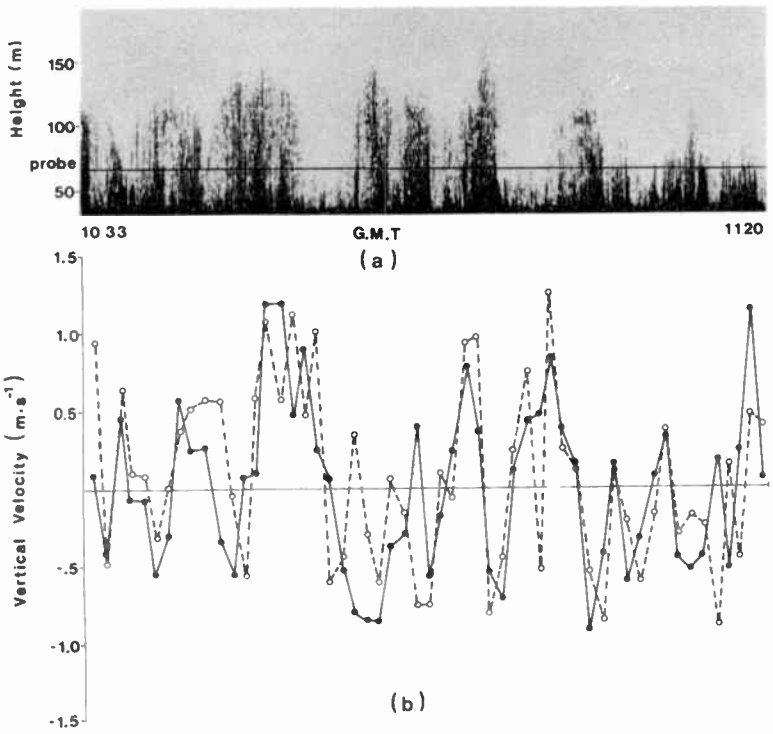


Fig. 14. (a) Monostatic facsimile recording of convective conditions between 1033 and 1120 G.M.T. on 18 June 1975.
(b) Comparison between sounder estimates ○ and direct measurements ● of vertical velocities averaged over 20 sounder pulses (~46 s).

The wavelength, $\lambda_m = \bar{u}/n_m$ (where n_m is the frequency at which the vertical velocity power spectral peak occurs and $\bar{u}$ is the mean wind speed) is widely used in assessing the dispersal of pollutants in the lower atmosphere. The present results indicate that λ_m varied between 250–600 m in the height interval 56–124 m for the first half of the run, which is in fair agreement with that expected from the free convection relationship, $\lambda_m \sim 5{\cdot}9\,Z$, obtained from other experiments.[25] For comparison purposes vertical velocity variances have been computed over the restricted frequency interval 2×10^{-4} to 2×10^{-2} Hz (the variance in this bandwidth is approximately 70% of the total variance). As expected, σ_w^2 shows a distinct increase with height in the markedly convective first run (see Fig. 17(a)). The second run in near neutral conditions yields shorter peak

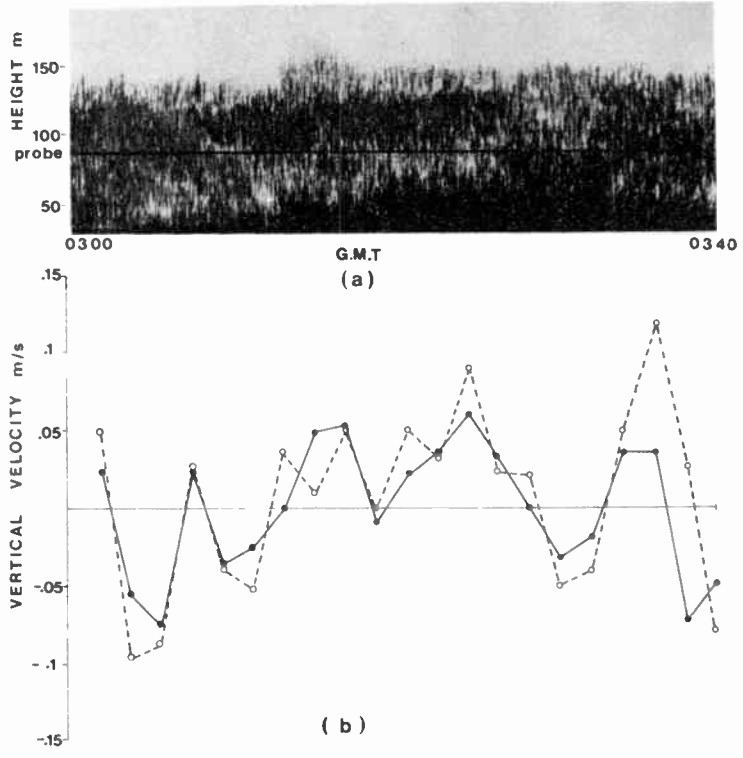


Fig. 15. (a) Monostatic facsimile recording of a surface-based nocturnal inversion between 0300 and 0340 G.M.T. on 26 June 1975.
(b) Comparison between sounder estimates ○ and direct measurements ● of vertical velocities averaged over 50 sounder pulses (~114 s).

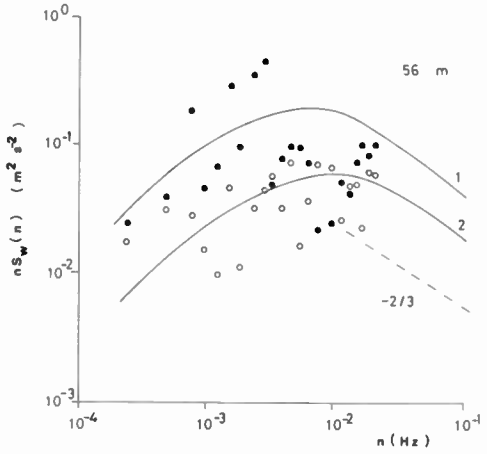


Fig. 16. Comparisons between the vertical velocity power spectra from the turbulence probe (solid curves) and the acoustic sounder on 29 April 1976 (● Run 1, ○ Run 2) at a height of 56 m in the convective boundary layer. The monostatic facsimile record indicated Run 1 to be considerably more turbulent than Run 2. Wind speed 1–3 m/s.

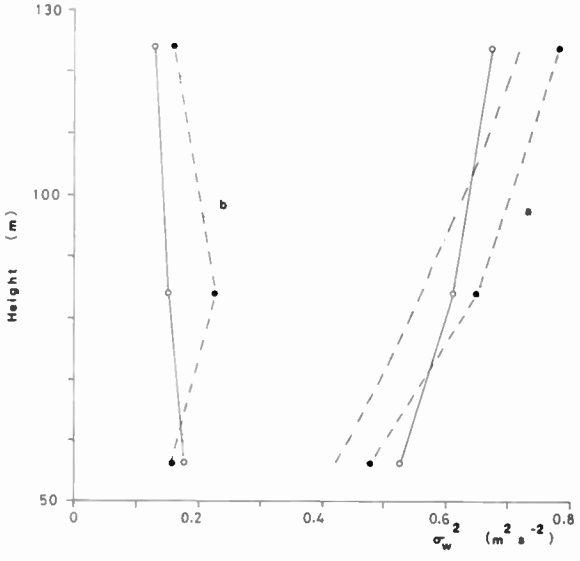


Fig. 17. Vertical profiles of the vertical velocity variance within the range $2 \times 10^{-4} < n < 2 \times 10^{-2}$ Hz from the turbulence probe ○ and acoustic sounder ● for (a) Run 1 and (b) Run 2 on 29 April 1976.

wavelengths (100–200 m) and the magnitudes of the overall vertical velocity variances also decrease significantly (see Fig. 17(b)). The profile shape now indicates a slow decrease of σ_w^2 with height and this relates well to numerical predictions[31] for neutral and slightly unstable planetary boundary layers.

So far only the vertical component of air motion has been discussed. Other experiments have investigated the feasibility of measuring the complete wind vector profile acoustically.[32,33] However, such systems are both sophisticated and expensive and it was considered worth investigating the feasibility of obtaining a component of the horizontal wind using a simple phaselock-loop as the Doppler shift detector.[34] The fan-beam bistatic system was employed so that in principle a profile of the wind component could be obtained. As in previous experiments turbulence probes, capable of measuring the complete wind vector, were flown on a tethered balloon to provide comparison data.

A comparison of the component of the horizontal wind in the plane of the bistatic system obtained in convective conditions with that measured directly is given in Fig. 18. The sounder estimates tend to show somewhat more variability than those from the probe, although the overall agreement is good. Similar results were obtained in stable conditions. Comparisons of the horizontal wind speed spectra were also carried out[34] and these showed quite good agreement for low frequencies and for heights < 100 m. At higher levels and with increasing frequency the acoustic data became less accurate due to poor signal-to-noise ratios.

The vertical and a component of the horizontal velocity variances from sounder and probe are compared in Fig. 19. In computing the variances the bandwidth used was 0·0003 Hz to 0·07 Hz to avoid contamination due to noise. Broad agreement is indicated, although clearly further long-term comparison studies are necessary before any exact specification of the accuracy and limitations of acoustic systems can be expected to emerge.

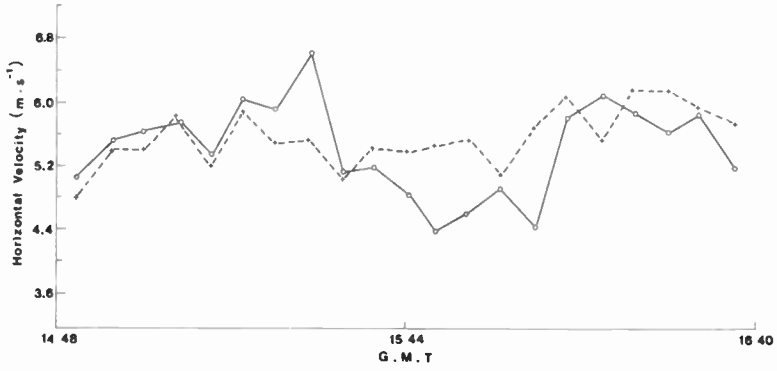


Fig. 18. A comparison of the component of the horizontal wind in the plane of the bistatic sounder obtained acoustically ○ and from the turbulence probe + on 21 June 1977 at a height of 137 m.

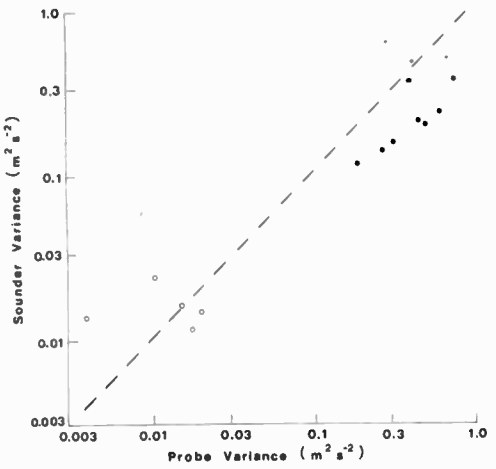


Fig. 19. Comparison of sounder and turbulence probe velocity variances. Vertical velocity, convective conditions (0·0003–0·07 Hz) ●. Vertical velocity, stable conditions (0·0003–0·02 Hz) ○. Horizontal velocity in the plane of the bistatic sounder, convective conditions + and stable conditions ■.

The performance of the phase-locked loop was found to be comparable to that of the spectrum analyser used in earlier experiments. Time histories of vertical velocities compared very well, as did the vertical velocity power spectra. However, as indicated in the earlier data (see Fig. 16), the acoustic data do not quite extend into the inertial subrange. This is because the upper limit to the acoustic data is set by the requirement to average over many pulses to achieve reasonable accuracy in the acoustic velocity estimates. However, with further refinement of this technique it may well be possible to extend these measurements into the inertial subrange. For vertical velocity power spectra obtained in stable conditions the agreement obtained is rather less good and the spectral peak has shifted to a much higher frequency (indicating that the predominant turbulent eddies are occurring on a much smaller scale). Clearly in this case the inherent errors of the technique are more significant since the magnitudes of the vertical velocities are comparable to the expected instrumental error.[29]

6 Concluding Remarks

This paper has provided information on the ability of some acoustic sounding systems to provide quantitative details of the structure of the atmospheric boundary layer. Precise measurement of the received signal strength confirmed that the predominant source of echoes is scattering from small-scale fluctuations of temperature and wind velocity. In this way the sounder could therefore be used to obtain profiles of the structure parameters for temperature and wind velocity and also of the rate of dissipation of turbulent kinetic energy. These quantities are useful to basic studies of the atmospheric boundary layer and are of significance in assessing the mixing capability of the lower atmosphere. Further long-term studies are required to investigate the discrepancies that occasionally occur, especially in stable conditions, and which may in part be attributable to an excess of attenuation over that expected from classical absorption and molecular attenuation.[10] Other possibilities include inhomogeneity and anisotropy in the turbulence field.[24]

The Doppler techniques are inherently more attractive since they do not require precise knowledge of the antenna characteristics or transmitter/receiver gains. Many techniques are available for extraction of the Doppler shift[35] and those reviewed here provide data which compare favourably with those directly measured. The question of an optimum Doppler processing technique is still open to some doubt. More sophisticated methods such as using a fast Fourier transform to compute the spectrum of the received sound can improve performance and reliability, especially in the presence of realistic background noise levels. Further refinements would also permit a greater range of scales to be sampled, principally by improving the signal-to-noise ratio. A computer-controlled acoustic system should be capable of providing near real-time profiles of mean wind speeds and turbulence velocity statistics to heights of several hundred metres. Such a system would find useful applications in detailed studies of the atmospheric boundary layer as well as in establishing a climatology of wind flow over various underlying surfaces.

It is worth noting one other possible area of application for acoustic sounders. This arises from the fact that the absorption of sound in air is a strong function of humidity, temperature and frequency. Suggestions have been made that multi-frequency absorption measurements using acoustic sounders could be used to obtain atmospheric humidity and temperature profiles.[36,37] This would have possible applications in the investigation of multi-path radio propagation effects caused by refractive index gradients, such as those associated with temperature inversions.[28,38] However, some recent multi-frequency sounding experiments have indicated large errors in the derivation of temperature and water vapour estimates.[22] It seems unlikely that temperature profiles can be obtained with useful accuracy. Nevertheless with certain improvements to the sounding system it is likely that water vapour pressure could be determined to within 1–2 mb at heights up to about 150 m which could be of use in studies of microwave propagation in the presence of strong humidity gradients.

The major limitations of acoustic systems will probably arise from the large variability in the level of small scale temperature and velocity fluctuations which occur naturally in the atmosphere. In quiet conditions echoes would fall to noise levels and the Doppler technique would therefore be inoperable. Complex (possibly anisotropic) structure in the turbulence field can also lead to difficulties in the interpretation of

quantitative data, as is already the case for stable conditions. Nevertheless, sounding has provided a valuable stimulus to boundary layer studies and offers a relatively simple and effective way of studying the structure of the lower atmosphere.

7 Acknowledgments

Thanks are due to the Science Research Council which provided financial support for much of the acoustic sounder work, including the design and building of three of the sounders at University College London, under a grant to UCL. Thanks are also due to the staff of the Meteorological Research Unit, Cardington, Bedfordshire, for assistance in the field experiments and computer processing of the data. One of the authors, T. J. Moulsley, is grateful for a Science Research Council CASE award in conjunction with the Meteorological Office.

8 References

1 Beran, D. W., Hooke, W. H. and Clifford, S. F., 'Acoustic echo-sounding techniques and their application to gravity wave, turbulence and stability studies', *Boundary-layer Met.*, **4**, pp. 133–53, 1973.

2 Crease, B. A., Caughey, S. J. and Tribble, D. T., 'Information on the thermal structure of the atmospheric boundary layer from acoustic sounding', *Met. Mag.*, **106**, pp. 42–52, 1977.

3 Cole, R. S., Asimakopoulos, D. N., Moulsley, T. J., Caughey, S. J. and Crease, B. A., 'Some aspects of the construction and use of atmospheric acoustic sounders', *The Radio and Electronic Engineer*, **50**, pp. 585–97, 1980.

4 Wesley, M. L., 'The combined effects of temperature and humidity fluctuations on refractive index', *J. Appl. Met.*, **15**, pp. 43–9, 1976.

5 Tatarski, V. I., 'Wave Propagation in a Turbulent Medium'. Translated from the Russian by R. A. Silverman (Dover, New York, 1961).

6 Monin, A. S., 'Characteristics of the scattering of sound in a turbulent atmosphere', *Sov. Phys. Acoust.*, **7**, pp. 370–3, 1962.

7 Harris, C. M., 'Absorption of sound versus humidity and temperature', *J. Acoust. Soc. Am.*, 40, pp. 148—59, 1966.

8 Sutherland, L. C., 'Review of experimental data in support of a proposed new method for computing atmospheric absorption losses', Rep. DOT–TST–75–87, Dept. of Transport, Washington DC, 1975. (Available from the US National Technical Information Service, Springfield, Va.)

9 Haugen, D. A. and Kaimal, J. C., 'Measuring temperature structure parameter profiles with an acoustic sounder', *J. Appl. Met.*, **17**, pp. 895–9, 1978.

10 Brown, E. H. and Clifford, S. F., 'On the attenuation of sound by turbulence', *J. Acoust. Soc. Am.*, **60**, pp. 788–94, 1976.

11 Neff, W. D., 'Beamwidth effects on acoustic backscatter in the planetary boundary layer', *J. Appl. Met.*, **17**, pp. 1514–20, 1978.

12 Thomson, D. W. and Coulter, R. L., 'Analysis and simulation of phase coherent acdar sounding measurements', *J. Geophys. Res.*, **79**, pp. 5541–9, 1974.

13 Moulsley, T. J. and Cole, R. S., 'A general integral radar equation for the bistatic acoustic sounder', *Boundary-Layer Met.*, 1980. (To be published.)

14 Caughey, S. J., Crease, B. A., Asimakopoulos, D. N. and Cole, R. S., 'Quantitative bistatic acoustic sounding of the atmospheric boundary layer', *Q. J. R. Met. Soc.*, **104**, pp. 147–61, 1978.

15 Asimakopoulos, D. N., Cole, R. S., Caughey, S. J., Moss, S. H. and Readings, C. J., 'A comparison between acoustic radar returns and the direct measurement of the temperature structure of the atmosphere', *Atmos. Envir.*, **9**, pp. 775–6, 1975.

16 Neff, W. D., 'Quantitative evaluation of acoustic echoes from the planetary boundary-layer', NOAA Tech. Rep. ERL 322–WPL–38, Boulder, Colorado, 1975.

17 Readings, C. J. and Butler, H. E., 'The measurement of atmospheric turbulence from a captive balloon', *Met. Mag.*, **101**, pp. 286–98, 1972.

18 Caughey, S. J., 'The Cardington turbulence instrumentation'. Unpublished UK Meteorological Office Turbulence and Diffusion Note, No. 82, 1977.

19 Haugen, D. A., Kaimal, J. C., Readings, C. J. and Rayment, R., 'A comparison of balloon-borne and tower-mounted instrumentation for probing the atmospheric boundary-layer', *J. Appl. Met.*, **14**, pp. 540–5, 1975.

20 Caughey, S. J., 'Boundary-layer turbulence spectra in stable conditions', *Boundary-layer Met.*, **11**, pp. 3–14, 1977.

21 Caughey, S. J., Wyngaard, J. C. and Kaimal, J. C., 'Turbulence in the evolving stable boundary layer', *J. Atmos. Sci.*, **36**, pp. 1041–52, 1979.

22 Moulsley, T. J., 'Acoustic sounding of the atmosphere', Ph.D. Thesis, University of London, 1979.

23 Caughey, S. J. and Readings, C. J., 'An observation of waves and turbulence in the Earth's boundary-layer', *Boundary-layer Met.*, **9**, pp. 279–96, 1975.

24 Moulsley, T. J., Asimakopoulos, D. N., Cole, R. S., Crease, B. A. and Caughey, S. J., 'Measurement of boundary-layer structure parameter profiles by acoustic sounding and comparison with direct measurements', *Q. J. R. Met. Soc.*, **107**, January 1981. (To be published.)

25 Kaimal, J. C., Wyngaard, J. C., Haugen, D. A., Cote, O. R., Izumi, Y., Caughey, S. J. and Readings, C. J., 'Turbulence structure in the convective boundary-layer', *J. Atmos. Sci.*, **33**, pp. 2152–69, 1976.

26 Caughey, S. J. and Palmer, S. G., 'Some aspects of turbulence structure through the depth of the convective boundary-layer', *Q. J. R. Met. Soc.*, **105**, pp. 811–27, 1979.

27 Caughey, S. J., Crease, B. A., Asimakopoulos, D. N. and Cole, R. S., 'A comparison of acoustic Doppler vertical velocities with direct measurements in the atmospheric boundary-layer', *Nature*, **262**, pp. 274–6, 1976.

28 Asimakopoulos, D. N., 'Atmospheric acoustic sounding and radio propagation', Ph.D. Thesis, University of London, 1976.

29 Spizzichino, A., 'Discussion of the operating conditions of a Doppler radar', *J. Geophys. Res.*, **79**, pp. 5585–91, 1974.

30 Asimakopoulos, D. N., Cole, R. S., Crease, B. A. and Caughey, S. J., 'A comparison of acoustic Doppler vertical velocity power spectra with direct measurements', *Atmos. Envir.*, **12**, pp. 1951–6, 1978.

31 Deardoff, J. W., 'Numerical investigation of neutral and unstable planetary boundary-layers', *J. Atmos. Sci.*, **29**, pp. 91–115, 1972.

32 Beran, D. W., Willmarth, B. C., Carsey, F. C. and Hall, F. F., 'An acoustic Doppler wind measuring system', *J. Acoust. Soc. Am.*, **55**, pp. 334–8, 1974.

33 Kaimal, J. C. and Westcott, J. W., 'An acoustic Doppler sounder for measuring wind profiles in the lower boundary layer', in Preprints of the 17th Radar Meteorology Conference, pp. 290–6, American Meteorological Society, Boston, Mass., 1976.

34 Moulsley, T. J., Asimakopoulos, D. N., Cole, R. S., Crease, B. A. and Caughey, S. J., 'Acoustic Doppler wind velocity measurements using a simple phaselock-loop', *J. Phys. E. (Sci. Instrum.)*, **14**, 1981 (To be published).

35 Brown, E. H. and Hall, F. F., 'Advances in atmospheric acoustics', *Rev. Geophys. & Space Phys.*, **16**, pp. 47–110, 1978.

36 Little, C. G., 'Acoustic methods for the remote probing of the lower atmosphere', *Proc. IEEE*, **57**, pp. 571–8, 1969.

37 Gething, J. T. and Jensen, D., 'Measurement of temperature and humidity by acoustic echo sounding', *Nature (Physical Science)*, **231**, pp. 198–200, 1971.

38 Ho, K. L., Mavrokoukoulakis, N. D. and Cole, R. S., 'Propagation studies on a line-of-sight microwave link at 36 GHz and 110 GHz', *IEE J. Microwaves, Optics and Acoustics*, **3**, pp. 93–8, 1979.

Manuscript first received by the Institution on 18th March 1980 and in final form on 8th August 1980
(Paper No. 1964/AMMS 104)

UDC 551.46: 621.396.969.029.55
Indexing Terms: Oceanography, waves, Radar, reflection

Remote sensing of the sea-surface by dekametric radar

Professor E. D. R. SHEARMAN,
B.Sc.(Eng.), S.M.I.E.E.E., F.R.Met.S., C.Eng., F.I.E.E.*

SUMMARY

The evolution of the technique of remote sensing of the sea-surface by dekametric radar is traced from the original observation of the Bragg resonant scattering phenomenon by D. D. Crombie in 1955 to its present applications to the measurement of surface wind-direction and wind-speed, wave-height spectrum, wave-directional spectrum and surface-currents. The central role played by the Doppler spectrum of the radar echoes is illustrated. The generation and spectral properties of wind-driven waves on the sea-surface are discussed including wave–wave interactions and the nature of electromagnetic scattering by first and second-order processes is described.

An account is given of the techniques used for UK experiments on ground-wave radar surveying the Celtic Sea from West Wales. Synthetic aperture experiments using a moving vehicle permitted the directional spectrum of the sea-surface to be determined. Sky-wave experiments from a radar in Southern England permitted comparisons with sea-state and surface winds measured in the Rockall Bank area of the North Atlantic by oceanographic vessels and buoys.

Finally, work in progress in the US and UK on improved narrow beam and interferometer radars and new methods for inversion of measured Doppler spectra to yield sea-wave spectra are reviewed.

1 Introduction

The extension of Man's engineering activities from sea coasts to progressively more distant areas of continental shelves for oil, gas and wave energy exploitation has underlined the need for improved techniques for gathering data on waves and currents. The oil-spillage hazards involved in the operation of large oil tankers constitute a further reason for improving such data, particularly on a real-time area-surveillance basis, since the lifetime and drift of oil slicks depend on waves and currents.

Until the recent introduction of remote sensing techniques, such data was obtained by point sampling using buoys or ships. Remote sensing by radar, whether from aircraft, satellites or by the shore-based h.f. radar techniques using dekametric wavelengths, discussed in this paper, has made possible the surveillance within a short time of wave conditions over a large area of the sea.

The aircraft and satellite radars use microwave wavelengths which interact with the small capillary ripples and foam on top of the longer gravity waves which carry the main energy. The deduction from such microwave radar echoes of the characteristics of the gravity waves is thus indirect, involving calibration against known wave conditions. This, and the limitations in time and space sampling associated with aircraft flights and with low-altitude satellite orbit geometry, are in contrast to the continuous direct observation of wave energy at long wavelengths possible with dekametric radar, and its ability to measure surface currents. Satellite radars, as evidenced by the short-lived *Seasat* satellite in 1979, have compensating strengths, in particular the world-wide survey possibility. The analysis of the radar altimeter, scatterometer and side-looking synthetic aperture radar in *Seasat* are at present yielding valuable ocean data, in many ways complementary to that provided by dekametric radar.

The basis of dekametric radar dates from the discovery by D. D. Crombie,[1] working in New Zealand in 1955, that at 13·56 MHz (22 metres wavelength) radar sea-echo detected at a coastal site had a characteristic Doppler spectrum with one strongly dominant spectral component. The Doppler frequency shift of this component, 0·376 Hz, corresponded to that expected from the velocity of sea-waves of one-half of the radio wavelength travelling directly towards the radar.† (Sea-waves are dispersive, so each wavelength has a characteristic velocity.) This was the first intimation that a selective process was occurring in the scattering mechanism at the sea surface. This mechanism is now known as Bragg resonant scattering from its relationship to the diffraction effect first identified by Bragg in the scattering of X-rays from the regular rows of atoms in a crystal. The process is illustrated in Fig. 1.

* *Department of Electronic and Electrical Engineering, University of Birmingham, P.O. Box 363, Birmingham B15 2TT.*

† Crombie's radar was not able to measure the sense of the Doppler shift; there may also have been a receding wave present.

0033–7722/80/110611+13 $1.50/0

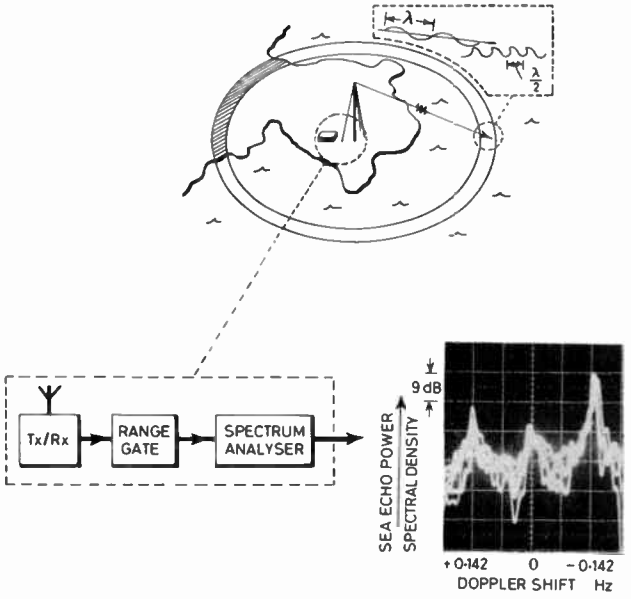


Fig. 1. Illustration of the basic technique of remotely observing ocean surface conditions using m.f./h.f. radar. The approach and recede resonant lines are a consequence of first-order Bragg scattering from the sea. Seven superimposed spectra, obtained during seven consecutive on-line spectral analysis periods, are shown. Data for: 1210 UT, 11.12.77, range 120 km, frequency 1·95 MHz, location as Fig. 9.

The importance of this mechanism is that the radar can be used to probe particular wave components in the chaotic pattern of waves on the sea surface, this pattern being made up of waves having a wind-dependent spectrum of wavelengths and spread of directions. By suitably framed experiments with multi-frequency radars having steerable beams, Crombie envisaged a coastal-based radar sea-wave spectrometer. Such radars employ a vertically-polarized transmission from a coastal site, so obtaining low-loss propagation over the sea by the Norton surface wave mode (the mode used in conventional medium frequency broadcasting), with radar sea-echo obtainable at ranges up to 100–500 km depending on the radar power and frequency. This is the 'ground-wave radar' mode of operation. In his experiments, Crombie also saw evidence of higher-order scattering effects, a precursor of later discoveries.

In 1969, Ward[2], working in Australia, presented characteristic sea Doppler spectra obtained from ocean areas at ranges covering the area from 1,000 to 3,000 km using sky-wave propagation by way of the ionosphere, a phenomenon also reported by Tveten.[3] This extended the potential area of surveillance of sea-state ten-fold in range, although with a requirement for good phase stability in the ionospheric path which restricts the usable time of this 'sky-wave radar' mode. (Back scatter from the Earth's surface had of course been used for many years in the study of sky-wave propagation and communication, and the factors which influence its propagation have been fully reviewed.[4,5]) In 1972 Long and Trizna[6] showed that the relative magnitude of the Bragg resonant lines from approaching and receding sea-waves (see Fig. 2) could be used to infer wind direction on the sea surface. Using a sky-wave h.f. radar at Chesapeake Bay, Maryland, USA, they mapped wind circulation around an Atlantic storm extending over an area between 1,100 and 2,000 km from the radar. Later work led to the determination of wind speed as well as direction.[7]

Returning to 1971 and Ward's sea Doppler spectra, Hasselman[8] made use of these experimental results to test the validity of a new analysis of second-order effects in the sea scattering mechanism, and this analysis, further refined by Barrick[9,10,11] and Lipa[12] has led to a new dimension in the potential of h.f. radar for probing sea-waves. This arises because the spectrum observed from one patch of sea at one radar frequency reveals in its second-order structure (see Fig. 2) the complete wavelength spectrum of sea-waves. This implies that it should no longer be necessary to measure sea-echo amplitude at different transmitter frequencies to probe different sea wavelengths (including inconveniently low frequencies to probe the longest and most energetic sea waves), as envisaged by Crombie. Furthermore, relative amplitude measurements of first and second-order components in the echo spectrum at one radar frequency are inherently more accurate than absolute measurements of radar returns at different transmitted frequencies. However, radar echo spectra of the quality needed to achieve this inversion process to yield sea spectra put high demands on the engineering of the radar and signal processing equipment and, for sky-wave data, on ionospheric stability. Nevertheless, the technique appears very promising.

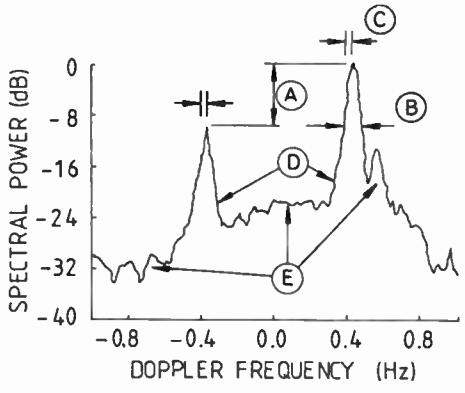


Fig. 2. Features of radar spectra used for sea-state measurement
A Ratio of two first-order Bragg lines: Wind direction
B −10 dB width of larger first-order Bragg line: Wind speed
C Doppler shift of first-order Bragg lines from expected values: Radial component of surface current
D Magnitudes of first-order Bragg lines: Ocean wave-height spectrum for one wave-frequency and direction
E Magnitude of second-order structure: Ocean wave-height spectrum for all wave-frequencies and directions
Sky-wave data for 1000 UT, 23.8.78; Frequency 15 MHz;
Data window, Hanning, FFT 1024 points, Averages 10; Slant range 1125 km; Sea area as Fig. 12.

Three other developments complete the background of key ideas and techniques from which h.f. radar oceanography has developed.

In the ground-wave radar technique Peterson, Teague and Tyler[13] of Stanford University, showed in 1970 how bistatic radar observations of sea echo resulted in the sorting by Doppler frequency of the returns from different directions. Provided that the delta-function-like Bragg resonant peaks are strong compared with the second-order continuum, the shifted Bragg peaks from different directions can be identified and used to map the amplitude of the probed wave components on the sea surface: this principle was used with good effect by the authors, who made use of Loran A transmissions at 2 MHz to map the ocean wavelength spectrum and direction of the sea aroused by the trade wind at Hawaii. It does, however, invoke an assumption of homogeneity of the sea-state over the area surveyed.

Another technique[14] introduced by the Stanford University group was the synthetic aperture radar method, particularly valuable in radar observations at low h.f. or m.f., as in experiments at 2 MHz with Loran A pulse transmitters. At 2 MHz ($\lambda = 150$ m) the achievement of a beamwidth of 8°, for example, requires an antenna array 1 km wide, an expensive installation and one for which a coastal site is not easily found. By transmitting from a fixed site, carrying a receiver in a vehicle along a straight path 1 km long and recording the sea-echo in amplitude and phase, the problem is converted from a major hardware task to an off-line computing task. This technique will be discussed further below in connection with work at Birmingham.

The last basic technique arose from observations described by Crombie[15] of ground-wave sea-echoes from the Gulf Stream with omnidirectional antennas. The effect arises because the radially-travelling Bragg resonant waves which are being observed on a particular patch of sea have a characteristic velocity *relative to the water surface* on which they are travelling. If this water surface has a velocity component, due to tidal streams or currents, towards the radar, the radial velocity of the approaching waves will be increased and therefore also their positive Doppler shift, while the radial velocity of the receding waves will be decreased. As a result the Doppler spectrum is displaced towards positive frequencies in the manner shown in Fig. 2. Over the whole field of view, the superposition of spectra, shifted by varying amounts by surface currents, as received from different directions, results in the effect seen in Fig. 3.[16] The spectra of the Bragg resonant lines are no longer narrow, but broadened out, the width corresponding to the Doppler broadening expected from the calculable approach and recede line-of-sight velocities of the water surface over the field of view.

In the experiment described by Crombie, two antennas spaced along the shoreline were used for reception. It was found possible to select for each of the two antennas that echo component having a particular range and Doppler shift and to measure the phase difference between the two received versions of this echo component. From this phase difference, knowing the radio wavelength and antenna spacing, the angle of incidence on the shoreline was computed.

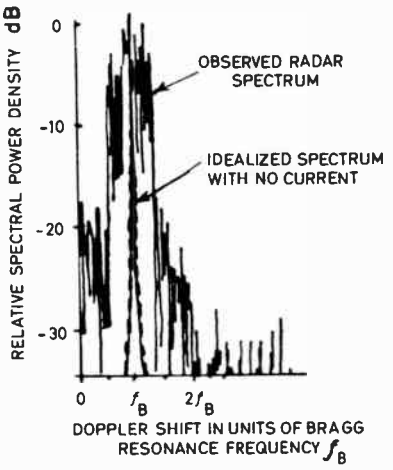


Fig. 3. Smearing of Doppler spectrum by surface current. Spectrum observed by ground-wave radar located at Fort Lauderdale showing smearing of Bragg resonance by Gulf Stream current shear (after Crombie[15]).

This concept has led to the evolution by the US National Oceanic and Atmospheric Administration (NOAA) of a current measuring radar,[16] now undergoing advanced development. This has produced very useful data on estuarial flow patterns, and since it exploits the simplification of the antenna system resulting from the sophisticated on-line computing now possible, achieves the virtue of transportability, permitting its deployment in a variety of locations. It should be noted that since the radar measures the radial component of current flow only, two such radars must be deployed to map current vectors. Also ambiguities in angle measurement can arise if two or more echoes return from different directions with the same Doppler shift.

It will be seen that h.f. radar oceanography has revealed some exciting possibilities of the mapping in real time of sea-state, surface wind and current flow patterns in coastal regions and more distant waters. A number of techniques have been mentioned and two new techniques (proposed at the University of Birmingham) will be discussed in this paper. The present need is for a clear assessment of the accuracy and freedom from ambiguity achievable with the different techniques available (and any new methods that emerge) and a selection of the best approaches to solving user needs. Progress in the UK towards achieving these aims is the main subject of this paper.

Work in this field in the UK began in 1975 at the University of Birmingham, building on earlier research

in the use of sky-wave radar for h.f. propagation analysis. Early sky-wave observations of Atlantic weather systems using an h.f. radar with a 4° beam steerable over ±30° (see Fig. 4) have been described elsewhere,[17] as have experiments on ground-wave radar for sea-state sensing in the Irish and Celtic Seas using Loran A pulse transmissions at 1·95 MHz.[18]

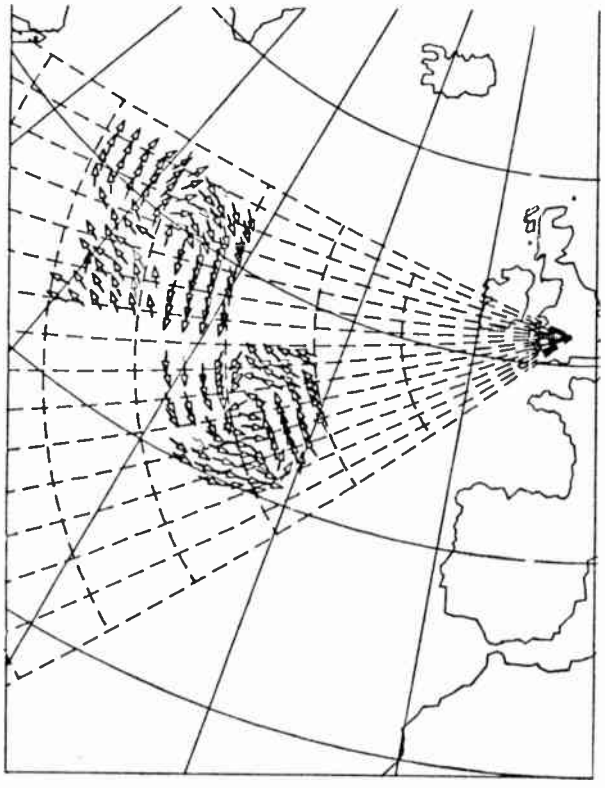


Fig. 4. Coverage of sky-wave radar used for observations of sea-state in the North Atlantic. Computer-drawn map with wind vectors representative of high and low pressure systems extracted automatically from input of simulated Doppler spectra.

This early work led in 1978 to a major joint project with the Science Research Council Appleton Laboratory, aimed at exploiting both the ground-wave and sky-wave techniques in a UK context. As a first step, advantage was taken of the opportunity for obtaining 'sea-truth' from the international Joint Air Sea Interaction project (JASIN) in the Summer of 1978, in which fourteen oceanographic research vessels were deployed in the Rockall Bank area of the North Atlantic. As the purpose of JASIN was to obtain a better understanding of the interaction of wind, waves and currents, a concentrated programme was carried out to measure these quantities from ships, buoys, balloons, aircraft and the *Seasat* satellite. It thus provided a unique opportunity for comparing remote sensing techniques, such as h.f. radar, with *in-situ* measurements: the analysis of these is still in progress. A special experiment was framed by the Birmingham University/Appleton Laboratory group of which an account is given below.

In the following Sections, a brief account of the oceanography of wind and waves and scattering theory is given, followed by an account of the UK experimental programme in ground-wave and sky-wave radar and its results, in which signal processing plays a key role. Finally, some considerations and possibilities in the design of future sea-state radars are discussed.

2 Wind-driven Waves and Swell on the Sea-surface

Before discussing the quantitative spectral description of waves on the sea surface, it may be useful to give a short descriptive account of the manner in which the wind gives rise to waves.

If wind blows across a previously undisturbed sea—a 'glassy calm'—an area of small ripples of wavelength from a fraction of a centimetre to a few centimetres is produced on the surface, described by the sailor as a 'catspaw'. For wavelengths of below 1·7 cm, the dominant restoring force for waves on water is capillary (surface tension), and such waves are therefore known as capillary waves. If the wind continues, the surface roughness produced by the capillary waves increases the friction experienced by the wind, more momentum is transferred to the water surface and longer waves are produced. For waves of length longer than 1·7 cm, the dominant restoring force is gravity, and such waves are known as 'gravity-waves'. The presence of these longer waves further roughens the surface and increases the transfer of momentum, progressively longer wavelength gravity-waves being set up. Some interchange of momentum between waves of different lengths also takes place by non-linear processes. Eventually, if the wind continues for long enough and over a large enough extent of water (the 'fetch'), an equilibrium is reached for which the amplitude of waves of each wavelength remains constant. This condition is known as a 'fully-developed sea' for the particular wind-speed existing.

As already mentioned, wave propagation on the sea-surface is dispersive, the dispersion relationship between wavelength and phase-velocity being as shown in Fig. 5 with a velocity increasing with wavelength for gravity-waves. For a given wind speed, momentum transfer can take place to waves of speed up to the wind-speed, but

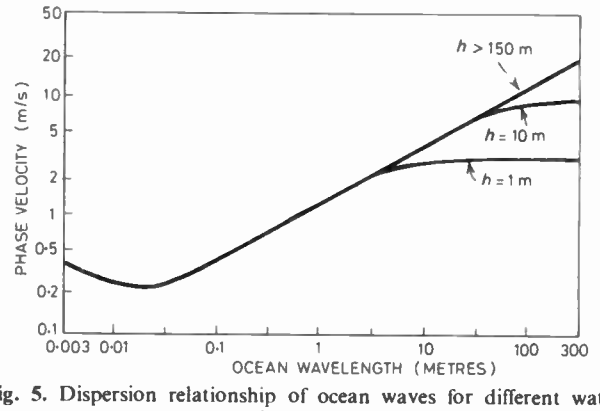


Fig. 5. Dispersion relationship of ocean waves for different water depths (h).

beyond this speed no mechanism exists.† A characteristic feature of sea waves is thus of a spectrum of wavelengths extending from capillary wavelengths up to a maximum wavelength, which for a fully-developed sea corresponds to a wave-speed equal to the wind-speed.

The form of this spectrum is influenced by the size of the water basin in which the sea is confined, even a basin as large as the North Sea having a significant modifying constraint. In the open sea a useful simple model for the spectrum is the Phillips model shown in Fig. 6, in which the spectral cut-offs for various wind-speeds are shown. It will be seen that the greatest wave energy is exhibited by the longer wavelengths, and it is these waves, which are comparable in size with ships, breakwaters and oil rigs, with which man has the most direct interest.

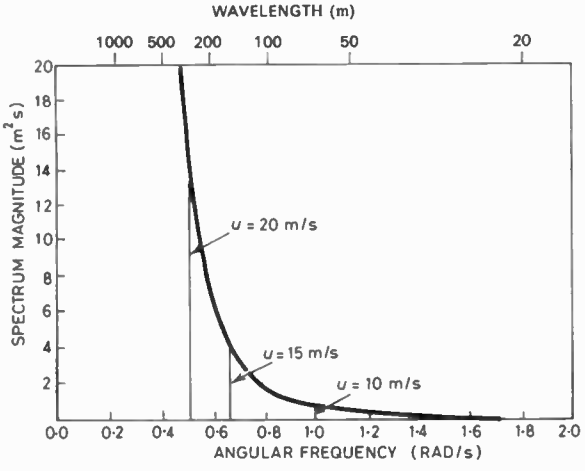


Fig. 6. Phillips wave energy-spectrum.

The other important quantity in addition to wave energy and wavelength is wave direction. In an area of sea over which a steady wind blows, the predominant waves propagate in the direction of the wind, but there is a spread of wave energy about this direction, which can be summarized for a particular wavelength in the form of a 'polar diagram' such as Fig. 7.

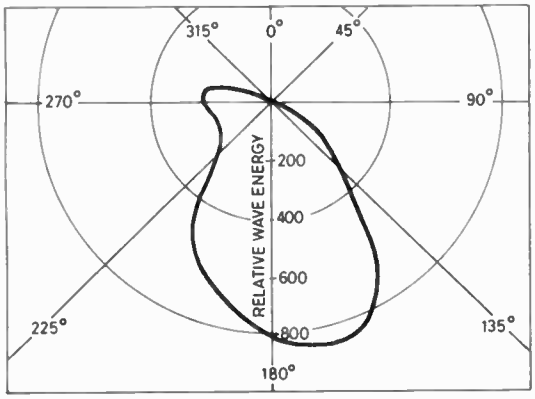


Fig. 7. Directional wave energy spectrum (polar diagram) from 'SWOP' aircraft stereo photographs. 60 m waves.[32]

† This statement must be qualified if a minor exchange of energy by 3rd order non-linear processes is included.

To summarize these concepts mathematically, we may describe the distribution of wave energy with wavelength and direction in the form of a spectral function.

In practice it is more common to express the energy in terms of the water wave frequency f_w rather than wavelength L, where the two are related by

$$f_w = V_w/L, \tag{1}$$

where V_w is the phase velocity of the wave, which for gravity wave is,

$$V_w = (gL/2\pi)^{\frac{1}{2}}, \tag{2}$$

where g is the acceleration due to gravity.

The water wave frequency f_w is of course the number of waves passing a given point per second, and its reciprocal, the wave period in seconds, is the quantity used by a sailor when he describes as a '10 second wave', a wave which takes 10 seconds to pass a fixed point.

The wave energy per square metre of ocean surface can be expressed in terms of a wind-speed dependent non-directional wave-height variance spectrum $S(f_w, u)$, whose dimensions are $m^2\ Hz^{-1}$. In terms of this,

$$(\text{wave energy})/m^2 = \rho g \int_0^\infty S(f_w, u)\,df_w \quad \text{joules/m}^2 \tag{3}$$

where ρ is the density of water and u is the mean wind speed. (Often the factor ρg is omitted, and the quantity $S(f_w, u)$ is referred to loosely as the temporal wave energy spectrum). This non-directional wave-height variance spectrum is the quantity measured by a buoy which floats up and down on the water surface as waves travel by in various directions. The form of the Phillips spectrum for various wind velocities is shown in Fig. 6.

To include direction we form a temporal, directional wave-height variance spectrum $S(f_w, \theta, u)$, whose dimensions are $m^2\ Hz^{-1}\ rad^{-1}$. In terms of this spectrum,

$$(\text{wave energy})/m^2 = \rho g \int_{-\pi}^{+\pi} \int_0^\infty S(f_w, \theta, u)\,df_w\,d\theta \quad \text{joules/m}^2 \tag{4}$$

where θ is the angle between the direction of travel of a particular wave component and the reference direction, usually the direction of the mean wind. (Other useful formulations of these spectra appear in terms of angular wave frequency $\Omega = 2\pi f_w$, wavelength L, wave number $K = 2\pi/L$, also expressible as a vector in the wave direction, $\mathbf{K}$, and the orthogonal components K_x, K_y of $\mathbf{K}$.[19])

The temporal, directional spectrum may be conveniently factored into temporal and directional expressions:

$$S(f_w, \theta, u) = S_0(f_w, u)\,G(\theta, u) \tag{5}$$

where $G(\theta, u)$ is dimensionless

$$\int_{-\pi}^{+\pi} G(\theta, u)\,d\theta = 1.$$

The directional factor $G(\theta, u)$ is difficult to measure physically, the commonest instrument used being a wave tilt buoy containing a compass as a directional reference. Airborne stereographic photography has also been used, although analysis is very laborious (Fig. 7). An expression, due to Munk,[19] used to fit measured directional spectra is

$$G(\theta, u) = [\varepsilon + (1-\varepsilon)\cos^s(\theta/2)] \qquad (6)$$

where s, the directional spreading parameter, is a function of u and L. Wave-tilt measurements are not in themselves sensitive enough to reveal the wave component ε which represents up-wind wave propagation; ε was included by Munk to take into account consistent radar-oceanographic evidence of such components.

The spread s of the directional spectrum of wind-driven waves varies with wave frequency, the spreading being greater for the higher frequency, shorter wavelength, gravity waves. When wind-driven waves travel out of the area in which wind exists, the shorter wavelength waves degenerate, leaving the longer wavelength components travelling onwards over great distances. Such waves are referred to as 'swell'. They become near-sinusoidal and, when a great distance from the source area, appear near-unidirectional with parallel wave fronts.

3 Scattering of Dekametric Radio Waves by the Sea

Figure 1 has shown a simplistic explanation of the Bragg resonant scattering mechanism; to compute the magnitude of the back-scattered radio waves or to take into account second-order effects, a fuller treatment is required. The most profitable theoretical approach has been the extension and adaptation to electromagnetic scattering by Rice[20] and Barrick[9] of Rayleigh's treatment of the scattering of sound waves from a corrugated surface.[21] In this 'perturbation' method, a plane radio wave is considered to be incident on a corrugated or rough conducting surface. The vector sum of the incident and scattered waves at the surface of the conductor must satisfy the electromagnetic boundary conditions, in particular that the tangential component of the electric field is zero. (A more complex boundary condition is needed when the finite conductivity and permittivity of sea water are taken into account.)

The scattered wave-field is represented as the summation of an angular spectrum of plane waves, these plane waves interfering together on the conductor surface to give a resultant electric field, whose component along the surface exactly cancels the component of the incident wave along the surface. The problem to be resolved is thus to determine the phase, amplitude and polarization of the constituent scattered plane waves making up the angular spectrum in order to satisfy this condition. This problem is closely related to that of determining the angular spectrum of waves emerging from a thin phase-changing screen illuminated by a plane wave, a situation examined by Ratcliffe.[22] (Ratcliffe's treatment is recommended for the physical insight it gives into the diffraction process.)

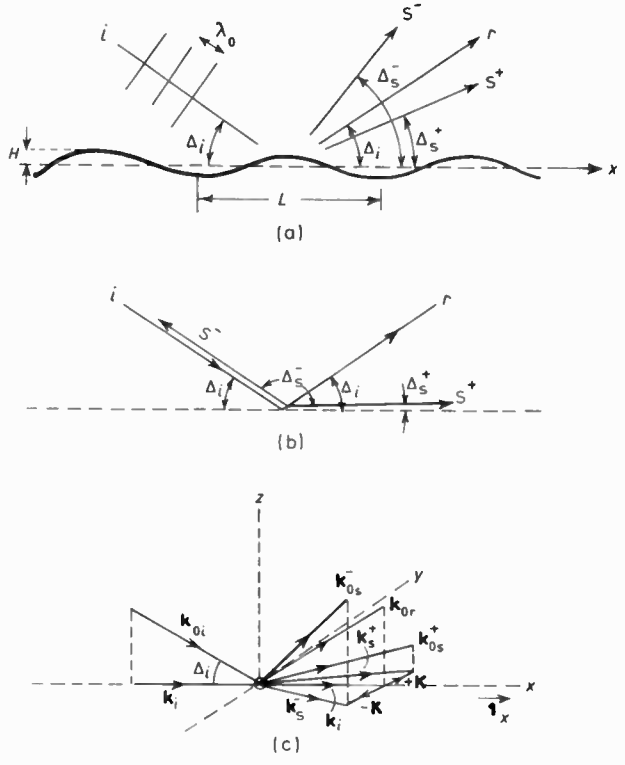


Fig. 8. (a) Scattering from a sinusoidally corrugated surface with $H \ll \lambda_0$, i, r, s indicate the incident, specularly reflected and first-order scattered waves.
(*b*) Back-scatter case, $\Delta_s^- = \pi - \Delta_i$.
(*c*) Showing general three-dimensional geometry with vector construction for scattered radio waves.

As a simple example, let us consider a downcoming plane radio wave, wavelength λ_0, incident at a grazing angle Δ_i (Fig. 8 (a)). We assume there exists on this sea surface a swell in the form of sinusoidal corrugations of height $H \ll \lambda_0$ and of wavelength L, travelling in the plane of incidence. The form of the angular spectrum of the scattered waves then simplifies to three components only, the specular component at grazing angle Δ_i, slightly reduced in amplitude relative to that for a flat surface, and two side waves at grazing angles Δ_s given by

$$\cos\Delta_s = \cos\Delta_i \pm \lambda_0/L \qquad (7)$$

We note that if the water wavelength L is such that λ_0/L is equal to 2 cos Δ_i, then the scattered component corresponding to the negative sign has

$$\cos\Delta_s = -\cos\Delta_i = \cos(\pi - \Delta_i).$$

This is the condition for backscatter along the incident ray (Fig. 8 (b)), the Bragg resonant scattering condition already discussed.

If the condition that the waveheight H is much less than λ_0 is removed, the angular spectrum of scattered waves will have additional side waves given by

$$\cos \Delta_s = \cos \Delta_i \pm n \lambda_0/L \tag{8}$$

(These higher-order side waves have a close analogy with the extra sidebands which appear in the spectrum of a phase modulated wave where the phase modulation is no longer small compared with one radian[22].) Barrick[9] has shown that for dekametric waves incident on the sea, this higher-order scattering is negligible except at the shortest radio wavelength and highest sea-states.

Inspection of equation (7) shows that the first and second terms are the horizontal components of two vectors lying along the propagation directions of the incident and diffracted waves. Rearranging the equation to associate the radio wavelength, λ_0, with the radio rays and multiplying through by 2π, we get

$$(2\pi/\lambda_0) \cos \Delta_s = (2\pi/\lambda_0) \cos \Delta_i \pm 2\pi/L \tag{9}$$

or

$$k_0 \cos \Delta_s = k_0 \cos \Delta_i \pm K \tag{10}$$

where k_0 is the radio wave-number and K is the water wave-number. The vectors being resolved are thus the wave-number vectors for the incident and scattered radio waves. The scattered vectors are found by resolving the incident wave-number vector (of magnitude k_0) along the surface, adding or subtracting the water wave-number vector to obtain a resultant and then finding the radio wave-number vector (also of magnitude k_0) which has this resultant as a horizontal component. Rewriting as a vector equation,

$$\mathbf{1}_x(\mathbf{k}_{os} \cdot \mathbf{1}_x) = \mathbf{1}_x(\mathbf{k}_{oi} \cdot \mathbf{1}_x) \pm \mathbf{K} \tag{11}$$

where $\mathbf{k}_{oi}$, $\mathbf{k}_{os}$ are vectors of magnitude k_0 in the directions of the incident and the two side waves respectively and $\mathbf{1}_x$ is the unit vector in the x direction. The resolved vectors along the sea plane in (11) are of sufficient importance to give new symbols, so we write,

$$\mathbf{k}_s = \mathbf{k}_i \pm \mathbf{K} \tag{12}$$

where $\mathbf{k}_s$, $\mathbf{k}_i$ now no longer have magnitude k_0. The backscatter condition is obtained by taking the negative sign and putting $\mathbf{K} = 2\mathbf{k}_i$.

Although the geometry has been described in terms of the special case for which the water-wave normal lies in the plane of incidence of the radio waves as shown in Fig. 8 (a) and (b), analysis shows that the construction implied by (12) is valid for water waves propagating at an angle to the plane of incidence of the radio waves. This situation is illustrated in Fig. 8 (c). An important special case is one in which the incident wave and one of the scattered waves lie in the plane of the sea. Since the scattered wave, to propagate along the surface, must have magnitude k_0 like the incident wave, the resultant vector triangle is isosceles with the implication that the scattered wave is 'specularly reflected' from the wave-fronts of the water wave, the angles of incidence and reflection on the wave-fronts being equal.

It should be noted that it is not always possible to find a direction of propagation for the scattered wave which yields a wave vector of magnitude k_0 projecting down on to $\mathbf{k}_s$. This condition of no freely propagating wave corresponds to $|\mathbf{k}_s| > k_0$, implying a wavelength shorter than free-space wavelength and hence an evanescent wave bound to the surface.

The relationship of the Doppler shift observed in the radar echo to the water wave parameters turns out to be very simple. We consider a generalized form of the Bragg condition for any angle of incidence, namely that the path length by way of one sea-wave crest is to be one radio wavelength longer than that by way of the previous crest. This implies that the radar echo will advance by 2π radians in the same time that the water wave advances by one wavelength, in other words that the scattered radio wave has frequency f_s, shifted from the incident radio wave, frequency f_i, by the water wave frequency f_w.

$$f_s = f_i \pm f_w \tag{13}$$

Equations (12) and (13) permit both the direction and Doppler frequency shift of the scattered waves to be computed. An extension of the theory to second-order electromagnetic scattering effects, that is to propagation involving successive scatter from two water waves having in general different wave directions and wavelengths, proves to be of considerable importance. The relationships (12) and (13) become, for second-order scattering,[9]

$$\mathbf{k}_s = \mathbf{k}_i \pm \mathbf{K}_1 \pm \mathbf{K}_2, \tag{14}$$

$$f_s = f_i \pm f_{w1} \pm f_{w2} \tag{15}$$

If we impose the additional constraint, used in a particular case above, that the scattered radio wave must constitute a freely-propagating wave of velocity c, then

$$f_s \lambda = 2\pi f_s/k_s = c \tag{16}$$

This results in scattering from two sea-waves travelling at right angles, analogous to the corner reflector in optics or microwave radar.

These equations permit the computation of the Doppler spectrum (echo power v. Doppler frequency) of sea-echo by both first-order and second-order mechanisms, given the sea wave-height variance spectrum in terms of wave number and direction. The corner-reflector effect results in a marked peak in the spectrum at $2^{\frac{1}{2}}$ times the Doppler shift for Bragg resonance. To compute fully the second-order contributions to the spectrum, however, the hydrodynamic properties of finite amplitude waves must

be considered. Two effects are significant. The first arises from the trochoidal shape of water waves of finite amplitude. Bragg resonant scattering occurs with the harmonics of the fundamental sinusoid, but since these harmonics travel at the speed of the fundamental and not at the speed corresponding to infinitesimal sinusoidal waves of the same length, they give an anomalous Doppler shift. This gives rise to peaks in the observed spectrum at $2^{\frac{1}{2}}$, $3^{\frac{1}{2}}$, $4^{\frac{1}{2}}$, etc., times the Bragg frequency.

The second effect arises when two finite amplitude waves of suitable wavelength and direction interact non-linearly to give an (evanescent) wave oriented in the direction of the radar. If the wavelength of this new wave is correct for Bragg resonance, a radar echo will be identified. This mechanism gives rise to a continuum in the Doppler spectrum.

The importance of understanding and quantifying these second-order effects lies in the opportunities they provide for measuring wave height, non-directional wave-height spectrum and directional wave-height spectrum by inversion processes. The features of the spectra which permit this will be indicated in the discussion of results of the UK experiments.

4 Ground-wave Experiments and Signal Processing Techniques

The early Birmingham experiments on ground-wave radar[18] utilized a 1·95 MHz Loran A transmitter at Angle, Pembrokeshire, (Fig. 9) and were aimed at demonstrating the technique for sea-state and surface-current surveillance over the Celtic and Irish Seas. Work now in progress aims at developing, at the same site, a transportable ground-wave radar which can then be used to carry out similar measurements wherever they are required.

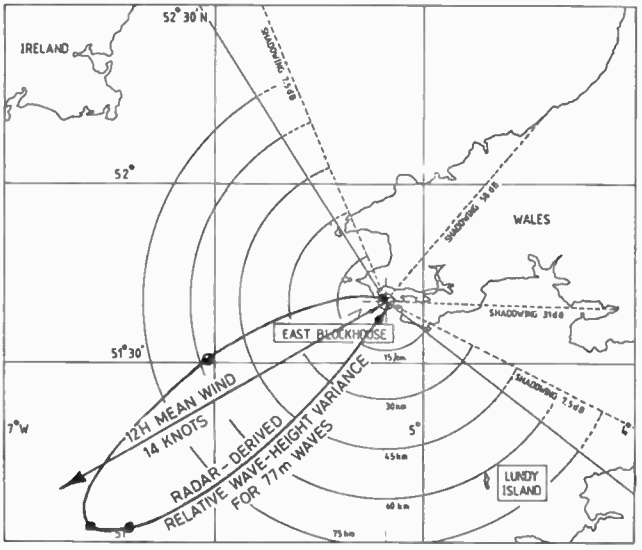


Fig. 9. Location of UK ground-wave sea-state radar at East Blockhouse, Angle, West Wales, showing azimuthal variation of wave-height variance of 77 m waves measured in the synthetic aperture experiment of 27.11.77. Range 30 km.

Figure 1 shows inset a sea-echo spectrum obtained in one of the early observations and highlights an important feature of such spectra. The photograph shows superimposed a number of successive spectral measurements, each made with coherent dwell of 15 seconds. The vertical scale is logarithmic. Three spectral lines can be identified, a central non-Doppler-shifted line from ground and the two Bragg resonant lines from approaching and receding sea-waves. In this case the former is 10 dB stronger, indicating the predominance of approaching wind-driven waves, or swell.

The decibel scale of relative echo power accentuates the noise floor at the skirts of the sea-echo spectrum (the signal/noise ratio being poor in this early experiment), and clearly this noise floor could be more accurately estimated by averaging successive spectra. However, the important feature to notice is the difference between the estimates of the spectral shape at the peaks of the spectrum, where the effect of noise was negligible. Individual spectral estimates may show a peak of low amplitude or high amplitude, possibly split, or with the nulls partially filled in. The incoherent average over several such power spectra, however, gives a completely unambiguous estimate of the spectral shape. With sufficient signal/noise ratio the second-order features on the skirts of the spectra are delineated without ambiguity.

The physical reason for this change of spectral shape in successive dwells is of some interest. If the spectrum being measured has individual pure monochromatic lines, successive spectra will show no variability, as can be verified by computer experiments. The variability arises with physical phenomena such as sea-echo, in which the monochromatic lines are replaced by narrow band noise-like processes. In sea-echo, the returns from different directions within the antenna beam are smeared in Doppler shift by surface currents, an extreme case of this appearing in Fig. 3. Since these spectral components can lie within the resolution of the spectral analysis (1/15 Hz in Fig. 1), the resultant estimate in a spectral resolution cell will vary in successive coherent dwells as the relative phase of the components changes and, therefore, their phasor resultant.

Figure 1 was obtained with a single range-gate hard-wired on-line discrete Fourier transformer using 8-bit digital processing. The echo waveform, after translation to 8 Hz, was also recorded on magnetic tape, digitized off-line for 64 range-gates to 16 bit accuracy on a minicomputer and the resulting digital tapes processed using seven 150 second coherent dwells averaged on a main-frame computer. Figure 10 shows spectra from such analysis for ranges 0 to 75 km at 15 km intervals, thresholded at −25 dB. The features evident include, as well as the approach and recede Bragg resonant lines from 77 m waves, non-Doppler shifted coast echoes at shorter range, a strong return from Lundy Island at

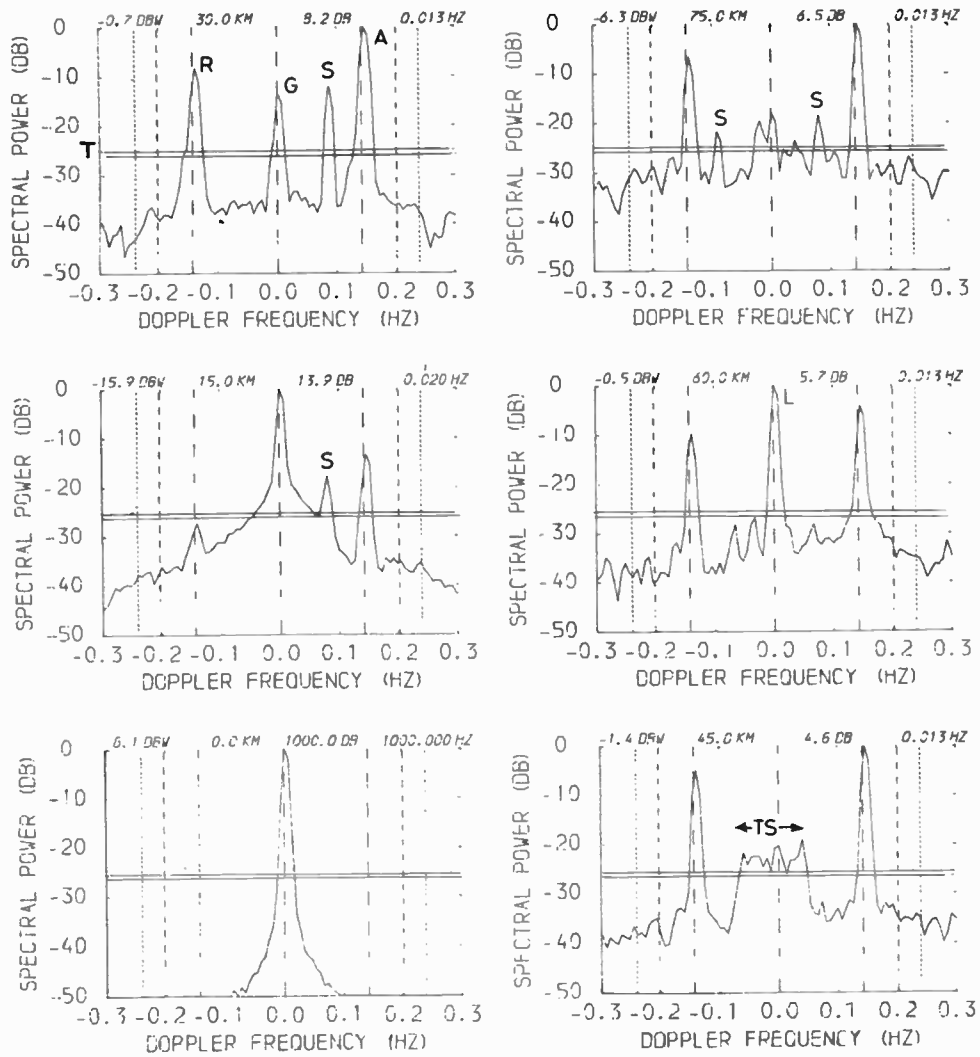


Fig. 10. Computer print-out of off-line processed spectra from East Blockhouse radar. Date 25.11.77, 1634 UT, 1·95 MHz, range resolution 7·5 km, antenna coverage as shown by ringed sector in Fig. 9. Average of seven 150-second spectra plotted for each 15 km increment, normalized to largest line with corresponding analyser input in dBW indicated. The two right-hand figures above each spectrum indicate the automatically extracted Bragg line ratio in dB and the width of the larger Bragg line in Hz. Returns above −25 dB threshold, T, are indicated as A approach and B recede Bragg resonant lines, S shipping, TS turning ship, G ground, L, Lundy Island.

60 km and discrete and spread spectra from shipping of velocity smaller than the waves, seen between the Bragg lines.

The above experiments were carried out with the existing omnidirectional Loran A transmitter and receiver at East Blockhouse, Angle. To obtain directional observations early in the programme, the synthetic aperture technique of Teague *et al.*[14] was employed, using a receiver with stable local oscillators carried in a vehicle along a straight road, the returns from different directions being identified by the resultant Doppler spreading of the Bragg lines. The velocity was kept low enough to prevent overlap of the spread Bragg lines and the spread ground echoes. The resultant polar diagram of echo power (proportional to wave height variance) from 77 m receding sea waves in the experiment of 1228 UT, 27.11.77 is shown superimposed on the map of Fig. 9. The aperture synthesized was ~ 1 km giving an azimuthal resolution of ~ 8°. The interesting feature is the narrow directional spread of the waves, which aligned well with the mean wind over the previous 12 hours.

5 Sky-wave Experiments

Observation of sea-echo by sky-wave greatly extends the area of coverage surveyed, as already explained, but introduces the variable factor of ionospheric propagation. To illuminate a particular area of ocean, the operating frequency must be changed with time of day and season to ensure that the frequency is not so high that the skip distance extends beyond the near edge of the zone to be surveyed and not so low that multipath propagation is possible or D layer absorption is excessive. It is thus not possible to choose frequencies of

operation on oceanographic considerations alone. There is thus considerable interest for sky-wave operation in the second-order spectrum inversion techniques of Section 3 by which complete oceanographic spectra can be determined from observations at a single radio frequency. Unfortunately, these are the features of the sea-echo spectrum which require the highest signal/noise ratio and the greatest phase-path stability to achieve freedom from contamination and smearing. Measurement of the relative height of the Bragg lines to deduce wind direction can, however, be carried out in the presence of quite severe contamination, while determination of wind speed has an intermediate requirement for spectral quality.

The Birmingham/Appleton Laboratory experimental programme on sky-wave radar is aimed at quantifying the accuracy with which the different oceanographic quantities can be measured in terms of the fraction of time for which different confidence levels can be achieved. Also since spectral quality also depends on antenna beamwidth, determination of the optimum cost-effective antenna aperture is sought. Antenna beamwidth enters for two reasons. The first is oceanographic, in that if the antenna beam is too wide, either the same wave pattern will be looked at with different aspects, or different wave patterns will be observed. The second reason is ionospheric, in that the presence of travelling large-scale wrinkles or Travelling Ionospheric Disturbances, TIDs, in the ionosphere will cause sea-echo at the edges of a broad beam to have different Doppler shifts, thus introducing Doppler smearing into the spectra. Typical scales for TIDs in the F layer are 100 km length, 3 km height and 100 m/s velocity. An example of the effect of a TID on forward and back-scatter propagation is seen in Fig. 11, obtained by Bagwell[23] using the same radar installation as used for sea-state observations, but at a frequency of 9 MHz with an antenna of 6° horizontal beamwidth subtending

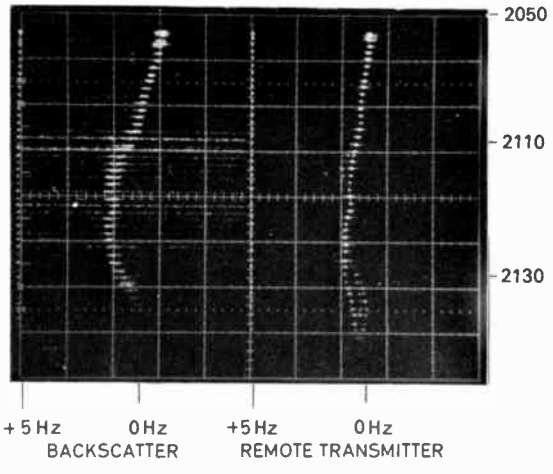


Fig. 11. Time sequence of Doppler spectra on 30.8.78, 2050 GMT, the backscatter at 9·25 MHz and the remote transmitter at 9·585 MHz. Each spectrum is analysed with a coherent dwell of 25·6 s.

70 km in the ionosphere directed towards Vienna. The two records show side-by-side the Doppler/time history of the received carrier of a broadcasting transmitter of very high intrinsic carrier stability, located in Vienna and the corresponding history of the received back-scatter from the land in the same locality during an unusually severe TID. It will be noticed that the back-scatter frequency undergoes twice the shift of the broadcasting transmission, corresponding to two-way propagation compared with one-way. In this case the shift over a 20 min interval is 2 Hz, compared with 0·6 Hz for the separation of the Bragg frequencies sea-scatter at the same frequency. This wander need not invalidate sea measurements since both Bragg frequencies would move together, but the consequent spectral spreading is more serious. The effect of this can be seen in Fig. 11, where the spread of the Doppler is appreciably wider at the time of rapid rate of change of Doppler, 2110, than at 2130, when the rate of change is small.

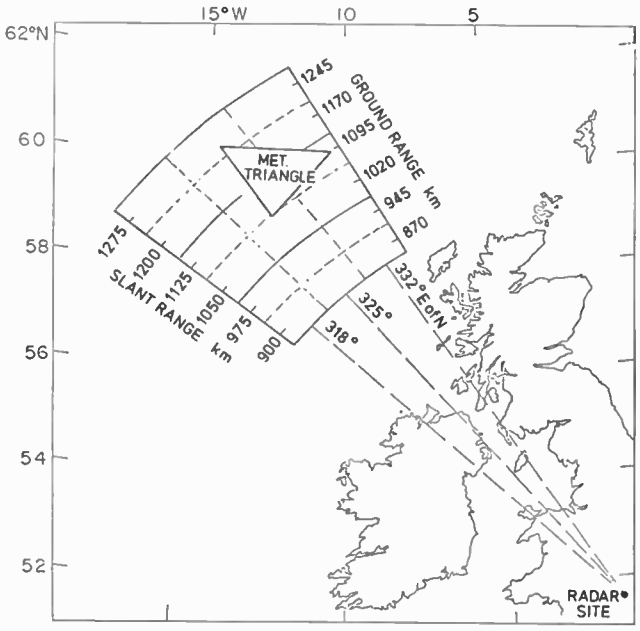


Fig. 12. Radar coverage during JASIN experiment.

When the UK joint programme was initiated, the first priority in an experimental programme was considered to be a requirement for 'sea-truth' of the best quality possible from the area of sea surveyed, as a cross-check on radar measurement and analysis techniques. An ideal opportunity for such comparison of radar and *in-situ* measurements presented itself in the JASIN project in the summer of 1978, mentioned in Section 1. A special rhombic transmitting antenna of beamwidth 21° illuminated the area around the 'meteorological triangle' (see Fig. 12) around which the ships, buoys and other measurements were concentrated. A steerable-beam receiving array provided three 7° beam positions within the transmit beam and this together with the range resolution of 75 km in slant path corresponding to the

0·5 ms pulse duration used, gave the matrix of 18 resolution cells of approximately 75 km × 125 km in cross-range, shown in Fig. 12. Propagation during the observations (all in daylight) was one-hop by way of the sporadic E layer at a height of about 100 km.

The spectra from this experiment were processed using coherent dwells of 50 s, ten such being averaged, this being found to give the optimum compromise between high spectral resolution and the reduction of the noise-like fluctuation of the spectral estimates already mentioned.

A first comparison between the radar-deduced and ship-measured wind directions and speeds has been described elsewhere.[24] Good agreement was found in radar wind direction determination using published relationships when the wind speed was about 10 m s^{-1}, poorer agreement when the wind speed was only 5 m s^{-1}, indicating that the relationships were then not valid. Wind speed measurements agreed with ship measurements at a time when consistent wind speeds were reported by the various ships. More complete local measurements are now becoming available and the comparisons are being pursued.

Meanwhile a comparison has been made between further radar data now available and surface isobaric maps.[25] A sample of the processed spectra (for 21.7.78 at 1012 UT at 325° E of N) is shown in Fig. 13. In this, some further reduction in the variance of the spectra has been achieved, by deriving 15 averages where 10 were achieved previously. This is done by overlapping the coherent dwells by 50%, taking advantage of the lower contribution made by the beginning and end of the dwell when it is subjected to Hanning weighting.

These spectra, though obtained by sky-wave, are of

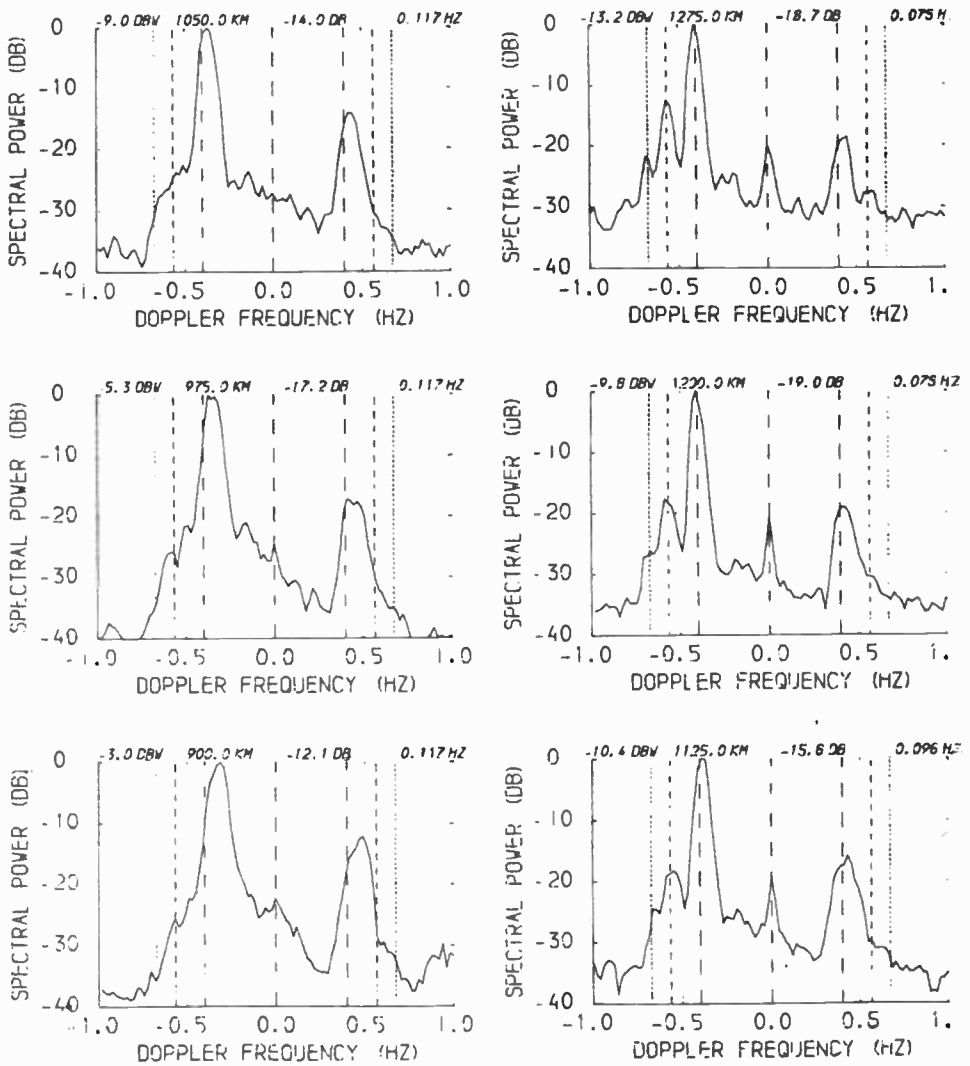


Fig. 13. Averages of 15 individual spectra for 6 different slant ranges. The dBW (for 0 dB), slant range, Bragg line ratio (dB) and width of higher Bragg line are indicated above each spectrum.
Date 21.7.78. 1012 GMT. Frequency 15·666 MHz. Bearing 325. Number of sample averages 15.

very good quality and show the second-order features at $2\frac{1}{2}$ and $2\frac{3}{4}$ times the Bragg frequency (see dotted lines) as discussed in Section 3. The trend of the continuum under the peaks is also of the nature found in good ground-wave data and of the quality capable of inversion, and non-directional spectra are being computed from them. The way in which the second-order peaks build up in amplitude at progressively greater ranges, suggests that the sea-state is significantly higher towards the North, although *in-situ* measurements in this area are not available.

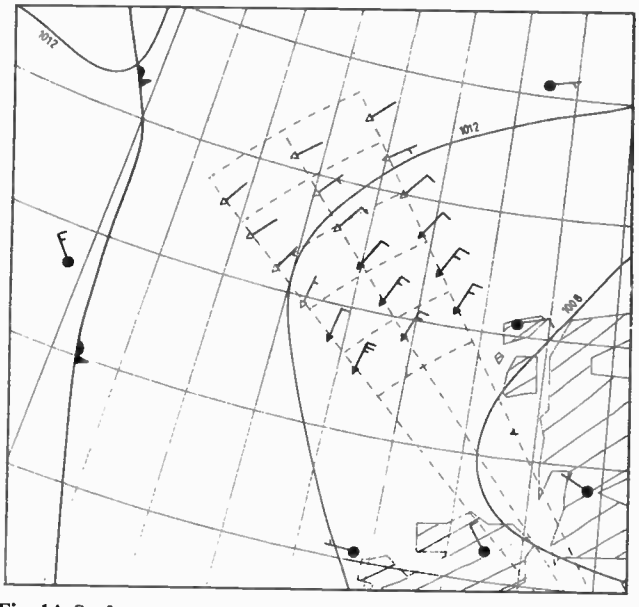


Fig. 14. Surface meteorological analysis for 1200 GMT on 3.8.78 with radar-deduced surface wind vectors superimposed. Radar-deduced vectors were obtained using the Long–Trizna technique.[6]

Figure 14 shows the comparisons with isobaric data for the area on 3.8.78 at 12 UT. The data here were processed by a specially-developed suite of computer programs, which computes spectra, extracts wind direction and wind-speed data automatically and plots the deduced wind vectors on a computer-drawn map. The only manual intervention was to resolve the left-right ambiguity in the vector direction on the basis of known meteorological trends. (If a wider sector of antenna beams had been available, this ambiguity would also have been resolved automatically on criteria of meteorological continuity.)

The manner in which the deduced wind vectors follow the isobars, as would be expected on Coriolis grounds, is very gratifying as a preliminary check on this result.

These comparisons will be carried further on the basis of a statistical study of confidence limits in the extraction of data from spectra developed by Bramley,[24] extending the work of Barrick and Snider.[26] Other work on the accuracy of deduced wind vectors from radar data by Maresca[27] is in progress at Stanford Research Institute, using a bistatic radar incorporating the Wide Aperture Research Facility (WARF) at Los Banos, California, a unique receiving antenna array of 2·5 km aperture yielding 0·5° beams. The particular emphasis in this work is the tracking of tropical storms, for which small-scale structures the high spatial resolution of WARF is invaluable. Dexter and Casey[28] have carried out similar studies over the Southern Ocean with the rotatable 24° beam-width radar at Townsville, Australia.

A study[29] of ionospheric effects on sky-wave sea-echo spectrum measurements is also in progress at the NOAA Wave Propagation Laboratory, Boulder, Colorado, with the aim of assessing the possibility of deconvolving ionospheric multipath effects from such measurements.

6 Future Developments

The use of h.f. radar techniques in oceanography appears already to have a secure place for certain applications. The use of ground-wave radar for surveillance of surface currents and tidal streams is an established technique with a number of potential applications and the development of a commercial transportable radar, using a very simple antenna system and on-line computer processing for this purpose, is in progress in the US at the Wave Propagation Laboratory of NOAA. The surveillance of wave-height and wave spectrum by ground-wave radar has been demonstrated, but the optimum realization of suitable systems is not clear. Studies are in progress in the US, in France[30] and in the UK. Whereas current measurement can be achieved at a single frequency (typically 26 MHz), wave measurement may require multi-frequency h.f. radars. Work at the University of Birmingham is at present directed to trials with an experimental system which will resolve these uncertainties. A synthetic aperture technique giving both angular and Doppler resolution is also under study. *Sea-sat* type satellites do not appear to compete in the Continental Shelf and estuarial waters surveyed by such ground-wave radars.

Sky-wave systems have been demonstrated to be of value in tracking frontal weather systems and hurricanes. Ionospheric propagation limitations will prevent their use on a 24 hours per day basis, but with the slow-moving character of weather and oceanographic features this may not be too great a drawback. A satellite samples only twice per day, so for hurricane tracking a sky-wave radar could give more continuous coverage.

A feature of sky-wave radars is the large antenna aperture needed. The present UK system uses a 300 m aperture and some reduction in beam-width is thought desirable. The Stanford Research Institute WARF has a 2·5 km aperture and some difficulty might be expected in funding many such systems. Part of the UK programme is to evaluate a technique suggested by Cooper,[31] whereby the beamwidth of a radar system is narrowed by the use of an effective transmission pattern involving coding of the radiation from spaced transmitting

antennas. The economic benefits of this could be crucial.

In conclusion, h.f. radar oceanography has already shown a considerable potential for remote sensing of currents, wave spectra and wind parameters and appears to be set to earn itself a place among the established tools of oceanographer and meteorologist.

7 Acknowledgments

The UK work described in this paper has been essentially a team enterprise and would not have been possible without the dedicated work of W. A. Sandham, D. J. Bagwell, S. Theodoridis, D. C. Cooper and G. D. Burrows of the University of Birmingham, P. A. Bradley, E. N. Bramley, F. D. G. Bennett, C. R. Carter and M A. Trower of the Appleton Laboratory and Sqdn Ldr J. Symons, B. W. Fursman and D. V. Tibble of the Ministry of Defence. Acknowledgment is also due to the Ministry of Defence and the Royal Signals and Radar Establishment for crucial support in the early stages of the work and to the Science Research Council for their continuing support.

The author also acknowledges his indebtedness to the many researchers in the UK, US, Australia, France and Germany, discussion with whom has formed the background of this paper.

8 References

1 Crombie, D. D., 'Doppler spectrum of sea echo at 13·56 Mc/s.', *Nature*, **175**, p. 682, 1955.

2 Ward, J. F., 'Power spectra from ocean measured remotely by ionospheric radio backscatter', *Nature*, **223**, p. 1325, 1969.

3 Tveten, L. H., 'Ionospherically propagated sea scatter', *Science*, **157**, p. 1302, 1967.

4 Wilkins, A. F. and Shearman, E. D. R., 'Back-scatter sounding: an aid to radio propagation studies', *J. Brit. I.R.E.*, **17**, p. 601, 1957.

5 Croft, T. A., 'Sky-wave backscatter: a means for observing our environment at great distances', *Rev. Geophys. and Space Phys.*, **10**, no. 1, p. 73, 1972.

6 Long, A. E. and Trizna, D. B., 'Mapping of North Atlantic winds by h.f. radar sea backscatter interpretation', *IEEE Trans. on. Antennas and Propagation*, **AP-21**, no. 5, p. 680, 1973.

7 Maresca, J. W. and Barnum, J. R., 'Measurement of oceanic wind speed from h.f. sea scatter by h.f. skywave radar', *IEEE Trans. on. Antennas and Propagation*, **AP-25**, no. 1, p. 132, 1977.

8 Hasselman, K., 'Determination of ocean wave spectra from Doppler radio return from the sea surface', *Nature, Phys. Sci.*, **229**, p. 16, 1971.

9 Barrick, D. E., 'Remote sensing of sea state by radar', chap. 12 in 'Remote Sensing of the Troposphere', V. E. Derr (Ed.), (US Govt. Printing Office, Washington, D.C., 1972).

10 Barrick, D. E., 'The ocean waveheight nondirectional spectrum from inversion of the h.f. sea-echo Doppler spectrum', *Remote Sensing of Environment*, **6**, no. 2, p. 201, 1977.

11 Barrick, D. E., 'Extraction of wave parameters from measured h.f. radar sea-echo Doppler spectra', *Radio Science*, **12**, no. 3, p. 415, 1977.

12 Lipa, B., 'Derivation of directional ocean-wave spectra by integral inversion of second-order radar echoes', *Radio Science*, **12**, no. 3, p. 425, 1977.

13 Peterson, A. M., Teague, C. C. and Tyler, G. L., 'Bistatic-radar observation of long-period, directional ocean-wave spectra with Loran A', *Science* **170**, p. 158, 1970.

14 Teague, C. C., Tyler, G. L., Joy, J. W. and Stewart, R. H., 'Synthetic aperture observations of directional height spectra for 7s ocean waves', *Nature, Phys. Sci.*, **244**, p. 98, 1973.

15 Crombie, D. D., 'Resonant backscatter from the sea and its application to physical oceanography', Proc. IEEE Ocean '72 Conference, IEEE Publn. 72 CHO 660-1 OCC, p. 173, 1972.

16 Barrick, D. E., Evans, M. W. and Weber, B. L., 'Ocean currents mapped by radar', *Science*, **198**, p. 138, 1977.

17 Shearman, E. D. R., Bagwell, D. J. and Sandham, W. A., 'Progress in remote sensing of sea-state and oceanic winds by h.f. radar', Proc. Conf. 'Radar '77', London, IEE Conf. Publn., No. 155, p. 41, 1977.

18 Sandham, W. A., Shearman, E. D. R. and Bagwell, D. J., 'Remote sensing of sea-state and surface winds in the Irish and Celtic Seas using h.f. radar', Proc. Conf. 'Antennas and Propagation', London, IEE, Conf. Publn., No. 169, Part 2, p. 125, 1978.

19 Tyler, G. L., Teague, C. C., Stewart, R. H., Peterson, A. M., Munk, W. H., and Joy, J. W., 'Wave directional spectra from synthetic aperture observations of radio scatter', *Deep Sea Research*, **21**, p. 989, 1974.

20 Rice, S. O., 'Reflection of electromagnetic waves from slightly rough surfaces', p. 351, in 'Theory of Electromagnetic Waves,' M. Kline, (ed.) (Interscience, New York, 1951).

21 Rayleigh, Baron, 'The Theory of Sound', 2nd edition, vol. 2, p. 89, (Macmillan, London, 1896; Dover, New York, 1945).

22 Ratcliffe, J. A., 'Some aspects of diffraction theory and their application to the ionosphere', *Reports on Progress in Physics*, **19**, p. 188, 1956. (The Physical Society, London).

23 Bagwell, D. J., 'Doppler processing of h.f. backscatter radar and mode structures of ionospheric propagation', Proc. IEE Conf. 'Antennas and Propagation', IEE Conference Publication No. 169, Part 2, p. 56, 1978.

24 Shearman, E. D. R., Sandham, W. A., Bramley, E. N. and Bradley, P. A., 'Ground-wave and sky-wave sea-state sensing experiments in the UK' Agard Conf. 'Special Topics in H.F. Propagation', Lisbon, AGARD Conference Proceedings No. 263, 1979.

25 Sandham, W. A. and Theodoridis, S., 'H.F. Radar Observations of Surface Winds in the Rockall area during the JASIN experiment of 1978', University of Birmingham, Department of Electronic and Electrical Engineering Memorandum No. 476, 1980.

26 Barrick, D. E. and Snider, J. B., 'The statistics of h.f. sea-echo Doppler spectra', *Trans. IEEE on Antennas and Propagation*, **AP-25**, no. 1, p. 19, 1977.

27 Maresca, J. W., Jr., 'Measuring rms wave height and the scalar ocean wave spectrum with h.f. skywave radar', *J. Geophys. Res.*, (To be published).

28 Dexter, P. E. and Casey, R., 'Ocean windfield mapping at long ranges with an h.f. radar', *Aust. Met. Mag.*, **26**, no. 2, p. 33, 1978. 1978.

29 Georges, T., 'Progress towards a successful skywave seastate radar', *Radio Science*, (To be published).

30 Broche, P., 'Estimation du spectre directionnel des vagues par radar decametrique coherent', Conf. 'Special Topics in H.F. Propagation', Lisbon, AGARD Conference Proc. no. 263, p. 31, 1979.

31 Cooper, D. C., 'The use of an effective transmission pattern to improve the angular resolution of within-pulse sector-scanning radar or sonar', *The Radio and Electronic Engineer*, **28**, no. 3, pp. 153 60, September 1964.

32 Kinsman, B., 'Wind Waves, their Generation and Propagation on the Ocean Surface' (Prentice-Hall, Englewood Cliffs N.J., 1965).

Manuscript received by the Institution on 8th July 1980
(Paper No. 1965/AMMS 105)

UDC 621.377.623.222: 621.391.83: 539.37
Indexing Terms: Magnetic tape, Mechanical properties, Elasticity

The influence of the elasticity of magnetic tape on some parameters of magnetic recording

WOLFGANG FELL. Dr.-Ing.*

Based on a paper presented at the IERE Conference on Video and Data Recording held at Southampton in July 1979

SUMMARY

A tape driven by an accurately running capstan between precisely aligned guides cannot be scanned exactly. Due to the elastic properties of the tape there are elastic displacements superposed on the nominal tape movement. These additional tape movements which are generated by tension disturbances or natural frequencies may affect the recording and the playback. This effect is considered and estimated

* *Robert Bosch GmbH, Television Equipment Division, D-6100 Darmstadt, West Germany.*

1 Introduction

The mechanical quality of a magnetic recording and play back is defined by the precision with which the play back head runs over the track which the recording head has written beforehand (Fig. 1). Deviations from the nominal tape speed in the direction of the track result in timebase and velocity errors, while deviations perpendicular to the track result in tracking errors.

Deviations like these can be produced by the capstan or the scanner or by the synchronization of both. These deviations are not considered here and in the following, capstan and head wheel are assumed to be running ideally and correctly synchronized. We shall consider deviations of the tape caused by the elastic property of the magnetic tape. Elastic material is deformed by forces, these produce tensions, which in turn produce local displacements.

2 Effects of Forces Acting on the Tape

First of all there is a force providing the normal tape tension which is required for smooth contact between tape, heads and guides and for obtaining a uniform build-up of tape on the reel. Figure 2 shows the parameters involved. For the purposes of calculation one can neglect the magnetic coating and consider only the thickness of the base film.

In video tape recorders particularly there will be a torsion effect between two guides (Fig. 3), resulting in non-uniform tape tension.

All recorders use tape guide pins or rollers with flanges for laterally confining the tape edges (Fig. 4). In connection with the tape movement, friction-forces are working on the tape edge producing unsymmetrical tape tension. Additionally, tension and pressure is produced in the tape surfaces by bending around the pin or roller. This tension, inversely proportional to the bending radius, will reach high values. A bending radius of $r = 0{\cdot}8$ mm produces in a mylar film of 25 μm a remanent plastic deformation.

Unsymmetrical tape tensions are also produced by geometrical errors in the tape path (Fig. 5). Even if the tape path is correct, but the tape edge has a curvature, bending tension over the high edge is produced in forcing back the tape into the correct path.

Additional tensions (Fig. 6) are produced during starting and stopping the tape, because the inertial mass

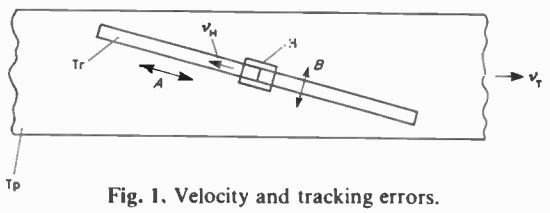


Fig. 1. Velocity and tracking errors.

Tp	tape	*A*	velocity errors
Tr	track	*B*	tracking errors
H	head	v_H	head speed
		v_T	tape speed

The Radio and Electronic Engineer, Vol. 50, No. 11/12, pp. 624–630, November/December

0033–7722/80/110624+07 $1.50/0

of all rollers driven by the tape must also be accelerated and decelerated. In extreme cases loops are generated for short periods, to be suddenly tightened at a later point, causing in this way a tensile shock in the tape.

All the tensions mentioned so far have been calculated or estimated for a certain recorder and the results are given in Table 1.

Even supposing worst case conditions, the resulting tension remains below the yield strength for the base film. One can see, there is no danger of permanent deformation of the tape, at least not in this recorder.

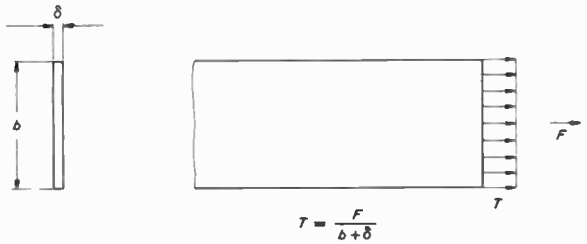


Fig. 2. Normal tension in the tape.

T = tape tension b = tape width
F = force δ = tape thickness (basefilm)

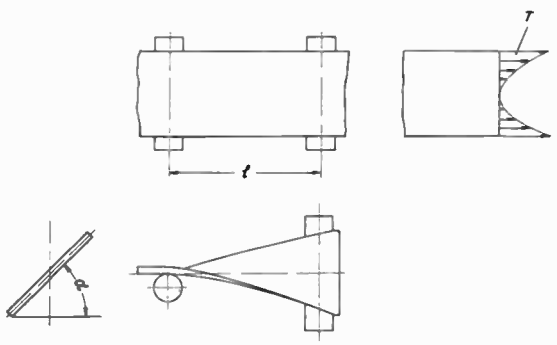


Fig. 3. Tension caused by torsion between tape guides.

$$T = \frac{\Delta l}{l} E, \qquad E = \text{modulus of elasticity},$$

$$\Delta l = \sqrt{l^2 + \left(\frac{b}{2}\tan\alpha\right)^2} - l$$

3 Elastic Deformations of the Tape Caused by the Resulting Tension

In the first approach one can resolve the resulting tape tension into a uniform average tension with a superposed moment (Fig. 7). The uniform tape tension T_m produces a displacement Δl and, if the uniform tape tension depends on time, it produces a displacement velocity.

A bending moment between two guides produces a lateral displacement, and a variation of the moment with time produces a lateral displacement velocity. The tape is very sensitive to bending moments especially if the tape path between the guides is long.

3.1 Effect of Unsymmetrical Tape Tension on the Tracks

The effect on the track of an unsymmetrical tape tension depends on the tracking angle (Fig. 8). The uniform tape

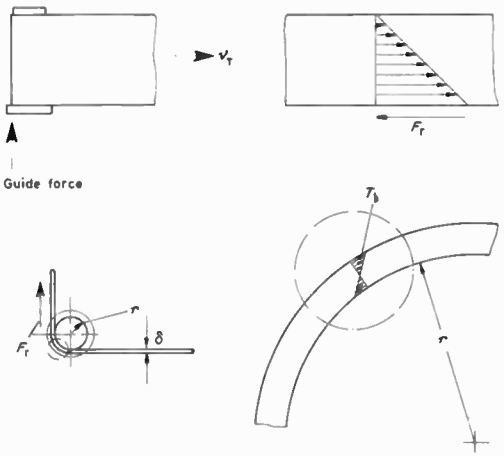


Fig. 4. Tape tension caused by bending, guiding and friction.

Pressure on tape edge:

$$\max T_p = \frac{2G}{\delta r \pi}$$

Tape tension resulting from friction:

$$\max T_f = \frac{F\mu}{\delta b/3}$$

Tape tension from bending:

$$\max T_b = \frac{E\delta}{2r}$$

tension produces displacements in the direction of the running tape. If the tension varies in a short time—faster than the head needs for one track—it can be possible, that, for instance, the recording head overlaps an adjacent track. This danger is serious for track 1, less so for track 2 and non-existent for track 3.

The bending moment produces a lateral displacement relative to the head path. This again results in tracking errors but now track 1 is not affected, track 2 only to a certain amount, while track 3 is very sensitive. If the moment varies rapidly with time, the danger of the head

Table 1

Tape tensions

Tape tension (N/mm²)	Static	Dynamic
Normal tape tension	4·14	
Torsion	1·75	
Guide force	2·39	
Friction by guide force	1·24	
Wrapping (r = 7 mm)	6·19	
Geometrical displacement	5·43	
Loops		34
Start, stop		8·3
Static yield strength	52	
Σ tension	21·14	
Dynamic yield strength		124
Σ tension		42·3

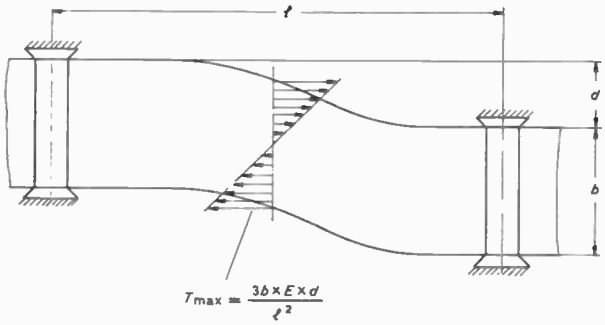


Fig. 5. Bending tension caused by geometrical errors (or curved tape edge).

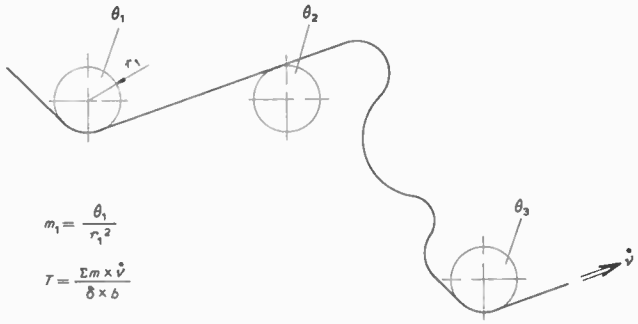


Fig. 6. Tensions caused by acceleration, deceleration and loops.

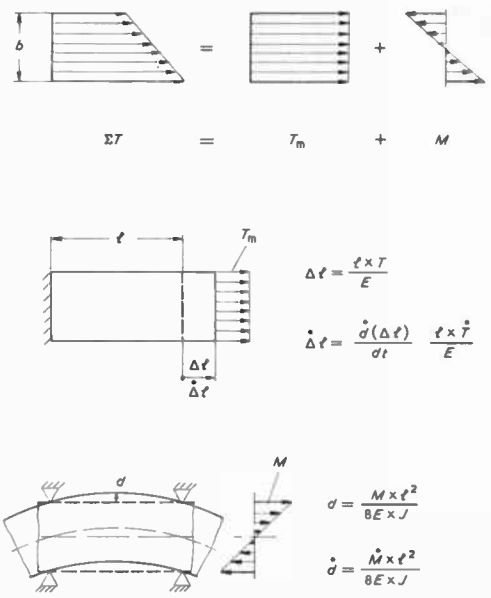


Fig. 7. Tape displacement caused by unsymmetrical tension.

overlapping an adjacent track is most serious for track 3, especially if the guide-to-guide tape path is long, non-existent for track 1.

Of course displacements produced by uniform tape tension and moment occur simultaneously and have been separated purely to give a better understanding of what is involved. So track 2 has tracking and velocity errors in each case, but is not subject to the *extreme* tracking and velocity errors, as tracks 1 and 3 are. One can try now to find an optimal tracking angle, if one assumes permissible tape deviations for the longitudinal and lateral displacements. Longitudinal displacements present no serious problems for most of today's video tape recorders with inclined tracks, because those errors can easily be corrected. In contrast to this, tape recorders are very sensitive to the lateral displacements of tape and track, because the track width is very small (40 to 200 μm). Assuming a ratio of 10 : 3, for permissible longitudinal to lateral displacement, results in an optimal tracking angle of 16° (very close to the B format).

4 Velocity Errors

External influences generate velocity errors: for instance, a single jerk when starting and stopping the tape, sudden friction variations (finger prints on the tape), or imprecisely rotating rollers. This results in tension pulses of different pulse width and slope or, if there happens to be a resonance, it results in an oscillation. In this case, it would be useful to predict the natural frequencies.

4.1 Natural Frequencies of a Discrete Spring-mass-system

For this purpose one can reduce the actual recorder to a spring-mass-system as shown in Fig. 9. The tape between two rollers is active as a spring, the rollers as the masses. To obtain a simple spring-mass-system the inertia of each rotating part is reduced to an equivalent mass with translational movement Fig. 10).

The natural frequencies and corresponding oscillatory modes of this system can be found by solving a set of 10 linear equations for its zero values. The results for a similar recorder are shown in Fig. 11. The natural frequencies have values between 20 and 900 Hz. Each natural frequency belongs to a certain oscillatory mode.

Knowing the oscillating amplitude of a corresponding frequency, the maximum velocity can be calculated.

Example:

Natural frequency	350 Hz
Amplitude	5 μm (more than worst case)
Tracking angle	14°

Maximal tape velocity $V_{Tmax} = 0{\cdot}005 \times 2\pi \times 350 \times \cos 14°$
$= \pm 10{\cdot}67$ mm/s

Assuming this oscillating velocity is directly affecting the video head to tape velocity of 24 m/s, one obtains a modulation of the signal at a frequency of 350 Hz, depth of modulation being 0·04% ($= 10{\cdot}67 \times 100/24 \times 10^3$).

That seems to be fairly uncritical, so that one can state: Frequencies produced by the spring-mass system cannot seriously affect the video signal, the frequencies being too low.

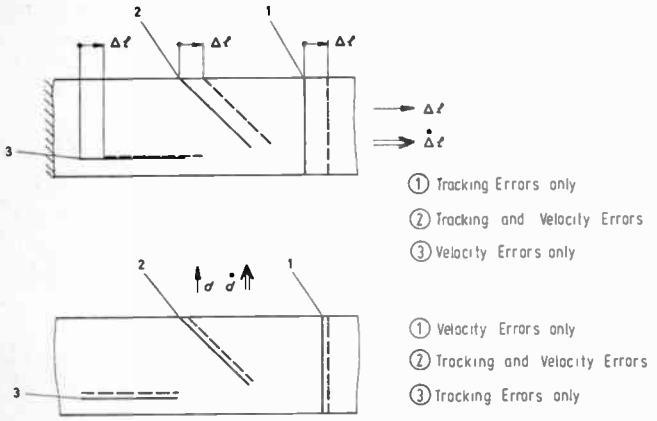


Fig. 8. Influence of an unsymmetrical tape tension on tracking and velocity errors.

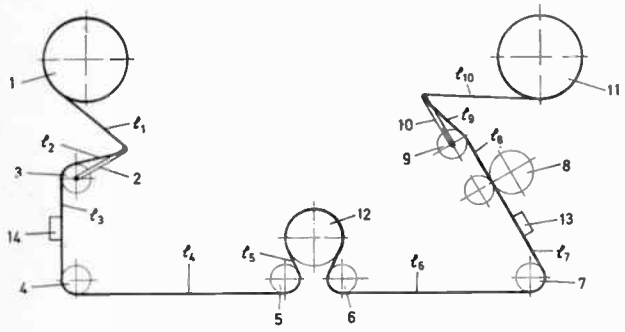


Fig. 9. Configuration of a tape deck.

1 Tape reel
2 Tension lever
3 Roller
4 Roller
5 Roller
6 Roller
7 Roller
8 Capstan and pinch roller
9 Roller
10 Tension lever
11 Take-up reel
12 Drum
13, 14 Audio and erase heads
$l_1 \ldots l_{10}$ Tape length between the rollers

4.2 Natural Frequencies of the Tape as a Continuum

At higher frequencies the rollers are not able to follow the tape displacements. They are then active as immovable masses and the tape can be considered to be oscillating as a taut strap clamped at both ends. In this case, for a video tape recorder, it is of primary interest to consider only the region of the tape being scanned, that is, the tape between the two guide rollers close to the scanner drum.

Figure 12 shows that tape, passing over the scanner drums, and beneath it the unrolled tape length l from one guide roller to the other. The natural frequencies of assumed longitudinal and bending oscillation can easily be calculated. They depend only on the modulus of elasticity, the density, width and thickness of the tape and on the guide separation (l). The example shows frequency values for an actual scanning system.

Is it possible that the high frequency of 6·5 kHz in the longitudinal direction can affect the video signal? To check this, the following example is given:

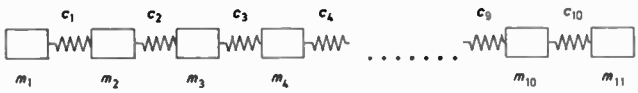


Fig. 10. Reduced spring-mass system.

$$c = \frac{EA}{l}$$

$$m = \frac{\theta}{r^2}$$

System of linear equations

$$M\ddot{x} + Cx = 0$$
$$(C - \omega^2 M)x = 0$$

E = modulus of elasticity of the tape
A = cross-sectional area of the tape
θ = momentum of inertia
r = radius of running tape
m = reduced mass

$\ddot{x} = -\omega^2 \ddot{x}$ substitutions

$$\frac{m\omega^2}{c} = \lambda$$

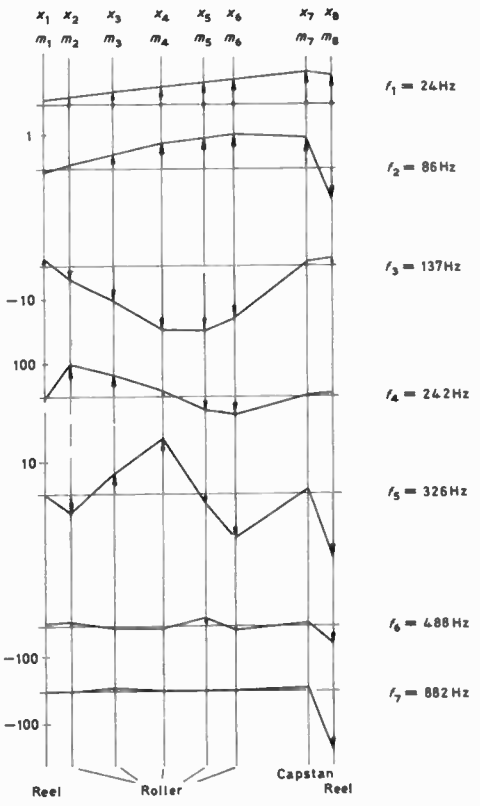


Fig. 11. Oscillating modes of a spring-mass system.

An oscillation, having an amplitude of 0·5 µm and a frequency of 6·5 kHz, for a nominal head-to-tape velocity of $V_H = 24$ m/s yields a deviation of velocity of 0·08% ($p = 0{\cdot}5 \times 10^{-6} \times 2\pi \times 6{\cdot}5 \times 10^3 \times \cos 14° \times 100/24 = 0{\cdot}083\%$).

This is a negligible velocity error and cannot seriously affect the video signal.

5 Effect of Shock Waves

Figure 13 shows a video head entering a magnetic tape and generating transversal deformation. This deformation runs across the tape with the specific

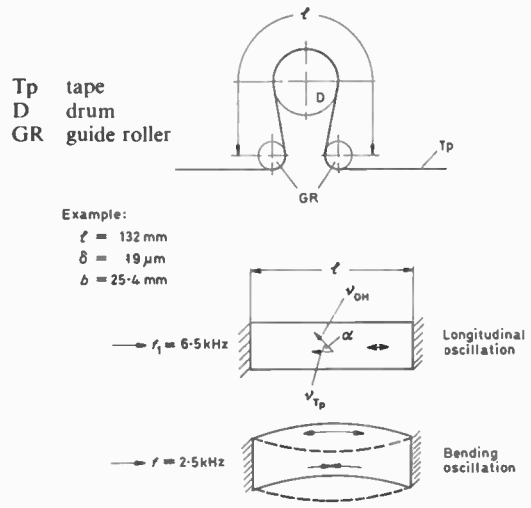


Fig. 12. Oscillations of a tape in the scanner area.

l = tape length
α = tracking angle
v_{Tp} = velocity of the oscillating tape
v_{DH} = component in the direction of the track

velocity c_w. It depends on the tape tension and the elastic properties of the tape.

Taking, for instance, a force for tape tension of 2 N and 1 in video tape—as used with the B format—c_w becomes 48 m/s, nearly double the value of the video head velocity. This means that the deformation an ingoing head generates can overtake the head in front, if the distance propagated is long enough. But the head leaving the tape also generates a wave of deformation running with specific velocity c_w against the direction of the video heads, causing a disturbance in head to tape contact. The magnitude of it depends on tape tension, and tape pressure on the head tip. In extreme cases, head-to-tape contact can be momentarily lost. This can severely affect the video signal.

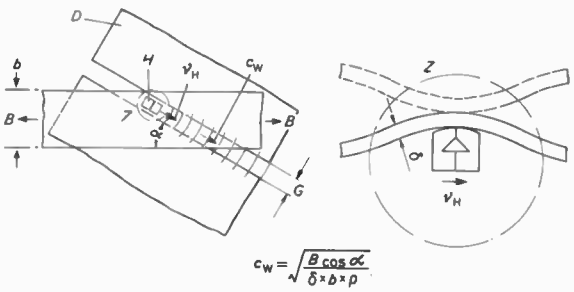


Fig. 13. Transversal oscillation of the tape in the scanner area.

D drum
G gap between drums
α track angle
B force for tape tension
δ thickness
b tape width
ρ tape density

Just as it generates transversal deformations, the head entering the tape also generates longitudinal deformations. For longitudinal propagation the specific velocity is the speed of sound in the tape, depending only on the modulus of elasticity and density of the tape (Fig. 14). Video tapes using a mylar base have a speed of sound of c_w = 1·6 km/s. This means that the deformation wavefront runs 67 times faster than video heads normally do (usual is 24 m/s).

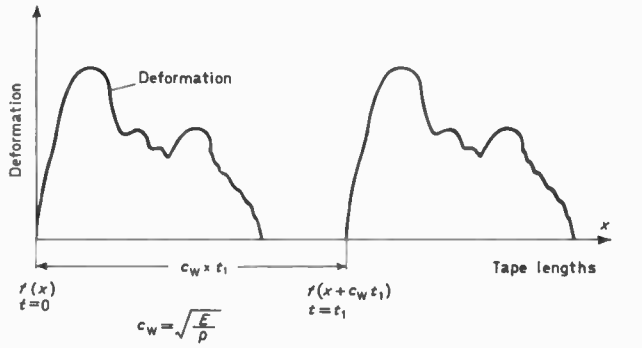


Fig. 14. Shock wave, running along the track.

5.1 Slope of the Wavefront

This slope can not be calculated exactly, but can be estimated by considering the geometry of scanner drum and head projection (Fig. 15).

Example:

Radius of drum	r = 25·165 mm
Head projection	p = 0·030 mm
Head velocity	v = 24·4 m/s

Rise-time to maximum deformation,

$$t = r \text{ arc cos } [r/(r+p)]/v$$

$$t = 51 \text{ µs}$$

After this time the wavefront will have run forward 82 mm ($x = c_w \times t = 51 \times 10^{-6} \times 1{\cdot}6 \times 10^6 = 82$ mm).

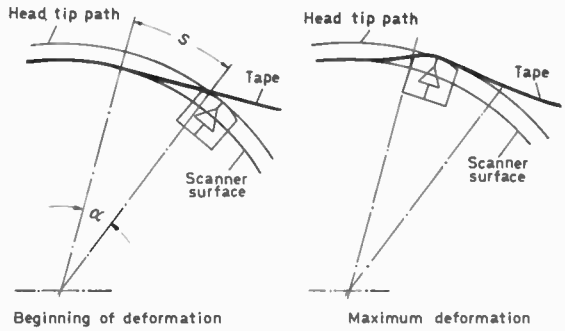


Fig. 15. Tape deformation caused by protrusion of the head.

The magnitude of the deformation can be estimated from the protruding head which may be considered to pitch a tent consisting of the tape upon the scanner drums. The estimation—aided by calculation and measurement—yields 0·3 µm. Figure 16 shows the deformation in detail. Curve 1 shows the calculated slope and line 2 shows a probable deformation if the entering head is tending to stick or catch on the tape.

A similar deformation step is generated by the head leaving the tape. In this case as well, a shock wave runs over the tape at the speed of sound (Fig. 17).

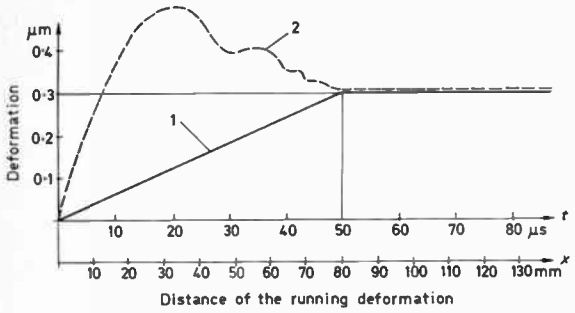


Fig. 16. Shape of deformation.

1 Head enters the tape, causing linear deformation
2 Head enters the tape with a short sticking or hooking

5.2 Effect of Shock Waves on the Video Signal

On running over a recording or reading head, the shock wave produces a deformation of the track of 0·3 µm for a time of 51 µs. This results in a disturbance of at least one line, probably two lines, because the disturbance period may not fall exactly within one line (64 µs).

The impact of this disturbance depends on the colour carrier frequency, which is recorded on the tape with a wavelength of $\lambda = 5{\cdot}42\,\mu m$ ($f_c = 4{\cdot}43$ MHz PAL, $v_H = 24$ m/s). The phase of this wave will play a large part in defining the colour reproduction in the video picture. A disturbance of 0·3 µm in this case produces a phase error of $360 \times 0{\cdot}3/5{\cdot}42 = 20°$. Permissible phase errors are in the order of 2°. Thus the shock waves will produce a severe colour error.

In reality, the phase errors are not so drastic, because in the video picture one can only observe the difference of deformation during recording and subsequent playback. The difference is very low or disappears completely if recording and playback immediately following this is performed on the same recorder. However, variations in head projection, tape tension or frictional damping will alter the shape of the deformation, so that phase errors can build up. It is a problem of v.t.r. compatibility and the number of possible tape copy generations.

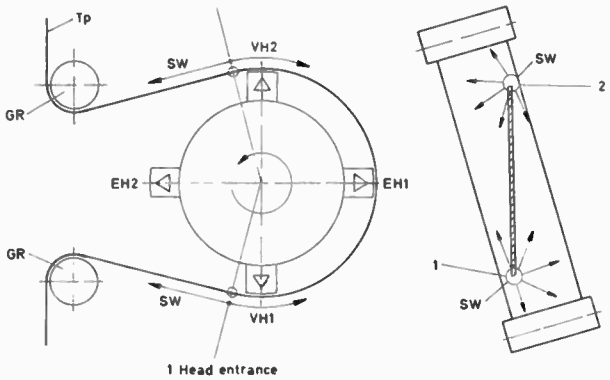


Fig. 17. Head wheel with tape.

VH video head GR guide roller Tp tape
EH erase head SW shock wave

Figure 17 shows a head wheel with four heads (two video and two erase heads) with a tape-wrap of more than 180° to provide overlapping. The actual length of the track corresponds to a wrap of 180°. During one revolution the four heads produce a total of four entrance shocks and four exit shocks, running to the guide rollers where they are reflected and returned.

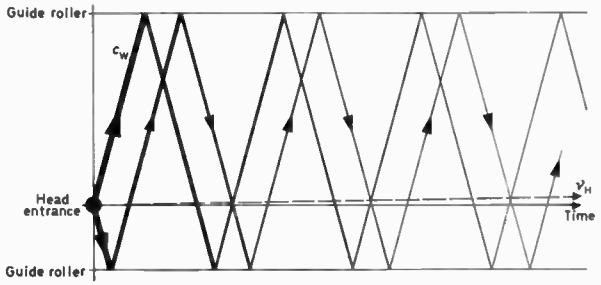


Fig. 18. Reflection of shock waves between the guide rollers.

Figure 18 shows a single shock wave, running to and fro between the guide rollers, in a displacement–time diagram. The nearly horizontal line represents the velocity of the video head, which seems to stand still. After each reflection the shock wave is damped until finally it disappears. If there is good damping in the tape and in the recorder, one can assume that the shock wave will only survive one reflection.

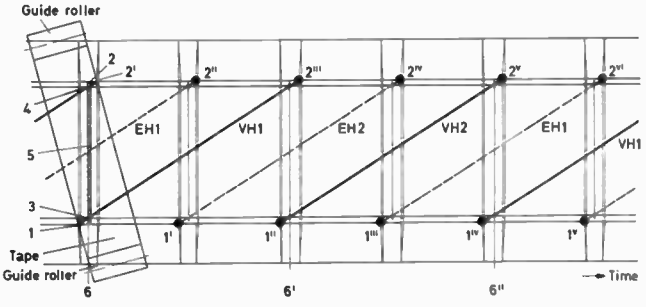


Fig. 19. Longitudinal shock waves running along the tape.

1 Head entrance
2 Head exit
3 Beginning of video track
4 End of video track
5 Video track
6 Switch point

Figure 19 shows all the propagated shock waves and the video heads in a displacement–time diagram. Every time a video head leaves the actual track length (180°) the playback signal is switched to the other video head which is just starting to read its track. Due to overlapping of the effective track length (190°), one video head enters the tape before the other has finished reading, and leaves the tape only after the machine has switched to the other video head. Between the video heads (90° offset) run the erase heads.

The steep lines are the shock waves produced at the points (1) and (2). Only one reflection is assumed to provide a clear presentation. Looking at the line of the

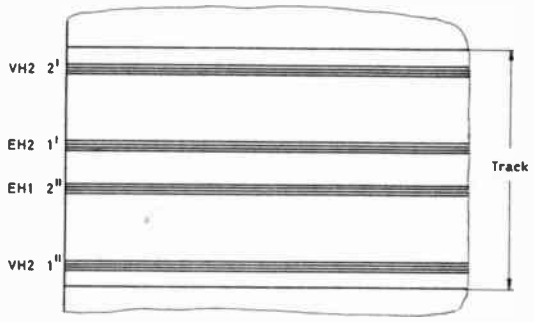


Fig. 20. Colour errors in video signal of one track.
VH video head EH erase head

video head VH 2, one can see the first disturbance a short time after head switching has occurred. It will have been produced by the video head VH 1 just leaving the tape. In the middle of the track there appear two new disturbances, the first one produced by the erase head EH 1 entering the tape, the second one produced by erase head EH 2 leaving the tape. A short time before end of track, another disturbance appears, this time from video head VH 1 just entering the tape. The recording or reading of a video head is disturbed in this way at least four times, corresponding to the number of heads and the number of shock wave reflections. This happens to all video tracks in the same way.

5.3 The Influence of Shock Waves on the Picture

Figure 20 shows the video signal of one track read by the video head VH 1. At exactly the same time the shock waves run over the track of video head VH 1, the picture will exhibit colour disturbance over some lines (naturally not in the black and white picture). Looking at the whole picture, one gets the impression of a pattern of stripes, slightly moving up or down.

6 Conclusions

To conclude, one can state:

1. An unsymmetrical tape tension can cause tracking errors, especially rapidly changing tape tension and moment, even if capstan and headwheel are running ideally.
2. The video signal is very sensitive to shock waves: transversal waves disturb the head-to-tape contact, longitudinal shock waves disturb the colour.

Sometimes it is hard work to find the origin of these errors, but if the parameters are known, one can take measures to reduce them.

7 References

1. Zwick, L., 'Untersuchung der Längsdehnung der Bandbreite mittels Bandzugmessung in beiden Hälften eines bewegten Magnetbandes', Dissertation, Technical University Braunschweig 1965.
2. Wittig, W., 'Untersuchungen am Laufwerk eines Magnetbandgerätes mittels Bandzugmessungen', Dissertation, Technical University Braunschweig 1963.
3. Steinhorst, W., 'Elastische Longitudinalschwingungen in Tonbändern speziell bei Stik-Slip-Anregung', Dissertation, Technical University Braunschweig 1965.
4. Fell, W., 'Das mechanische Laufverhalten von Magnetbändern auf Magnetband-Digital-Speichern (MDS)', Dissertation, Technical University Braunschweig 1970.
5. Zahn, H., 'Das BCN-System zur magnetischen Aufzeichnung von Fernsehprogrammen', *Bosch Technische Berichte*, **6**, no. 5/6, pp. 176–85, 1979, or 'The BCN system for magnetic recording of television programs', *Soc. Mot. Pict. Telev. Engrs J.*, **88**, no. 12, pp. 823–31, 1979.

Manuscript received by the Institution in final form on 17th June 1980.
(Paper No. 1966/Comm 209)