AD NUMBER
ADC950176
CLASSIFICATION CHANGES
TO: **unclassified**
FROM: **restricted**
LIMITATION CHANGES
TO: **Approved for public release; distribution is unlimited.**
FROM: **Distribution authorized to DoD only; Foreign Government Information; JUL 1975. Other requests shall be referred to British Embassy, 3100 Massachusetts Avenue, NW, Washington, DC 20008.**
AUTHORITY
DSTL ltr dtd 15 Feb 2007; DSTL ltr dtd 15 Feb 2007

JOURNAL OF NAVAL SCIENCE

Vol. 1 100,157 ★ JULY 1975 ★ No. 3

RESTRICTED

NOTICE TO READERS

The Editor extends to all readers of the Journal and particularly those of the younger generation, a cordial invitation to contribute articles of naval or general scientific and technical Interest.

Authors are requested particularly to note that as the Journal is in the Restricted category the printing of the material within its pages does not constitute open publication and does not prejudice the subsequent use of the material, following any necessary security clearance, in the journal of a learned society or institution, or in the scientific and technical press.

Manuscript should be typewritten if possible, in double spacing and on one side of the paper only. It should include, in addition to the title, the name of the Author together with initials, degrees and department, establishment or ship. Pseudonyms are acceptable only in special circumstances. A convenient length is between 4,000 and 7,000 words, plus essential illustrations, but other lengths are acceptable and no contribution would be debarred on this ground alone. Illustrations are in most cases a desirable addition. Photographs should be of good quality, glossy, unmounted and preferably between two and three times the size of the required final picture. Graphs and line drawings should be made on a similar large scale, with bold lines to permit reduction in block making. A recent photograph and biographical note of the Author(s) will also be welcomed.

Views and opinions expressed in the *Journal* are not necessarily endorsed either by the Ministry of Defence or the Editor.

JOURNAL OF NAVAL SCIENCE

Communications to:
The Editor, J.N.S.
Defence Research Information Centre
Station Square House
St. Mary Cray, Orpington,
Kent BR5 3RE
Tel: Orpington 32111. Ex. 345. Telex: 896866

RESTRICTED

Volume 1
Number 3
July 1975

CONTENTS

H. L. R. Hinkley — An Appreciation

On 31st May, 1975, Mr. H. L. R. Hinkley retired from Government service and so brought to an end a period of 22 years as Editor of the *Journal of the Royal Naval Scientific Service* (recently re-cast as the *Journal of Naval Science*). As Chairman of the Editorial Board of both of these *Journals*, I cannot let this occasion pass without paying tribute to Harry Hinkley's achievements throughout his long term of office; a tribute with which I am sure all readers and contributors to the *Journal* would wish to be associated.

Building, with drive and enthusiasm, on the solid foundations of his predecessors in the editorial chair, Dr. J. P. Lawrie (from 1945 to 1947) and Mr. B. H. Adkins (from 1947 to 1953), Harry Hinkley was instrumental in making significant improvements to the *Journal* thereby enhancing its reputation among Defence scientific communities in many countries. Under his editorship, the content was expanded not only by an increase in the number of major scientific papers published, but also by the introduction of such features as the Notes and News from the various Establishments, and Technical Notes. He also developed the format of the *Journal* to the elegant and professionally produced volume that it is today, and gave it a steadily widening distribution.

The original aim of the *JRNSS* was to provide corporate identity, and a channel of communication, between scientists, engineers and Service personnel serving within the RNSS in geographically scattered locations. It also became a vehicle for promulgating the results of the work of UK Naval scientists to those working in similar fields in many countries throughout the world. This wider dissemination has brought reciprocal benefits in the form of contributions from scientists engaged in work of common interest in other countries, and in the exchange of journals.

With the termination of the RNSS, the *Journal* under its new title, became the vehicle for common identity of scientific activity in the Naval environment. Harry Hinkley was almost solely responsible for effecting this transition in the least painful way, and as a result of dedicated activity has set the scene for the years to come.

The *Journal* is graded Restricted, but the inclusion of some of the articles within its pages does not preclude their subsequent publication in the open press, subject always to the usual clearance procedures. In this respect, Harry Hinkley has always been untiring, and often successful, in his efforts on behalf of authors in getting such papers published in the scientific and technical press. It must be remembered also that at no time were his duties as Editor full-time, his other responsibilities being manifold, culminating latterly in his being Head of the Naval Scientific and Technical Information Centre, and subsequently a Deputy Head of the Defence Research Information Centre.

Perhaps few of his acquaintances know that Harry Hinkley can claim no less than 45 years service within Government Service. This, together with his long years as Editor has given him renown in Naval Science circles, and many will wish to claim him as a friend.

The qualities which made him such a successful editor were his ideas, initiative, enthusiasm, loyalty and resource. We will indeed be fortunate if we can be so well served over the next 22 years. I am sure that everyone who has enjoyed being associated with the *Journal* in its past and present form will be eager to join me in wishing Harry, and his wife Ethel, a full and eventful retirement.

N. L. Parr
Chairman, Editorial Board
Journal of Naval Science

DESIGN AND PERFORMANCE OF A CERAMIC PISTON

Commander B. D. Gibson, B.Sc., M.Sc., R.N.
and
Lieutenant Commander R. B. Stone, B.Sc., R.N.
Royal Naval Engineering College, Manadon

Introduction

Over the past few years, consideration has been given to the incorporation of ceramic materials, particularly silicon nitride, into the design of diesel engine pistons with view to improving their thermal performance. In general the consensus has been that a one-piece silicon nitride piston of any size would be impracticable, and designs have tended towards a composite piston with a silicon nitride crown bonded and/or keyed to metal.

A project which formed part of a post-graduate marine engineering course at the Royal Naval Engineering College, Plymouth, undertook the design of a piston, rings and gudgeon pin in reaction bonded silicon nitride to fit an existing diesel engine. While the rings were rejected as unfeasible at the design stage, the piston and gudgeon pin went on to be manufactured at the Admiralty Materials Laboratory, Holton Heath. During subsequent testing the gudgeon pin fractured, but the piston still survives after many hours running and several rather traumatic experiences. A precis of the mechanical aspect of the design and testing is given in this article.

Equipment and Design Data

The engine used in the trial was a *Gardner* 1L2 single cylinder, four stroke, direct ignition reciprocating water cooled compression ignition engine, having a compression ratio of approximately 13 : 1. The design is pre-war, and is essentially massive.

Technical Data

Cylinder Bore	= 108 mm
Piston Stroke	= 152·4 mm
Con rod length	= 311·15 mm
Piston height	= 165 mm
Piston Mass	= 1·73 kg
Brake power output (max)	= 9 kW
Max speed	= 1100 rpm

Piston material — aluminium alloy
Gudgeon Pin material — High Tensile Steel
Ring material — Cast Iron (3 pressure + 1 scraper)

Power output is absorbed by a GEC hydrostatic dynamometer, which can also be used for starting. An identical experimental engine was used to produce an indicator diagram (combustion space pressure *versus* crank angle). A computer program was written to calculate piston and gudgeon pin loads throughout the four-stroke cycle.

Temperature measurement was done with commercially obtained *Templugs*. These are small threaded steel plugs which are tapped into a component and returned to the manufacturer (Shell Mex) after a run for analysis. Hardness relaxation allows the operating temperature to be accurately estimated. However, two major problems presented themselves when the component was made from RBSN.

(a) The fine threads (6BA) could not be tapped into the brittle partly-nitrided material.

(b) The drilled sockets would create disastrous stress-raisers if located in areas of high tensile stress.

An alternative fixing method was required. After a series of furnace trials with various cements and adhesives it was found that *Autostic* (a ceramic cement), suitably applied, would seal a *Templug* into a smooth drilled socket and also attach it to a flat, machined RBSN surface. This method was tested up to 700°C. Adequate tolerance had to be left round the *Templug* in the drilled socket to allow for differential expansion.

It was intended that the surface temperatures of the original aluminium piston would be measured at full power as a basis for the stress calculations and design of the RBSN piston. When fitted, the new piston would first be run with the engine decompressed, then under full compression and finally with fuel admitted. Surface temperature would again be taken at full power. The heat paths from the combustion space are so cross-coupled and the difference in thermal conductivity so great that it is impossible to forecast accurately the RBSN surface temperatures from the measurements in aluminium[1].

A two hour full power run with the aluminium piston fitted yielded the results shown below. *Templug* location numbers are shown in Fig. 1.

Location No.	*Temperature (°C)*
1	153
2	157
3	157
4	149
5	140
6	126
7	117
8	112

Thus the temperature difference across the crown can be taken as (157 - 126) = 31°C.

Stress Analysis and Design

The successful design of a hot, stressed component in an IC engine is invariably a compromise between the requirements of the thermodynamicist, the metallurgist, the stress analyst, the production shop etc. The relevant considerations for this study are:

(a) A satisfactory stress analysis.

(b) The elimination of all corners and sharp fillets to avoid stress concentrations.

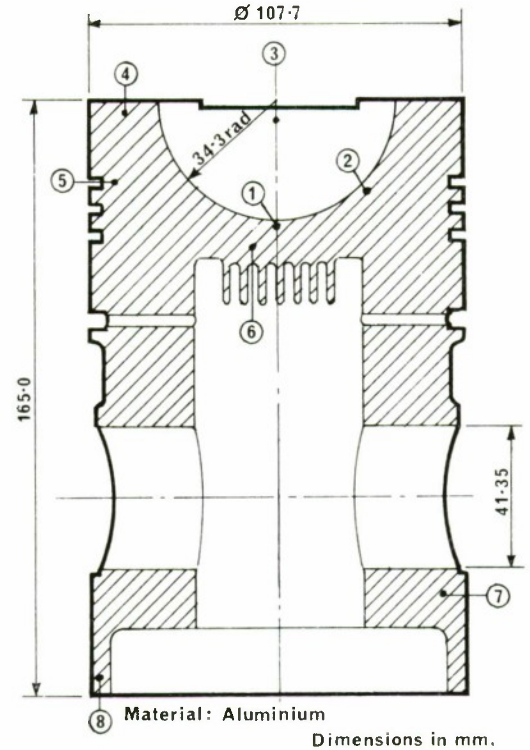


FIG. 1. Section through original piston.

(c) The choice of a suitable thickness of section so that effective nitridation can take place. The reaction is exothermic, and if the section is too thick the generated heat cannot escape, causing the silicon powder to melt. This renders the component useless.

(d) Differential expansion between mating or adjacent components in metal and RBSN. It is advisable to locate this interface in as cool a region as possible.

(e) The possibility of assembling the engine once the components have been manufactured. Springing or shrinking would prove difficult with RBSN.

(f) The frictional performance of the materials used. Little is known of highly loaded RBSN bearings.

Piston

The original piston is a single piece, cast aluminium component with a hemispherical dish in the crown surrounded by high, solid walls (Fig. 1). The design is very tall and massive by modern standards, and the four piston rings (three pressure and one scraper)

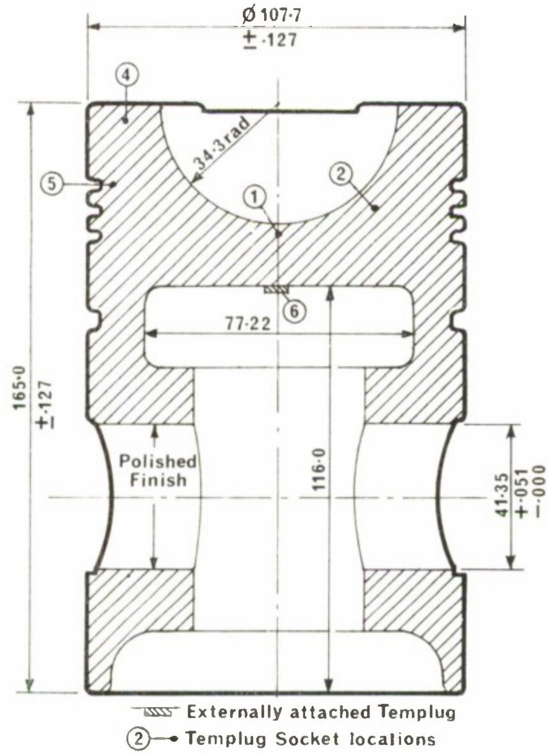


FIG. 2. Section through final piston design.

are all fitted above the gudgeon pin boss. The skirt is comparatively short. Cooling is by an oil spray from the connecting rod onto the underside of the piston crown, which is finned. Radial drainage holes are drilled from the oil scraper groove to the piston void.

Stress conditions in a running piston are of a very complex nature due to the complicated and discontinuous shape, and the rapidly changing environment. They can be categorised into:

(a) *Pressure Stress.* The stresses due to combustion pressure are repeated periodically, and at any instant are proportional to the instantaneous gas pressure. A calculation performed using the maximum combustion pressure will produce the maximum pressure stress at any point in the piston.

(b) *Thermal Stress.* This can be divided into two types, steady and fluctuating. The temperature in the combustion space is varying throughout the cycle, but this variation is so rapid, that the resultant temperature changes in the piston penetrate to a very small depth Eichelberg[2] has estimated that the surface fluctuation is of the order of 20°C, and that the maximum depth of penetration is about 3 mm. For stable materials behaving elastically the stresses raised by these fluctuations are insignificant. By far the largest component of the thermal stress is due to the steady temperature gradient between the exposed upper surface and the cooled underside of the piston crown. This gradient varies with position throughout the crown, and will change when the power output of the engine is changed. Transitions take a certain time to complete, during which period supplementary stresses are generated.

(c) *Mechanical Stress.* As a result of local contact between components. Piston side thrust, ring drag and Hertzian stress at the gudgeon pin boss are examples. These are usually small in magnitude and tend to be localised.

(d) *Inertial Stresses.* These are caused by the acceleration of the piston in the cylinder bore. These stresses are less than 5% of the pressure stress, and so can be ignored.

Several authors have produced analytical solutions of the stress distribution in a running piston, and all have ignored everything except the steady thermal stress and the fluctuating pressure stress[1, 3, 4].

Comparisons of these methods, and other minor variations[5] reveals that none of them gives results in accord with experiment, except in certain specific areas as the crown centre and the horizontal junction between the crown and the wall. Because it is the most comprehensive, and tends to give the most pessimistic results the basic Fitzgeorge and Pope[3] theory is used in this article. It is felt that this choice is justified in that stress levels are here required as a design guide, and not as an end in themselves.

A computer program was written to calculate the stresses at the underside of the crown centre and at the inside of the horizontal junction between crown and piston wall, where the tensile stresses are highest. The data required are the temperature conditions at the top and bottom surface of the piston crown, the maximum combustion pressure, the wall thickness and the crown thickness. The calculation is repeated for various conditions and dimensions until an optimum configuration is indicated. Because compression pressure fluctuates, the resultant maximum tensile stress at the two chosen locations do not occur at the same time. The crown centre reaches its maximum stress at TDC of the firing stroke while the horizontal joint maximum stress occurs at BDC of the induction stroke, the resultant values being the sum total and the temperature bending stress respectively.

Radial oilways at the horizontal junction were left out because of their effect on peak stresses in this highly loaded area. The effect of this omission on piston ring performance is to be ascertained. The bottom surface of the piston crown was not finned, as this would prove difficult to machine in RBSN. It is felt that the low coefficient of thermal expansion obviates the need for optimising the heat path through the crown.

The internal dimensions of the piston were initially chosen by:

(a) Inserting estimated surface temperatures into the program and selecting the thickness which tended to equalise and minimise the stresses at the crown centre and horizontal junction.

(b) Consideration of this selection from the point of view of the local shape ring grooves etc.) and the suitability of the section for nitridation.

Dimensions chosen were:

Wall thickness = 15·2 mm (0·6 in.)
Crown thickness = 15·2 mm (0·6 in.)

A section of the resultant design is shown in Fig. 2 (c.f. Fig. 1).

Piston Rings

Generally made of cast iron, these components must have sufficient elasticity to be installed without undue distortion, must retain this elasticity under severe working temperatures and pressures, and must be highly resistant to wear. It has been proved that neither the tensile strength nor the hardness of cast iron bears any relation to its piston ring performance. The microstructure is all important, *i.e.* a totally pearlitic structure with large flake graphite, total absence of free ferrite and a large grain size gives best wear results. Although there is no direct structural analogy with reaction bonded Si_3N_4, the material does have good wear properties, and it is considered that silicon nitride rings would perform well. (Tip seals in rotary engines are commonly made of Si_3N_4).

Using relationships listed by Harrington[6] and BS 3267/63[7], dimensions of a silicon nitride pressure ring were calculated to be:

Radial thickness = 2·18 mm (0·086 in.)
Axial width Ω ·6t = 1·32 mm (0·052 in.)
Free gap = 43·6 mm (1·72 in.)

These calculations assumed a working stress of 103 MN/m² (15000 lb/in.²) and that the ring gap must increase by eight times the radial thickness to pass over the piston crown[6]. These proportions are shown in Fig. 3.

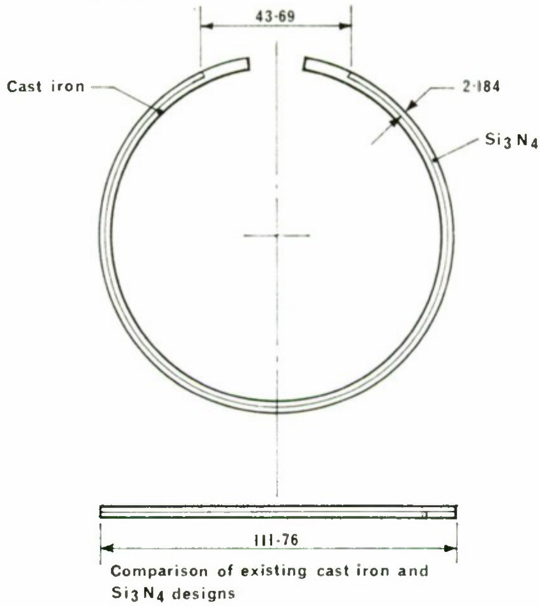


FIG. 3. Piston rings. Comparison of existing cast iron and silicon nitride designs.

It was considered that the production and use of such a piston ring is not feasible because:

(a) The fabrication and flawless machining of the non-nitrided ring would be extremely difficult.

(b) The small radial thickness would result in an unsatisfactory fit in the piston groove.

(c) The large free gap would make it difficult to compress the ring into the cylinder bore when the piston was offered up. The brittleness of the material would make this operation a fitter's nightmare.

It was therefore decided not to manufacture solid piston rings in Si_3N_4.

A sectioned ring with steel backing springs is a possible alternative, but this solution was not pursued. The original components, both of the pressure and oil scraper types, are retained.

Assuming maximum temperature rise of the top ring = 220°C.

Therefore longitudinal differential expansion in ring groove

= 2·413 (14-2) 10^{-6} (220 = 0·00635 mm (0·000251 in.).

This is small enough to be taken up in the sliding fit.

Gudgeon Pin

The original gudgeon pin is made of high tensile steel and is hollow, with a D/d of 1·8. It is fully floating in the small end bush and piston boss, and is not restrained by circlips. Lubrication of the small end is *via* a circumferential groove from a drilling in the connecting rod, while the boss relies on oil working its way from the piston void.

Gudgeon pins carry comparatively light loads, but are subjected to rapid changes in the magnitude and direction of that load, and to the hammering effect as the gudgeon pin moves in its clearance.

Ref. 8 gives an empirical relationship.

$$\frac{\sigma l D}{W} = (\beta + 1)\left[0{\cdot}559\beta\left(\frac{6\beta-1}{2\beta-1}\right) - 0{\cdot}5396\right]$$

where σ = maximum tensile principal stress
l = overall length of gudgeon pin
W = maximum load on gudgeon pin
D = outside diameter
β = (D + d)/2(D − d)
d = inside diameter

For the original pin:

D = 41·275 mm
d = 23 mm
β = 1·74
W_{max} = 49200 N
l = 106 mm
Hence σ_{max} = 96·5 MN/m² (14000 lb/in.²)

It is apparent that an Si_3N_4 component of these dimensions would survive if the maximum tensile stress were the only consideration. However, history shows that the usual mode of failure is by cracking parallel to the axis, which suggests a crushing rather than a bending action. Because of this, and a sceptical conservatism, it was decided that the gudgeon pin would be solid. The above empirical relationship then no longer applies, as when $\beta = \frac{1}{2}$ it gives $\sigma \rightarrow \infty$. So a more traditional bending stress and crushing deflection analysis was involved.

This gave a max bending stress for the silicon nitride piston of 84·8 MN/m² (12300 lb/in.²) (13550 lb/in.² for the steel pin) and a deflection of the pin as 0·008636 mm (about a factor of 10 less than for the steel pin).

This small crushing deflection is disadvantageous, in that the gap created ensures that the oil film between the pin and the piston boss is retained. As the normally accepted deflection is about 0·06 mm lubrication problems seem likely, but very little work has been done on heavily loaded Si_3N_4 bearings so a forecast of performance is impossible.

Operating the test engine with the modified gudgeon pin fitted was considered to be a somewhat hazardous business, and so it was to be installed after the modified piston had proved satisfactory. This meant that allowance had to be made in the piston boss for the differential expansion of the steel pin as the engine warmed up. 0·762 mm was allowed for this, equivalent to a temperature rise of about 130°C. The Si_3N_4 pin was to be machined to a better fit. A section of the design is shown in Fig. 4.

Failure Probability

A striking and fundamental distinction between the mechanical strength characteristics of ductile and brittle materials is the unusually high degree of variability exhibited by the latter. This is not a result of badly controlled specimen preparation or test procedure; it is an inherent characteristic of a material which cannot deform plastically under increasing stress. As a result of this uncertainty the average strength value may not be an adequate basis for specification or design. The standard Weibull probability function can be expanded

$$P_f = 1 - \exp\left\{-\left[\frac{(\sigma - \sigma_u)}{\sigma_o}\right]^m\right\}$$

to give an estimate of the failure probability of a component (or part of a component) subjected to a 3 dimensional stress system[9, 10]. The performance of this calculation requires a full-bodied stress analysis and considerable computational effort. As the stress calculations were here performed in a traditional two dimensional way, albeit by computer, the results were not suitable for this type of manipulation. Consequently, in the absence of such a failure probability criterion, the usual design technique of minimising and equalising the maxima was applied.

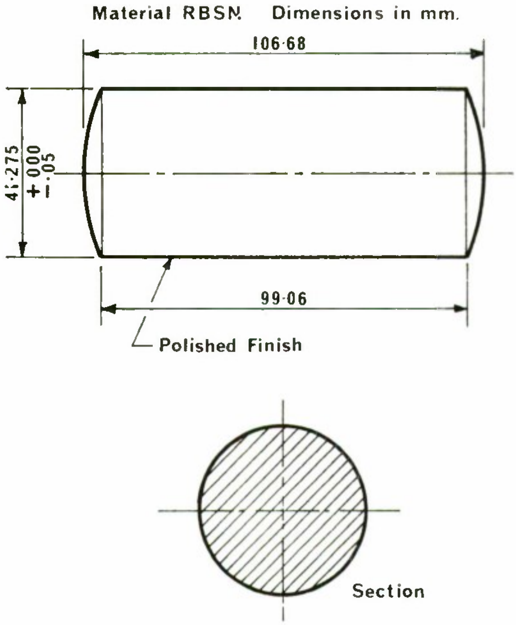


FIG. 4. Gudgeon pin design in RBSN.

FIG. 5. Silicon nitride piston after trials.

Piston Trials

The trials were conducted using the silicon nitride piston with the original steel gudgeon pin. Initially a number of runs were made with no externally applied torque and then the torque was gradually increased up to 45 Nm and then finally increased to full power torque at 91 Nm.

Most of the tests consisted of short runs of between 2 and 5 hours and over a period of 2 months during 38 separate tests, the total running time amounted to approximately 95 hours of which 54 hours was at full power.

Throughout the trials, the cooling water temperature remained at approximately 83°C the lubricating oil pressure was constant at 166 kN/m^2 and the engine revolutions varied between 1010 and 1060 rpm. The trials are summarised in Table I.

FIG. 6. Fractured silicon nitride gudgeon pin.

TABLE 1.

Si_3N_4Piston—HTS Gudgeon Pin No. of Tests 38.

Torque	*Total Running Time*
0	9 hrs
$\frac{1}{8}$ (11 Nm)	$\frac{3}{4}$ hr
$\frac{1}{4}$ ($22\frac{1}{2}$ Nm)	3 hrs 5 mins
$\frac{3}{8}$ (34 Nm)	1 hr
$\frac{1}{2}$ (45 Nm)	26 hrs 40 mins
1 (91 Nm)	54 hrs 5 mins

Total of 94 hrs 35 mins.

Throughout the trials, the piston behaved normally, and there was no evidence of any increased wear in the cylinder bore. For one of the trials, *Templugs* were fitted and the readings obtained are compared with those for the original aluminium piston (Table 2).

On one occasion, the engine was stopped after unexplained noises were heard even though the engine seemed to be functioning normally. On stripping down, it was discovered that a *Templug* had become detached from the piston crown and had apparently worked its way between the piston and the cylinder base. Apart from a few scratches, the only damage to the piston was a slight chip out of the top edge. The trial was resumed and the piston seemed unaffected.

TABLE 2.

Templug Temperatures.

Location	*Aluminium Piston*	*Silicon Nitride Piston*	*Temp. Increase*
Behind top piston ring	140	178	38°C
Underside of piston crown	126	204	78°C
Piston crown (near inlet valve)	157	199	42°C
Piston (on edge near crown exhaust valve)	149	207	58°C

At the conclusion of the trials, examination of the piston showed no dimensional changes and there was no evidence to show that the piston could not have run for many hundreds of hours.

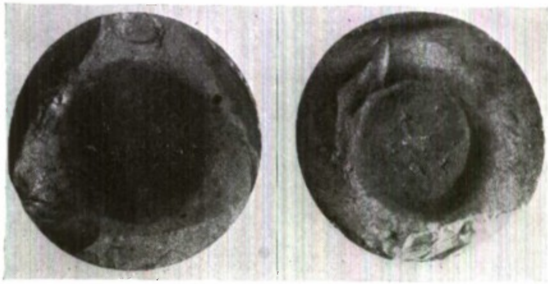


FIG. 7. Fracture surfaces of silicon nitride gudgeon pin.

Gudgeon Pin Trials

The Si_3N_4 gudgeon pin was tested using the original A1 piston. As in the piston trials, the engine was motored with no externally applied torque for about 8 hrs. and then the torque was increased to 11·5 Nm (approx. $\frac{1}{8}$ full power torque). After 15 minutes a loud knock was heard but oil pressure remained at 193 kN/m^2 (28 psi) and there was no excessive vibration. Some 15 minutes later, a consistent knock was heard and the oil pressure dropped suddenly to 124 kN/m^2 (18 psi) at which point the engine was stopped.

On stripping down the engine, it was seen that the gudgeon pin had sheared in a line with the small end bearing and the piston boss. Subsequent examination of the pin showed that nitriding had been incomplete, and that the centre core was still in a relatively soft condition.

Conclusions

Engine trials have shown that it is possible to manufacture and run a silicon nitride piston in a diesel engine and that it is sufficiently robust to stand the presence of foreign metallic bodies within the cylinder. These trials were however at relatively low powers in a single cylinder engine. It is hoped that future trials will take place at higher operating powers and temperatures in multi-cylinder engines.

Acknowledgements

The authors wish to thank the Admiralty Materials Laboratory and in particular Dr. D. Godfrey for all their help in manufacturing the silicon nitride components.

References

[1] H. F. P. Purday. 'Diesel Engine Designing' Constable.

[2] G. Eichelberg. 'Some New Investigations on Old Combustion Engine Problems' *Engineering*, **148**

[3] D. Fitzgeorge and J. A. Pope. 'An Investigation of the Factors Contributing to the Failure of Diesel Engine Pistons and Cylinder Covers'. *Trans. N.E. Coast. Engrs. Shpbldrs.* **71**, 5 (Feb. 1955).

[4] R. A. Dennis and J. M. Radford. 'Piston Stresses—Theoretical and Experimental Developments' I.Mech.E. Symp, 1964. Thermal loading of Diesel Engines.

[5] M. G. Holt. 'Pressure Stress Analysis of Three Single-Piece Pistons'. English Electric Report 69-RE-3, 1969.

[6] W. A. Harrington. 'Piston Rings, their Design, Functions and Manufacture'. *Trans I.Mar.E.* **66,** 12 (Dec. 1954).

[7] 'IC Engine Piston Ring and Groove Dimensions'. BS. 3267/63.

[8] B. Lawton and D. Crutcher. 'Mechanical Stresses in Pistons, Gudgeon Pins and Connecting Rods'. I.Mech.E. Diesel Symp. (March 1967).

[9] W. Weibull. 'A statistical Distribution Function of Wide Applicability'. *J. Applied Mech.*, **18,** 3 (1951).

[10] P. Stanley, H. Fessler, A. D. Sivill. 'An Engineer's Approach to the Prediction of Failure Probability of Brittle Components' (Nottingham Univ.) Science Convention. Cambridge Univ. 4 - 6 July 1972.

ROYAL NAVY'S PRESENTATION TO NELSON'S BIRTHPLACE

The Nelson memorial plaque made by the Royal Navy for the village of Burnham Thorpe, Norfolk, was presented, unveiled and dedicated there at 12 noon on Saturday, 14th June.

The plaque, which will stand in the centre of the village, takes the form of a sign with the inscription " Burnham Thorpe, Birthplace of Horatio Nelson, 29th September 1758 " and pictures of Admiral Lord Nelson and his father's parsonage.

Admiral Sir Anthony Griffin, Controller of the Navy, attended on behalf of the Admiralty Board, and the plaque was presented by Captain D. D. N. Long, captain of the Royal Navy's Marine Engineering School, Gosport, where it was made.

The dedication service included the prayer written by Nelson in his diary before the Battle of Trafalgar.

PHYSICAL METHODS IN SEA-BED SEARCH

S. J. Cocking
Marine Technology Support Unit
AERE Harwell

and

B. K. Gazey
Dept. of Electronics and Electrical Engineering
University of Birmingham

Introduction The famous search for the lost nuclear weapon off Palomares, Spain, in 1966 and searches for two lost nuclear submarines in 1963 and 1968 stimulated investigation of improved methods and instruments for sea-bed search in the United States (Patterson and Brundage, 1965 and Buchanan, 1970). Archaeologists (Bass and Katzer, 1968 and Hall, 1966) have advanced development of new search techniques since wrecks on the sea-bed offer a sample of ancient life isolated from confusion by later developments.

These methods should be increasingly making an impact on search and survey of the sea-bed for commercial purposes. Discovery and identification of specific objects such as wrecks is of increasing importance, as large ships operate with ever-decreasing bottom clearance. Re-location of expensive hardware in the sea has often proved surprisingly difficult, and even such massive objects as oil wellheads have proved difficult to find. The location and inspection of sea-bed pipelines and cables is a rapidly developing need. Detailed knowledge of the state of the sea-bed is required for routeing cables and pipelines, for general geological survey and in particular for discovery and assessment of manganese nodule deposits. Fisheries research and studies of sea-bed transport further involve detailed sea-bed surveys. These applications justify some study of the available physical techniques for in-water search.

The total problem in search for an object on the sea-bed may be separated into means of detection and identification of the object, and also means of accurate location in the search area which may be large or small, depending on previous information. Space does not permit coverage of the latter, interesting problem, for which a wide range of methods of surface radio navigation, and subsurface acoustic navigation are under development (Penzias and Goodman, 1973). Instead, the present discussion is limited to acoustic, optical and magnetic methods of detection and identification, areas of current development being emphasised.

Electro-magnetic waves suffer considerable attenuation in water over the whole wavelength spectrum from gamma rays to radio waves, the best transmitted wavelengths being in the visible region. Light attenuation is extremely variable from place to place due to the widely differing loads of suspended or dissolved matter (Williams, 1973). Thus visual detection is limited to about 50 metres in clear oceanic waters, and is often as low as one metre in coastal regions. It is well known that longitudinal sonic waves are well transmitted in water, and that the sonic reflection from objects of differing acoustic impedance to their surroundings is widely used for detection at ranges from centimetres to many kilometres. Thus at medium and long range, sea-bed search is the province of acoustic methods. Nevertheless identification, particularly if the target lies among rock, may require the more detailed information available with visual observation. An object detected by acoustic means is thus often identified by visual means so that the methods complement one another.

Basic Properties of Light and Sound in the Sea

Intensity decrease, I_0 to I with range, R (for R $\gg$ the source size) may be expressed as

$$I = I_0 \exp(-\alpha R)/R^2$$

the parameter α, is known as the attenuation coefficient, $\frac{1}{\alpha}$ as the attenuation length. Figs. 1(a) and 1(b) show the variation of attenuation coefficient with wavelength for light and sonic waves.

The wavelength for minimum attenuation of light shifts from 480 nm in pure or clean sea water towards 550 nm in typical coastal waters; the deep blue and green of these water masses is well known. X-rays and gamma rays are so rapidly attenuated that reactor cores may be viewed in swimming pool reactors through a few meters of water, while at the other extreme radio waves have effectively zero range in water. Visible light in a relatively narrow wavelength band is the only em radiation available for detection or communication over a range of tens of metres. Optical detectors thus require high sensitivity in this region and light sources with high emission in this narrow band are required.

In particle-free water, optical detection range would be limited only by light absorption. In practice, suspended matter limits identification range to about 50 metres even in the clearest oceanic waters. Suspended matter consisting of plankton, sand and silt add to the light attenuation by further absorption and by scattering the light. These processes lead to a rapid decrease of natural illumination with depth. Artificial illumination must supplement or replace natural light in all but shallow water. Power supply at depth proves costly and efficient application of available power is necessary. Discharge lamps which offer high optical efficiency, together with the possibility of maximum emission at the best transmitted wavelengths have been developed. Mercury discharge lamps with thallium iodide additives yield a dominant emission at 540 nm, a wavelength which is well transmitted and which also corresponds to the maximum detection efficiency of vidicon tubes used with conventional TV cameras. Due to wavelength dependent absorption (Fig. 1) colour information at ranges in excess of say 10 m is not obtained.

Classical calculation of sound attenuation in liquids due to viscous damping and thermal conduction gives an attenuation coefficient

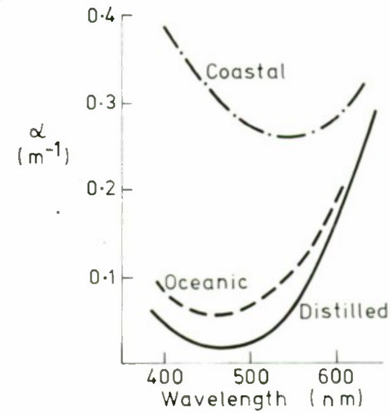


FIG. 1(a)

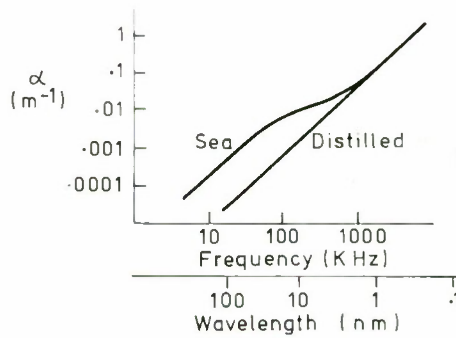


FIG. 1(b)

FIG. 1(a). Attenuation coefficient for electromagnetic radiation in the visible region. The data for sea water is typical but wide variations occur (Williams 1973).

FIG. 1(b). Attenuation coefficient for sound in water. For distilled water $\alpha \propto$ (frequency)^2 This is enhanced by resonant absorption by dissolved sea salts.

proportional to the square of the sonic frequency, f. For fresh water, thermal conduction loss may be neglected but the measured value of $\alpha/f^2 = 25\cdot10^{17}$ $cm^{-1}s^2$ at 20°C can only be explained using a 'volume' viscosity approximately four times higher than the normal shear viscosity. The considerably higher attenuation of sea water up to 1MHz has been explained by relaxation phenomena associated with dissolved salts, principally magnesium sulphate (Stephens, 1970). At 1MHz the attenuation coefficient is comparable with that for light in clear water so that lower frequencies are used to extend detection range. Usable sonic wavelengths are between 10^3 and 10^6 times longer than those of light waves, for example at 1MHz the wavelength is $1\cdot5\ 10^{-3}$m. Sonic waves thus offer lower resolution of target detail. Piezo-electric transducers are most commonly used for transmission and reception of sound in water. Maximum power per unit area of transmitting face is limited by cavitation to about 1/2 Wcm^{-2} near the surface; this however increases with depth and if pulse excitation is employed. Transducers may be shaped according to the required transmission pattern, for example a fan-shaped transmission pattern can be achieved with elongated transducer geometry.

The simplest detection problem arises with an 'active' target, that is one which emits a recognisable signal. Objects whose recovery is essential are frequently marked with acoustic pingers or, less frequently, optical flashers. 'Passive' targets must be detected by their characteristic reflection of an imposed signal. Maximum detection range depends on the attenuation of the imposed signal and on the difference of reflecting properties of the target compared with those of the background against which it is detected. The sources of background which limit optical and sonic detection are thus of central importance.

Sonic Methods

Two basic types of sonar (Sound Navigation and Ranging) are currently in use: the most common employing the pulse-echo technique, Fig. 2, in which a short pulse of acoustic energy is transmitted in a well defined 'beam', echoes received are used to infer the presence of targets and the time elapsed between transmission and reception used to measure range.

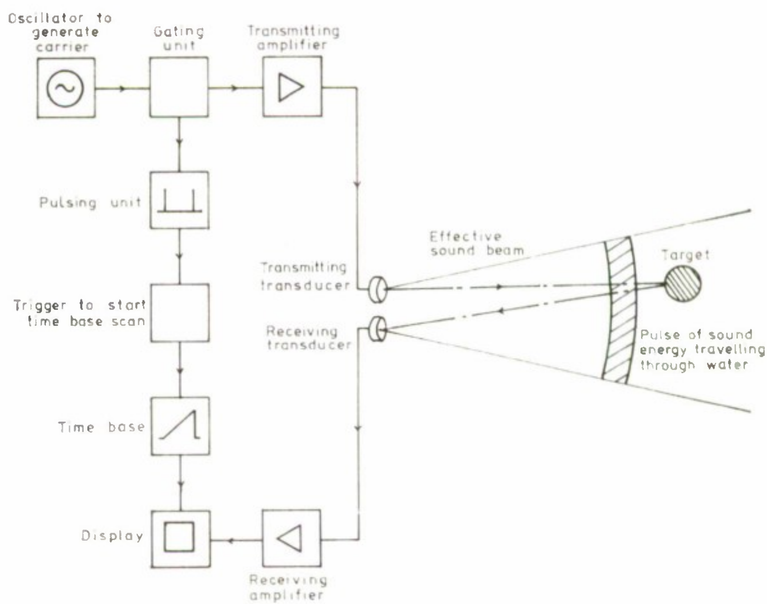


FIG. 2. Schematic arrangement of a typical pulsed sonar system.

The display medium may be aural or, more frequently, either a chart recorder or a cathode ray tube display, the latter often provided with a variable or long persistence phosphor to permit some integration of successive traces to improve the detection capability. (Skolnik, 1962).

Less commonly, Continuous Tone Frequency Modulation (CTFM), sonar is used in which the carrier is swept in frequency during a relatively long pulse. The echoes, which are delayed versions of the original signal, are heterodyned with the outgoing signal and the frequency difference is used as a measure of the range of the target.

Vertical, Forward and Side Scanning Search

The pulse-echo method may be used in a a variety of modes. The simplest, currently used in hydrographic survey, is the depth sounding mode in which transmitter and receiver are directed vertically to the sea-floor to measure depth. As the carrying vessel moves forward a narrow contour of the sea-bed is obtained; the probability of detecting a small target in this mode is of course low. If high transmitted powers and low carrier frequencies are used sound may penetrate into the sedimentary layers of the sea-bed to give information about the sub-bottom structure, buried pipelines etc. (Belderson, *et al.*, 1972).

In forward search the transducer is designed to transmit a fan-shaped beam inclined to impinge upon the sea-bed at an acute angle. The time varying echo at the receiving transducer thus represents the sea-bed reflectivity as a function of oblique range over the narrow insonified sector. Transmitting and receiving transducers are scanned mechanically back and forth so that as successive echo traces are laid radially on the display medium, a view of a larger sector of the sea-bed is built up. The speed of sound in sea water being about 1500m/s, considerable time is needed to investigate a wide sector. A transducer must maintain direction until echoes are received from targets at maximum range before being stepped round. A typical maximum range of 300m occupies 1/5s, a 45° sector scanned with a 3° wide beam requiring three seconds. Thus for high angular resolution the search rate is low.

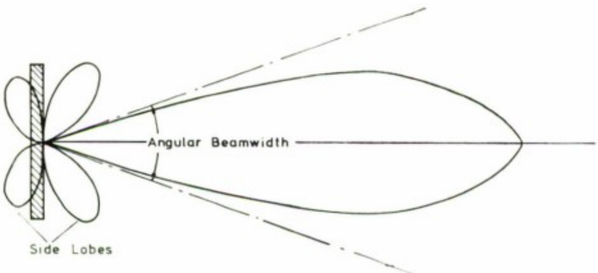


FIG. 3. Directional response (in polar form) of a disc transducer.

The transducers must maintain directional stability in spite of the motion of the carrying vessel, and costly stabilisation equipment is often required. If the transducers are mounted to transmit to the side of the vessel, its forward motion allows successive echoes to be laid side by side on the display medium, usually a paper chart recorder, an oblique view of the sea-bed to one side is thus built up without the need for mechanical scanning. Accurate surveys of vast areas of the sea-floor have been made by allowing for variations of ship's speed, compensating for distortion of the record caused by the ship's motion (Chesterman, 1971). Directional stability may also be obtained by mounting the transducer in a submerged towed body. This also reduces the distance from transducer to sea-bed allowing a higher resolution over a smaller area. Simultaneous scanning to port and starboard is sometimes used. A typical commercial side scan sonar uses a transducer mounted in a towed 'fish' operating at 120kHz giving maximum range on each side of about 300m. The fan-shaped acoustic pulse 1° wide by 20° high occupies 0·1ms. The towing speed is less than eight knots.

Resolution

The definition of position, or resolution, of detected targets on the display is determined in angle by the sonic beam's angular width and in range by the pulse length.

A single disc transducer of diameter a, will have a three-dimensional directional function with the cross section shown, Fig. 3, characterised by the function

$$D(K) = \frac{2\,J_1\,(Ka/2)}{Ka/2}$$

where $K = \frac{2\pi}{\lambda}$ Sin θ and $J_1(\ )$ is the Bessel function of first order and first kind.

This complex form, having side lobes in addition to the main peak, arises from constructive and destructive interference of the coherent waves from the elements comprising the full disc.

The angle between the first zeros in this pattern, termed the beam-width, $D\frac{1}{2}$ is given by 2 Sin^{-1} $1{\cdot}22\lambda/a$. It is clear that narrow beams require large diameter transducers, $a>>\lambda$ for which $D\frac{1}{2} \approx \frac{150\lambda^\circ}{a}$, which may prove costly and unmanageable if low frequency carriers are used to achieve long range operation.

The resolution in range is $c\tau$ where τ is the duration of the transmitted pulse and c is the velocity of sound. The lower pulse duration limit is usually set by the fractional bandwidth of the transducer; consequently high resolution requires a high carrier frequency, with a relatively short operational range because of higher absorption losses.

Improvements in range resolution at low frequencies can be obtained by sweeping the transmitted energy from low to higher frequency during a relatively long pulse. A received echo signal is passed through a dispersive delay line in which the low frequency components travel more slowly than the higher frequencies. If the dispersive characteristics of the delay line are matched to the transmitted signal, then all frequency components emerge from the delay line together to produce a much shorter pulse giving a higher range resolution.

The inherently non-linear pressure-density relationship of all acoustic media has been exploited to improve resolution in range and bearing (Muir *et al.*, 1969). If two high frequency primary signals at slightly different frequencies are radiated simultaneously from the same array, non-linear interaction in the water produces a virtual array of sources at the intermodulation frequencies. In this way it is possible to produce a narrow beam at a low frequency from small transducers albeit at the expense of high primary intensities.

Many advantages result from subdividing the radiating face to form an array and adjusting the radiated power or the receiving sensitivity of each element, this process being referred to as 'tapering' or 'shading'. For example, the side lobes of the directional function D(K), which give rise to ambiguous target detection, may be partially suppressed. A cosinusoidal taper results in reduction of the first side lobe from 21% of the main lobe for a uniformly excited array, to 9%. A phase difference between elements of an array may similarly be controlled (phase-taper). When phase-taper is used with constant phase differences between adjacent elements of the array the polar diagram is displaced from straight ahead. A quadratic phase taper can be used to focus the sound as if the array had been mechanically deformed into a concave arc. Such an arrangement allows high angular resolution albeit over a limited depth of field. When the constant phase difference between elements is time-varied the beam is deflected over a sector. This, 'electronic sector scanning' has been used in increasing search rate.

High Search Rate Sonar

Three high search rate sonars are in use, achieving an increase of search rate between 30 and 100 times those of simple systems. In each a transmitted pulse insonifies (the sonic equivalent of illuminates) a sector, typically 30 - 45° wide; each employs a sectionalised receiver array, but the output signals are processed differently. Welsby *et al.* (1963), use a long sectionalised receiver array to provide a narrow (1°) directional response. This response is scanned across the insonified sector during the time taken for the sound pulse to move its own length in water, thereby ensuring that no target information is lost. The high deflection rate typically 10^4 scans per second is obtained by electronic sector scanning. Complete 360° search can be obtained by using a cylindrical array. Target echoes modulate the intensity of a long persistence cathode ray tube upon which a bearing *versus* range raster is generated, the former being synchronised with the rapid angular deflection of the receiver's directional response, and the latter to the slower passage of the sound pulse through the water.

In 'Pre-formed Beam' sonar (Nesbitt *et al.*, 1971) a number of linear phase tapers in parallel are applied to the array to generate a number of overlapping directional responses in space from which target information is available continuously as the sound pulse propagates. A scanning switch is used to sample the outputs of the beam forming processor and the resulting serial target information modulates the display as in the previous case. With the previous two methods, phase

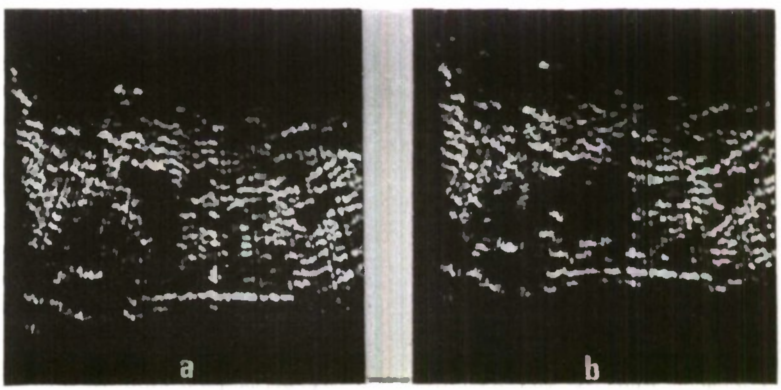


FIG. 4(a). Problems of display interpretation—mechanically scanned forward search sonar display showing pistol lying on a sandy bottom. The pistol is clearly seen against the smooth bottom.

FIG. 4(b). Ditto—but in (b) bottom roughness confuses the shadow cast by the pistol.

shifts are imposed between adjacent array elements. Alternatively the phase difference between signals received on adjacent array elements can be measured. The mean phase difference across the array gives target bearing and variation between adjacent pairs are used in the signal processing unit to distinguish between a target and noise (Creasey *et al.*, 1969). Signal amplitude information is not used and the processor is made relatively cheaply using digital circuit techniques. Although phase only sonar is capable of displaying only one target per range increment, experience has shown that this is not a serious limitation.

An electronic sector scanning sonar developed at the Admiralty Research Laboratory at Teddington, England (Cook and McKay, 1971) is currently being developed commercially. One version insonifies a 60° sector which is scanned at a rate of 7500 scans per second. A 300 microsecond pulse yields a nominal range resolution of 0·1m and beam width 0·5° equivalent to a linear resolution of 0·3 metres at 175 metres range. With operating frequency of 180kHz ranges to 600m are possible.

Problems of Propagation and Interpretation

The sea is an imperfect sonic transmission medium. Echo signals fluctuate in amplitude, phase and envelope due to thermal microstructure generated by surface and internal waves (Mintzer, 1955). In addition, indirect signal paths between sonar and target *via* a rough sea surface or sea floor give rise to pulse lengthening or even a false detection range. Variations of sound velocity result in non-rectilinear propagation; rays bending away from sound velocity maxima. Such continuous refraction can cause severe geometrical distortion of a sonar record and in extreme cases a sound velocity maximum can cause a shadow zone into which sonar transmissions cannot penetrate directly at all.

Suspended matter, plankton and fish act as unwanted targets, causing a background against which target echoes have to be detected (Urick, 1967). Reverberation can also arise from a rough sea surface and sub-surface air bubbles produced by wave action. Noise is generated within the sea (Wenz, 1962) by breaking waves, by distant ships and storms as well as the noise produced by the biological population and such transitory phenomena as rain squalls. Machinery and propeller cavitation of the vessel carrying the sonar is a major noise source, as is water flow over the receiving transducer. Mounting the transducer in a towed, streamlined 'fish' can alleviate these problems which are of central importance in designing the signal processor.

A detected target may be identified if the sonar's resolution is sufficient for some characteristic features of the target to be distinguished. However, the pattern of reflection from a target is usually complex. This, together with propagation problems described, produces a display which generally bears little relationship to the geometrical shape of the target and raises serious interpretation problems. Location and recognition is easiest when a target casts a shadow against an otherwise uniformly scattering bottom, and this shadow gives the projected cross-sectional area of the target. However, any large scale bottom roughness or bottom contours can completely nullify the chances of location or recognition, Fig. 4. Unambiguous identification can often only be made visually.

Image Forming Systems and Holography

Acoustic waves may be focused using lenses and reflectors to form a sonic image in space. The problem is then to transform this image into a visible display. Acousto-electric and acousto-optic converters have been demonstrated (Berger, 1969) but at present are too insensitive and have apertures too small for most field applications. Acoustic transducers emit coherent waves, acoustic holography (Hildebrand *et al.*, 1972) thus appears a natural choice for imaging, but a suitable sonic to optical image converter operating in real time is lacking.

These methods, though limited in range, offer the possibility of imaging in conditions where high turbidity would prevent visual search.

Optical Methods

Visual search involves either direct human vision, photographic cameras or television, the latter two being inevitable for extended search or survey due to their storage and re-run capability. Television is particularly valuable in showing remotely, in real time, the underwater view. However, in the most time-extended searches and in surveys it can be advantageous to photograph the area systematically for later processing and study.

The higher resolution of photographic images compared with television (with the exception of special systems which are bulky and expensive) can be turned to advantage in allowing a wider area of sea-bed to be imaged with each traverse of the camera. Since the imperfect optical properties of typical sea-water limit the viewing range, wide angle lenses for in-water use have been developed to maximise the viewed area. Viewing range is limited by simultaneous loss of image intensity, resolution and contrast. Some approaches to alleviation of these problems are indicated here.

The attenuation of light in sea water limits the depth of unaided human vision or imaging by standard television systems or photography often to less than 50 metres. However considerable increases of detection sensitivity can be achieved with the use of electro-optic devices which offer intensification of received images by factors up to 10^5. These can be used either for direct viewing or placed in front of a camera. Sensitive TV imaging tubes now available allow viewing of scenes illuminated by the equivalent of starlight only. Such sensitive camera tubes are currently being introduced into underwater applications.

The light scattering by suspended matter limits the advantage which can be gained solely by overcoming problems of light intensity; scattering also limits optical resolution. Image carrying light is scattered predominantly through small angles, and part is received together with unscattered light. The blurring of the resulting image increases with target range. Inhomogeneities of refractive index in natural water bodies arising from temperature and salinity variations also limit resolution.

A leading additional problem is the scatter of ambient light by suspended matter into the receiver, resulting in a reduced apparent contrast of the target image. The detection range for a dark object viewed horizontally in free water is $4/\alpha$ (Duntley, 1963) based on detection of a contrast difference of 2%, appropriate for well lighted conditions.

Addition of artificial illumination can, under some circumstances, make detection or recognition worse, since light back-scattered from suspended particles into the receiver can mask the target image. A variety of special techniques to restore image contrast, all seeking to exploit some difference between wanted and unwanted light, recently reviewed by Cocking (1975), are outlined here.

Reduction of backscatter: geometrical, gating and polarisation methods

Fig. 5 shows a typical arrangement of lamp, target and optical receiver. The volume of water and hence suspended particles which is both illuminated and viewed is responsible for the veiling backscatter. A simple method to minimise this volume requires only a narrow angle light displaced from the receiver, directed to illuminate only the region of the target. This geometrical isolation, method showed promising results when tested in 1971 on a submersible by one of the present authors. A variable beam angle lamp was developed to allow a search for targets in an uncertain position with a wide beam, the narrow beam being used for final identification.

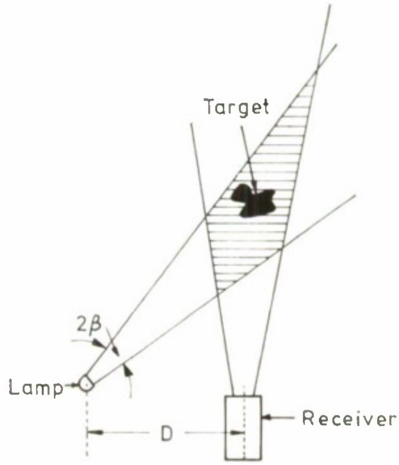


FIG. 5. A typical arrangement of lamp, target and receiver. The displacement D, and cone angle of the lamp, β, are defined.

The basic principle of reducing the scattering volume by reducing beamwidth can be extended to such narrow beams that only an element of the target would be illuminated. The total viewed area is then built up by scanning this light beam, a mechanically coupled photo-detector gives a time varying signal which, when displayed on a television monitor with synchronous scanning, builds up the total image. Laboratory demonstrations of this method have been reported (Angelbeck, 1966).

Light sources can be pulsed with an on-time of a few nanoseconds, using Kerr or

Pockels cells. Similarly optical switches or electronic switching of imaging tubes allows the use of gated receivers. Range selection can thus be achieved by timing the light pulse on its path from emitter to receiver. This method, so closely analogous to radar and sonar, is appropriately termed 'lidar'. Lidar allows the light from a target at pre-selected range to be isolated, eliminating the back-scatter from other distances. Fig. 6 gives an example of the fall-off of contrast with range which can be achieved with range gating compared with those from the application of geometrical isolation. This example assumes 20 and 40 nanosecond wide gating pulses for both illuminator and receiver. For short ranges (less than $2c\delta t$ where c is the light velocity and δt the gating time), the range gating method achieves less than the basically simpler isolation method. However at greater ranges considerable and often impractical separations would be needed to achieve the same contrast improvement. Range gating can, in principle, be used to reduce backscatter effectively to zero, since techniques to achieve pulses as short as 10^{-12} seconds have been demonstrated; however, at present the low mean optical power available limits the practicable viewing range. Keil, Immarco and Kerpchar (1968) demonstrated an improvement with range gating compared with continuous lighting at ranges up to 39m corresponding to 4·5 attenuation lengths. Several undetailed claims of extension of range of up to five times the range for unaided vision have appeared. Range gating involves relatively complex and massive gear, and these features must restrict the potential applications.

Since a polarised light beam is largely depolarised on reflection from a matt target, while the scatter from small particles retains the original polarisation, crossed polarising filters over light source and receiver should enhance target contrast. Following earlier reports of great promise for the method, Gilbert (1970) showed that improvement or deterioration of contrast was strongly dependent on the size range of the particulate suspension; only for particles of less than 0·8 μm diameter was contrast improved for all particle concentrations.

All of the above methods offer some improvement in image contrast but the benefit of any one has not been adequately demonstrated in a real operational situation. Only if the target is itself the light source and the viewing path is not illuminated by ambient light can the problem of unwanted scatter be eliminated. Image quality, limited by degraded resolution, can then be optimised by a range of techniques grouped together as image processing.

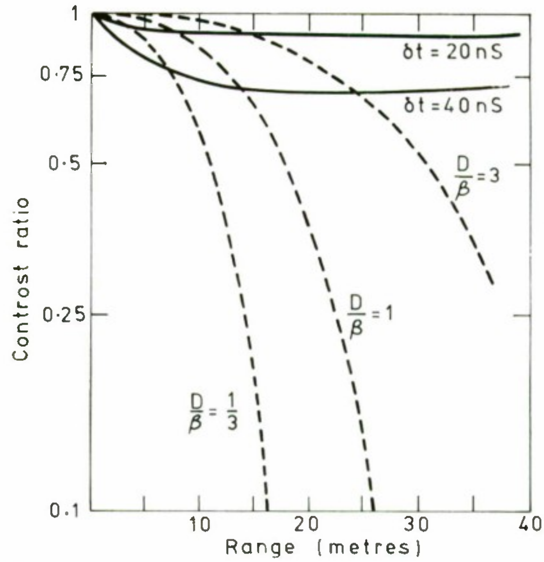


FIG. 6. Fall-off of target contrast with range in water with $\alpha = 0{\cdot}2\text{m}^{-1}$. The contrast is improved using the geometrical isolation method by increasing the ratio $\frac{D}{\beta}$ — dashed curves. Using range gating (full curves) the contrast can be maintained at a level determined by the gating time δt.

Image Processing

Image processing comprises a wide variety of modifications which can be used to diminish or enhance particular features of an image. Of special interest in the present context are reduction of backscatter and enhancement of detail reproduction. Further, known target shapes can be selectively sought and where a number of images of the same subject from different viewpoints are available, correlation between these images can give further selective detection.

An image can be regarded as a set of two dimensional Fourier components covering a spectrum of spatial frequencies. In terms of these components backscatter represents an unwanted addition at low spatial frequencies while resolution loss is seen as attenuation of the higher frequencies. Processing of through-water images should therefore tend to diminish low and enhance high frequency components; the limitation on diminution being that components of the wanted image are reduced while enhancement is limited by simultaneous enhancement of image noise. Mertens (1970) described the preparation of an optical filter having transmission minimum at low spatial frequencies and increasing at higher frequencies by preparing a blurred photographic negative from the original image. The original image when transmitted through the filter shows enhanced detail reproduction. This method illustrates one of the techniques which can now be applied to an optical image (Mazurowski, 1969) or in electronic form in an electro-optic storage tube (Munsey, 1968). A powerful method used in the U.S. space programme (Billingsley, 1970) depends on digitising the image information, this digital information being processed at will in a computer. Removal of backscatter has been demonstrated with the last two techniques.

Television scanning yields a two-dimensional image as a time-varying one-dimensional signal. Continuous processing of this signal can be advantageous. However, since different parts of a picture are separated in time, complete image processing is only possible with total image storage and access. Processing techniques to enhance edge definition and increase picture contrast photographically are well known, and are used by experienced underwater photographers. The present authors have demonstrated improved image recognition by electronic enhancement of contrast (γ enhancement). The human eye-brain combination performs the most sophisticated image processing, including time correlation. The great potential in optical and elecronic image processing lies in presentation of images in a form which can best utilise this remarkable human interpretive ability.

Metal Detectors and Magnetometers

Certain materials such as metals offer the possibility of selective detection. For example ferrous metallic objects disturb the earth's magnetic field and thus maintain a distorted local field which can be detected, albeit over relatively short ranges of the order of metres.

Sensitive magnetometers are commonly used offshore in exploration for mineral deposits (Gotla, 1970). The widely used proton magnetometer measures the Larmor precession frequency of protons about the local magnetic field. This precession frequency is accurately proportional to the local magnetic field. From the difference reading of two such sensors, the local field gradient may be accurately measured. Such magnetometers have therefore been used to locate wreckage and also lost pipelines. Flux gate magnetometers and optically pumped magnetometers based on the Zeeman effect have also been developed and offer high sensitivity with continuous readout.

Metallic objects, not necessarily ferrous, even when buried at shallow depths, can be selectively detected since they respond to an imposed magnetic field. Such instruments have been developed for use in the sea. A brief current pulse in a coil induces an eddy current in metallic objects within the range of the magnetic field. The resulting magnetic field which persists beyond the imposed pulse time may be detected using the same coil. The detection probability varies as d^{-3} for small objects where d is the distance of the field coil from the target. Typical detection ranges are 0·5m for a copper coin or 2·5m for a steel gas pipe.

Ore deposits exhibiting nuclear radiation have been detected on the seabed by towed scintillation counters. Gamma ray back scatter methods and X-ray fluoroscopy are also techniques being applied for identification of specific elements on the seabed.

Practical Examples

Technical possibilities frequently await an application before they are realised. This is certainly true of sea-bed search techniques. The emergencies of lost submarines and apparently the great interest of archaeological exploration have so far been responsible for leading the way. Whether commercial exploitation of the sea-bed will contribute to further development may perhaps depend on how clearly the potential benefits of new technologies can be spelt out. Thus the challenge may lie less in new physical techniques than in detailed analysis of costs and benefits.

Bass, *et al.* (1968) developed a pressure-proof vessel with a transparent dome to allow a man to visually search the sea-bed for ancient wrecks while being towed from a surface craft. Continuous observation time was, however, greater using a towed television camera. Bass concludes that the combination of side-scan sonar for potential target detection with television and finally divers for final identification offers the optimum system for archaeological search.

Continuous development of a towed sensor package has been stimulated by the loss of a nuclear bomb and submarines in the U.S.A. (Patterson, 1968 and Buchanan, 1970). An aluminium alloy frame, roughly 1m × 1m × 3m in size carried three photographic cameras with flash illuminators, two side-scan sonars (operating at 35 and 40 KHz), scintillation counter for radiation monitoring and a proton precession magnetometer on a boom to separate it from the magnetic effect of the other instruments. Experience in the use of this system (Buchanan, 1970) in the particular conditions of a search for a nuclear submarine indicated that their side-scan sonar detected so many targets with insufficient identification that the sensor was not of primary importance, further, the nuclear radiation detector gave null results. Hence optical and magnetic detectors have been emphasised in further development of the package. Photographic cameras were used rather than television since the higher resolution allowed target identification at greater distances from the bottom and hence of greater area coverage in a single instrument sweep. Duplication of the instrument package allowed processing of one set of photographs while another package continued the photographic recording. A photomosaic of the seabed was built up in this way and two lost submarines successfully found at depths to 2,400m.

Exploitation of the manganese nodule deposits on the sea-bed has led to the development of instrumented sledges which, being towed behind a vessel, either slip along the sea-bed or are lifted by hydro-dynamic forces to travel several meters above it. Such a sled, towed by the German research ship *Valdivia* carries both television and photographic cameras. Illumination is provided for vidicon television cameras, but to conserve power in illumination, low light television cameras are carried; powerful flash affords illumination for still photographs while the sled is in motion.

The above instrument packages are towed behind a surface vessel. Propeller-driven, surface-controlled sensor packages are under development. Manned, autonomous vehicles are proving their value in inspection of fixed sea-bed structures, but speed and endurance limit their value in wide area search.

An essential part of an area search operation is a local navigational system. When search operations are repeated at widely separated times, a local system must be referred to a more permanent framework, such as one of the hyperbolic radio chains or by satellite navigation. Several sub-surface location systems based on range-range measurements using acoustic transponders have been proposed or tested. A comprehensive system based in the U.S. Navy's specially adapted vessel *Mizar* was used in the submarine searches. *Mizar* carried a sonic transmitter and three sonic detectors mounted at the corners of an isosceles triangle on the hull. The instrument package carried an acoustic transponder which, on receiving a sonic pulse, replied with a pulse which was in return detected by the three sonic receivers. Using the known sound velocity the range of the transponder from the three receivers is determined and hence the vector of the transponder relative to *Mizar* was calculated using a shipboard computer. A second transponder was fixed to the sea-bed in the search area, so that the position of the package relative to this fixed sea-bed position was continuously plotted. Further, the computer received an input from the navigational system of *Mizar* which allowed the sensor package position to be determined relative to north-south, east-west and geographically vertical axis. More modest systems for commercial operations are being developed.

This fascinating subject, together with consideration of the optimum search operational procedure, deserves separate review outside the scope of the present article.

References

Angelbeck, A. W. (1966). Application of a laser scanning and imaging system to underwater viewing. *Underwater Photo-optics; seminar proceedings* pp. B-VII-1-12. Redondo Beach, Calif., Society of Photo-optical Instrumentation Engineers.

Bass, G. F. and Katzev, M. L. New tools for Underwater Archaeology, *Archaeology* 21 (1968) 164-173.

Belderson, R. H., Kenyon, N. H., Stride, A. H., and Stubbs, A. R. Sonographs of the sea floor. Elsevier, Amsterdam (1972).

Berger, H. 'A Survey of Ultrasonic Image Detection Methods', Acoustical Holography, **1**, Eds. A. F. Metherell, H. M. A. El-Sum and L. Lamore, Plenum Press, New York, Chapter 2 (1969) 27-48.

Billingsley, F. C. Applications of digital image processing. *Applied Optics*, **9** (1970) 289-292.

Buchanan, C. L. Optical and Magnetic Instruments for Mapping the Ocean Floor. Rep. I.E.E.E. Trans. on Geoscience Electronics GE.8, 4, 320 - 325 (1970).

Chesterman, W. D. Sediment mapping with isometric side scan sonar. *Proc. BAS* **1** (1971) 1.

Cocking, S. J. In 'Underwater Research', Editors Drew *et al.* (1975). Academic Press, London.

Cook, J. C. Mackay, J. M. National Electronics Review 7, No. 1 (1971), 5.

Creasey, D. J. and Braithwaite, H. B. Experimental results of a sonar system with a digital processing unit. *Applied Acoustics* 2 (1969 39 - 57.

Duntley, S. Q. Light in the sea. *J. Opt. Soc. America,* **53** (1963) 214 - 233.

Gilbert, G. D. The effects of particle size on contrast improvement by polarization discrimination for underwater targets. *Applied Optics,* **9** (1970) 421 - 428.

Gotla, P. The magnetometer in oceanology. *Oceanology International* (Feb. 1970) 36 - 39.

Hall, E. T. Use of a proton magnetometer in underwater *Archaeology, Archaeometry* **9** (1966) 32 - 44.

Hildebrand, B. P. and Brendan, B. B. An introduction to acoustic holography. Hilger, London (1972).

Keil, T., Immarco, A. and Kerpchar, M. Recent underwater range-gated measurements. *In* Underwater Photo-optical Instrumentation Applications: *SPIE seminar proceedings,* **12** (1968) 21 - 32. Redondo Beach, Calif., Soc. of Photo-optical Instrumentation Engineers.

Mazurowski, M. J. Mathematics with light. *New Scientist,* **44** (1969) 636 - 639.

Mertens, L. E. In-water photography: theory and practice. New York (1970). Wiley Interscience.

Mintzer, D. Wave propagation in a randomly inhomogeneous medium 1. JASA. **25,** 5 p. 922 - 927 (1955).

Muir, T. G. *et al.* (1969). Non-linear acoustics. *Proc. Conference* at Applied Research Lab. Austin, Texas.

Munsey, C. J. An electro-optical underwater search and visibility enhancement technique. *In* Underwater photo-optical instrumentation appl: *SPIE seminar proc.* **12** (1968) 175 - 182. Redondo Beach, Calif. Soc. of Photo-optical Instr. Engineers.

Nesbitt, R. and Berktay, H. O. An electronic tracking system for acoustic telemetry, *Proc. BAS,* **1,** 2 (1971).

Patterson, R. B. and Brundage, W. L. Deep sea photographic search techniques. Proc. of Ocean Science and Ocean Engineering conf., Marine Technology Soc. (1965) 1178 - 1197.

Patterson, R. B. Photographic Techniques for Ocean Floor Search and Research. In Marine Science Instrumentation, **4.** Plenum Press (1968) 354 - 363.

Penzias, W. and Goodman, M. W. (1973). Men Beneath the Sea. Chapter 6 'Communications, Navigation and Sonar'. Wiley Interscience, New York.

Skolnik, M. I. (1962). Introduction to radar systems, McGraw Hill, Chapter 2.

Stephens, R. W. B. in 'Underwater Acoustics' Chpt. 1 Editor: Stephens, Wiley-Interscience (1970).

Urick, R. J. (1967). Principles of underwater sound for engineers, McGraw Hill, Chapter 8.

Welsby, V. G. and Dunn, J. R. High resolution sector-scanning sonar. *J. Brit. IRE* **26,** 3 (1963) 205 - 208.

Wenz, G. M. Acoustic ambient noise in the ocean, Spectra and sources. *JASA,* **34,** 12 (1962) 1936 - 1956.

Williams, J. Optical properties of the ocean. *Rep. Prog. Phys.* **36** No. 12 (1973) 1567 - 1608.

OXIDATIVE ATTACK ON AN EPOXIDE RESIN

Commander G. Emmons, B.Sc., M.Sc., Ph.D., F.C.S., M.I.E.E., R.N.
Royal Naval Engineering College, Manadon

Introduction

The results presented and discussed in this article derive from a study of the epoxide system which is formed when the diglycidyether of diphenylol-propane (DGEBA) is co-polymerised with the stoichiometric concentration of diaminodiphenylmethane (DDM) in an organic solvent.*

DGEBA : $H_2C—CH—CH_2—O—Ar—C(CH_3)_2—$ (epoxide O bridging H_2C and CH)

$Ar—O—CH_2—CH—CH_2$ (epoxide O bridging CH and CH_2)

100 p.b.w.

DDM : $H_2N—Ar—CH_2—Ar—NH_2$

27 p.b.w. (50 p.b.w. of solution)

A cured sample examined in an electron spin resonance (ESR), X-band, spectrometer gives rise to a first derivative spectrum which is, typically, of the form shown in Fig. 1. The changes in amplitude of the line show that the stable radical population increases with the time elapsed after cure, and reaches a steady state in about one month. Changes in line shape during this period suggest that the spectrum is not due to a single radical species.

* The solvent used in this work was γ - butyrolactone. The solvent molecules are not part of the resultant cross-linked structure, and the full amount of solvent used can be recovered from the cured epoxide system.

In this work the stable radicals which produce the spectrum are identified, the conditions which lead to their formation are established and reaction mechanisms which govern radical formation and destruction are proposed.

Experimental

Sample Cured and Examined in Vacuo

A sample was prepared by heating 100 p.b.w. DGEBA to about 80°C, in order to reduce its viscosity, adding 50 p.b.w. of a solution of DMM in γ-butyrolactone, which contributed 27 p.b.w. DDM to the mix, and stirring thoroughly. A quantity of this mix, sufficient to give a column of about 50 mm in length and 3 mm in diameter, was introduced into a spectrosil tube which was then connected to a high vacuum line. The system was pumped to 10^{-1} mm Hg by use of a rotary pump and then a diffusion pump was used to take the pressure to 10^{-6} mm Hg. During the initial stages of pumping, whilst bubbles formed, judicious application of heat reduced the viscosity, allowing bubbles to escape and sufficiently easy flow to maintain the column. Since flow occurred readily at a stage when no more bubbles were able to form, it seems likely that most of the dissolved air was removed. When the pressure reached 10^{-6} mm Hg the diffusion pump was isolated and the tube sealed. A second sample was prepared in a similar way. Both samples were examined at intervals, over a period of a year, in an ESR spectrometer operated at maximum gain.

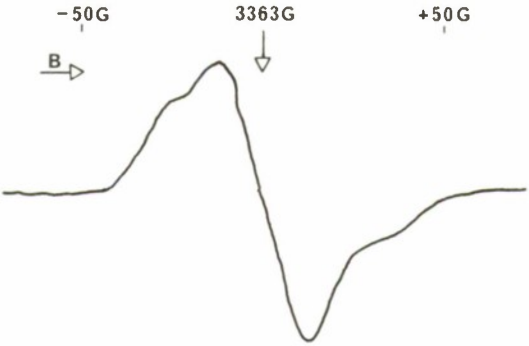


FIG. 1. Typical Spectrum from DGEBA cured with DDM.

DGEBA Cured With BF_3 Monoethylamine (BF_3MEA)

A mix was prepared by heating 100 p.b.w. DGEBA to about 80°C and stirring in 5 p.b.w. BF_3 MEA ($F_3B : NH_2C_2H_5$) until the mix appeared homogeneous. Rod samples were cast and cured for 4 hours at 105°C, followed by 4 hours at 180°C. The rods were examined in an ESR spectrometer immediately after the above treatment and then at intervals until the radical population reached a steady state.

Oxidation of a Solution of DDM In γ - Butyrolactone

A solution of DDM in γ - butyrolactone was oxidised by the addition of about 20% by volume, 100 vol H_2O_2, and spectra recorded at RT (room temperature) and 77K, respectively.

A similar solution was deuterated by warming with D_2O and D_2SO_4. An oxidising solution was prepared by dropping (carefully) small quantities of KO_2 into D_2O standing in ice. About 20%, by volume, of the oxidant was added to the deuterated solution, and a spectrum recorded at RT.

Cures Using Non-Stoichiometric Concentrations of DDM

Seven sets of rod samples were cast; each set consisted of 5 samples which had, respectively, 30, 40, 50, 60, 70 p.b.w. DDM in solution. Three sets were cured in air at RT, and the other 4 sets were cured by heating for 4·5 hours in air at 105°C. Spectra were recorded at RT, when all samples had reached the steady state.

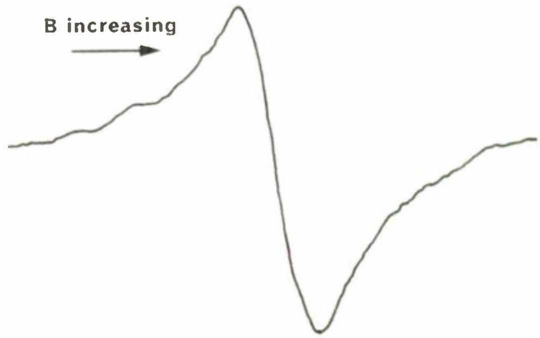


FIG. 2. DGEBA cured with BF_3MEA.

All the samples previously described were heated in air at 125°C for 1 hour, restored to RT and spectra recorded.

Two samples, one containing 30 p.b.w. and the other 50 p.b.w. DDM in solution, were placed in spectrosil tubes which were then connected in turn to a high vacuum line. Pressure was reduced to 10^{-6} mm Hg, the pump isolated, the tubes sealed and spectra recorded. Both tubes were maintained at 125°C for 1 hour, restored to RT and spectra recorded.

Results

Sample Cured and Examined in Vacuo

No radical was present in a detectable concentration.

DGEBA Cured with BF_3 MEA

A spectrum of the form shown in Fig. 2 was recorded. Cross-over occurs at a value of B corresponding to $g = 2{\cdot}0045 \pm 0{\cdot}0004$ and the linewidth, ΔB_{pp}, is $(13{\cdot}0 \pm 0{\cdot}1) \times 10^{-4}$ T.

Oxidation of a Solution of DDM in γ - Butyrolactone

At RT a many line spectrum, extending over rather more than 47×10^{-4}T, was recorded. Analysis shows the radical responsible to have the form $R\ CH_2ArNH\dot{O}$. At 77K the spectrum changed its form to that shown in Fig. 3 in which the hyperfine structure, used to identify the radical in solution, is quite unresolved.

The oxidised, deuterated solution gave a many line spectrum which extended over 35×10^{-4}T. Analysis shows that this spectrum may be assigned to a radical of the form $R\ CH_2Ar\ ND\dot{O}$.

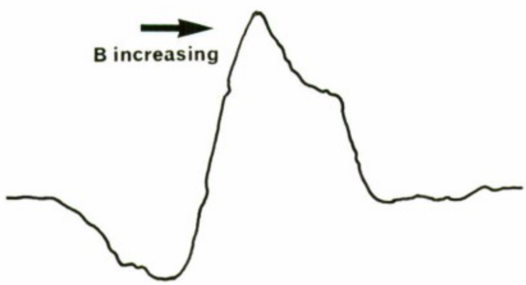


FIG. 3. Spectrum from DDM dissolved in γ-butyrolactone and oxidized with H_2O_2 recorded at 77K.

Cures Using Non-Stoichiometric Concentrations of DDM

A typical set of spectra is shown in Fig. 4. Examination of the spectra in all such sets reveals nothing to suggest that the two different cure cycles affected the defining parameters of the spectral lines in any systematic way. Also, the displacement of the cross-over point in the low-field direction, and the gradual change of line-shape, from that of Fig. 2 to that of Fig. 3, with increasing concentration of DDM are repeated within each set.

In striking contrast to the results described above, the spectra from *all samples* after heating are much alike. The cross-over points and line-shapes are close to those shown in Fig. 3. Also, there are no great differences between the radical populations in the different samples, and what differences there are do not seem to vary in a systematic way with the concentration of DDM.

The spectra recorded before heating in vacuo were typical of the many observed from samples having those concentrations of DDM. After heating the radical populations had decreased by two orders of magnitude and both lines were broad and featureless.

FIG. 4. Spectra from DGEBA cured with 70, 60, 50, 40, 30 p.b.w. DDM in solution, respectively.

Discussion

The mechanism of the curing reaction between DGEBA and primary amines is fairly well understood, and it may be stated with confidence that free radicals play no part. That is, the radicals responsible for the spectrum of Fig. 1 are not formed as the result of reactions between DGEBA and DDM*. Moreover, the spectrum is not changed if the sample is made from newly prepared, relatively pure samples of DGEBA and DDM in solution, neither of which contains free radicals in detectable concentrations. The radical population reaches a detectable concentration about two days after mixing and reaches a steady state of typically (1 to 4) $\times 10^{16}$ spins in a 25 mm length of 3 mm diameter rod, in about one month. That steady state situation remains substantially unchanged whether the sample is cured at RT or at, say, 100°C.

It is interesting to note that, if an accelerator** is used in the curing reaction, the radical population reaches a detectable level sooner—the time to detection decreasing as the concentration of accelerator is increased from zero to the maximum recommended value of 10 p.b.w. Also, during the initial development of the spectrum, the ratio of the maximum amplitude to the amplitude at some value of field in the wings is significantly greater. The line shape in the steady state is, however, that of Fig. 1, but the steady state population is typically rather smaller than that attained in the absence of accelerator.

The absence of a detectable concentration of radicals, in the samples cured and examined in vacuo, demonstrates clearly that reaction with atmospheric oxygen is a necessary condition for establishing the radical population responsible for the spectrum of Fig. 1.

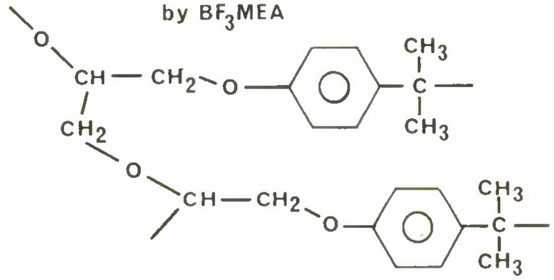
The use of BF_3 MEA to promote cross-linking allows the examination of a simpler network in which the molecules of DGEBA are linked to each other as shown at the top of page 217.

The results show that DGEBA, when cross-linked by an agent which simply promotes the formation of a network of identical molecules, develops a free radical population during cure.

* It is easily shown that the solvent plays no part in radical formation.

** Reactions between the accelerator and the other components do not lead to the formation of radicals.

DGEBA cross linked by BF_3MEA



This result is quite significant in that no such radical population develops in uncross-linked DGEBA exposed to the atmosphere—even if it is heated to the temperature used in this cure. In fact, a spectrum which very closely resembles that of Fig. 2 in cross-over and unresolved structure in the wings, but which has a somewhat narrower line, may be obtained by heating DGEBA and PbO_2 at 100°C and stirring vigorously. Taken together, these results imply that oxidation of DGEBA leads to a stable free radical population, but that it is necessary for bonds to be stressed if attack by atmospheric oxygen is to succeed. It is well-known[1], that rather large internal stresses are established during the cure of epoxide systems and, evidently, atmospheric oxygen is able to react with a molecule of DGEBA when interatomic potentials are modified by the inclusion of these stresses.

If, as seems very likely, the spectrum of Fig. 2 is a component of the spectrum of Fig. 1, the mechanism suggested in the previous chapter is consistent with the earlier appearance of radicals in an accelerated cure of DGEBA with DDM. Cross-linking and the associated internal stress occur earlier in an accelerated cure. However, the results show that internal stress alone does not lead to the formation of stable radicals.

Examination of the spectrum rules out a radical in which the unpaired spin is localised on a carbon atom with even one associated α– or β– proton. The necessity for stress implies a radical formed by scission of the chain backbone and the stability of the radical, together with the unresolved structure, imply delocalisation of the unpaired spin. These factors, taken together, suggest fission of the other bond to give a radical of the form R Ar$\dot{O}$. That is also consistent with the fact that ethers possessing hydrogen atoms in the position α to the oxygen atom are easily oxidised, and that benzyl ethers are particularly prone to attack, due to the additional resonance stabilisation[2]. A possible reaction sequence is shown below.

$$RCH_2OArR \xrightarrow{O_2} R\overset{\alpha}{C}H(OOH)OArR \longrightarrow RCH{=}O + \dot{O}ArR$$

Consideration of the parameters[3], of the spectrum, produced by the phenoxyl radical in solution, allows some estimate to be made of the spectrum to be expected from the postulated radical observed in a rigid, amorphous, hydrocarbon matrix. The observed spectrum, shown in Fig. 2, is quite consistent with its being due to the phenoxyl radical RAr$\dot{O}$.

It is proposed that, during cure of DGEBA, phenoxyl radicals are produced, are stable and are responsible for a broad singlet contribution to the spectrum of DGEBA cured with DDM, Fig. 1.

The result of oxidising selectively* deuterated DDM in solution at RT provides a useful check of the analysis of the spectrum from DDM, oxidised under the same conditions, in that the splitting of $12{\cdot}5 \times 10^{-4}$T, attributed a^{H}_{NH}, is absent.

It is hardly possible to make a worthwhile prediction of the line-shape to be expected from $RCH_2ArNH\dot{O}$, when observed in a rigid, amorphous, hydrocarbon matrix. On the other hand, it seems reasonable to suppose that the spectrum recorded at 77K, at which temperature motion is very hindered, reproduces very closely the line-shape of interest. Inspection of Fig. 3, coupled with measurement of the line parameters, show a quite striking resemblance to the spectrum of Fig. 1, and suggest that the former, or something like it, may well be a component of the latter.

It remains to consider whether any other radical likely to be formed by oxidation in the solid may have a spectrum similar to that of Fig. 3, or may have a spectrum which can co-exist undetected with that of Fig. 3. Clearly, radicals may result from oxidation of singly reacted amine groups. Loss of water, followed by oxidation, would result in the formation of a radical

* The hyperfine splitting shows that only the amino hydrogens were replaced.

$$ArN = CH\dot{C}HCH_2OR \longleftrightarrow Ar\dot{N}CH = CHCH_2OR$$

However, consideration of the likely hyperfine splitting shows that the spectrum from this radical is likely to be quite distinct from that due to $RCH_2ArNH\dot{O}$ and, in particular, to extend over a considerably greater range of magnetic field. Radicals may result, also, from oxidation of the CH_2 group, but this possibility may be ruled out at once. When cure is effected by use of 4,4′-diaminodiphenyl (benzidine) $H_2NArArNH_2$ in solution, and a similar cure cycle, oxidation results in a spectrum of the same line-shape as that of Fig. 3**.

It is proposed that radicals of the form $RCH_2ArNH\dot{O}$ are produced by oxidation of unreacted NH_2 groups in DGEBA cured with DDM, are stable and are responsible for a broad feature with characteristic shape in the spectrum of Fig. 1. A possible reaction sequence is shown below.

$$RCH_2ArNH_2 \xrightarrow{\dot{R}} RCH_2Ar\dot{N}H \xrightarrow{O_2} RCH_2ArNH\dot{O}_2$$

$$RCH_2ArNH\dot{O}_2 \xrightarrow{RH} RCH_2ArNOOH \longrightarrow RCH_2ArNH\dot{O}$$

The difficulties and at least some of the pitfalls encountered in assigning components to a complex spectral line, produced by radicals observed in a rigid, amorphous, hydrocarbon matrix, are well known. Nevertheless, it is reasonable to suggest that the spectrum of Fig. 1 may be synthesized by combination of the spectra of Figs. 2 and 3. Synthesis is aided by knowledge of cross-over points with an accuracy of $\pm$ 1 part in 10^4, and the availability of well-defined line shapes. However, the relative amplitudes of the proposed components can only be estimated. Use of a Hewlett Packard 9810A calculator and its associated graphical display unit 9862A make it easy to examine the effects of varying relative amplitudes and (very slightly) cross-over points. The conclusion is that the synthesis proposed can certainly be effected.

** The line shapes are identical but the population is slower to reach its steady state value in this case: presumably because the absence of the CH_2 group makes the amine groups, attached to the same rings, less reactive[4].

The *maximum permissible* amplitude of the spectrum of Fig. 1 is about 16% that of the composite line, which implies that phenoxyl radicals may account for up to 10% of the total radical population.

It is interesting to note that a very plausible alternative analysis can be made by increasing the amplitude of the component of Fig. 2 up to 90% that of Fig. 1. The difference may be made up from a symmetrical broad feature combined with the very characteristic spectrum of a peroxy radical having $g^* = 2{\cdot}022$ and $g^\dagger = 2{\cdot}014$. The presence of detectable concentrations of peroxy radicals in an epoxide matrix would have interesting implications. However, observing the spectrum from a spectrometer operating at Q-band allows this possibility to be ruled out at once—the difference in g-values between the peroxy and the other components would result in very considerable shift of components and a completely changed spectrum. In fact, the Q-band observations are consistent with the proposals made above.

It remains to see what more may be said about the formation and destruction of the phenoxyl and the nitroxide radicals, and it is here that the results from those cures in which non-stoichiometric concentrations of DDM were used are helpful. It seems reasonable to suppose that the concentration of unreacted NH_2 groups in the cured matrix may be varied by varying the amount of DDM used in the cure, and that such variation may change the relative concentrations of phenoxyl and nitroxide radicals. The spectra shown in Fig. 4 fully support that speculation: the changes in line-shape are obvious, and the movement of baseline cross-over points in the low field (high g) direction, in the sequence 30, 40, 50, 60, 70 (p.b.w. DDM solution contained in sample), also indicates an increasing concentration of nitroxide radicals, in that sequence.

Clearly, the shapes of the observed composite lines depend upon the relative populations of nitroxide and phenoxyl radicals present. Because of the very different line-shapes of the components due to those radicals it seems likely that amplitudes may be taken directly off the derivative curves and used to calculate, to a reasonable approximation, the relative populations. This analysis turns out

* g parallel.
† g perpendicular.

to be possible, with an error which does not exceed 3% in the worst case. The results are shown in Fig. 5: on both axes quantities have been normalised with respect to stoichiometric values. If it is assumed that the concentration of unreacted NH_2 groups is simply related to the concentration of DDM used, the nitroxide radical steady state concentration is seen to be related to the concentration of its substrate by a sigmoidal curve. That is, there is a rapid increase in the radical population as unreacted amine becomes available but, as more substrate becomes available, a level is reached after which it no longer controls the radical population. The decrease in the steady state population of the phenoxyl, which mirrors the increase in population of the nitroxide, is particularly interesting. The apparently well-correlated, opposite tendencies suggest that the phenoxyl may be instrumental in producing the nitroxide. That is, the phenoxyl either acts as, or leads directly to, the initiating radical $\dot{R}$ (postulated above).

Three results, taken together, throw further light on the reaction mechanisms which lead to the observed radical populations.

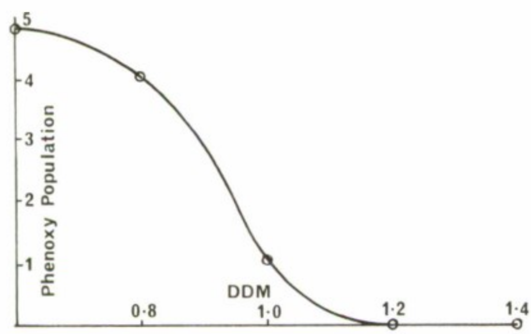
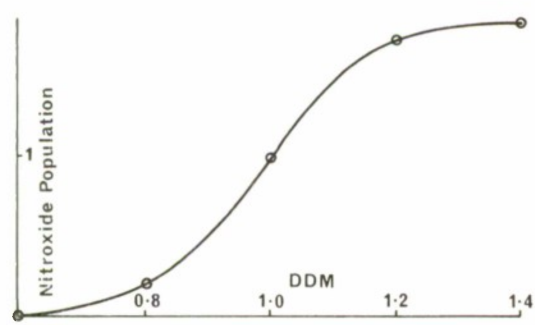


FIG. 5 (a) above, (b) below. Radical populations *versus* DDM concentrations, all scales normalised to stoichiometric concentrations.



(a) The above mentioned levelling off in nitroxide population with increase in DDM concentration.
(b) The changes, after heating, in the lineshapes of the spectra from the two samples containing less than the stoichiometric concentration of DDM.
(c) The approximate equality, and the reduction in values, of all radical populations after the samples had been heated.

These results are most easily explained by making the assumption that, in stoichiometric samples, diffusion of oxygen is the rate limiting process in radical formation. Then, for formation of nitroxide radicals, oxygen must be able to produce phenoxyl radicals in regions where unreacted NH_2 groups exist. The result in (a) is then due simply to diffusion becoming the controlling factor, once there exists a sufficient concentration of NH_2 groups. The result in (b) is due to the enhanced diffusion, at higher temperature, making oxygen available in NH_2 containing regions, which are virtually inaccessible at RT. The result in (c) is again due to availability of oxygen at suitable reaction sites, but it shows also that speeding up of destruction mechanisms is dominant at the higher temperature. That is, the increase in temperature is more effective in promoting destruction than diffusion. The only numerical values uncovered refer to polyethylene[5]. These give activation energies of destruction, *via* radical migration, and diffusion of oxygen as 28 k cals per mole and 12·5 k cals per mole, respectively. If these figures are taken as a guide to relative values in epoxides, destruction is enhanced relative to diffusion by a factor of the order 10^3, which offers modest support for the assumption.

The result of heating samples in vacuo simply demonstrates that radical destruction is, as expected, dominant in the absence of oxygen, and that the existing radical populations are not able to set up branching chain reactions.

The foregoing discussion may be considered to lead up to the proposal of a reaction scheme which includes all the reactions that play important parts in determining steady state radical populations. Clearly, a very large number of reactions is probably involved and it is not easy, *a priori*, to select those which are dominant. Nevertheless, it is considered that the points outlined below suggest a probable reaction sequence.

(a) Phenoxyl radicals are formed, *via* a hydroperoxide by direct reaction with atmospheric oxygen.
(b) Nitroxide radicals are formed by reaction of phenoxyl radicals with unreacted NH_2 groups.
(c) Concentrations of phenoxyl and nitroxide radicals reach steady states.
(d) Diffusion of oxygen is the rate determining process.
(e) Although very many less stable radical species may be produced, they can be present in only small concentrations. Although that by no means implies their lack of importance, it is clear that the effects of branching chain reactions and rapid autoxidation are not observed in this rather stable epoxide system. Chain termination may be assumed to dominate chain generation, so that unstable loops are quenched.

Since no rate-constant is known, the analytical possibilities are restricted to showing that the phenoxyl and nitroxide populations exhibit the observed functional dependence. Of the many reaction schemes examined, that which follows seems the most satisfactory. The cured network which forms a substrate for the formation of phenoxyl radicals is denoted Sp, and, the radical $\dot{P}$. The unreacted amine is denoted S_N, and the nitroxide radical $S_N\dot{O}$. All other radicals are denoted $\dot{R}$, and molecular products M.

Included with the assumption that only these reactions enter into the production of $\dot{P}$ and S_NO is, of course, the implicit assumption that all the radicals $\dot{R}$ are destroyed rather rapidly.

$$S_P + O_2 \xrightarrow{k_1'} S_POOH \xrightarrow{k_1''} \dot{P} + \dot{R} + M$$

$$\dot{P} + S_N \xrightarrow{k_2} \dot{S}_N$$

$$\dot{S}_N + O_2 \xrightarrow{k_3} S_N\dot{O}_2$$

$$S_N\dot{O}_2 + S_P \xrightarrow{k_4} S_NOOH + \dot{R}$$

$$S_NOOH \xrightarrow{k_5} S_N\dot{O} + \dot{R}$$

$$S_N\dot{O} + S_P \xrightarrow{k_6} M + \dot{R}$$

This scheme leads to

$$[\dot{P}] = \frac{k_1'[S_P][O_2]}{k_2[S_N]}$$

and

$$[S_N\dot{O}] = \frac{k_1'[O_2]}{k_6}$$

If $[S_N] \longrightarrow 0$, some reaction such as

$$\dot{P} + S_P \xrightarrow{k_7} \dot{R}$$

must be included,* to give

$$\dot{P} = \frac{k'_1[O_2]}{k_7}$$

Although this approximation represents, presumably, a point of view which is much over-simplified, the expressions for the radical populations have the observed dependence. It is interesting to reflect that aromatic amines, even when present in very small quantities, are known to be very effective inhibitors of autoxidation in polymers, and the formation by the inhibitor of stable nitroxide radicals, in oxidising hydrocarbon systems, seems to be a fairly general phenomenon. Thus, the formation of stable nitroxide radicals, accompanied by the decrease in phenoxyl radicals, observed in this epoxide system is consistent with the unreacted curing agents acting as an inhibitor.

* This reaction may be omitted in the presence of S_N, since $[\dot{P}]$ is then very much less than $[\dot{P}]$ in the absence of S_N.

References

[1] Lee and Neville. Handbook of Epoxy Resins; McGraw Hill (1967).
[2] Ingold, C. K. Structure and Mechanism in Organic Chemistry; 2nd Ed; Bell and Sons Ltd. (1969).
[3] Stone, T. J. and Walters, W. A. *Proc. Chem. Soc.;* 253 (1962).
[4] Brownlie, I. T. and Ingold, K. U. *Can. J. Chem.*, **45**, 2419 (1967).
[5] *Encyclopaedia of Polymer Science and Technology*, **5**, 683; *Interscience* (1966).

REINFORCEMENTS for PLASTICS intended for ELECTRICAL and ELECTRONIC APPLICATIONS

J. Cook, M.Sc.
Explosives Research and Development Establishment

This article deals with the properties of reinforcing fibres; chiefly those of low conductivity such as glass, asbestos and cellulosic fibres and of the reinforced plastics made using them. Usually the requirement is for some combination of electrical and mechanical properties; and representative values for these are given for both the fibres themselves and the resulting composites.

Some of the elementary theory of fibre reinforcement is presented to show the limits within which the arrangement of the fibres in a composite can be adapted to control its behaviour.

Introduction

The use of reinforced plastics in the electrical and electronics industries is widespread and is on the increase. Some typical applications are insulators for high voltage supply lines, printed circuit boards, radomes, and miscellaneous uses in generators and switchgear. There are in addition other important areas of use within these industries, but the feature that the applications listed have in common is that both the electrical and mechanical properties of the material are significant to some degree, and it is such materials that are our concern in this article. The properties of the resin matrix materials will be dealt with in this Journal,[1] while here we will be describing the various reinforcements that are available. Particulate fillers and flake-like reinforcements such as mica are used quite commonly but this article will be confined to a discussion of fibrous reinforcements only.

Usually the electrical properties sought are low conductivity and low dielectric loss over a particular frequency range and in addition a low value for the permitivity is often also desirable. Dielectric breakdown and resistance to tracking tend to be governed principally by the polymeric matrix and are consequently adjudged to come outside the scope of this article. However, these subjects, together with the resistive and dielectric properties of polymers in general are well covered in an excellent review by Baird.[2]

On the mechanical side, the strength, stiffness and dimensional stability of the reinforced plastic are obviously of general importance and the factors governing these will be covered in some detail. Other properties such as impact strength and resistance to erosion and weathering are extremely important in some applications (*e.g.* radomes) and virtually insignificant in others; so these will receive moderate but not undue attention.

For most purposes we are not concerned with composite behaviour at elevated temperatures, higher than 100°C or so, so the commonly quoted room temperature data should suffice. There are exceptions, but these come outside the scope of a general survey of this sort.

In the next section, the properties, origins or manufacture, availability and variability of the fibres are discussed at some length, and this is followed by a short section on surface treatments. Later the various ways in which fibres and matrix contribute to the composite behaviour are briefly touched upon, and in the final section data on the various reinforced plastics; unidirectional, random, partially random and cross plies; is presented together with qualifying remarks.

Fibre Properties

Table 1 lists some of the relevant properties for a number of commonly used types of fibre. One of the difficulties in compiling a table of this sort is that some properties, notably the tensile strength and the resistivity, may depend very strongly on factors that need to be specified in some detail. For example, it is well known that glass fibres drawn singly under clean conditions and tested within ten minutes or so of drawing will have considerably higher strengths than the same type of fibre that has been drawn as a strand, sized, wound on to a reel and stored for six months. Again, the resistivity of many fibres depend very much on extrinsic factors such as water content or surface treatment.

In the table the strength figures quoted will normally be those pertaining to the type of material that is commercially available, or in the case of properties dominated by water absorption the likely range of values is given.

Having adopted this philosophy one can be reasonably confident about comparisons between fibres within the same class, *i.e.* between one type of glass fibre and another, or one type of asbestos and another; but when it comes to class to class comparisons there is no truly objective way of giving accurate comparisons between fibres whose nature and origins differ so widely.

The next task is to examine types of reinforcing fibre in more detail.

Glass Fibres

Easily the most important type of glass fibre used for reinforcing purposes is 'E' glass fibre. 'E' glass is a low alkali composition that was originally developed as an insulating coating for fine metal wires. It replaced the earlier 'A' glass which had a high alkali content and a relatively low resistivity due to the mobility of the sodium ions. 'E' glass was found to have superior mechanical properties to 'A' glass as well as being better electrically, and it is also relatively simple to produce in fibre form.

The manufacture of glass fibre products is described in detail by Loewenstein.[3] Briefly, the method is to extrude molten glass through a bushing containing a number of holes, often 200 or a multiple thereof. The emergent ends of glass are drawn, sized, bunched into a strand and wound on to a removable sleeve on a high speed winding head (Fig. 1). When a convenient quantity of material has been wound the winding head is stopped and the package removed from the machine. Such a package is known as a cake.

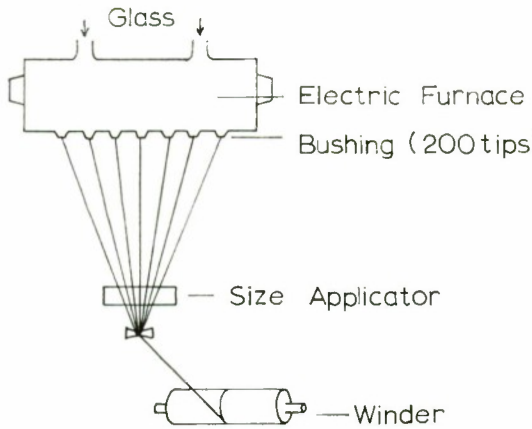


FIG. 1. A typical fibre forming arrangement.

The range of fibre diameters commonly produced is roughly from 5 to 16 micrometers. The manufacturers use a labelling system whereby letters of the alphabet denote the fibre diameter, *e.g.* G refers to fibres of nominal thickness $9{\cdot}53 \pm 0{\cdot}63$ μm. This is not to be confused with the notation A, E, D, S etc. that describes the composition; in this article only the lettering denoting composition will be used. The number of fibres in a strand is often, but not necessarily, the same as the number of fibres emerging from a single bushing.

Starting from a fibre cake, a number of glass fibre products can be made. For example:—

(i) Chopped Strand Mat: a non-woven fabric made from strands chopped to lengths usually between 20 and 50 mm and held together by a binder.

(ii) Rovings: bundles of strands (*e.g.* 30 strands) wound on to a reel.

(iii) Roving Cloth: produced by weaving rovings.

(iv) Yarns: twisted strands. A doubled yarn is a pair of yarns twisted in opposite senses.

(v) Glass cloth woven from yarns: This is not very common, but is used for some of the higher quality printed circuit boards.

TABLE 1.

Fibre Properties.

	UTS *GNm^{-2}*	*E* *GNm^{-2}*	*SG*	*Thermal Expansion Coefficient (deg. C^{-1})*	*Resistivity ohm cm*	*Permittivity*	*Loss Factor (Tan δ)*
Glasses							
E	2·5	76	2·56	$4{\cdot}9 \times 10^{-6}$	10^{15}	6·33 (1 MHz) 6·11 (10 GHz)	0·001 (1 MHz) 0·006 (10 GHz)
D	2·4	52	2·16	–	–	3·56 (1 MHz) 4·0 (10 GHz)	0·0005 (1 MHz) 0·0026 (10 GHz)
S	3·2	88	2·45	5×10^{-6}	–	5·2 (1 GHz)	0·0068 (10 GHz)
Asbestos					*Resistivity megohm cm*	Approximate Composition	
Chrysotile	1·3	160	2·55		0·003 - 0·15	3 MgO, 2 SiO_2, 2 H_2O	
Amosite	1·0	160	3·43		up to 500	5·5 FeO, 1·5 MgO, 8 SiO_2, H_2O	
Crocidolite	1·5	190	3·37		0·2 - 0·5	Na_2O, Fe_2O_3, 3 FeO, 8 SiO_2, H_2O	
Anthophyllite	1·0	150	2·85 - 3·1		2·5 - 7·5	7 MgO, 8 SiO_2, H_2O	
Little data is available on permittivities, but the author has measured values in the range 5 - 10.							
Cellulosic Fibres					Resistivities of cellulosic fibres vary within the range 10^4 - 10^{18} ohm cm depending on the relative humidity. Their permittivities are close to 5 at low frequencies.		
Flax	0·42 - 1·1	86 - 110	~1·5				
Hemp (Dry)	0·9	57	~1·5				
Ramie (Dry)	0·76	51	~1·5				
Wood Fibre	0·9	72	~1·5				
Other Fibres							
Aromatic Fibres	2·7	130	1·45	0 to -6×10^{-6}	high		
Carbon Fibres (Type II)	2·8	240	1·7	-1×10^{-6}	10^{-3}		
Steel Wire	1 - 4	210	7·8	$1{\cdot}2 \times 10^{-5}$	$1{\cdot}5 - 8 \times 10^{-5}$		

Permittivities and loss tangents do not seem to be available for the aromatic fibres in isolation, but values of about 3·2 and 0·005 respectively (at 10 GHz) can be inferred from data on composites.

The prices of some of the more sophisticated glass fibre products, particularly the cloths, may be an order of magnitude higher than the basic fibre; an indication of the lower useage and the involved technology.

Loewenstein[3] gives the compositions and properties of a number of common glasses. For the present purpose, apart from 'E' glass, 'S' and 'D' glasses[4] are likely to be of most interest. 'S' glass is a high strength glass developed by Owens-Corning in the early 1960s for such applications as rocket motor cases. It contains a high percentage of alumina and magnesia and no borides or alkalis. 'S' glass fibre appears to be more susceptible to surface damage than 'E' glass fibres so that a precise comparison of strengths is difficult. Were we to compare the freshly drawn fibres, the superior strength of 'S' glass would be more pronounced: figures as high as 5 - 6 GNm^{-2} have been quoted for it.

'D' glass is a composition developed specifically for radomes. It has a dielectric constant and loss tangent at X-band frequencies (~ 10 GHz) significantly lower than 'E' glass, but also has a lower strength and substantially lower Young's modulus.

The total production of both 'D' and 'S' glasses is very small compared with that of 'E' glass. A comparative newcomer is 'R' glass which appears to have properties intermediate between those of the 'S' and 'E' types. In addition, silica fibres[4] which have a lower dielectric constant than 'D' glass and a remarkably low loss tangent of 0·0002 at 10 GHz, may be a possibility for some specialised applications. The difficulty with these appears to be that they are difficult to produce and are extremely sensitive to surface damage.

Asbestos Fibres

Asbestos is a general name for naturally occurring silicate minerals that have a very fibrous form. As opposed to glass fibres asbestos minerals are crystalline, and this fact has a pronounced effect on their physical properties. The structure, the physical and chemical properties as well as the geological and mining aspects of asbestos are described in an excellent booklet by Hodgson.[5]

Six species of asbestos minerals are known, of which one, chrysotile, forms a group on its own, while the other five are crystallographically similar to each other and are known as the amphiboles.

Chrysotile, which is familiar white asbestos, is the most common type, and constitutes about 94% of the total annual world production of four million tons. Chrysotile deposits occur fairly widely throughout the world and the largest are in Canada and the U.S.S.R. Chrysotile is a hydrated magnesium silicate whose structure is such that the silicate sheets curl into fibrils resembling miniature swiss rolls some 300 Å or so in diameter. Because neighbouring "swiss rolls" are relatively weakly bonded to each other, most chrysotiles are able to subdivide readily, and this property gives them a soft texture.

Asbestos minerals of the amphibole group, which includes amosite, crocidolite, anthophylite, tremolite and actinolite, have two sets of cleavage planes and thus break down into needle-like fibres. In general this subdivision occurs less readily than with chrysolite and so the fibres have a harsher spikier appearance. Amphibole fibres have greater resistance to chemical attack than chrysotile and are used widely for filters in chemical plant.

Most types of asbestos occur as narrow seams of thickness from 1 to 20 mm in rocks having a chemical composition very close to that of the fibres. The fibres occur tightly packed in parallel formation, often running across the seam, in which case the fibre length is the same as the seam width. Fibres carefully removed by hand from a rock sample usually have strengths of the order of $3GNm^{-2}$ [6] but the methods used to extract the fibres from the rock in quantity, which involve crushing, hammer milling and screening the material, inevitably lead to a degradation of the fibre strength. The fibres used in asbestos cement products probably have an average strength of no more than 0·2 GNm^{-2} and are mixed with quite a lot of particulate matter. For reinforced plastics something between these extremes is needed, and the table gives some idea of what is typical of some of the higher grades of asbestos that are available. Well under 1% of the total asbestos production finds its way into reinforced plastics, but the proportion seems to be increasing.

A range of methods for classifying short fibres has been developed at E.R.D.E. over the last few years. Briefly, they involve sorting the fibres for length by passing a suspension of the fibres in water through rotating screens, with automatic removal of the retained fibre, and sorting for thickness using hydrocyclones. Both of these two basic techniques are useful for removing particles from the fibres, and the application of a suitable combination of them enables the average fibre strength of a high grade asbestos to be improved, or a good quality fraction to be extracted from low grade material.

The risk to health involved with inhaling asbestos dust has received a good deal of attention recently. Normal asbestosis can result from breathing in quantities of any type of asbestos, but there have, in addition, been a number of cases of lung cancer, some of them in people who have undergone relatively light exposure. These carcinogenic effects have been associated with blue asbestos (crocidolite) and the use of this material is now largely avoided.[7] In any event, the development of wet handling methods should help to remove these problems, at least at the fibre processing end of the business.

With regard to electrical properties, there is little information available on asbestos fibres or on rock samples. If the fibres are dry one can expect a dielectric constant a little higher than most glasses and a resistivity that varies considerably between one type of fibre and another, but even for amosite, which seems to be best in this respect, the value is rather lower than for the least suitable glasses.

However, more serious than this is the tendency of asbestos fibres to absorb water, and this leads to a further deterioration of the electrical properties.

To summarise, asbestos fibres are a cheap reinforcement with lower strength than glass, but more than twice the stiffness. Their electrical properties are not suitable for demanding applications, especially in humid conditions, but they find use in thin cloths and papers used as insulating separators or sleeves.[8]

Cellulosic Fibres

Cellulosie is the reinforcing fibre found in all kinds of vegetation, although the most useful sources are wood and various grass-like plants.

Glucose molecules, which are formed in the green parts of a plant by photosynthesis are transported in aqueous solution along the internal passages to the growing areas. Here they link by means of a condensation reaction to form long polymeric chains running roughly along the axis of the fibre. Cellulose fibres invariably take the form of hollow tubes a few micrometres in diameter with the walls consisting largely of this polymeric material. X-ray diffraction reveals that a substantial proportion of the cellulose is crystalline with the crystallites or fibrils elongated in one direction and embedded in an amorphous matrix of disordered chains. The direction in which the fibrils are oriented is usually not parallel to the fibre axis, but at some small angle, so that the fibrils run along helical paths within the wall of the fibre tube. All this is described by Gordon.[9]

The crystalline cellulose is comparatively stiff and strong in the axial direction[7] and is dimensionally stable in the presence of moisture, but the amorphous material is not very stiff or strong and will absorb anything up to 50% of its weight of water and will swell accordingly. For example, wood fibres can take up 25% of their weight in humid conditions and this results in an increase in thickness in the region of 2 - 9% but an increase in length constrained by the long crystalline fibrils to perhaps 0·1%.

This absorption of water causes loss of mechanical properties, accelerates creep and lowers the resistivity by orders of magnitude. According to Murphy[10] the resistivity changes by a factor of 10^{12} between 1% and 99% R.H. The low frequency dielectric constant of water is about 80, and the permittivity of substances containing it are usually increased roughly in proportion to the amount present[11]. The dielectric loss is extremely high at frequencies below 10 KHz due to rotation of the water molecules, while at higher frequencies small currents may be induced in the short conducting paths along the fibres leading to dielectric loss over a broad frequency band.

The tensile strength of natural cellulosic fibres has been measured by a number of authors, *e.g.* El Hosseiny and Page[12] the upper limit being in the region of 1·7 GNm^{-2}.

Other Fibres

The high modulus aromatic fibres that emerged around 1968, namely Kevlar (formerly PRD-49) produced by Du Pont and the PABH-T series produced by Monsanto, combine high strength with a reasonable modulus and low density. A range of fibres exists, differing chiefly in their Young's moduli but having similar strengths. The fibres are abrasion resistant and can be handled and woven readily with very little degradation. They are highly anisotropic and can be split relatively easily under transverse or shear loading without being broken. This property coupled with viscoelastic properties under high loads, means that the fibres and the composites made from them give high mechanical damping and are also relatively tough, being good under impact loading, the lower modulus versions of the fibre being better in the latter respect.

One drawback of these fibres is their relatively low strength under prolonged loading. The worst figures quoted have given a life of only a few hundred hours in a Kevlar structure loaded to 50% of its short-term U.T.S., but there is some disagreement between the results from various sources.

A further difficulty is that the fibre compressive strength is poor, being only about 20% of the tensile strength. This is presumed to be due to the fibrils comprising a fibre splaying out in a buckling type of compression failure. For structural applications hybrid composites containing Kevlar together with boron or carbon fibres to bolster the compressive strength are being considered.

The low frequency dielectric loss of the aromatic fibres appears to be higher than that of glass. This is because they contain dipoles which are partially free to rotate in the electric field. At high frequencies the dipoles cannot move fast enough to respond to the field and the loss tangent becomes similar to that of 'E' glass. The dielectric constant is lower than that of glasses at all frequencies. The resistivity seems to be at least as high.

The Kevlar range of aromatic fibres have water regain figures between 1·5% and 4%. In view of this the good electrical properties are a little surprising, but much must depend on the way in which the water molecules become attached to the fibres. The properties of the Kevlar fibres are given in detail in a series of short reports issued by Du Pont and in some articles by Stratton, *e.g.* [14]. The Monsanto fibres are dealt with in a book by Black and Preston[13] which covers similar ground and also has lengthy sections on the chemistry of the fibre materials.

The properties of carbon fibre and steel wire are included in Table 1 for purposes of comparison. Obviously their conductivity is such as to preclude their use except in applications in which conducting properties are a positive asset, or irrelevant, as perhaps in supporting structures.

Fibre Surface Treatments

The subject of surface coatings is an important ancillary to the production of glass fibres, although is not of great significance for most of the other fibres considered.

As has been mentioned it is usual to apply a size to glass fibres immediately after drawing. The size serves many purposes[3] of which the most important are that it protects the fibre from damage during subsequent processing, keeps the strand intact, for example during chopping, and it may be given conducting additives to avoid the electrostatic charge build-up that can cause problems with the higher resistivity glass fibres. The use of anti-static additives obviously conflicts with the requirement for low conductivity in the reinforced plastic, and they would not be used if electrical properties were of prime importance.

From the point of view of composite properties the use of coupling agents to improve the bonding between fibre and resin is of some interest. The surface of a glass fibre is normally fairly well covered with hydroxyl groups, either absorbed from the air or derived from the size. Coupling agents, which are usually organo-silicon compounds, are substances capable of forming strong bonds with these and also with reactive groups in the resin. For example, polymerisation of polyesters involves breaking and crosslinking $-CH=CH-$ groups, so a good coupling agent for a polyester would ideally attach to hydroxyls on the glass leaving a $CH=CH_2$ group projecting that can readily link into the polyester. There is some dispute over the exact mechanism by which coupling agents operate; for example one might expect epoxy resins to attach readily to hydroxyl groups just as well as a coupling agent which would therefore be unnecessary. In fact the use of suitable coupling agents can significantly improve the glass-epoxy bond, and being especially effective in delaying the deterioration after prolonged soaking in water, so the position is not entirely clear. One explanation put forward for this is that the coupling agent forms a coating which is more than a molecular monolayer thick, but more like 10 layers. Consequently there is a definite intermediate zone around the fibre with distinct properties that are beneficial to the overall properties. There are a number of coupling agents in existence, most of which are quite specific to the type of resin and to a lesser extent to the type of glass.

Most of these observations should apply in a general way to asbestos fibres which have similar hydrated surfaces to those of glass. For asbestos, however, the bonding problems are less severe because the rougher surface texture gives a good mechanical keying.

TABLE 2.

Fibre-Epoxy Interlaminar Shear Strengths (MNm^{-2}) measured on undirectional bars ($v_f = 60\%$)

	Before 3 day water boil	*After* 3 day water boil
'E' glass (treated with coupling agent)	63	38
Kevlar III (untreated)	62	55
Carbon type II (treated)	66	37

Cellulosic fibres likewise have large numbers of surface hydroxyls available and there is usually no difficulty in obtaining a strong bond.

It does not appear to be common practice to use coupling agents in conjunction with aromatic fibres, but it is recommended that for a good bond the fibre should be dried prior to incorporation in the matrix. The fibres themselves, being extremely anisotropic, have a relatively low transverse strength, so that fibre splitting under transverse load will occur if the bond is good. This sets an upper limit to the degree of adhesion that is worth aiming for.

An oxidation treatment for carbon fibres is commonly used to give improved bonding to the matrix, but the effect of this appears to be nothing more than a physical roughening of the surface giving an improved mechanical key.

Tables 2 and 3 have been included in order to give some idea of the relative ease with which the various fibres can be bonded to a resin, and also of the effectiveness of coupling agents. Table 2 (taken from du Pont literature) indicates that Kevlar fibres form about as good an interfacial bond with epoxies as do treated glass or carbon fibres, but this conclusion has been disputed. The general opinion seems to be that while some epoxies bond better to Kevlar than others, the Kevlar-epoxy bond is inferior to that between a properly treated glass and an epoxy.

Simple Theory of Fibre Reinforcement

It is usual, when developing the theory of fibre reinforcement, to start by analysing the unidirectional composite (Fig. 3). The properties of cross-plied laminates or composites containing woven fabrics can then be calculated with fair accuracy by combining two or more basic unidirectional laminae.

To deduce the complete set of elastic and electric properties of unidirectional continuous fibre composite from a knowledge of the properties of the two phases is a difficult task. Some of the composite properties can be calculated exactly, but others cannot be written down explicitly; and if the fibres are randomly distributed in the transverse plane one has to be content with finding upper and lower bounds within which the true value must lie. Such problems are dealt with in great detail in a book by Hashin[15]. However, for the present purpose, it is sufficient to broadly describe the nature of these results and this can be conveniently done by grouping the composite properties according to the way in which fibre and matrix contribute to them.

TABLE 3.

'E' Glass-Epoxy Flexural Strengths (MNm^{-2}) measured on glass cloth laminates.

Manufacture and Distribution		*Chemical Name*	*Before* 3 day water boil	*After* 3 day water boil
		None	550	204
Du Pont	Volan A	Methacrylato chromic chloride	589	408
Union Carbide	A 1100	γ-aminopropyltrimethoxysilane	635	380
Union Carbide	A 186	β-(3,4-epoxycyclohexyl) ethyltrimethoxysilane	712	465

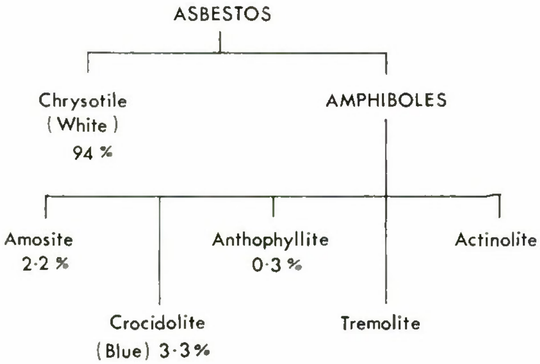


FIG. 2. Asbestos family tree with percentages of total amount mined (1964 figures).

Longitudinal Properties

The electrical conductivity and permittivity measured in the longitudinal direction are given exactly by the mixture rule:

$$\sigma_c = v_f \sigma_f + v_m \sigma_m \qquad \text{CONDUCTIVITY}$$

$$\varepsilon_c = v_f \varepsilon_f + v_m \varepsilon_m \qquad \text{PERMITTIVITY}$$

Where v_f and v_m are the respective fibre and matrix volume fractions ($v_f + v_m = 1$) and the subscripts c, f and m refer to eomposite, fibre and matrix.

Similarly we have for the Young's Modulus:

$$E_c = v_f E_f + v_m E_m \qquad \text{YOUNG'S MODULUS}$$

an expression which is exact if the Poisson's ratios of the two phases are identieal and is not greatly in error if they are not.

The longitudinal tensile strength of a composite cannot normally be predicted with any accuraey; factors which may influence it inelude the relative strength of the two phases and their ductility, the nature of the interfacial bond, and the presence of notches and defects[7, 10]. Nevertheless for reinforced plastics it is a reasonable working rule to use the mixtures relationship for strength also. Since in these materials the fibre strength and stiffness normally greatly exceed the matrix strength and stiffness, this means that the longitudinal meehanical properties are almost entirely governed by the fibres.

FIG. 3. Ideal unidirectional composite.

The compressive strength of reinforced plastics measured in the fibre direetion is less predictable. Usually it is lower than the tensile strength, particularly if the fibres are thin or not very straight. In relation to the tensile strength, the comprehensive strength of GRP is usually between a half and three quarters, and of cellulosic materials sometimes higher and sometimes lower. For plastics reinforeed with aromatic fibres, however, it is very low, a fifth to a quarter of the tensile strength, due to the low compressive strength of the fibre itself.

The longitudinal thermal expansion coefficient, α, is given by a modified mixtures rule in which allowance is made for the fact that the phase with the higher modulus tends to impose its expansion coefficient upon the composite:

$$\alpha_c = \frac{E_f v_f \alpha_f + E_m v_m \alpha_m}{E_f v_f + E_m v_m} \qquad \text{LINEAR THERMAL EXPANSION}$$

Normally we are not concerned with thermal expansion except in that interfacial stresses will be set up during fabrieation if the fibre and matrix expansion coefficients do not match and that warping can take place if irregular cooling occurs after curing. However, length ehanges due to other causes, such as water pickup, will also be governed by an equation similar to the above, that is with the change in dimension weighted towards that of the stiffer phase.

Transverse Properties

Here the situation is a little more complieated because the transverse properties are not described by single analytic expressions. Broadly, however, they are governed by the properties of the continuous phase, that is the matrix.

Fig. 4 shows the transverse Young's Modulus plotted as a function of fibre volume fraction for an E glass—epoxy composite. Upper and lower bounds are given together with a number of experimental points.

The transverse conductivity and permittivity are similar functions of the fibre volume fraction although there are differences in detail.

The transverse strength of most composites is lower than that of the matrix; often it is about half, but the nature of the fibre matrix interface is partieularly important here.

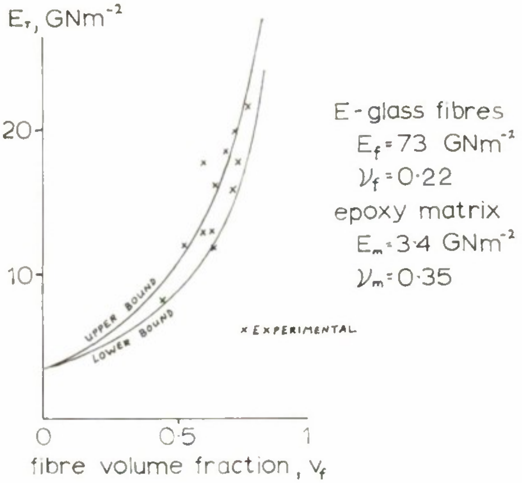


FIG. 4. Transverse Young's modulus for glass fibre reinforced epoxy.

The transverse expansion coefficient of a fibrous composite is close to that of the matrix if this is the stiffer phase, but if the fibres are stiffer, as is normally the case with reinforced plastics then the ordinary mixture law

$$\alpha_c = v_f \alpha_f + v_m \alpha_m$$ TRANSVERSE THERMAL EXPANSION

gives a rough idea of what to expect.

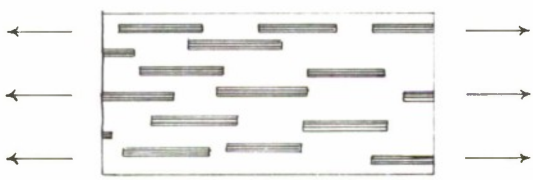


FIG. 5. Loading of a short fibre composite.

Other Properties

Composite shear strengths and stiffnesses are governed by much the same considerations as are the corresponding transverse properties. The hardness of a reinforced plastic is governed largely by that of the matrix, and likewise the resistance to erosion, although for splitable fibres such as asbestos and Kevlar there may be some benefit to be gained from splayed out ends of fibre covering the surface and shielding the soft matrix underneath. A strong interfacial bond also aids the erosion resistance by preventing protruding fibres from being easily dragged away.

It is difficult to make general predictions about dielectric loss, there are too many contributory effects each of which is frequency dependent. If both phases are good insulators then the loss in the composite will be the mean value weighted to take account of volume fractions and the tendency of the phase with the greater permittivity to attract the greater share of the electric flux. It is possible to have losses due to conduction in one of the phases or at the interface. If this is due to the presence of water then the frequency of the resulting loss peak will be strongly independent on the geometries of the conducting paths and on their resistance, hence its size and position will be extremely uncertain. The physics of the dielectric properties of solids is expounded in a book by Böttcher[17].

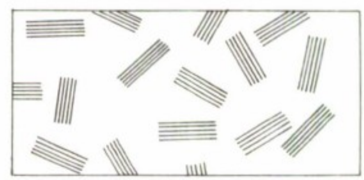


FIG. 6. Composite containing randomly oriented chopped fibre bundles.

Extra Considerations Resulting from the Use of Short Fibres

It is quite common to make reinforced plastics using discontinuous fibres, either through necessity, because asbestos and cellulosic fibres are only available in this form, or through convenience, as would be the case with chopped glass fibres. There are several differences with respect to continuous fibre composites.

Firstly the state of alignment may vary, so that all states of alignment from complete angular randomness in three dimensions to perfect unidirectional alignment are possible. In papers and felts the fibres may be confined more or less to a single plane and these constitute an important class of material[18, 19]. Unless the alignment is close to unidirectional locally or the fibres are very short indeed it is impossible to attain a large fibre volume fraction.

Composites containing fibres whose alignment is not perfect naturally have Young's Modulus, permittivites, etc. measured in the longitudinal direction that are somewhat lower than for a perfectly aligned material with the same volume fraction. Reference[20] gives some idea of the magnitude of the effect in the elastic case, and reference [7] deals with the amount of strength lost.

Another effect associated with short fibres is the relationship between the composite longitudinal properties and the fibre aspect ratio (length/diameter). Taking the Young's Modulus as an example it is evident that if we have an array of discontinuous stiff fibres in a less stiff matrix (Fig. 5) where the fibres are intended to take most of the load, there is some difficulty in achieving effective transfer of stress from fibre to fibre.

The stiffer the fibres in relation to the matrix the larger their length to diameter ratio must be if the Young's modulus is to approach that of a continuous composite with the same fibre volume fraction[18].

A similar situation arises with other properties such as permittivity and conductivity; the length of fibre required for satisfactory transfer of flux, current etc. increasing with the ratio of the property (*i.e.* permittivity, conductivity) between the two phases.

In calculating the strength the relevant parameter is the fibre tensile strength relative to the matrix shear strength or the interfacial shear strength, whichever is the lower.

In reinforced plastics one can say that for most fibres an aspect ratio of a few hundred, or length of the order of a few millimetres (for 10 μm diameter fibres), is sufficient to make a composite having a strength very close to that obtainable with a corresponding continuous composite; while a satisfactory Young's modulus is attainable using rather shorter lengths.

It is perhaps worth pointing out that these remarks are only valid in modified form for composites containing chopped bundles of fibres (Fig. 6). In this situation it is clearly possible to have a high packing fraction of high aspect ratio fibres arranged in what is macroscopically a directionally random fashion.

However, one finds that it is difficult to transfer load (or flux, current etc.) into the central fibres of a bundle, so the arrangement is not so efficient as might appear at first sight. As far as the composite properties are concerned the aspect ratio and other properties of the bundle become more important than those of the individual fibres themselves; for example, when the composite fails in tension, the mode of failure will almost certainly be the pulling out of bundles rather than the breaking of fibres. Since the bundle aspect ratio is usually small, this means that the strength and stiffness of these materials are generally not very impressive.

Laminated Composites

Usually the purpose of producing a laminated composite is to obtain adequate mechanical properties in more than one direction within the plane of a sheet. This is necessary because the transverse and shear strengths and stiffnesses of a unidirectional composite are usually insufficient when it is to be used in the form of a relatively thin sheet or shell structure.

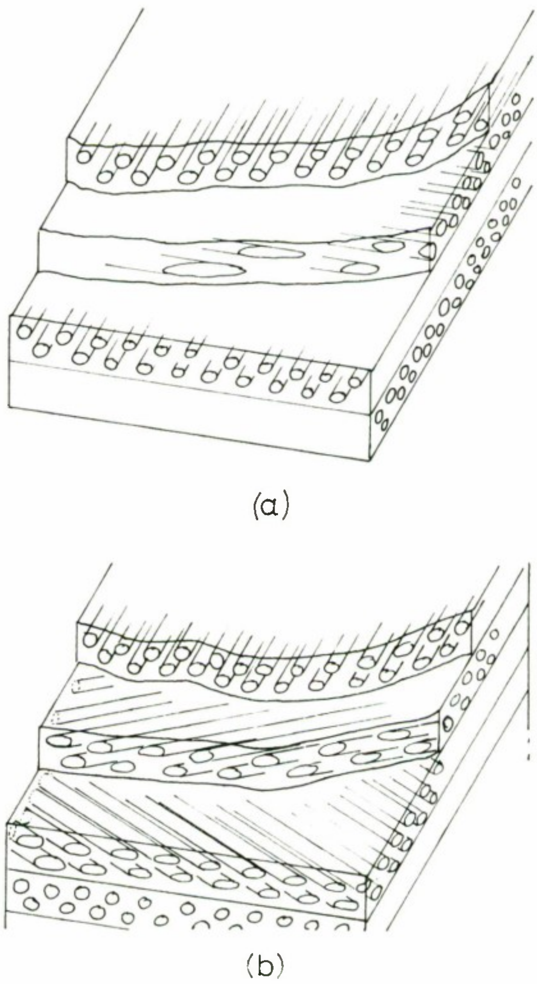


FIG. 7. Schematic diagrams showing the stacking sequence in (a) a simple crossply. (b) a 0°, ± 60° laminate.

A simple cross plied material (Fig. 7a) or one made from woven fibre mats, which behave similarly, is often satisfactory for general purposes and a good deal of laminated sheet is marketed in this form. However a (0°, 90°) crossply has poor shear stiffness in the plane of the sheet and a relatively low Young's modulus and tensile strength in the ± 45° directions, and, to bolster these up, fibres running in three or more directions are needed. For example a balanced (0, ± 60°) laminate (Fig. 7b) has the same Young's modulus in every direction in the plane and very nearly the same tensile strength. A composite made from random in-the-plane fibre felt is similarly isotropic in the plane of the sheet, but has much lower mechanical properties due to the lower fibre volume fractions possible.

The method of calculating the properties of these laminated structures is well established and straightforward, although tedious[21, 22]. Fig. 8 shows roughly the type of directional properties to be expected from glass-epoxy laminates, but to make accurate predictions or to deal with asymmetric or unbalanced fibre layups there is no substitute for the full sequence of calculations described in the references.

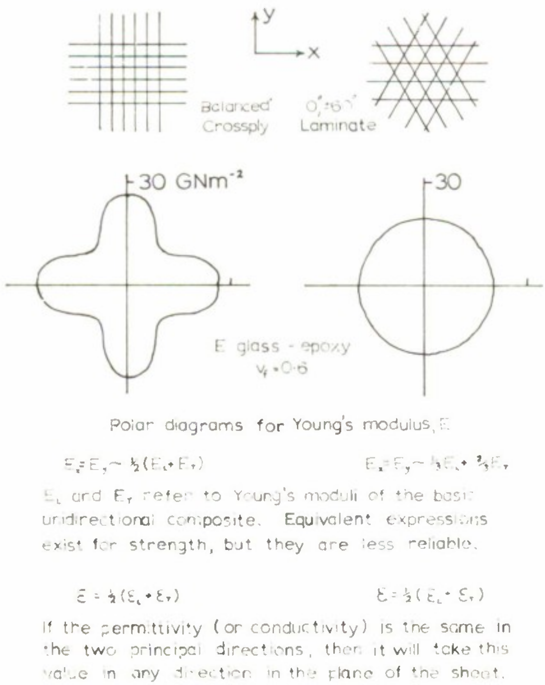


FIG. 8. Directional properties of various fibre configurations.

Composite Properties

Table 4 lists the properties of a number of reinforced plastics. It is incomplete because electrical properties are not normally quoted for materials containing asbestos. The table shows what sort of properties are quoted for currently available commercial materials as well as illustrating what can be done, particularly to the mechanical properties, given care and effort.

For the unidirectionally reinforced plastics only the longitudinal properties have been given, but for the cross-plied and imperfectly aligned materials transverse values are also quoted where these are available.

Asbestos fibres, particularly the soft chrysotiles, can withstand a fair amount of pressure without being seriously degraded. Durestos is a material moulded at low pressure from partially (60% - 40%) oriented chrysotile[7] while Pyrotex is a very similar material moulded at 4 MNm^{-2}. The increased mechanical properties result from the higher volume percentage of fibre incorporated by doing this. At pressures higher than this the fibre packing improves rather slowly and eventually the fibres break up and the composite properties are degraded. Usually the optimum pressure is between 10 and 100 MNm^{-2} depending on the grade of asbestos.

To obtain significantly better properties from short fibrous materials such as asbestos the fibres have to be aligned a good deal better than they are in Durestos. When this is done one can obtain much higher volume percentages for the same pressure, and this is reflected in superior mechanical properties as can be seen from the test results on the material classified and aligned by the E.R.D.E. processes[23].

Gordon Aerolite was a composite material produced early in World War II containing a high volume fraction of carefully aligned flax fibres in a phenolic resin. It was not in production for very long, but is included in the table because it represents something fairly close to the ultimate that can be expected with a cellulosic reinforcement. Modern commercially available materials, *e.g.* 6F/45, bearing in mind that they are bidirectional and produced in quantity, do not compare badly with Aerolite.

The properties of the aromatic fibre composites have been compounded from published figures on Kevlar 49 and PABH-T X-500

TABLE 4. Typical Composite Properties.

Material	UTS GNm^{-2}	Young's Modulus GNm^{-2}	Resistivity ohm cm	Permittivity	Loss Factor (tan δ)	References
Unidirectional	0·55 - 1·72	45	$2{\cdot}5 \times 10^{15}$	4·7 (1 MHz)	0·025 (1 MHz)	1, 2, 5
E glass-epoxy				4·5 (1 GHz)	0·02 (1 GHz)	
Unidirectional	0·5 - 1.5	35		3·4 (1 GHz)	0·004 (1 GHz)	2
D glass-silicone resin						
Press moulded Durestos	0·17 (L)	21 (L)				5
(Chrysotile-phenolic)	0·12 (T)	17 (T)				
Pyrotex mat	0·38 (L)	30·7 (L)				5
pressed at 10 MNm^{-2}	0·24 (T)	20·6 (T)				
Chrysotile-epoxy	0·55 (L)	95 (L)				5, 21
Classified and aligned by ERDE processes						
Anthophyllite-epoxy classified	0·54 (L)	84 (L)		6 (1 KHz)	0·15 (1 KHz)	
and aligned by ERDE processes	0·09 (T)	12 (T)				
Glass-epoxy Laminate	0·46 (L)	17·9 (L)		5·0 (1 MHz)	0·021 (1 MHz)	22
(Tufnol Grade 10G/41)	0·45 (T)	17·9 (T)				
Cellulosic Fabric-Epoxy Laminate	0·124 (L)			4·4 (1 MHz)	0·040 (1 MHz)	22
(Tufnol Grade 6F/45)	0·124 (L)					
Gordon Aerolite	0·38 (L)	40 - 60 (L)				5
(Unidirectional Flax-phenolic)	0·014 (T)	5·5 (T)				
No longer available						
Aromatic Fibre (Kevlar 49 - Polyimide)	0·9	87	10^{16}	5·2 (100 Hz) 4·4 (1 MHz)	0·028 (100 Hz) 0·029 (1 MHz)	12 Low frequency values inferred from 11
Unidirectional)				3·4 (10 GHz)	0·007 (10 GHz)	

Fibre volume fractions in the region of 60% for glass and aromatic fibres, 50% for cellulosic reinforcements, and 30 - 50% for asbestos.

(Monsanto). It is to be hoped that they give some idea of what is achievable. Apart from the low compressive strength of these materials, their chief drawback appears to be that they are difficult to cut and difficult to machine to give a smooth finish.

With regard to electrical properties, as far as one can generalise, aromatic fibres have a lower dielectric constant than glass at all frequencies, slightly greater dielectric loss at low frequencies but a loss factor similar to the best glasses at high frequencies. The stability of dielectric constant with respect of temperature, which is a consideration for radomes that reach elevated temperatures during service, seems to be marginally better for glass fibres. Asbestos and cellulosic fibres can be expected to be less satisfactory under most conditions. The figures claimed for linen fabric-epoxy are admittedly not too bad, but much must depend on the severity of the environment. In humid conditions it must be purely a question of time before the reinforcing fibres are modified by water pickup. Obviously the water permeability of the resin and the extent to which the fibre ends are sealed off are important factors here.

References

(1) Wright, W. W., *J.N.S.*, **1**, 4 (October 1975).

(2) Baird, M. E. Electrical Properties of Polymeric Materials. The Plastics Institute, 1973.

(3) Loewenstein, K. L. The Manufacturing Technology of Continuous Glass Fibres. Elsevier. 1973.

(4) Walton, J. D. Radome Engineering Handbook, design and principles. Marcel Dekker. 1970.

(5) Hodgson, A. A. Fibrous Silicates. The Royal Institute of Chemistry, Lecture Series 1965, No. 4.

(6) Burman, D. The Tensile Strength of Asbestos Fibres, Paper 2 - 8, Conference on the Physics and Chemistry of Asbestos Minerals, Oxford. 1967.

(7) Parratt, N. J. Fibre Reinforced Materials Technology. Van Nostrand-Reinhold. 1972.

(8) Rosato, D. V. Asbestos; Its Industrial Applications. Reinhold. 1959.

(9) Gordon, J. E. The New Science of Strong Materials. Pelican. 1968.

(10) Murphy, E. J. *J. Phys. Chem. Solids*, **16**, 1 - 2 (Nov. 1960) 115 - 22.

(11) Hearle, J. W. S. and Peters, R. H. Moisture in Textiles. Textile Institute and Butterworths. 1960.

(12) El Hosseiny, F. and Page, D. H. Mechanical Properties of Wood Pulp Fibres. *Fibre Science and Technology*, **8**, 1 (Jan. 1975), 21 - 32.

(13) Black, W. G. and Preston, J. High Modulus Wholly Aromatic Fibres. Marcel Dekker, New York. 1973.

(14) Stratton, W. K. Evaluation of du Pont's High Modulus Organic Fiber PRD 49 Type 1. 16th National SAMPE Symposium and Exhibition. April 22, 1971.

(15) Hashin, Z. Theory of Fiber Reinforced Materials. NASA-CR-1974 (March 1972).

(16) Cooper, G. A. Micromechanics Aspects of Fracture and Toughness. Chapter 10 of Composite Materials Vol. 5. Fatigue and Fracture. Ed. L. J. Broutman. Academic Press. 1974.

(17) Böttcher, C. J. F. Theory of Electric Polarisation. Elsevier. 1952.

(18) Cox, H. L. *Brit. J. Appl. Phys.* **3** (1952), 72.

(19) Cook, J. The Dielectric and Conductivity Properties of Imperfectly Aligned Fibrous Materials. ERDE - TR - 142 (1974).

(20) Cook, J. The Elastic Constants of an Isotopic Matrix Reinforced with Imperfectly Oriented Fibres. *Brit. J. Appl. Phys.* (J. Phys. D), **1** (1968) 789 - 812.

(21) Tsai, S. W., Halpin, J. C. and Pagano, N. J. Composite Materials Workshop; *Progress in Materials Science, Technomic* **1** (1968).

(22) Ashton, J. E. and Whitney, J. M. Theory of Laminated Plates: *Progress in Materials Science, Technomic* **4** (1970).

(23) Bagg, G. E. C., Evans, M. E. N. and Pryde, A. W. H. The Glycerine Process for the Alignment of Fibres and Whiskers. *Composites* **1**, 2 (Dec. 1969) 97 - 100.

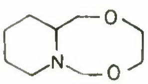
(24) Tufnol Industrial Plastics Literature. 1974.

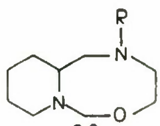
STEREOCHEMISTRY OF PERHYDROPYRIDO [1,2-c] [1,6,3] DIOXAZOCINES AND 2-ALKYL-PERHYDROPYRIDO [1,2-c] [1,3,6] OXDIAZOCINES

Part 2

G. C. Jackson, A.R.I.C., Ph.D.
Exposure Trials Station
Central Dockyard Laboratory

Introduction A series of perhydropyrido [1,2-*c*] [1,6,3] dioxazocines (3.1) and 2-alkyl-perhydropyrido [1,2-*c*] [1,3,6] oxdiazocines (3.2) have been synthesised as part of the antifouling research programme at the Central Dockyard Laboratory[1]. Stereochemical studies of these compounds were carried out to enable biological activity to be related to chemical structure.


3.1


3.2

Structural studies on perhydropyrido [1,2-*c*] [1,6,3] dioxazocines

As described above[1] all the perhydropyrido [1,2-*c*] [1,6,3] dioxazocines, with the exception of 3,4-dimethyl-perhydropyrido [1,2-*c*] [1,6,3] dioxazocine, were highly unstable and rapidly underwent polymerisation. However, it was possible to obtain the n.m.r. spectrum Fig. 1 of the parent unsubstituted compound, but before an expansion could be run polymerisation had occurred. The spectrum showed an AB quartet for the C(6) methylene protons (δ 4·79 and 4·16 p.p.m. and J_{gem}–10·5 Hz), a six proton multiplet at δ 3·5 to 4·0 p.p.m. for the protons adjacent to the oxygen atoms, and a three proton multiplet at δ 2·6 to 3·1 p.p.m. corresponding to the protons adjacent to the nitrogen atom. The absorption of the *N*-CH_2 protons below δ 2·0 p.p.m. suggests a fused ring conformation[2], since an α–CH protons *trans* diaxial to the nitrogen lone pair would absorb to higher field of δ 2·0 p.p.m. A *cis*-fused ring conformation is also suggested by the absence of Bohlmann bands in the i.r. spectrum, but is not conclusive due to unreliability of this criterion for medium sized ring compounds.

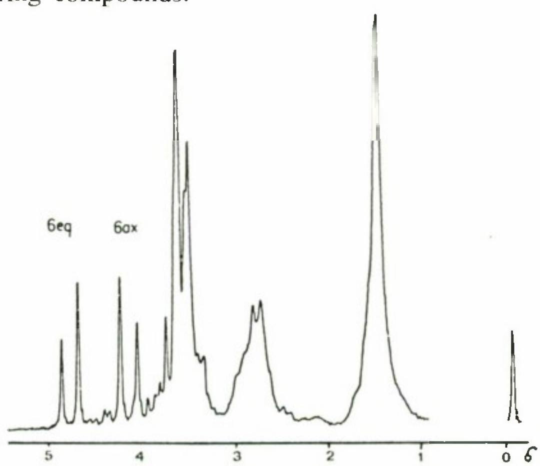
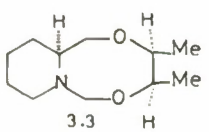


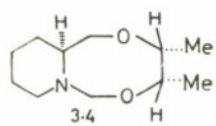
FIG. 1. 60 MHz n.m.r. spectrum of perhydropyrido [1,2-c] [1,6,3] dioxazocine in CCl_4

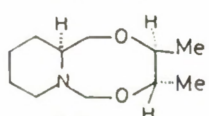
The n.m.r. spectrum of 2-methyl- (or 3-methyl)-perhydropyrido [1,2-*c*] [1,6,3] dioxazocine showed the presence of two isomers (two methyl doublets at δ 0·97 and 0·95 p.p.m.), but it was impossible to assign the position of substitution of the methyl groups since the compound was polymerising during the running of the spectrum. A J_{gem} of −10 Hz was observed for the C(6)-methylene protons similar to that obtained for 3·1, indicating a similar conformation for this compound.

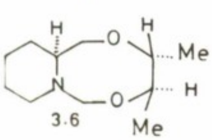
The n.m.r. spectrum of 3,4-dimethyl-perhydropyrido [1,2-*c*] [1,6,3] dioxazocine Fig. 2 showed the presence of two isomers when prepared from the mixture of isomeric 1 2-dimethyl-2-(2-piperidylmethoxy) ethanols. When prepared from 6-*p*-nitrophenyl-3,4-dimethyl-perhydropyrido [1,2-*c*] [1,6,3] dioxazocine as described previously[1] a mixture of isomers was again isolated Fig. 3, but the ratio had altered as indicated by the change in the relative area of the N-CH_2-O AB quartets. The unreacted 1,2-dimethyl-2-(2-piperidylmethoxy ethanol from the above preparation of 3,4-dimethyl-perhydropyrido [1,2-*c*] [1,6,3] dioxazocine was isolated from the reaction mixture, and condensed with aqueous formaldehyde solution at room temperature. The n.m.r. spectrum of this product Fig. 4 showed a single low field quartet for the N-CH_2-O protons (δ 4·96 and 4·17 p.p.m., J_{gem} −10 Hz), absorption for the four protons adjacent to oxygen at δ 3·07 to 3·65 p.p.m., and for the three protons adjacent to nitrogen at δ 2·43 to 2·83. This, and the lack of Bohlmann bands in the i.r. spectra indicates a similar conformation for this compound and perhydropyrido [1,2-*c*] [1,6,3] dioxazocine.

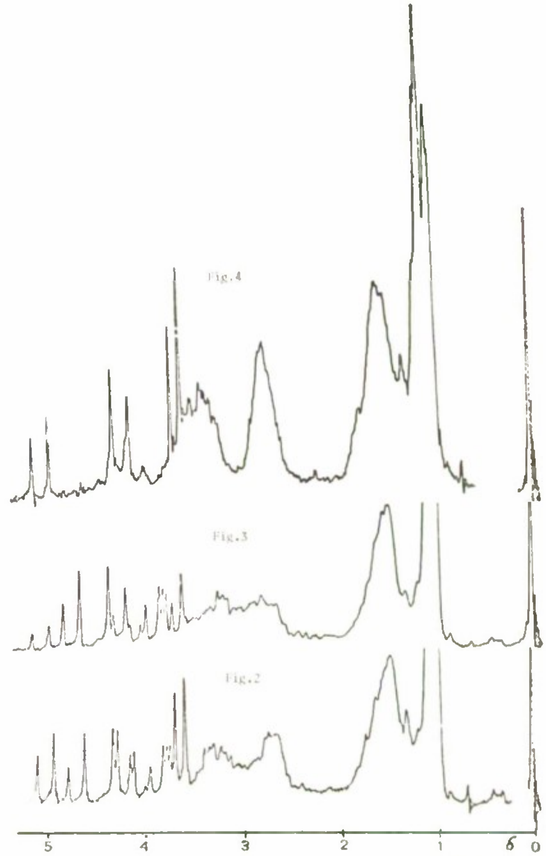
The 3,4-dimethyl-perhydropyrido [1,2-*c*] [1,6,3] oxdiazocine was prepared using *meso*-butane-2,3-diol, and its n.m.r. spectrum was the same as that prepared from the mixture of *meso*- and *dl*-butane-2,3-diol, *i.e.* with two quartets for the N-CH_2-O protons. It seems


3.3


3·4


3.5


3.6



FIGS. 2 to 4. 60 MHz n.m.r. spectrum of mixed isomers of 3,4-dimethyl perhydropyrido [1,2-*c*] [1,6,3] dioxazocine (Fig. 2), and of the isomer prepared from its 6-*p*-nitrophenyl derivative (Fig. 3), and that prepared from the 1,2-dimethyl-2-(2-piperidylmethoxy) ethanol that did not react with *p*-nitrobenzaldehyde (Fig. 4).

that only the *meso*-form reacts to give the cyclised products, which must be 3·3 and 3·4 rather than 3·5 and 3·6.

To gain more information regarding the stereochemistry of this system the 6-*p*-nitrophenyl-3,4-dimethyl-perhydropyrido [1,2-*c*] [1,6,3] dioxazocine was examined. Consideration of Dreiding models suggest three possible structures for this compound ignoring at present the configuration about C(3) and C(4). These are shown in Fig. 5,* where the *trans*-ring fused 3·7 is interconvertible by nitrogen

* The compound exists as racemates, and the structures are drawn for ease of representation and do not necessarily represent the same enantiomer throughout.

inversion and ring inversion with the *cis*-ring fused 3·8. It is also possible to obtain the *cis*-ring fused 3·9 which is a diastereoisomer of 3·7 and 3·8. For each of the structures shown in Fig. 5 the eight-membered ring can adopt a chair-chair conformation (shown by —) or a boat-chair conformation (shown by . . .) and the inter-conversion does not affect the stereochemistry of the rest of the molecule. Both 3·7 and 3·9 are destabilised by the generalised anomeric effect, whereas 3·8 is not. This is illustrated in Fig. 5 by bold arrows indicating dipoles, interactions between these forming part of the anomeric effect.

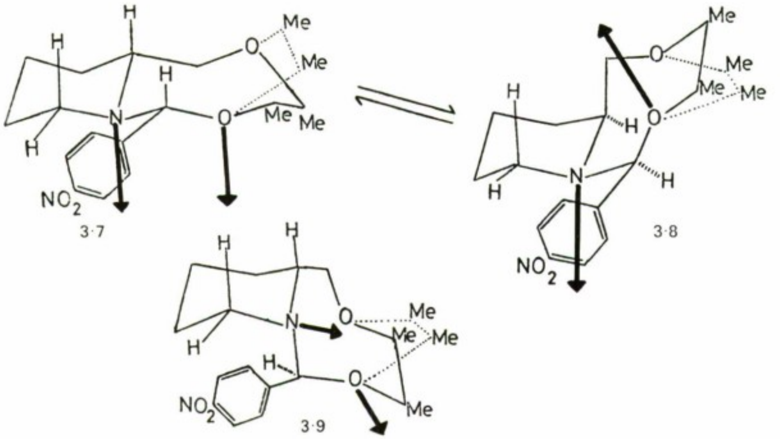


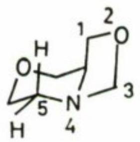
FIG. 5. Possible structures for 6-p-nitrophenyl-2, 3-dimethyl-perhydropyrido [1,2-c] [1,6,3] dioxazocine from examination of Dreiding models.

The 220 Mhz n.m.r. spectrum of 6-p-nitrophenyl-3,4-dimethyl-perhydropyrido [1,2-*c*] [1,6,3] dioxazocine showed multiplets of δ 2·76 and 2·34 p.p.m. corresponding to H8ax and H8eq respectively, see Fig. 6 and analysis gave coupling constants characteristic of a chair piperidine ring, see Table 1. The relatively high field absorption of H8eq (δ 2·34 p.p.m.) compared to that normally observed in piperidine (δ 2·8 p.p.m.) must be a result of shielding by the aromatic ring. This is possible in all three structures in Fig. 5 if the plane of the aromatic ring is parallel to the H-C(8)-H plane. Perhaps the most interesting feature of the spectrum is the very low field absorption of H8ax (2·76 p.p.m.). In quinolizidine[3] the corresponding axial proton absorbs to much higher field (δ 2·07 p.p.m.) than the equatorial proton (Δ ae 0·93 p.p.m.), whereas in (3·20) H8ax actually absorbs 0·42 p.p.m. to lower field of H8eq. This is not explicable in terms of structures 3·7 and 3·9, but in 3·8 the H8ax should be deshielded by the *syn* axial C(1) methylene[4], and by the *syn* axial C-O(5) bond. A similar deshielding effect has been noted in 3·10 with chemical shifts of δ 2·62 and 2·23 p.p.m. for H5ax and H5eq respectively[5].

TABLE 1.

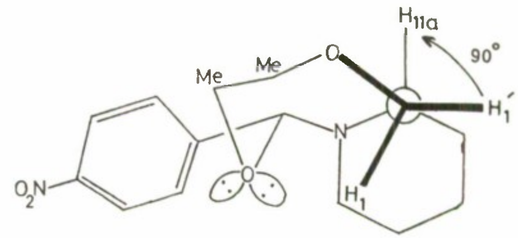
Analysis of the 220 MHz n.m.r. spectrum of 6-p-nitrophenyl-3,4-dimethyl-perhydropyrido [1,2-c] [1,6,3] dioxazocine.

Protons	*J(Hz)*	*Protons*	*J(Hz)*	*Protons*	*Chemical Shift (δ p.p.m.)*
H8ax H8eq	−12	H8eq H9eq	2·5	H8eq	2·33
H8ax H9ax	12	H1 H1′	−12	H8ax	2·76
H8ax H9eq	3	H1 H11(a)	9	H1′	3·91
H8eq H9ax	3·5	H1′ H11(a)	1·5	H1	4·28
				H3 & H4	3·50



3·10

The C(1) methylene protons absorb as an AB part of an ABX multiplet at δ 4·25 and 3·90 p.p.m., and if these were strictly axial and equatorial with respect to H11a coupling constants of 11 and 5 Hz would be expected[6], but the observed values Table 1 suggest dihedral angles of approximately 140° and 100°. Examination of the Dreiding model of 3·8 in fact shows angles of 90° and 150° Fig. 7 for the dihedral angles involving CH11a-CH1′ and CH11a-CH1 respectively which is consistent with the observed $J_{vicinal}$ values. In addition the H1 proton absorbs at much lower field (0·37) than H1′, and this too is in accord with structure 3·8 as it can be seen in Fig. 7 that H1 is close to the oxygen atom which will result in a deshielding of this proton. Thus structure 3·11 is consistent with the n.m.r. data, but assignment of the methyl groups at C(3) and C(4) is not possible from the n.m.r. spectrum.



3·11

FIG. 7. Proposed structure 3·11 of 6-*p*-nitrophenyl-2,3-dimethyl-perhydropyrido [1,2-*c*] [1,6,3] dioxazocine showing dihedral angles of the H11a proton with the C(1) methylene protons, and the close approach of H(1) to the 5-oxygen atom.

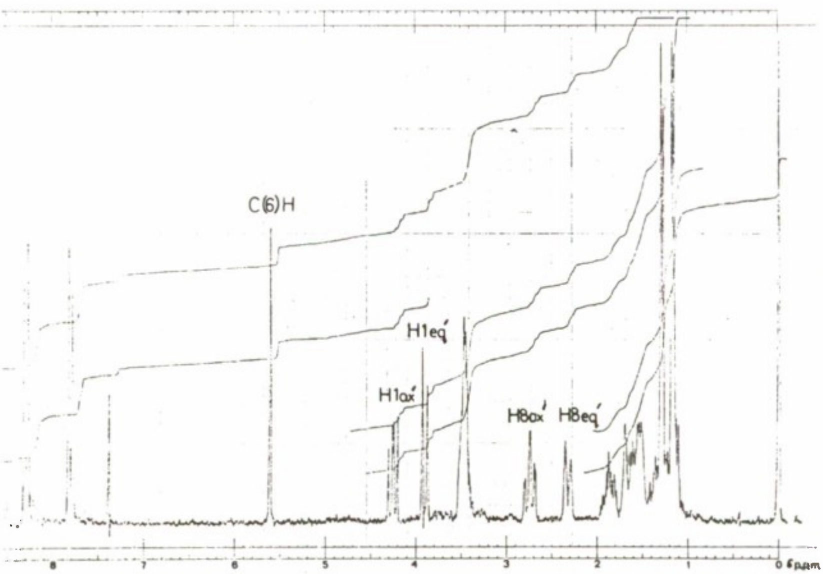
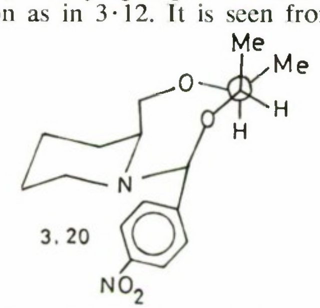


FIG. 6. 220 MHz n.m.r. spectrum of 6-*p*-nitrophenyl-2,3-dimethyl-perhydropyrido [1,2-*c*] [1,6,3] dioxazocine.

The conformation of the eight-membered ring *i.e.* boat-chair or chair-chair cannot be deduced from the n.m.r. spectra since alteration around C(3) and C(4) does not affect the relative disposition of the 5 oxygen atom with respect to H8ax, or the dihedral angles between the C(11a) and C(1) methylene protons.

The assignment of the relative configurations of the C(3) and C(4) methyl groups was made by synthesising 3,4-dimethyl-perhydropyrido [1,2-*c*] [1,6,3] dioxazocine from *meso*-butane-2,3-diol when as discussed above only 3·3 and 3·4 were produced. However 3·5 and 3·6 derivable from *dl*-butane-2,3-diol were not obtained from a mixture of *meso*- and *dl*-butane-2,3-diol. Study of Dreiding models of the *cis*-fused conformations show that the boat-chair 3·16 and 3·18 and chair-chair 3·17 and 3·19 conformations of the eight-membered ring in compounds derivable from *dl*-butane-2,3-diol contain serious non-bonded interactions. Conformations 3·17 and 3·18 have non-bonded interactions arising from diequatorially substituted methyl groups, and the axial methyl groups at C(3) and C(4) have serious non-bonded interactions with the C(1) and C(6) methylene protons. The possible *cis*-fused conformations with boat-chair and chair-chair conformations of the eight-membered ring obtainable from *meso*-butane-2,3-diol are 3·12 to 3·15. The structure 3·14 has serious non-bonded interactions between the C(3) axial methyl group and the C(1) methylene and H11a protons, and consequently the equilibrium favours 3·15 where non-bonded interactions between the C(4) axial methyl group and the C(6) methylene protons can be relieved by slight distortion of the eight-membered ring. The other preferred conformations possible for compounds derived from *meso*-butane-2,3-diol 3·12 and 3·13 appear free of non-bonded interactions, and from previous studies on the conformations of eight-membered rings, the boat-chair conformation 3·12 is probably preferred.

Further studies of Dreiding models show that the single isomer of 6-*p*-nitrophenyl-3,4-dimethyl-perhydropyrido [1,2-*c*] [1,6,3] dioxazocine prepared from the isomeric mixture of 1,2-dimethyl-2-(2-piperidylmethoxy)ethanols will have the structure 3·20, where the C(3) and C(4) methyl groups have the same configuration as in 3·12. It is seen from Dreiding



3. 20

models that the reaction of the precursor of 3·15 with *p*-nitrobenzaldehyde would yield a compound that was sterically crowded. Consequently the single isomer of 3,4-dimethyl-perhydropyrido [1,2-*c*] [1,6,3] dioxazocine synthesised from the 1,2-dimethyl-2-(2-piperidylmethoxy) ethanol that did not react with *p*-nitrobenzaldehyde was assigned structure 3·15. The n.m.r. spectrum of 3·15 shows an AB quartet for the N-CH_2-O protons Fig. 4 at lower field than those arising from similar protons in 3·12. This is due to the deshielding effect of the C(4) axial methyl group in 3.15. It was observed from Dreiding models that interconversion of the boat-chair and chair-chair conformations of the eight-membered ring did not seem to affect the N-CH_2-O geometry, as was shown by the constant J_{gem} of −10 Hz observed for these systems.

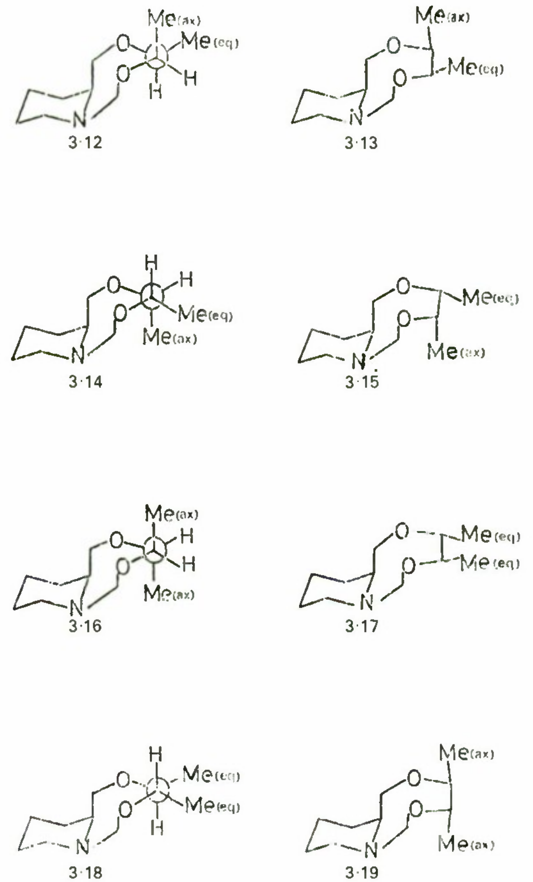


FIG. 8. Possible conformations of *cis*-ring fused 3,4-dimethyl perhydropyrido [1,2-*c*] [1,6,3] dioxazocine.

The 60 MHz n.m.r. spectrum of the mixture of 9-ethyl-perhydropyrido [1,2-*c*] [1,6,3] dioxazocines showed two AB quartets of approximately equal intensity (δ 4·77 and 4·16 p.p.m., J_{gem} -11·2 Hz, and δ 4·75 and 4·18 p.p.m. J_{gem} −10·0 Hz), and these signals were assigned to the C(6) methylene protons situated between the nitrogen and oxygen atoms, see Fig. 9. Attempts at separation of the isomers by column or thin layer chromatography, and by picrate formation were unsuccessful, but partial separation of the 5-ethyl-2-(2-piperidylmethoxy) ethanols was achieved. On storage of the liquid 2-(5-ethyl-2-piperidylmethoxy) ethanols at −30° for several weeks, partial solidification occurred, and the liquid fraction, decanted off and reacted with aqueous formaldehyde solution, yielded a single isomer of 9-ethyl-perhydropyrido [1,2-*c*] [1,6,3] dioxazocine. The 60 MHz n.m.r.

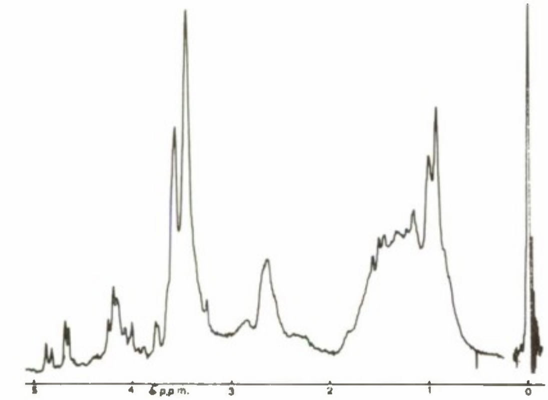


FIG. 9. 60 MHz n.m.r. spectrum of the mixed isomers of 9-ethyl-perhydropyrido [1,2-c] [1,6,3] dioxazocine in $CDCl_3$.

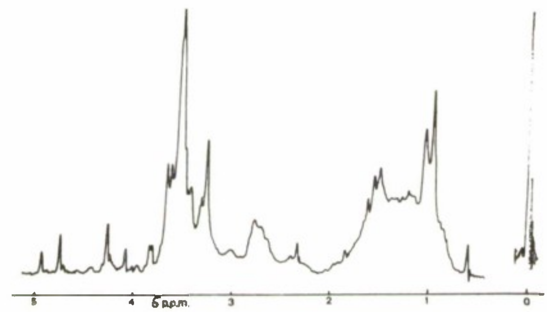


FIG. 10. 60 MHz n.m.r. spectrum of the single isomer of 9-ethyl-perhydropyrido [1,2-c] [1,6,3] dioxazocine obtained from the liquid fraction of 2-(5-ethyl-2-piperidylmethoxy)ethanol in $CDCl_3$.

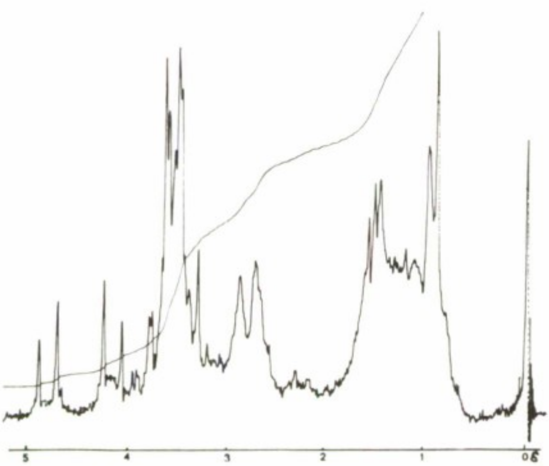
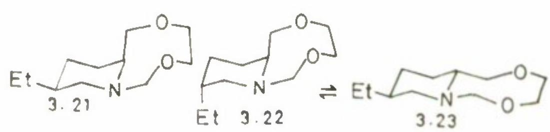


FIG. 11. 60 MHz n.m.r. spectrum in $CHCl_3$ of the mixed isomers of 9-ethyl-perhydropyrido [1,2-c] [1,6,3] dioxazocine after the decomposition of one isomer. The peak at δ 2·95 p.p.m. was removed by shaking the solution with deuterium oxide.

spectrum of this single isomer Fig. 10 in deuterated chloroform showed a single AB quartet for the $N\text{-}CH_2\text{-}O$ protons (δ 4·77 and 4·16 p.p.m., J_{gem} −11·2 Hz). A six-proton multiplet was observed centred at δ 3·60 p.p.m. and this was assigned to the protons adjacent to the oxygen atoms. A three proton multiplet centred at δ 2·75 p.p.m. was assigned to the protons adjacent to the nitrogen atom, with a broad multiplet equivalent to ten protons at δ 0·75 to 1·90 p.p.m. The spectrum was thus very similar to that obtained from the unsubstituted perhydropyrido [1,2-c] [1,6,3] dioxazocine.

The n.m.r. spectrum of the mixed isomers was found to change gradually with time, with the AB quartet for the $N\text{-}CH_2\text{-}O$ protons (δ 4·75 and 4·18 p.p.m.) becoming less intense, and a broad peak at δ 2·95 p.p.m. becoming more intense. The final spectrum obtained Fig. 11 showed a single AB quartet, and the peak at δ 2·95 p.p.m. was removed by shaking the chloroform solution with deuterium oxide, indicating the decomposition of one isomer to a product containing a hydroxyl group.



The similarity between the n.m.r. spectrum of the single isomers obtained by 9-ethyl-perhydropyrido [1,2-c] [1,6,3] dioxazocine Fig. 11 and that of the parent unsubstituted compound indicates a similar stereochemistry for the two compounds. This suggests the equatorially substituted *cis*-fused structure 3·21 for the 9-ethyl-perhydropyrido [1,2-c] [1,6,3] dioxazocine. The second isomer would possess an axial ethyl ethyl group in the *cis*-fused conformation 3·22, and would tend to adopt the *trans*-fused conformation 3·23 in which the ethyl group is equatorially substituted. Such *trans*-fused derivatives have not so far been encountered, and might well be unstable relatime to the *cis*-fused conformations.

Structural studies of 2-alkyl-perhydropyrido [1,2-c] [1,3,6] oxdiazocines

This series of compounds were all obtained as colourless mobile liquids which were stable at room temperature. The use of Bohlmann bands to deduce the ring fusion of these com-

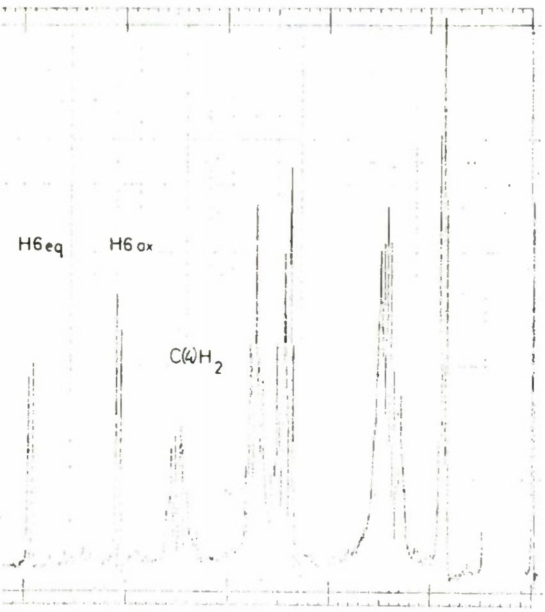
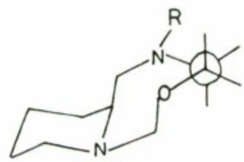
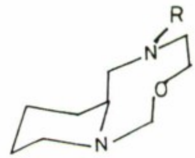


FIG. 12. The 220 MHz n.m.r. spectrum of 2-iso-propyl-perhydropyrido [1,2-c] [1,3,6] oxdiazocine in carbon tetrachloride.

obtained for the perhydropyrido [1,2-c] [1,6,3] dioxazocines are typified by that for 2-*iso*-propyl-perhydropyrido [1,2-c] [1,3,6] oxdiazocine Fig. 12. As seen in Table 2 the chemical shifts of the AB quartets for the C(6) methylene protons are all very similar, as is the value of J_{gem} for these protons (*i.e.* –9·9 to –10·1 Hz). The chemical shifts of the multiplets for protons adjacent to oxygen (3·30 to 3·77 p.p.m.) and to nitrogen (2·20 to 3·07), are very similar to those observed for the oxygen analogues as is the value of J_{gem} for the N-CH_2-O protons. The 220 MHz n.m.r. spectra of the 2-alkyl-perhydropyrido [1,2-c] [1,3,6] oxdiazocines are not analysable, as the region that would yield detailed information on the structure *i.e.* δ 2·2 to 3·1 p.p.m. for the absorption of protons adjacent to nitrogen, contains many overlapping multiplets.



3·24



3·25

TABLE 2.
Chemical shift and J_{gem} values obtained from the 60 MHz n.m.r. spectra of 2-alkyl-perhydropyrido [1,2-c] [1,3,6] oxdiazocines in CCl_4

2-alkyl group	*6eq p.p.m.*	*6ax p.p.m.*	*6ae p.p.m.*	J_{gem} *6ax6eq Hz*	*O-CH_2 protons (p.p.m.)*	*N-CH_2 protons (p.p.m.)*
methyl	4·76	4·04	0·72	–10·0	3·27 - 3·70	2·40 - 2·93
ethyl	4·87	4·07	0·80	–10·0	3·30 - 3·77	2·33 - 3·00
n-propyl	4·90	4·07	0·83	–10·1	3·30 - 3·73	2·30 - 3·07
iso-propyl	4·91	4·08	0·83	–9·9	3·30 - 3·73	2·20 - 3·00
n-butyl	4·91	4·07	0·84	–9·9	3·32 - 3·70	2·30 - 3·00
iso-butyl	4·95	4·15	0·80	–9·9	3·30 - 3·70	2·30 - 2·90
t-butyl	5·09	4·13	0·96	–10·0	3·30 - 3·70	2·30 - 2·90

pounds is complicated by absorption in this region from the alkyl substituent on the 2-nitrogen atom. A *cis*-ring fusion may be indicated by the lack of Bohlmann bands in the i.r. spectrum of 2-*t*-butyl-perhydropyrido [1,2-c] [1,3,6] oxdiazocine, but this criterion is unreliable for medium sized rings.

The n.m.r. spectra of this series of compounds which showed similarities to those

It may be assumed, however, from the similarities between the 60 MHz n.m.r. spectra of these compounds and those of the perhydropyrido [1,2-c] [1,6,3] dioxazocines, that the nitrogen analogues possess a similar structure with a *cis*-ring fusion. The eight-membered ring may adopt the boat-chair conformation as shown in 3·24, or the chair-chair conformation 3·25.

Studies of the chemical shifts of the low field 6-proton (designated 6 ax in Table 2), shows that as 2-alkyl substituent increases in bulk the proton becomes increasingly shielded, *i.e.* δ 4·76 for 2 methyl to 5·09 for 2-*t*-butyl. Examination of Drieding models for the 2-*t*-butyl compounds shows that only (3·25 R=*t*-butyl is relatively free of non-bonded interactions. However, the 2-methyl compound can exist both as 3·24 and 3·25 where R is in pseudoequatorial orientation. In 3·25, H-6ax is close to the nitrogen lone pair and so should be deshielded, but this is not the situation for 3·24.

An explanation for the chemical shifts of H-6ax of these compounds, is that the 2-methyl compound is an equilibrium between 3·24 and 3·25 with 3·24 predominant. The proportion of 3·25 in the equilibrium mixture increases as the steric requirements of the alkyl group increases, and with the 2-*t*-butyl compound is almost entirely 3·25.

Acknowledgements

I would like to thank Mr. D. R. Houghton, officer in charge ETS and Dr. T. A. Crabb of the Portsmouth Polytechnic for their invaluable help and encouragement in this work.

References

(1) Jackson, G. C. *J.N.S.* **1**, 2 (April 1975).
(2) Crabb, T. A., Newton, R. F. and Jackson, D. *Chem. Rev.*, **71** (1971) 109.
(3) Hamlow, H. P., Okuda, S. and Nakagura, N. *Tetrahedron Letter* (1964) 2553.
(4) Booth, H. *Tetrahedron*, **22** (1966) 615.
(5) Hall, M. J., Ph.D. Thesis (C.N.A.A.), 1973, p.96.
(6) Karplus, M. *J. Amer. Chem. Soc.* **85** (1963) 2870.

Graham C. Jackson joined C.D.L. in March, 1964, as a Scientific Assistant having been previously employed at Cyanamid of Great Britain Ltd., at Gosport. He was assigned to the Exposure Trials Station and worked on seawater analysis, paint inspection and copper leaching rates. In 1967, he was promoted to A.X.O. He obtained his graduateship of the Royal Institute of Chemistry, by part time study, in 1970. Since that time he has been concerned with the evaluation and synthesis of antifouling agents. He was awarded a Ph.D. under the Council for National Academic Awards regulations in 1974 and attended the Lieutenants Course at the Royal Naval College, Greenwich.

THE AMATEUR SCIENTIST AND THE ROTATION OF THE EARTH

G. J. Kirby, B.Sc.
Admiralty Underwater Weapons Establishment

What is the connection between the rotation of the Earth and a group of amateur astronomers, led by the author travelling to a deserted beach at Hayling Island in the middle of winter? I'll tell you but first let us consider what we understand by TIME.

The Nature of Time

We think instinctively of time as a steady flow along a dimension over which we have no control. This concept has been expressed for over two thousand years by man's attempts to provide a mechanical means of recording the passage of time and this has led to the continuous search for ever more regular mechanical devices. It has been estimated[(1)] that the Greeks, over two thousand years ago could have measured time to an accuracy of about one minute over very long periods (typically years) by astronomical observations made with simple instruments and indeed, Hipparchus was easily able to determine the time of the equinoxes to better than a few hours. However, the Greeks, unlike some modern races, were not obsessed with a desire for an accurate knowledge of time and the accuracy of time measurement was not improved on a large scale until water clocks were refined to give an accuracy of a few minutes in a day. The first verge and foliot clocks introduced around 1340 gave an accuracy of about 15 minutes in a day and steady improvements eventually led to Harrison's chronometers which, by the mid-18th century were accurate to half a second a day even under sea-going conditions. The ultimate in pendulum clocks probably achieved an accuracy approaching 10 mS in a day.

A common characteristic of all mechanical clocks is that they have to be calibrated and reset from time to time against a standard; they are no more than devices for interpolating on the fundamental passage of time. Until 1955 the standard of time measurement was the rotation of the Earth. The old observatory at Greenwich was erected for the express purpose of time measurement and the associated problem of determining longitude at sea, and this year the Observatory celebrates its tercentenary. Time was determined by the observation of the transit of stars through the eyepiece of a special telescope. The axis of this telescope, which is now redundant but still on view to the public, forms the prime standard of longitude for the entire world with the exception of a small group of French scientists who doggedly kept to the Cassini standard based on Paris!

Thus, the rotation of the Earth was, for over three hundred years, used as the prime standard for time and Greenwich set the world's clocks, initially by the visual sighting of the dropping of a ball on the old observatory building and later by the broadcasting of the famous Greenwich 'pips'.

There are other clocks which can be used in this way. The motion of the Moon around the Earth can be used as a clock but this is not as simple as using the rotation of the Earth since the latter was regarded as unvarying and the former extremely complex. The application of his newly discovered Laws of Universal Gravitation allowed Newton to predict the position of the Moon to an accuracy of about 1·5 mrad† but great advances were made in lunar orbital theory spurred on by the importance of lunar sightings for determining longitude at sea. The motion of the Moon is very complex on account of the many forces acting upon it due to the Earth's non-spherical shape and the Sun and planets. It is little short of incredible that in 1787 the French mathematical genius Laplace was able to announce that the predicted position of the Moon and the observed position had deviated by 50 μrad† over a period of a century (cy). Consider that the Moon moves no less than 75·4 rad in a year and we see that this discrepancy is no more than 1 part in 15×10^7; and this in the days when computation was all done by hand. The modern calculation of just one component of the Moon's motion involves the summation of over 1700 algebraic expressions!

The discrepancy was attributed to a secular perturbation due to the planets but subsequent refinements by such men as J. C. Adams in 1853, and Newcomb in 1878 showed that the deviation between prediction, based on Newton's Laws, and observation could not be reconciled. The deviation took the form of an apparent acceleration of the Moon in its orbit. Then similar accelerations were found in the motions of the Earth and other planets around the Sun and these unaccountable accelerations were proportional to the apparent mean angular motion of the celestial bodies around the Earth. The truth dawned late in the 19th century. Newton's Laws *were* valid; it was the system of time reference used in the theory that was not constant. In other words the rate of rotation of the Earth is not constant. The average rate of slowing down of the Earth is very small such that the length of the day increases by only 2·0 mS in a century. This was far too small to be detected by Earth-bound clocks; at least until 1955 when atomic time (AT) was introduced.

Time based on the Earth's rotation, as recorded at Greenwich is referred to as Universal Time (UT) and the system of time defined as uniform in which Newton's Laws of Gravitation (as modified by Einstein) correctly predict the motion of the Moon and planets is called Ephemeris Time (ET). The change in UT relative to ET is a measure of the variable rotation of the Earth and is constantly monitored around the world. There is ample evidence[2, 3, 4, 5] that the slowing down of the Earth has continued throughout recorded history. Detailed predictions of ancient solar eclipses do not agree with observations recorded by the Chinese, Greek and Arab civilisations. For example, an eclipse recorded in the annals of the Former Han Dynasty on July 17th 187 bce (before common era) occurred several hundred miles east of the predicted path of totality. Hundreds of such records show beyond doubt that the spin-down of the Earth has produced a slip between UT and ET of several hours over the last two thousand years.

Over longer timescales it has been found that fossil corals show evidence of annual, monthly and daily growth rings. By counting these rings it has been possible[6, 7] to define the number of days in a year back to about 5×10^8 years in the past. Fig. 1 shows data accumulated[18] recently. The broken line is an extrapolation of the mean rate of slowing down of the Earth based on modern observations such as those plotted on Fig. 2.

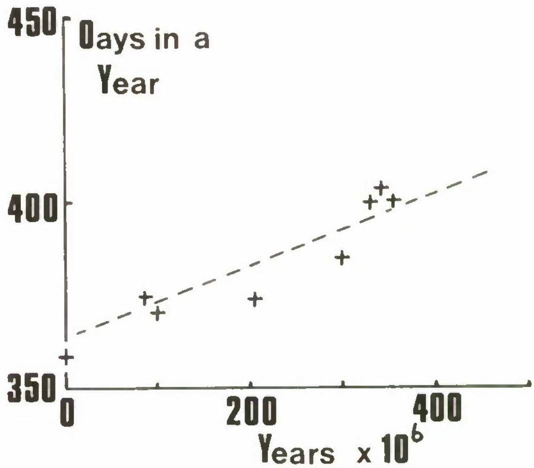


FIG. 1. Coral fossil evidence of lengthening day.

Why is the Earth Slowing Down?

The overall rate of change in the length of the day has many fluctuations superimposed upon it. Certainly, one cause is the transfer of angular momentum from the Earth to the

† 1 arcsecond is approximately 5 μrad.

Moon by tidal action on the Earth's oceans. This effect can be deduced by measuring the relative apparent accelerations of the Moon and planets since the latter are not affected by tidal interactions in the Earth-Moon system. Even when this effect is eliminated the variations in the length of day are complex[8]. Harmonic analysis shows several regular fluctuations. There is a long term variation having a period of about 10 years. The origin is unknown but is suspected to be due to processes in the Earth's core. The Chandler Wobble has a period of 12 to 14 months with an amplitude of about 25 ms in the length of the day. This is almost certainly due to climatic changes in atmospheric density but might possibly be due to pulsations in the Earth's fluid core. There are periodic fluctuations with amplitudes of about 1·5 ms which occur every two weeks and various other small changes occur as a result of the wandering of the Earth's axis of rotation within the main body.

Most of these fluctuations, even if of unknown cause, can be predicted by virtue of their regular nature. When these effects are eliminated we are left with the curve shown on Fig. 2. It can be seen that the excess length of day has two components; a steady increase of about 2·0 ms/cy and irregular fluctuations. This steady rate of increase is consistent with the fossil evidence that the day was about 10% shorter in the epoch of about 5×10^8 years ago. The cause of this slowing down is partially due to tidal interaction between the Earth's seas and the Moon but there is a large part still unaccounted for and various theories have been proposed to match observation with the laws of physics. Theories include:

1. A settling of iron from the mantle to the Earth's core.

2. An expansion of the crust as a delayed reaction to the melting of glaciers after the last ice age.

3. A redistribution of the water between the polar regions and the temperate regions.

4. A long term expansion of the Earth from an initial sphere approximately half its present diameter.

This latter idea, whilst sounding wildly unorthodox, has many attractions. There is a growing body of palaeomagnetic and geological evidence pointing towards a general rate of expansion of the Earth of 0·5 mm/yr; indeed, Creer[9] has assembled no less than twelve separate palaeological estimates stretching back to an epoch $4{\cdot}5 \times 10^9$ years ago. There can be little doubt that the Earth was indeed once a small sphere completely covered with sial and Egyed[10, 11] has shown that subsequent expansion can account for the present proportions of land and sea masses as well as the continental drift. How can the Earth expand in this fashion? (During a recent lecture on this subject a rather frightened member of my audience asked me whether the Earth was due to explode soon!). Fred Hoyle explains this phenomenon by reviving a theory of Dirac[12] in which the force of gravitation attraction, as embodied in the Gravitational Constant G, is assumed to decrease inversely with time since the creation of the universe. This has enormous consequences on the development of the universe, the solar system and life on our planet; this latter because a higher gravitational force in the past would have resulted in a much hotter sun than otherwise expected at about the time that life is currently thought to have originated.

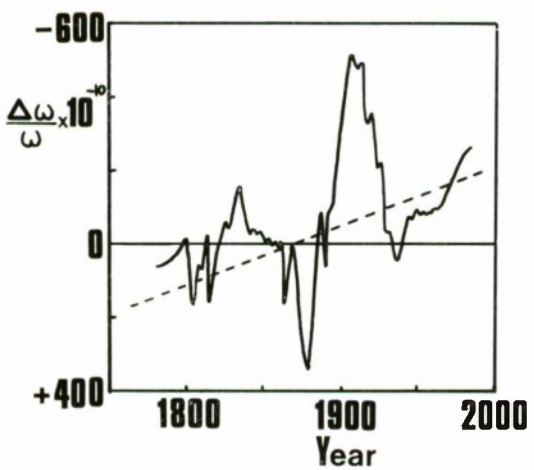


FIG. 3. Results of author's measurements of Moon's motion compared with Theory.

The random fluctuations on Fig. 2 have largely been correlated with earthquake activity[13] and the large deviations on the curve around the end of the 19th century are due to a high activity of explosive seismic disturbances in Asia, of which the explosion of Krakatoa was an example. In 1972 a huge solar flare was observed[14] and it was subsequently found that when the flare erupted the length of the day changed by 10 ms and the Earth rotated at a slower rate whilst the flare was active but reverted to normal when the flare vanished. How can activity on the face of the sun affect the rate of rotation of the Earth? No satisfactory answer is yet available.

We see therefore that a study of the changes in the rate of rotation of the Earth not only has great geological, palaeological and cosmological consequences but also gives us insights into the seismic activity and solar activity cycles, the latter being of increasing importance in the future planning of agriculture[15].

The data used to obtain the curve of Fig. 2 was almost entirely from data collected by amateur astronomers, like myself, who observe with relatively simple equipment from their back-yards. How can this be?

Lunar Occultations and the Amateur

We saw earlier that by comparing the observed and predicted position of the Moon in its orbit using the time scale based on the rotation of the Earth we can arrive at estimates of the difference between UT and ET. This is the basis of Fig. 2. The most accurate method of fixing the position of the Moon against the fundamental coordinate system of the stars is by timing the precise moment that a star either disappears behind the limb of the Moon or reappears after suffering an occultation. The Moon appears to move about 2·5 μrad/s through the stars so that a timing accurate to 0·1 s (achievable by experienced amateurs) fixes the relative position of the Moon and a star to an accuracy of 0·25 μrad. This is far more accurate than any other method except the use of photoelectric timing equipment. (So few photoelectric observations are made that the great mass of amateur visual timings actually carry an overall greater statistical significance[16]. Photographic and micrometer observations are plagued by an effect called irradiation and the only other accurate method of measurement, laser ranging, whilst achieving accuracies of 200 mm in determining the distance of the Moon[17], cannot be used to fix the longitude of the Moon in its orbit relative to the fundamental FK4 celestial coordinate system. Thus, the amateur astronomer reigns supreme.

Keen amateurs are supplied by the Royal Greenwich Observatory with very extensive predictions of occultations for their geocentric coordinates which have to be known to better than ± 50m on the Earth's surface. Accurate timings of events are sent to the RGO where a comparison is made between the predicted positions of the Moon and the observed position. Fig. 3 shows the distribution of residuals obtained from the author's observations over the last six years. It can be seen that the standard deviation is about 3 μrad.

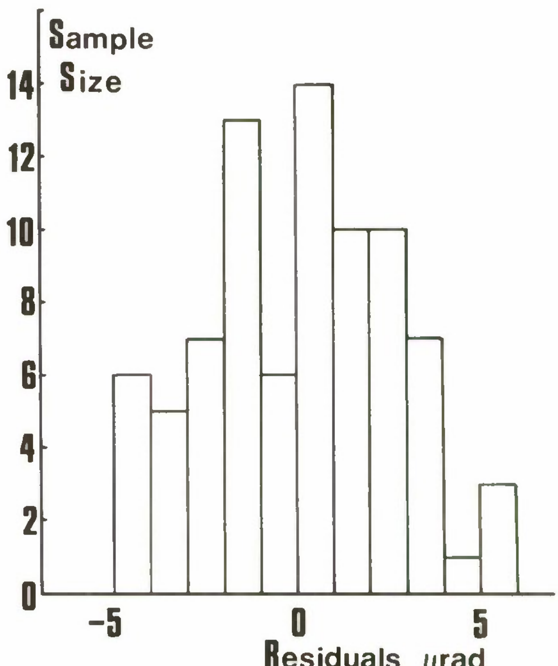


FIG. 2. Recent variations in rotation rate of earth.

This is significantly greater than the error due to the method of measurement which we have seen is about 0·25 μrad. This shows that the technique works. The overall error, after eliminating the acceleration of the Moon which is retrospectively subtracted, is made up of four components. These are[18]

(a) Errors remaining in the numerical integration of the equations of motion of the Moon (± 0·5 μrad).

(b) Errors in the accuracy of maps of the lunar edge profile (± 1 μrad).

(c) Errors in the position catalogued for the various stars (this varies with the brightness of the star, the bright 'fundamental stars' are probably catalogued to an accuracy of better than ± 0·5 μrad but fainter stars, such as form the majority of the author's measurements, are in error by as much as ± 5 μrad).

(d) Errors of measurement, which we have seen are negligible.

By making repeated long-term observations these four factors can be separated allowing the three parameters of interest to be studied separately.

Of great interest are grazing occultations. If an observer places him or herself in exactly the right location on the Earth at the right time a star may be seen being occulted intermittently by the mountains and valleys on the lunar limb rather than by the main body of the Moon. Observers separated by only a few metres may see quite different events since the pattern of flashes of the star is determined by lunar features which may be no larger than boulder size. Obviously such events can only be seen from very restricted parts of the Earth's surface but the RGO make great efforts to issue predictions to enable amateurs to travel to the chosen areas. It was to observe such an event that the author led several groups of keen observers to the wind-swept deserted beach at Hayling Island in Hampshire one freezing winter's night in 1974. Portable telescopes were set up separated by a few hundred metres and the multiple appearances of the star on the darkened lunar limb were timed using tape-recorders and portable shortwave radios tuned to a radio station in East Germany which transmits one-second pulses fixed to the UT time system. The observations were successful and the long preparations and freezing conditions on the night were all worthwhile. The RGO have not, at the time of writing, completed the analysis of the observations but Fig. 4 shows results obtained in 1972. Each point represents an observed disappearance or reappearance of the star from behind the lunar limb and thus a profile of the limb can be built up projected against the fundamental stellar coordinate system. The full curve shows the predicted position of the limb using the best numerical integration of the orbital motion. It can be seen, apart from obvious errors in the mapping of the limb features, that there is a gross error of about 2·5 μrad in the position of the Moon corresponding to a physical displacement of about 1,000 m in the actual position of the Moon in space. The uncertainty in this figure is about 150 m; not bad for amateur observers!

Observations such as those described not only allow the factors described above to be studied, with all their consequences on geology, cosmology, etc., but the specialised observation of grazing occultations has allowed the moments of inertia and the figure of the Moon to be revised showing that the density gradients in the lunar interior are different from those commonly assumed[19].

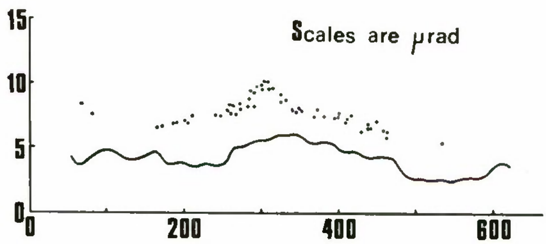


FIG. 4. Comparison of predicted and observed position of lunar limb from Grazing Occultation.

Nature, in characteristic fickleness, does not show regard for the areas of land over which grazing occultations can be observed and in 1972 I led a team onto the fringe of a tank training area in Dorset. Unhappily the Army were carrying out night manoeuvres and we were accosted, not only by the police, but by a tank whose crew were somewhat sceptical of our intentions! I have also noted an aggressive attitude to develop in courting couples when their darkened cars, deep in the country, are suddenly surrounded by a lot of scruffy enthusiasts with shortwave radios, tape recorders and telescopes. However, as long as we get the results who cares about the danger? Amateur astronomers are not easily deterred!

References

(1) The Application of Ancient Astronomy to the Study of Time. Newton, R. R. *Endeavour*, **33**, 118 (1974) 34.

(2) The Secular Accelerations of the Moon's Orbital Motion and the Earth's Rotation. Morrison, L. V. *The Moon*, **5** (1971) 253.

(3) Ancient Astronomical Observations and the Accelerations of the Earth and Moon. Newton, R. R. John Hopkins Press (1970).

(4) The Earth's Acceleration as deduced from Al Biruni's Solar data. Newton, R. R. *Memoirs Roy. Astron. Soc.*, **76** (1972) 99.

(5) The Earth-Moon System. Marsden, B. G. and Cameron, A. G. W. Plenum Press, New York (1966).

(6) Coral Growth and Geochronolgy. Wells, J. W. *Nature*, **197**, 948.

(7) The Implications for Geophysics of Modern Cosmologies in which G is Variable. Wesson, P. S. *Qtly. J. Roy. Astron. Soc.*, **14** (1973) 9.

(8) The Rotation of the Earth. Munk, W. and MacDonald, G. Cambridge University Press (1966).

(9) Earth, Possible Expansion of, Creer, K. M. *International Dictionary of Geophysics*, p.383, Pergamon Press Ltd. (1967).

(10) Paleomagnetism and the Ancient Radii of the Earth. Egyed, L. *Nature*, **190** (1961) 1097.

(11) The Expanding Earth ? Egyed, L. *Nature*, **197** (1963) 1059.

(12) The History of the Earth. Hoyle, F. *Qtly. J. Roy. Astron. Soc.*, **13** (1972) 328.

(13) Earthquakes and the Rotation of the Earth. Anderson, D. L. *Science*, **186** (1974) 49.

(14) A Solar Storm that Shook the World. Gribben, J. *New Scientist* (10 May 1973) 339.

(15) Agriculture and Sunspots. Anon. *Nature*, **252** (1974) 2.

(16) A Comparative Study of Visual and Photo-electric Timing of Occultations. Morrison, L. V. *Highlights in Astronomy* (1971) 589.

(17) The Lunar Laser Ranging Experiment. Bender, *P. L. et al. Science*, **182** (1973) 229.

(18) Evolution of the Moon's Orbit and the Origin of Life. Turcotte, D. L. *et al. Nature*, **251** (1974) 124.

(19) Secular Changes in the Lunar Elements. Martin, C. F. and Van Flandern, T. C. *Science*, **168** (1970) 246.

Geoffrey J. Kirby graduated from London University in 1960 and worked for six years with the Atomic Energy Authority. In 1966 he transferred to his present establishment within the Ministry of Defence where he now heads a simulation and system modelling group as a Principal Scientific Officer.

Commissioning of H.M.S. *Antelope*

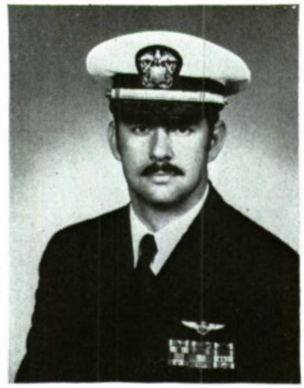
H.M.S. *Antelope*, second of the Royal Navy's new Type 21 frigates, was commissioned at Woolston, Southampton, on Saturday, 19th July, 1975.

The Commanding Officer is Commander Nicholas John Hill-Norton, R.N., of Alton, Hampshire, only son of Admiral of the Fleet Sir Peter Hill-Norton and Lady Hill-Norton.

After commissioning the ship will be based in Devonport to carry out sea trials. She is the 10th ship to bear the name. The first, a 38-gun galleon, saw action against the Spanish Armada in 1588, and the ninth, a destroyer, saw World War II action against the *Bismarck*, in the Malta convoys and in the North Africa landings.

The new H.M.S. *Antelope* was built by Vosper Thornycroft at their Woolston shipyard, laid down on 23rd March 1971 and launched in March 1972. Designed by Vosper Thornycroft Ltd. in collaboration with Yarrow (Shipbuilders) Ltd., H.M.S. *Antelope* is powered by gas turbines and armed with the *Seacat* missile system, 4·5 in. gun, and will carry a Lynx helicopter. H.M.S. *Antelope* has a crew of 158 officers and men.

SMART AIRCRAFT, SMART WEAPONS —A COMPARISON



Lieutenant Commander Robert Smith, U.S.N.

" Because of the enormous quantum jump in technology, the past decade has witnessed a virtual revolution in the art of waging conventional war ".

Air Vice-Marshal Stewart Menaul.[1]

" . . . that final arbiter of effectiveness (is the) ability to hit the target ".

Air Vice-Marshal P. De L. Le Cheminant.[2]

The object of offensive air warfare is to destroy the target. The number of bombs required to destroy a target is inversely proportional to the lethal radius of high explosive in the bomb and the delivery accuracy. The aim of this article is to review some recent developments in weapons system technology and to compare their effectiveness in various aspects of offensive air warfare.

During World War II, bomb delivery was relatively inaccurate. As a result, large numbers of bombs were required to achieve the desired degree of destruction of area targets such as German industrial complexes. Point targets such as bridges were even more difficult to knock out. The large number of aircraft required to deliver these bombs resulted in many aircraft losses. As weapons delivery accuracy was limited by the unsophisticated aircraft and delivery techniques of the period, larger bombs were developed to increase effectiveness. The 25,000 lb " block buster " bomb was produced and proved valuable in many instances. The effort to further increase the lethal radius of the explosives in a bomb continued until eventually the nuclear fission weapon was produced. Although the destructive capability of this weapon was proved conclusively in Japan in 1945, political and humanitarian considerations seem likely to restrict any future use of nuclear weapons to a deterrent role.

Despite the deterrent effect of nuclear weapons, conflicts between nations have continued and have resulted in limited wars in which offensive air power has become increasingly important. In the Korean conflict of 1950 - 1953, conventional bombs, rockets, and napalm were used extensively. However, the iron bombs of this period were relatively unchanged from World War II and delivery techniques were still inaccurate. Very little progress was made in improving the accuracy of air to ground weapons delivery during the remainder of the 1950's and early 1960's. In fact when United States Air Force and Navy pilots began conducting offensive air operations over North Vietnam in 1965, weapons and delivery methods had changed very little from the days of the Korean conflict. Large number of conventional bombs were required to achieve the desired destruction of interdiction targets such as bridges, POL storage areas, and truck convoys. Intense Anti-Aircraft Artillery (AAA) and Surface to Air Missile (SAM) fire resulted in high losses and many aircraft were required in supporting roles such as AAA and SAM suppression in order to permit bombing aircraft to penetrate these defences. These support requirements meant that fewer aircraft were available to carry bombs on each strike. Restrikes were often required to destroy the target. This approach was clearly inefficient and costly and some method had to be found to increase effectiveness.

As a result, the United States Air Force and Navy directed considerable effort toward improving the accuracy of weapons delivery and the destructive power of the weapons themselves. Missile warheads of expanding rod and fragmentation design were produced to increase the destructive capability of small amounts of explosive. A new explosive technique was developed using a Fuel Air Explosive (FAE) in which a fuel/air mixture was allowed to form and was later detonated. FAE has been particularly effective against soft targets. In addition, the U.S. developed a family of area weapons called " cluster bombs " designed to be effective despite relatively inaccurate delivery. Although some increase in effectiveness was gained through warhead improvements, the greatest potential for the destruction of the target with fewer aircraft was to improve the accuracy of weapon delivery.

Weapons accuracy is usually expressed as a 50% probability of a weapon hitting within a particular distance from the target. This value is referred to as the Circular Error Probable (CEP) and is usually indicated in feet or yards. When the American bombing began in Vietnam, the U.S.A.F. cited in Congressional testimony its ability to achieve combat CEP's of 750 ft. By 1968, pilots of *Republic* F-105 fighter bombers were achieving CEP's of 365 ft against lightly defended targets, 540 ft against moderately defended targets, and 575 ft against heavily defended targets. The same pilots and aircraft were able to get CEP's of 125 ft while training at Nellis Air Force Base in Nevada[3]. By 1971, aircraft CEP's had decreased to 250 ft in combat with aircraft such as the *A4E Skyhawk* and *A7A/B Corsair* II. Although this increase was a step in the right direction, CEP's measured in tens rather than hundreds of feet were desired.

One of the first highly accurate air to ground weapons in the United States was the guided missile *Bullpup*. This missile was operational in the early 1960's and was guided to the target by the pilot using radio signals from the aircraft to the control section of the missile. A flare in the tail of the missile allowed the pilot to visually track the missile to the target. The missile was very accurate in peacetime firings (20 - 30 ft CEP). However, it had a short range and the pilot was required to fly his aircraft nearly wings level during guidance leaving the aircraft vulnerable to target defences in combat. As a result, the accuracy of *Bullpup* was poor when used in North Vietnam in 1965 - 1967. Subsequent missile work was directed at achieving greater stand-off range and/or some sort of terminal guidance to improve aircraft survivability and weapon accuracy. This weapon research resulted in a new group of weapons called " smart weapons ".

One of the most successful " smart weapons " was developed through the use of lasers, an acronym for light amplification by stimulated emission or radiation. Soon after Dr. Theodore H. Maiman invented the first laser in 1960, he called it, " a multi-million dollar solution in search of a problem ".[4] Many military and commercial problems have subsequently been solved using lasers. Military applications exist for both the high energy and low energy lasers. High energy gas and chemical lasers are currently being studied for possible use as a " death ray ". Such a laser would produce an intense burst of power to ignite, melt, or vaporize an object.[5] However, high energy lasers are only experimental and it is the low energy laser which is currently of most value to the military. Low energy lasers are excellent for range finding and target marking. The principle involved in laser weapons is that the target is first illuminated (designated) with laser energy and then a laser seeker in the nose of the weapon receives the energy reflected from the target. The weapon homes in on this reflected energy which provides a very accurate form of terminal guidance. Two types of laser weapons used in Vietnam were the *Bulldog* missile and the *KMU*-351/*B* Laser Kit. The *Bulldog* is a

Bullpup missile modified with a laser seeker head. The *KMU*-351/*B* kit which consists of a seeker head and control section becomes a highly accurate glide bomb when fitted on a standard *Mk*. 83 1000 lb or *Mk*. 84 2000 lb bomb. In practice, these laser weapons have been very accurate usually scoring a direct hit.

The improved accuracy of bombing using laser weapons meant that a greatly reduced number of bombs were needed in Vietnam to destroy most targets. In addition, a number of important targets which had been "off-limits" were made available for attack. For example, the Lang Chi Hydroelectric Power Plant previously "off-limits" due to its proximity to an adjacent dam, was successfully attacked using laser bombs.[6]

A second group of "smart weapons" used in North Vietnam is called the electro-optical weapons. This group which includes the *Maverick*, *Walleye*, and *HOBOS*, homes on the video contrast of a target. Both *Maverick*, which is rocket powered, and *Walleye*, a glide bomb, use a television camera mounted in the nose of the weapon to view the target. The pilot first locates the target visually and then orients the aircraft to bring the target within the field of view of the weapon. The picture as seen by the weapon is also displayed in the aircraft cockpit. The pilot centres the camera's cross-hair sight on the target and launches the weapon when lock-on is achieved. After release, the weapon provides its own guidance to the target. The *HOBOS* (Homing Bombing System) is a kit similar to the laser kit described earlier. It consists of a TV guidance section and a control section which can be attached to either the *Mk*. 84 or the U.S.A.F. *M*118*E*1 3000 lb bomb. As in the other electro-optical weapons, the aircraft is free to manoeuvre after release due to the weapon's terminal guidance.

The electro-optical weapons, like the laser guided weapons, are very accurate. They proved especially effective against railroad bridges which have traditionally been difficult to destroy.[7] One particular advantage of these weapons is that guidance is internal and the delivery aircraft is free to manoeuvre once release has occurred. The use of magnification in connection with the TV camera permits a good stand-off capability. However, these advantages are partially offset by other factors. Since these weapons rely on visual means to acquire and track the target, their performance is significantly degraded by adverse weather. This same factor also limits their use to daylight hours. However, a night capability is possible through the use of low level light TV or infra-red guidance.[8]

Electro-optical guidance has also been selected for use in the Anglo-French *Martel* missile. This missile significantly differs from the other TV weapons in that guidance is accomplished from the delivery aircraft using data-link. Data-link guidance permits a good stand-off capability and a low altitude launch capability. A second version of *Martel* homes on radar signals from a target radar. This technique was also used in North Vietnam by a series of U.S. anti-radiation missiles.

The introduction of anti-radiation missiles such as *Shrike* and *Standard Arm* into North Vietnam marked the beginning of a trend toward the development of specialised "smart weapons". One of these specialised weapons is *Harpoon* which is designed for anti-ship attack. *Harpoon* is intended to be launched from a surface ship, submarine, or aircraft and to fly a powered low altitude profile using a radar altimeter to monitor missile height. An active frequency agile radar guidance system is included to provide all weather target acquisition and terminal tracking. The missile's radar frequency agility and computer logic are designed to give a counter counter measures capability. In the last few seconds before striking the target, the missile is programmed to leave its wave skimming altitude and pitch up so that it can hit the target in a shallow dive. This close in manoeuvre should allow *Harpoon* to better avoid target defences and also enhance the effectiveness of the warhead.[9]

The anti-radiation and anti-ship missiles are only two examples of the trend toward specialization in "smart weapons". The wire guided *Tow* anti-tank missile is another. The exceptional accuracy of these weapons does make them valuable for destroying very important targets. But, they are also very expensive. The *KMU*351/*B* laser kit costing about $3,000 per kit is the least expensive of the "smart weapons". The *HOBOS* costs about $15,000 and the *Maverick* and *Walleye* are in the $20,000 to $30,000 range.[10] This high cost of "smart weapons" limits their use to only the most critical of targets reducing flexibility. If "smart weapons" are to be more flexible, methods must be found to decrease their cost.

During the initial phases of the development of "smart weapons", the Department of Defence permitted separate development projects by each of the military services. This approach ensured that many of the possible applications of "smart weapons" were explored and also reflected the Pentagon's desire to get "smart weapons" into action in North Vietnam as quickly as possible. This resulted in some duplication of effort and in a few cases similar weapons were produced. The U.S.A.F. *Maverick* and the U.S. Navy's *Walleye* are a case in point. However, individual services are currently acting as project managers for particular types of weapons and U.S. government studies are being conducted to determine which uses of "smart weapon" technology show the greatest promise. One approach which should help reduce the cost of these weapons is called the "Modular Weapons Concept".

The "Modular Weapons Concept" recognises that the trend toward developing specialised weapons is expensive, inflexible, and has immense logistical implications. The "Modular Weapons Concept" would provide a group of warheads, fusing components, guidance seekers, and aerodynamic components which are interchangeable and are optimised for differing combat conditions. This system would provide a wide range of target acquisition methods coupled with the most effective warhead and an appropriate method of delivery.[11] These weapons could be manufactured as "all-up" rounds or as separate components which could be assembled at the station or ship level much like the laser kit.

"Smart weapons" rely on sophisticated and expensive components which are destroyed with the weapon. Another approach to improving the accuracy of weapons delivery is to concentrate the "smart" portion of the system in the delivery aircraft rather than in the weapon. In this way, the weapons could be less expensive and if delivery accuracy is very good, standard iron bombs could be used effectively.

In the past iron bombs were dropped using a manual type of delivery. With the manual technique, the pilot had to decide what delivery he would use (*i.e.*; 45° dive, 5000 ft release, 450 kts airspeed) and then calculate a sight setting for the aircraft optical gunsight. At the target, he had to fly the aircraft to the proper roll in position to establish the aircraft at the predetermined dive angle. When actually in the dive, he had to control the aircraft to reach the desired release height at the proper airspeed and with the gunsight positioned correctly in relation to the target. Any deviations from the desired dive angle, airspeed, or release height resulted in considerable bomb errors. Strong or unexpected wind in the target area also degraded accuracy. Ground attack pilots needed continuous training to develop the expertise necessary to perform manual weapons delivery at an acceptable and predictable CEP. Poor weather often resulted in aborted runs or degraded accuracy. In addition, the accuracy achieved in training could not be expected in combat due to the difficulty of target acquisition and weapons delivery in a hostile environment. The "smart aircraft" increases the accuracy of iron bomb delivery by reducing the load factor on the pilot. The aircraft system provides penetration aids such as accurate navigation information to the target, solves the weapons delivery problem, and releases the bombs. The aircraft also presents the navigation and weapons delivery information in a simple manner. As the aircraft takes over the tasks of navigating to the target and delivering the weapons, the pilot is free to concentrate on tactical flying, target acquisition, and avoiding defensive fire.

The basic requirements for a "smart aircraft" are an inertial measurement unit which senses aircraft movement (including the effects of wind) and a computer which solves the navigation and weapons delivery problems. A Heads Up Display (HUD) usually serves as the method for the pilot to designate the target for the computer and displays the navigation and weapons delivery information. Other components such as doppler, forward looking radar, and a projected map display give the aircraft additional mission capabilities. The "smart aircraft" is not restricted to any predetermined type of delivery as the computer can quickly and accurately solve for a wide range of dive angles and airspeeds. This means that the pilot can manoeuvre around adverse weather, avoid AAA fire and attack targets of opportunity without degraded accuracy.

The *A7E Corsair* II, deployed with the U.S. Navy, was the first " smart aircraft " to see combat. During training, the *A7E* demonstrated that iron bomb delivery could be extremely accurate and could in fact consistently achieve hits measured in tens rather than hundreds of feet. Inexperienced *A7E* pilots attained CEP's of 25 ft on their first bomb runs during training at Yuma, Arizona[12]. The *A7E* was able to achieve ten mil accuracy while deployed in combat in Vietnam. Ten mil accuracy means that an iron bomb dropped at 8,000 ft slant range would hit within 80 ft of the target. Half of this error (40 ft in this case) is related to bomb dispersion from irregularities due to the manufacture of iron bombs and not inaccuracy of the aircraft system.

The improved accuracy of " smart aircraft " has greatly increased conventional weapons effectiveness. The use of the *A7E* in North Vietnam resulted in a significant decrease in the number of bombs required to destroy hard to hit targets such as bridges and power plants. Therefore, fewer aircraft were needed for each strike decreasing losses due to hostile fire. In addition, the ability of an aircraft such as the *A7E* to carry a wide selection of weapons to a variety of targets with a minimum of preplanning provides the flexibility so desirable in offensive air warfare.

A major disadvantage of the " smart weapons " is their lack of flexibility. Experience of U.S. Navy pilots in North Vietnam demonstrated that the use of " smart weapons " requires meticulous preplanning. Good target intelligence is required especially for the electro-optical weapons. This means that " smart weapons " are not well suited for targets of opportunity such as are attacked on armed reconnaissance missions.

Another important role of offensive air operations is close air support in the battle area. Targets normally encountered in close air support are groups of personnel, gun positions, and individual armoured vehicles. " Smart weapons " are not suitable for these operations. Laser designation is difficult for battlefield targets and electro-optical weapon preplanning is impossible in a fast moving tactical situation. Considering the cost of "smart weapons", their use against personnel or gun positions seems hardly cost effective. Conversely, the " smart aircraft " is particularly effective in close air support. In South Vietnam, the *A7E* proved its ability to deliver an assortment of weapons very accurately even in adverse weather.

" Smart weapons " cannot be used in large numbers against any single target of average size. Once weapons start going off, the target is obscured by smoke and dust. The smoke cloud, which can be quite large, makes laser designation inaccurate and also destroys target contrast increasing the CEP's for both laser and electro-optical weapons. For this reason, " smart weapon " strike groups must achieve all impacts simultaneonsly. Targets such as runways, fuel dumps, and aircraft on the ground require accurate weapon delivery but also need some area coverage. In this case, accurate strings of bombs are more effective than a few large bombs. An important aspect of the flexibility of " smart aircraft " is their ability to deliver either large bombs such as the *Mk.* 84 or strings of smaller bombs such as the 500 lb *Mk.* 82.

How " smart " each part of a weapons system should be is a point of unresolved contention among military planners striving to improve the accuracy and effectiveness of aerial weapons delivery. At issue is the question of how much capability, hence cost, should be invested in the aircraft or the weapon itself. Both " smart aircraft " and " smart weapons " can consistently achieve exceptional accuracy which means that in the future, fewer aircraft will be required to destroy targets. Even though " smart weapons " tend to be less flexible, their accuracy may be desirable for high priority targets such as ships which justify the use of a costly weapon. However, the operational flexibility of the " smart aircraft " is invaluable in all types of offensive air operations, especially in quick reaction situations such as armed reconnaissance and close air support.

References

(1) *Defence* (January 1973).
(2) *Army Quarterly* (July 1970).
(3) *Aviation Week* (3 May 1972),. 49.
(4) *New Scientist* (13 July 1972), 86.
(5) *Aviation Week* (14 August 1972), 13.
(6) *Ibid* (3 July 1972), 15.
(7) *Ibid* (29 May 1972), 16.
(8) *Ibid* (7 May 1973), 44.
(9) *Ibid* (7 May 1973), 18.
(10) *Ibid* (22 May 1972), 17.
(11) *Air Force* (August 1972), 32.
(12) *Aviation Week* (5 April 1972), 20.

TIME FLOW IN THE DIVISIONAL BATTLE: THE RARDE COMPUTER-ASSISTED WAR GAME

R. Beresford
Royal Armament Research and Development Establishment

Introduction The Divisional War Game at the Royal Armament Research and Development Establishment was originally designed to generate data for weapon assessment purposes and, in particular, to investigate the effects of introducing new weapon systems with novel characteristics into the battlefield environment, where the influence of tactics and terrain play an important part. It is a closed, two-sided game, manually implemented, but with computer support for record and limited housekeeping purposes, and is played with symbolic pieces on a 1/25,000 scale map model.

To facilitate management and control, the game originally progressed in time-slices of five minutes' duration, and the rules, including, for example, vehicle movement, fire unit deployment, target acquisition and weapon/target interactions, were modelled as accurately as possible within this time interval.

Whilst this game was adequate for the purpose of examining the characteristics of weapon systems against a battlefield background, using attrition as a criterion of effectiveness, it became apparent that the flow of time as the game progressed was of a complex nature and contained anomalies which made the interpretation of specific time aspects of the battle unreliable; for example, weapon response time or the delaying effect of obstacles. When a period of gaming was complete, it was invariably found that the military judgement of the players who took part indicated that, say, 30 five-minute game cycles (*i.e.* a 2½ hour period of game time) would be more like 12 hours in real battle based on movement of the FEBA and the general magnitude of military operations carried out.

The reasons for the discrepancy were established as follows:—

(a) Command, Control and Communications effects were not present.
(b) Activity times were "active times" *e.g.* bridge-building or minefield breaching times were based on the basic time for carrying out the operation, with only a rudimentary allowance for the "dead time" resulting from causes such as casualties, fatigue, equipment breakdown and the general fog of war.
(c) In the game, one man carried out the role of all the commanders in his force *e.g.* division, brigade, battle group and combat team, and there were no delays based on the actual time it takes to assimilate and sort information and then make a decision to carry out a particular course of action.
(d) The "time-slice" aspect of game progress contributed to the time discrepancy in that it was possible for the player/commander to look ahead and plan exactly to synchronise future events which, in practice, would have an element of uncertainty in the time of their occurrence.

To improve the representation of time, two activities are in progress at RARDE:—

First, studies of the occurrence and magnitude of the delays contributing to the extension of battle time, and, second, development of the techniques necessary to handle these delays. Feasibility studies followed by Project Definition have shown that modern computer technology offers the possibility of introducing the additional complexity necessary with, in particular, the essential changeover from "time slice" control to event-sequencing, where the game time flows evenly, events being implemented as they occur, delays being automatically included after sampling from the appropriate time distribution. A pilot game on a small computer has confirmed that this process can be carried out.

Computer Assistance

The second of these two activities is planned to be complete in the Autumn of 1975 when a new Divisional War Game with massive computer support will become operational. Utilising event-sequenced control, the game will use live commanders, since the tactical impact and lessons learned from the exercise of the command function have provided invaluable information throughout the history of war gaming. The commanders, it is considered, should operate in an environment which closely duplicates their operational function, *i.e.* their information and situation displays should be the same as on the battlefield. This is particularly important since the evaluation of ADPS and Command and Control techniques appears within the scope of future applications of the game (the training aspects, also, must not be forgotten). To this end, provision is being made in the game system for the utilisation of players at various command levels with the appropriate filtering and aggregation of orders and information. The resolution of the game will be down to Combat Team level.

The following assessment and simulation models are being included in the system:—

Force Movement (including terrain representation)
Order Input
Intervisibility
Reports
Barriers
Direct-Fire
Air-Ground-Air (incl. helicopters)
Surveillance and Target Acquisition
Artillery.

The main problem of studying and applying the occurrence of time delays will now be considered. The project falls into two parts; first, studies of the order and information flow structure requirements and their implementation, and second, collection of appropriate quantitative data.

Order and Information Flow

A programme of work is in progress with the following requirements:—

(a) to increase realism in the game order and information flow structure:
  (i) The various command and control chains have been described, *e.g.* Infantry, Armour, Artillery, Engineers.
  (ii) Appropriate communications systems have been defined in terms of their operational characteristics.
  (iii) A Command, Control and Communications Information Handling model is being developed incorporating these chains and systems with appropriate delays for use in the war game and taking into account the preparation, queuing, transmission and assimilation of messages.
(b) a game-produced information management system is being explored which will possess a filtering capacity so that game commanders will receive reports which resemble military messages and are in the correct timescale.

Software specifications will be developed for correlating and summarising information.

As a result of this work, therefore, a facility for implementing delay and time extension will operate at the centre of the game structure. However, this is only part of the problem. The remainder resides in the collection of data for use in the model.

Data Collection

The basic problems of data assembly lie in the sources available. Stated quite simply, the requirement is to obtain measures of the times taken to carry out various battlefield activities, ranging from signals traffic and the organisation of movement to the laying of minefields and their breaching, and many other tasks. Some of these activities are, on the surface, straight-forward, such as the building of a bridge or the execution of an artillery shoot, and can be measured in peace time. Others present more difficulty, for example, the withdrawal of a combat team under fire.

The modification of these timings in the real battle environment, where fatigue, poor training and loss of morale can have an effect, is the crux of the matter. There are various sources of information that can be explored, such as:—

(a) The Field Army and Arms Schools (*e.g.* Infantry, Armour, Artillery), which can provide practical estimates modified by military experience.

(b) Establishments working on future equipment problems (*e.g.* communications systems, armour and anti-armour interactions).

(c) Field Training Exercises, which provide a rich source of information which is, however, unavoidably lacking in the realism of war and which require close contact with the particular exercise to distinguish between real effects and the artificial constraints introduced to achieve the training aim.

(d) Field Experiments, which do allow a higher degree of control, but which are limited in size and which still have some artificiality.

One way of obtaining a measure of this elusive realism is by carrying out "market research" and submitting a questionnaire defining the basic military activities, such as combat team withdrawal, occupying a defensive position, preparation and issue of operational orders, to a large number of military experts who possess wartime experience.

In the present project, data collection is regarded as a continuous process where the various values are under constant review. Typical examples of information collected are shown in the following tables.

Communications Delay

From	*To*	*Event*
Unmanned Sensors	Monitoring	Sighting
Observation Post	Agency	
Recce Reports	,,	,,
Combat Team	H/Q	,,
Forward Air	Battle Group	,,
Controller	Brigade H/Q	,,
Brigade H/Q	Division H/Q	SITREP
Division H/Q	Brigade H/Q	TAC ORDERS

The distributions obtained, as Mean, Minimum and Maximum values, were derived from simulation of typical communication systems and can be modified according to priorities and means of communication.

Artillery Timing

These are divided into basic unit activities and time effects, the first derived from discussions with appropriate experts and the second from communications simulation.

Deploy (before action)
Prepare to move (after action)
Move to alternate position
Response Times (Request—Time on Target)
,, (Registered Target)
Registration Time.

Barriers and Engineers

(a) *Bridges.* No bridge construction is allowed to commence until a stated time interval after the arrival within 1 Km of the site of reconnaissance units. The purpose of this rule is to allow for reconnaissance, and the time distribution was established in discussion with Engineer specialists. The time for a specific bridging operation is obtained by sampling from this distribution.
The construction time is then estimated for various widths of river and type of site, whether "easy", "difficult" or "very difficult" (a distinction established as part of the permanent terrain description within the computer).

The problem of a standard speed for vehicle movement and the resulting ease of prediction of arrival time was referred to earlier in this article. In the search for improved realism and time flow, a new approach has been adopted where movement is based on the orders input by the game commanders, who may specify a change of unit position from the currently-occupied square to any other square on the game area. The distance covered is calculated by straight-line interpolation, and an average value for the time taken to cover this distance is derived from the "going" in the squares traversed and a table of the speeds of the appropriate unit under the particular "going" conditions. The actual time to cover the calculated distance is then obtained by sampling from a rectangular distribution centred on the previously-calculated average and of a spread to be established later. Initially, this spread of the distribution will be arbitrarily selected as 30% of the mean value and will be set up at the start of a game series.

Battlefield Activities

Discussions were held with appropriate Arms Schools (Armour and Infantry), and a questionnaire was submitted to the army in the field to obtain information on the timing of unit activities on the battlefield. Examples of information obtained are shown in the following table:—

Unit Attack
Contact Report
Commanders Reconnaissance
Warning Order
Orders
Forming up
Deploy

Attack
Fight through
Reorganise
Combat Team defence against an Advanced Guard
Combat Team defence against an armoured battalion
Battle Group on Combat Team halt and camouflage
Battle Group on Combat Team moving off
Movement speeds—fire units
—headquarters
Preparation and Issue of Operation Orders.

Conclusion

This article has attempted to describe the extent and complexity of the work which is in progress at RARDE to develop a war game facility where the flow of battle time is realistic. It cannot be claimed that when the exercise is complete, the time flow representation will be comprehensive, for how can this be evaluated? However, by providing a system in which appropriate planning and communication delays are included in the simulation of military activity, it is considered that the quality of the results obtained from the game will be greatly increased and an understanding of the time-orientated interactions between various strata of the battle will become possible.

Finally, one aspect of military operational analysis which has emerged again during the course of this work is the strong interaction between game development and the need for information describing various aspects of military operations. The requirement for game rules generates proposals for the collection of specific operational data or even for specially arranged field trials to determine key inputs. At the same sime a close watch has to be kept on the results of field trials carried out for other purposes so that rules can be kept under review and updated as the experimental data is improved. This close liaison between modelling combat operations and field studies is essential, and must continue.

Swedish King Becomes Honorary Admiral

The appointment of King Carl XVI Gustaf of Sweden as a honorary admiral in the Royal Navy with effect from 25th June was announced in the London Gazette on Tuesday, 8th July.

King Carl Gustaf wore the uniform of a British admiral for the first time on Tuesday when he arrived in Edinburgh for a State Visit to Britain.

The only other honorary admiral in the Royal Navy is King Olav V of Norway, who was granted the rank in 1958.

The last King of Sweden, King Carl Gustaf's grandfather, also held the rank, and so did his predecessor. The kings of Sweden have long been known as " Sailor Kings ".

The present king, who is also a Swedish admiral, served in the Royal Swedish Navy while he was Crown Prince.

REMOTELY PILOTED VEHICLES THE MILITARY AIRCRAFT OF THE FUTURE?

Squadron Leader D. A. H. Edwards, C.Eng., M.I.E.R.E., R.A.F.

Introduction

For political reasons, United Kingdom defence expenditure is limited, at the time of writing, to approximately 5½% of the gross national product (GNP). The Defence Estimates for 1973 - 74[1] total £3,365 million and represent 5·75% of the GNP. Of this sum, 37% is allocated to Expenditure on Equipment and 47% to Expenditure on Personnel. The latter figure reflects a trend of increasing manpower costs over recent years, and the indications are that these costs will continue to rise; this situation is compounded by the increasing cost of military equipment, particularly in the case of military aircraft as improvements in performance and weapon delivery accuracy are sought. Furthermore, the trend among the West European countries has been towards a progressive decline in defence expenditure, and it is a reasonable assumption that, following Britain's entry into the European Economic Community, economic pressures will tend to push down defence spending towards the French, German, and Italian level of 2·6 - 3% of the GNP[2]. Consequently, unless there is a change of political emphasis on defence expenditure, our military capability will be considerably reduced over the next few decades. If we are to retain a credible defence capability new ways must be found to achieve the same, or better, results for less money. Military aircraft, with their high capital, operating, and maintenance costs, are an obvious area for improvement.

Over the past decade, the United States has had some success with the development of target drone aircraft as vehicles for operational reconnaissance and electronic warfare in South East Asia. More recently, improvements have been made in the capability and reliability of these vehicles, largely as a result of the technological 'spin-off' from the aerospace industry, to the extent that the remotely piloted vehicle (RPV) now appears to be capable of development as a practical alternative to the manned aircraft. More importantly, however, American studies have indicated that new materials and methods of manufacture, together with the cost benefits of mass production, will make it possible to build RPVs for some types of mission much more cheaply than manned aircraft.

It is therefore an opportune moment to look at the RPV as a possible replacement for, or adjunct to, manned military aircraft. This article examines first the case for, and then the practicality of, the RPV, suggests applications appropriate to defence requirements, and then indicates the way in which the RPV should be developed. The discussion is confined to the context of British and European defence needs.

The Case for the RPV

The case for the RPV is based on the limitations of manned aircraft in terms of performance, the necessity to provide life support and survival equipment, and weapon delivery accuracy. These factors will be discussed from the cost and effectiveness points of view.

The limit of human tolerance to 'g' forces is the biggest constraint on aircraft manoeuvrability. The removal of the pilot from an aircraft would permit greatly increased effectiveness: in air-to-air combat much tighter turns could be made, so improving combat capability; in dive-bomb attacks a later pull-out would be possible, with consequent increase in weapon delivery accuracy; and more violent escape manoeuvres would increase the difficulty of tracking by enemy guns or missiles.

Aircrew survival and life support needs have a considerable impact on the design of modern aircraft. The requirements for ejection seats, pressurization and breathing systems, duplication of equipment and services, and demanding safety standards contribute significantly towards their complexity and cost. Furthermore, the size of these systems, together with the volume of space necessary to accommodate the aircrew, influences the ultimate size of an aircraft and hence its vulnerability. The removal of aircrew and their associated support systems would therefore offer substantial cost and weight savings, and permit smaller aircraft to be built.

Considerations of aircrew safety necessitate low loss rates for offensive operations, not only for humanitarian reasons but also because of the heavy capital investment both in aircraft and in aircrew training. The significance of loss rate figures can perhaps be understood if it is considered that a daily loss rate of 3% means that a pilot and aircraft have only a 50% chance of surviving for about three weeks; clearly, such a loss rate would be unacceptable save in extreme circumstances. However, the achievement of a low loss rate requires that weapons be delivered from ranges beyond target defences, with a consequent reduction in accuracy. To compensate for this reduction more weapons must be delivered; this requirement means a move towards the saturation type of raid, which is both expensive in terms of capital equipment and inefficient in the use of weapons and weapon delivery systems. An indication of the high cost of target destruction by manned aircraft is given by the following example: an attack by Phantom aircraft carrying eight 1000 lb bombs against a point target, using a circular error probability (CEP) of 100 yds and with a loss rate of 1%, would cost 1·1 million dollars for a 99% probability of destruction[3]. It may be deduced from this example that it would be worth spending the same amount of money on an expendable RPV to achieve the same result. It is also evident that such a move would have a significant advantage in removing the risk to aircrew.

In theory, however, the economics of the RPV appear much more favourable than a simple break-even cost-effectiveness comparison. For example, it has been calculated that an RPV dropping two 1000 lb bombs at 1000 ft slant range from a 45° dive would achieve a CEP of 30 yds and an 85% probability of target destruction. The target area of such a vehicle would be between 1% and 10% of that of a manned aircraft; by executing a 10g pull-out at bomb release and overflying the target at 400 ft, the loss rate to local defensive fire would be comparable to that of manned aircraft at 10,000 ft bombing range[4]. If a pessimistic loss rate of 25% is assumed for this attack, an RPV cost of four times the price of equivalent manned missions could be tolerated, and costs would still break even. However, the removal of the aircrew permits an entirely new approach to weapon delivery in terms of a functional system which is optimized in size, performance, and weapon load capacity, and RPVs could be produced for between 25,000[5] and 50,000[6] dollars, while re-usable vehicles would cost in the order of 10 times these figures[7]. If the cost of a Phantom aircraft is taken to be three million dollars, these figures represent 60 - 120 expendable RPVs, or 6 - 12 multi-use RPVs, for the same amount of money.

On the basis of this data, the RPV is an extremely attractive proposition in economic terms. We now need to examine the practicality of these vehicles.

The Practicality of the RPV

An RPV is composed of components which differ in requirements of either function or form from those of manned aircraft, and the practicality of the RPV depends to the extent to which these requirements can be met. Particular areas of interest include visual sensors, communications equipment, navigation systems, airframes, engines, and launch and recovery methods; these will be examined in terms of performance requirements and existing technology.

An essential function of the pilot in an aircraft is his ability to observe, discriminate, and decide, particularly in the target acquisition and attack phases. Human control in real time is an essential element of a combat aircraft. An RPV must therefore contain a visual sensor of comparable performance to the human eye in terms of discriminatory ability and field of view. The limiting criterion is the need to discriminate in time to take any necessary control action. For example, a useful objective would be the ability to recognise an object the size of a tank at a range of a mile. The field of view requirements have been assessed by experiments to be 5 to 6° for navigational tracking and target recognition at medium to short range, and between 20 and 30° for initial search and orientation[8]. These requirements can be met for daylight systems by standard commercial television cameras costing as little as 450 dollars[9]. It is conceivable also that a low-light television system could give a clear-night capability.

The sensor output must be transmitted to the control station, and a one-way video communications link is therefore necessary between the vehicle and the remote pilot. The pilot will also require a 'down' link to carry information on vehicle speed, height, and heading, and an 'up' link to carry the control data resulting from his discrimination and decision processes. Communication systems are already available or under development in the U.S.A. which would serve these purposes, including a video transmitter weighing only 3 lb[10] and a variety of control station equipments;[11, 12] equipment of comparable function but shorter range is also used with the *Martel* TV missile. The success of an RPV mission will depend to a large extent on the integrity of the communications links, and the effect of the jamming threat must therefore be given serious consideration. Operational tactics could provide a measure of protection by virtue of surprise, but the primary counter must be provided by technological means. Existing systems will therefore require considerable improvement by the provision of, for example, higher power, reduced bandwidth, frequency agility, and high directional sensitivity using narrow beamwidths and steerable antennaes. The risk of detection could also be substantially reduced by using the communications links for the shortest possible periods. RPV communications systems also suffer from two further constraints; firstly, bandwidth and frequency considerations limit their operation to 'line-of-sight' and, secondly, the power outputs available at the moment limit ranges to the order of 240 miles. Ground station control at long range is therefore only possible with RPVs at high altitudes; at extremely low level, ranges will be limited to the order of one or two miles. Low level operation at sensible ranges is thus possible only with control from an airborne station, or *via* a high altitude link to a distant ground station.

Navigation system requirements for an RPV are identical to those for a manned aircraft: the system must give the pilot the assistance he needs to locate the target with the right degree of accuracy. The simplest form for low level, clear weather, daylight operation, for example, could rely largely on the visual sensor output and the pilot's ability to recognise navigation checkpoints. A more autonomous system might rely on continuous position measurement, using area navigation aids such as *Decca* or *Loran*. Alternatively, a simple inertial platform coupled to a small digital computer could provide both navigation and flight control functions. In the light of the requirement for the least use of the communications link, this system is the most appropriate because it would allow the link to be switched on only when necessary to confirm navigational accuracy by referring to visual reference points and to acquire, identify, and attack the target. The first two types of system would present little technical problems, but no such system as the last exists at present. However, some of its elements are available, and the estimated cost of both platform and computer, in large production quantities, is about 10,000 dollars[13].

Airframe construction techniques have changed little over the past decade, and emphasis has been placed on strength and performance rather than on cost considerations. New materials have tended to be used by direct substitution, and there has been little exploitation of, for example, the wide range of plastics now available, possibly with glass fibre or carbon filament reinforcement, which offers enormous scope for a new approach to airframe design and manufacturing methods. In 1970, studies by the Martin Marietta Corporation showed that, for an expendable RPV, the cost of a conventional aluminium sheet and stringer airframe could be reduced by 70 - 80% by using modern materials and fabrication techniques[14]. The extensive use of plastics could also offer a significant reduction in radar reflection area. However, the lower structural strength of plastics necessitates their use in foam or honeycomb form in order to achieve comparable rigidity to conventional airframes. In this form, plastics are extremely vulnerable to gunfire, where a single hit could cause disintegration. In this type of high risk environment, the conventional aluminium airframe offers the best survivability from the effects of gunfire by virtue of its great structural strength and the good 'anti-tear' properties of the skin. Thus the use of modern plastics will largely be confined to those RPVs which will operate in a relatively safe environment or which are expendable.

The size, weight, and thrust parameters of RPV engines will vary according to the vehicle performance and weapon load requirements. The increased weapon delivery accuracy and consequent reduction in weapon load of the RPV, coupled with the saving in weight of aircrew and life protection systems, point towards a considerable saving in engine performance requirements. Several engines are available in the thrust range appropriate to RPV needs[15], which is presently judged to be between 125 and 5000 lb. However, none of these engines exploits the full potential for low cost design which is afforded by the RVP concept. New engine designs are notoriously expensive, but this is invariably due to the need for extended performance. The RPV, on the other hand, offers a regression to earlier, proven types of engine, with opportunities for cost reduction in simplified ancilliaries, commercial standards of inspection and record keeping, mass production, and short but reliable life-cycles related to expected RPV lifespan. The need for a new, special purpose engine design will ultimately depend on the cost equation, in which numbers and variation in types of engines will probably be the deciding factor. It is interesting to note that existing American engine costs range from 10 dollars per lb of thrust upwards.

Several approaches are possible to the launch and recovery of RPVs. The U.S. Naval Weapon Center has already demonstrated conventional runway use by these vehicles[16]. Other possibilities include launch from a mother aircraft or from a mobile rail; this latter method is used by the British Army for their battlefield reconnaissance drone. Retrieval can be achieved either by mid-air recovery, as practised by the Americans, or by parachute, with an inflatable air bag to cushion landing shocks. The air launch of RPVs is an attractive method, since the mother aircraft can then operate as a control station; this approach does, however, necessitate a permanent base for the mother aircraft, with attendant vulnerability problems. A launch rail on a mobile platform, on the other hand, offers an operational flexibility and security for the RPV comparable to that of the Harrier. Control of the RPV in this configuration could be exercised either by an airborne control station or by a ground station *via* an RPV acting as a communications link. On balance, the cheapest and most effective launch method is probably the mobile platform. Mid-air recovery is an unnecessarily costly and complex process if other means can be found of getting the RPV back on the ground. Runway landings involve the use of permanent bases, with the vulnerability problems already mentioned; additionally the vehicle will be complicated by the necessity to add landing gear. The parachute/inflatable bag method is likely to be the cheapest and most effective, especially if used with a mobile launch facility.

The foregoing examination indicates that the RPV is a practicable proposition in terms of both technology and cost, and that vehicles could be built to perform a wide range of functions, either as weapons delivery or sensor platform. However, one important factor in achieving low-cost RPVs is the necessity to mass produce; this means large numbers and few types. It is therefore necessary to constrain the inventory of RPVs to the smallest number compatible with defence needs. This requirement will now be examined in terms of possible RPV configurations.

RPV Roles

Britain's primary defence responsibility lies with her commitment to NATO in Europe. NATO tactical aircraft are numerically inferior to those of the Warsaw Pact by a factor of 2 : 1[17] and it is here that the most logical application for the RPV exists. In this context, the ideal vehicle should be capable of performing offensive counter-air, interdiction, close air support, and reconnaissance missions.

The lethality of the ground defences likely to be encountered in a war in Europe necessitates that aircraft in offensive tactical roles fly at very low level and high speed. The speed at which a tactical RPV could fly would be limited by drag (which rises steeply as Mach 1 is approached), low level buffeting, and visual orientation and acquisition requirements. A sensible compromise on these limitations is therefore a high subsonic speed at 200 - 250 ft altitude. Range requirements would depend on the furthest penetration required for counter-air offensive and interdiction missions and on the launch and recovery site locations; a radius of action in the range of 200 - 400 miles would cover most requirements. The vehicle size will depend largely on the weapon or sensor load to be carried. For weapon release, it has already been indicated that the smaller the vehicle the better the chance of survival and the greater the degree of accuracy possible. A dive-bomb release gives the most accurate method of attack, and the practical limitation on the optimization of weapon load, vehicle size, and accuracy will be the lowest height at which a weapon can be released without damaging the aircraft by weapon blast effect. The smallest practicable weapon load is probably two 1000 lbs or cluster weapons; such a capability would also give a useful reconnaissance equipment payload. This suggestion is, however, extremely tentative, and a practical solution in operational terms would require a very careful analysis. Nevertheless, RPV in this configuration would appear capable of carrying out the full range of tactical offensive missions.

Although the primary application of the RPV is seen in the tactical role, a possibility exists for its use in a high altitude long range/endurance form, to perform electronic intelligence gathering and reconnaissance/surveillance tasks, and to act as a communications relay for other long range RPVs. A particularly useful function of this vehicle would be maritime reconnaissance, both to locate areas of interest for Nimrod operations, and to provide protective surveillance for the fleet. Such a vehicle is envisaged as having an operating height of 50 - 70,000 ft, and an endurance of up to 30 hours; this endurance would permit either lengthy station keeping over the fleet or extremely long-range reconnaissance to be performed. For real-time intelligence, data from the sensors would have to be converted to digital form and transmitted to the control station.

A further possible role for the RPV is in air defence, where advantage could be taken of the vehicle's potential for improved manoeuvrability. The RPV would probably need to be equipped with advanced missiles, an airborne interception (AI) radar, and a visual sensor for recognition purposes. It is envisaged that guidance to the attack point would be accomplished by an extension of existing ground control facilities, and the remote pilot would then locate the target by AI, effect any visual identification necessary, and then release the missiles. The AI radar output could be plot-extracted using similar procedures to those available for ground radar. The great reliance placed on communications in this concept means, however, that the system could be very susceptible to jamming, and for this reason this is the least attractive of the options examined.

The three possible roles suggested for the RPV are those which are at first sight the most appropriate to Britain's and NATO's defence needs; however, much work will be necessary to confirm the RPV potential, including definition of requirements, detailed feasibility studies, and extensive development work — particularly in communications — before these vehicles will be capable of supplementing, and eventually replacing, the existing range of aircraft. How, then, should these activities be pursued ?

Collective Action

Britain, in common with her European allies, is facing an increasing strain on her defence budget, and most of the available money is taken up by existing defence projects. When funds are stretched, research and development programmes must inevitably be extended. If a timely extension of our defence capability is to be achieved in terms of supplementing our front-line aircraft then development costs will have to be shared. The RPV, in the tactical role in particular, offers a common solution to European defence needs and is therefore ideally suited for a collaborative approach. The basis for this approach has already been well established by the Martel, Lynx, Jaguar, and MRCA projects, and Western European technology should certainly be adequate for the successful development of the RPV.

Collaborative projects result in lower unit costs primarily because of the sharing of research and development costs and economies of scale, but it is generally accepted that unit costs are still higher than they would have been had one country produced the total number. The increase can be ascribed to those costs resulting from conflicting national requirements and those resulting from the difficulties of control and co-ordination. The lessons are obvious.

If the greatest benefit is to be gained from the economies of scale, it is vitally important that common requirements be specified. Once this goal is achieved, a collaborative project has an excellent chance of success. Thus far, Britain has shown no overt sign of interest in the RPV concept, but it has recently been reported that a study of the roles, applications, and technological requirements for unmanned aircraft is being undertaken by the NATO Advisory group for Aerospace Research and Development (AGARD), with American, British, French, German, and possibly Italian participation[18]. The AGARD members would be wise to consider carefully the possibilities offered by a common European RPV programme.

A project of this potential size will require an effective central agency to co-ordinate and control member-nation participation, so that national technological resources can be exploited to the full and so that no unnecessary duplication of work is performed. This agency should be able to identify areas of technological uncertainty in order to concentrate effort where it is most needed, and plan development to take the best advantage of developing technology. Only in this way are the potential benefits of the RPV likely to be realised in the foreseeable future.

Conclusion

The RPV offers great promise for the future. If the twin advantages of low cost and improved accuracy of which these vehicles appear to be capable can be realised, then the present problems of increasing defence equipments costs will be considerably reduced. The successful development of RPVs is well within the bounds of possibility, since existing technology can already match many of the requirements. However, in some areas, notably communications and navigation aids, further improvement will be necessary.

The application of the RPV will be dictated by defence requirements. In Europe, Britain's needs are identical to those of her Continental allies. The primary requirement is for a tactical vehicle for offensive and reconnaissance missions, while secondary possibilities centre on a high altitude, long endurance vehicle, and an air defence fighter. Much work is required to determine optimum RPV configurations for these roles, and to define the conditions to determine for which roles expendable and re-usable vehicles are best suited.

It is an encouraging sign that NATO's AGARD has begun to study RPV applications. If the study can produce a common rationale for the RPV, the door will be open to co-operation in Europe on a scale never before achieved. Certainly, the present economic pressures on defence spending will be a powerful incentive towards achieving commonality of requirements. The potential improvements of a more integrated and collective European defence are enormous, both in organizational effectiveness and in the economies of scale that would accrue from a common equipment programme. The opportunities afforded by the RPV for this approach deserve much more attention in Britain and in Europe than has hitherto been apparent.

References

(1) Statement on Defence Estimates 1973, Cmd 5231, February 1973, Annex A, Table 3.
(2) The Military Balance 1972 - 73, International Institute for Strategic Studies, p.70.
(3) Graham, W. B. Rand Corporation, RMVs in Aerial Warfare, *Astronautics and Aeronautics* (May 1972) 39.
(4) Graham, W. B. *op. cit.*, p.40.
(5) *Aviation Week and Space Technology* (22 January 1973) 39.
(6 7) Graham, W. B. *op. cit.*, p.40.
(8) Graham, *op. cit.*, p.44.
(9) *Aviation Week* (28 September 1970) 21.
(10) *Ibid.*, p.21.
(11) Graham, W. B. *op. cit.*, p.45; systems by Northrop Electronics, Teledyne-Ryan, and Goodyear Aerospace are quoted.
(12) *Microwaves* (October 1971) p.30; systems by Epsco, Motorola, and Sperry Rand are discussed.
(13) Graham, W. B. *op. cit.*, p.43.
(14) Graham, W. B. *op. cit.*, p.40, quoting A. B. Price *et al.*, 'Low Cost Airframe Design Studies for an Expendable Air-Launched Cruise Vehicle', Martin Marietta Corporation, Report No. ER14920, April 1970.
(15) *Aviation Week* (22 January 1973) 74 - 75.
(16) *Aviation Week* (28 September 1970) 21.
(17) The Military Balance, *op. cit.*, p.90.
(18) *Aviation Week* (22 January 1973) 93.

Correspondence

To The Editor

Journal of Naval Science

Dear Sir,

I should like to say how much my colleagues and I enjoyed the first issue of the *Journal of Naval Science*.

One article of particular interest was H.M.S. *Andrew's* farewell, and the firing of her gun a 4in. Mk 23 (not a 4·5in.) for the last time on 3rd December 1974. We still have one in number 4in. Mk 23 guns at this Establishment, the Standard Propellant, and ammunition necessary for the proof of propellant for this gun. These are being retained, as although now obsolete in the R.N. two overseas customers still have an interest in this equipment. When the first of the *A* class submarines H.M.S. *Amphion* commissioned in August 1945, I was a spare crew submarine chief torpedo gunner's mate on board H.M.S. *Forth* in the Holy Loch at Dunoon, and had the interesting job (always a tricky one with a first of class) of loading her first outfit of torpedoes. The first captain was a very distinguished submarine captain called Commander Dewhurst, and his first lieutenant was also an equally distinguished submarine captain, Lieutenant-Commander Gatehouse who had commanded *T* and *S* boats. This I think demonstrates how important the *A* class submarines were thought to be as smaller types like the *U's* and *V's* would usually have lieutenants as captain and first lieutenant.

I was also interested in the photo of oil pumping trials at Lyness, as I had to visit Lyness on one occasion and was given valuable help by the staff of the Oil Fuel Depot there.

It was in June 1959, when in the capacity of Scotland Command Clearance Diving and Bomb and Mine Disposal Officer I had occasion to visit the Shetland Isles to deal with some mines. Just before leaving, I had received a letter *via* the Joint Services Bomb Disposal School at Horsham, from a wartime sergeant in the Royal Engineers. He stated in his letter that he had been stationed near Lyness in November 1939, when a lone German bomber had dropped a stick of bombs about two miles from Lyness and one of the bombs had failed to explode. He enclosed a detailed sketch of the area. On arrival we found the area exactly as depicted in the sketch, the area was still roped off, and a Red Unexploded Bomb notice still legible despite 20 years buffeting by the elements. The area is very marshy, but probing to a depth of 10 feet over the site failed to locate the bomb. (See Fig. 1. of the writer and one of his valiant team. We were even colder and more wind-blown than we look). As far as I know the bomb remains there to this day. To attempt recovery would require heavy locating equipment, dewatering plant, and concrete shafting. The cost of all this would have to be balanced against the necessity. The bomb in all probability is a 50 K/G with a 55 or 55A fuze.

This stick of bombs constituted the first ever bombs of the Second World War to be dropped on British soil. The only casualty was one rabbit killed and the incident is said to have inspired the famous wartime song of 'Run Rabbit Run Rabbit, Run, We'll get by without our Rabbit Pie, So run rabbit run rabbit, run'.

FIG. 1. Probing for German UXB, Lyness, Orkney. June 1959.

FIG. 2. Burning out main charge of an old German U mine found at the door of Burristowe House Walls, Shetland. June 5 1959.

Fig. 2 shows the main filling of a German U mine being burnt out at Walls in Shetland and was dealt with on the same trip in June 1959. The mine was one of two that had been used as front door ornaments at the house shown in the background of the photo. The house had been unoccupied since the death of the owner Colonel Foster, R.A. in 1937, but was still fully furnished with clean curtains at all the windows and looked after by a local caretaker. The colonel, also the local laird, and a gallant gentleman of the old school had been retired during the 1914 - 18 war, but had then obtained a temporary commission as a Lieutenant in the R.N.V.R. and acted as the local mine disposal officer. He had expertly removed the detonators and primers from these two mines when they came ashore in 1918, and then kept the two mines as front door ornaments. The mine filling of 300 lb TNT was still in excellent condition.

W. Y. McLanachan

Commander R.N.

Proof and Experimental Establishment
Inchterf, Milton of Campsie

Correspondence

To the Editor

Journal of Naval Science

Dear Sir,

I have read with great interest the excellent account in your Journal given by J. D. S. Rawlinson of the early days of the development of radar for the Royal Navy.

It is gratifying to see this fascinating period documented.

It is, however, a pity that this record stops short of two memorable and significant phases. The first being the use of radar in *Operation Overlord* and the second the development from this by H. E. Hogben's division of the (then) A.S.E. during the latter days of the war of a prototype radar for merchant navy use. This development, together with the drafting of a specification which has lasted wih little change for 30 years, played a significant part in laying a sound foundation for the world's civil marine radar business which is now running at a volume of over 10,000 new radars each year.

You may, Sir, be interested to see some details of this phase which are given in the enclosed paper delivered to the Institute of Navigation in 1948.* This work stands greatly to the credit of the Royal Naval Scientific Service and to the Admiralty Signal Establishment and it would be a pity if it were missed by some future historian taking J. D. S. Rawlinson's otherwise excellent account as a complete summary of A.S.E's major contribution to the history of radar.

Yours faithfully,

R. F. Hansford

Decca Radar Ltd.

* The Development of Shipborne Navigational Radar
R. F. Hansford. Journal Inst. Nav. Vol. 1 No. 2 April 1948.

New Research Vessel for Navy

A new kind of research ship which combines the roles of oceanographic trials and specialised cable laying was launched on Wednesday, 25th June, at the Cartsburn Dockyard of Scott's Shipbuilding Company Ltd. at Greenock.

The ship, which will be operated by the Royal Maritime Auxiliary Service, was named R.M.A.S. *Newton* after Sir Isaac Newton. It has four laboratories for scientific work, seven different winches for overside experiments, and a high definition echo sounder. It will be capable of laying, recovering and repairing submarine cables.

The ship was launched by Mrs. Watson, wife of Vice-Admiral P. A. Watson, Director General Weapons (Naval), and the religious service was conducted by the Rev. Colin Anderson, Industrial Chaplain for the Greenock area.

R.M.A.S. *Newton* has a deep displacement of 3,940 tonnes, an overall length of 98·6 metres, a beam of 16 metres, and diesel-electric propulsion. It will be manned by 12 officers and 37 ratings of the Royal Maritime Auxiliary Service and will have accommodation for 12 trials and scientific staff.

The Royal Maritime Auxiliary Service operates a variety of civilian-manned vessels providing services both to the Fleet and in support of naval research.

NOTES AND NEWS

Senior Appointments

The Secretary of State for Defence has appointed **Mr. H. R. P. Chatten,** C.B., R.C.N.C. to succeed Rear-Admiral F. C. W. Lawson, C.B., D.S.C. and bar as Chief Executive, Royal Dockyards. Rear-Admiral Lawson will be relinquishing this appointment on August 31st on completion of a three year tour of duty.

Dr. J. D. Lewis has been promoted to Deputy Chief Scientific Officer and appointed Principal Superintendent (Development), Rocket Motor Executive, from 17th February, 1975.

Captain M. L. Stacey, A.D.C., R.N., to be promoted to Rear-Admiral on 7th July, 1975 and to be Assistant Chief of Naval Staff (Policy) in December, 1975 in succession to Rear-Admiral P. E. C. Berger, M.V.O., D.S.C.

Mr. R. Meredith to be promoted to Deputy Chief Scientific Officer and appointed Director Defence Science 2 from 17th March, 1975.

Mr. R. J. Monaghan to be appointed Director Defence Science 5.

Senior Naval Appointments

Admiral Sir Terence Lewin, K.C.B., M.V.O., D.S.C. to be Commander-in-Chief Naval Home Command in succession to Admiral Sir Derek Empson, K.C.B., A.D.C. in November, 1975.

Vice-Admiral A. S. Morton to be Flag Officer First Flotilla in succession to Vice-Admiral H. C. Leach in November, 1975.

Admiralty Surface Weapons Establishment

Dr. K. F. Woodman read a paper on " The concept of impedance and its measurement " at a Vacation School organised by Imperial College, London on the subject of Antenna Measurements on 16th April, 1975. Dr. J. A. Betts of Southampton University and the Establishment's Dr. M. Darnell read a paper on 'real-time channel estimation and its application to Naval HF communications' at the recent IEE Colloquim on Spectrum Allocation Management and Engineering for Radio Communications and Dr. Darnell also gave 'A Review of Channel Estimation Techniques for HF Communications' on the same occasion. Dr. P. A. Payne of the University of Wales' Institute of Science Technology and ASWE's

Dr. R. D. Dawson presented a paper on 'Automatic Testing Applied to a Massive Aerial System' at the recent IERE Conference on Advances in Automatic Testing Technology held at the University of Birmingham. At the same conference, Professor D. R. Towill and Mr. H. V. Harley read a paper on the Automation of large Electromagnetic System Testing.

Mr. D. N. Keep, SPSO, has left the Establishment to take up a post as Assistant Director under the Director of Long-range Surveillance and Command and Control Projects (DSCP). Dr. M. S. Ribbands, HSO, has resigned to take up an appointment concerned with cancer research at Manchester University. Mr. G. Marwood, SSO, has emigrated to Canada and holds a post as a radar systems engineer in the Communications Research Centre at Ottawa. Mr. G. H. Glover, HSO, of the Sensors and Techniques Division is attending the one-year's Guided-Weapons Systems Course at the Royal Military College of Science at Shrivenham. Mr. D. P. Valler, PSO, is returning to the Establishment following a tour of duty with the Naval Future Policy Staff.

Recent visitors to the Establishment have included Professor Shearman, of the University of Birmingham, who delivered a lecture on 'Ionospheric Characteristics as measured by HF radar', Mr. A. A. Law (Fleet Scientific Adviser) and Dr. C. J. Reading (Technical Staff Officer to CER). The Establishment has received two parties of students from the Royal Naval Engineering College at Manadon, and an informal visit from Councillor Mr. A. G. Dann who has just retired from his post of Lord Mayor of Portsmouth.

Admiralty Research Laboratory

Mr. J. E. Conolly has recently been appointed Assistant Director at A.R.L. Most of his career in the R.N.S.S. has been spent at A.E.W. which he joined in 1951 as a Scientific Officer. A Specialist in problems relating to the stability and control of ships he was promoted to Senior Scientific Officer in 1956 and to Principal Scientific Officer in 1961. He moved to Ship Dept. in 1966 but, on promotion to Senior Principal Scientific Officer in 1968, he returned to A.E.W. where he was Chief Scientist until his present appointment.

It is announced with regret that **James (Jim) Moss,** PSO died suddenly on 22nd June whilst swimming with his family.

An appreciation will appear in the next issue of this Journal.

Central Dockyard Laboratory

Exposure Trials Station

Three of the nine members of staff at E.T.S. have been awarded Ph.D's within the last year:

Dr. David J. Tighe-Ford, who joined E.T.S. from A.M.L. with an Honours degree from Sheffield University was awarded his Ph.D., in December 1974.

Dr. Graham C. Jackson came to E.T.S. just prior to obtaining O.N.C., then with part time study progressed to H.N.C., L.R.I.C. and G.R.I.C. He was accepted as an A.R.I.C. whilst studying for his Ph.D., part time which he was awarded in July 1974.

Dr. Sean A. Gage was enlisted with four 'O' levels. After obtaining three 'A' levels in 1966 he went to Hull University on special leave where he gained his degree in Botany. Undertaking a Ph.D./M.Phil., course part time, he was awarded his Ph.D., in June 1975.

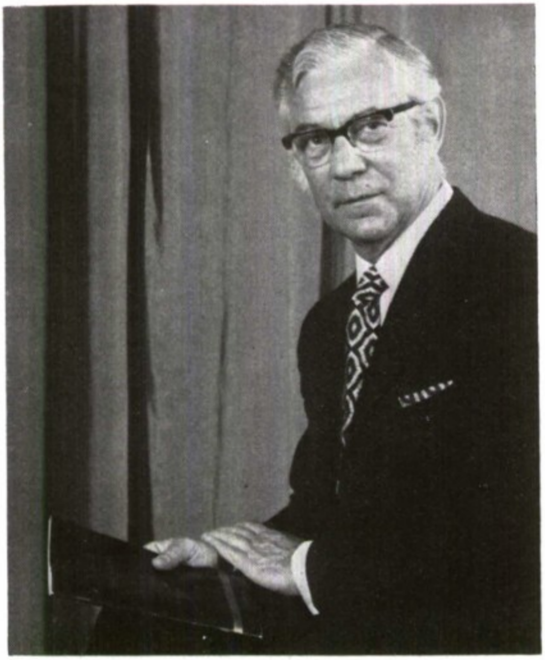
These three Ph.D's were obtained in conjunction with Portsmouth Polytechnic with **Mr. David R. Houghton** (who gained his B.Sc. (Lond.) by part time study in 1959) Officer in Charge E.T.S., acting as one of the supervisors in each case. Mr. Houghton is presently Deputy Superintending Scientist C.D.L.

Stephen Talbot, joined with four 'O' levels, gained two 'A' levels and after attending Derby Polytechnic on special leave was awarded his B.Sc. (Lond.) by external examination.

In addition two junior members of the staff, **Alex Hill** and **Tony Cullen** are in the process of taking their O.N.C. examinations.

Defence Research Information Centre

Mr. H. L. R. Hinkley, PSO (and latterly a dis-established HSO) retired on 31st May after 45 years' service.



Harry Hinkley entered Chatham Dockyard as an engineer apprentice in 1930 and subsequently worked on submarine construction. The outbreak of war found him in the Engineering-in-Chief of the Fleet's Department of the Admiralty and throughout the war he was engaged on main gearing design and providing estimates of the performance of R.N. and enemy ships for use by the operational staff.

In 1945 his career took a new turn when he formed and ran a library for the E-in-C Department. Other directorates such as DNC, DNO and DEE followed suit. These libraries were subsequently amalgamated to form the DG Ships Library.

Harry was to be engaged in library and scientific information work for the rest of his career. In 1948 he joined the RNSS and was given the task of setting up the headquarters library organisation within the Admiralty Centre for Scientific Information and Liaison (ACSIL). He was a moving spirit in many of the developments which contributed to the increasingly important rôle that ACSIL, later renamed the Naval Scientific and Technical Information Centre (NSTIC), played in giving a technical information service to Navy Department scientists and engineers. Harry's career prospered similarly, culminating in 1968 with his promotion to Chief Experimental Officer and his appointment as Head of NSTIC.

In 1971 when NSTIC was incorporated in the newly constituted Defence Research Information Centre, he became one of the two Deputy Heads of this organisation.

In the early 50's Harry was responsible for the R & D establishment's participation in the annual Physics Exhibition and from the 60's until the demise of NSTIC, was closely involved in Naval science publicity, the organisation of Open Days at establishments, Navy Days at the Dockyards etc., and the production of press releases. These activities brought him a host of friends throughout the Navy Department and in contact with people from the world of television, radio and the newspapers.

For the last 22 years of his service too, he was Editor of the *Journal of the Royal Naval Scientific Service*, a journal which under his editorship grew in scope and importance. It was typical of his zeal for all his official duties that when he reached the official retiring age last November, he consented to stay on for a further six months to see the successor to *JRNSS*, the *Journal of Naval Science* successfully launched.

At an informal gathering at the Defence Research Information Centre on 22nd May, Mr. Colin Schuler, Head DRIC, presented Harry with a Capo-di-Monte figurine, a collectors piece, on behalf of his many friends throughout MOD. Those present included Mr. Norman Parr, DR/Materials and Chairman of the Editorial Board of JNS, representatives from R & D Establishments and HQ Branches, and many DRIC colleagues.

Throughout, Harry's long career was characterised by his qualities of friendliness, enthusiasm, resilience and resource. These qualities he retains in full measure and they will serve him well in his retirement in which his many friends and associates wish him well.

Services Electronics Research Laboratory

John Frederick Gittins, O.B.E.—Obituary

With the death of John Gittins—Individual Merit Deputy Chief Scientific Officer at the Services Electronics Research Laboratory at Baldock—the Ministry of Defence has lost one of its most outstanding research scientists.

John Gittins, who has died at the early age of 51, was a man who combined great talent with a wonderful personal charm and humour and these he maintained in the face of a slowly increasing disability during the last 20 years of his life—a disability which he always overcame with such unbounded determination as to be an unforgettable example to all who knew and worked with him.

John read Physics at the Royal College of Science where he was awarded a B.Sc. in 1942 and then joined R. W. Sutton's group working on valves at Bristol (an outstation of the then Admiralty Signal Establishment). When SERL was set up in 1945 he joined the new Establishment and was appointed SO in the Reconstruction Competitions for the Scientific Civil Service. He then began the work which produced the first of his outstanding achievements in the research and development of microwave valves with the development of a high power magnetron. John was promoted to Senior Scientific Officer early in 1951 and appointed to the B.J.S.M. (Naval Staff) in Washington where he was the Scientific Adviser on valves and other electronic devices. On his return to U.K. at the end of 1952 he rejoined SERL and started research on high power travelling wave tubes — this programme was highly successful and led to many devices used in radar and communications, including the transmitter tube for the Goonhilly Downs Station at the time of the first Telstar satellite experiments. For this work he was promoted to S.P.S.O. on individual merit in 1961 and awarded the O.B.E. in 1964. His well-known book " High Power Travelling Wave Tubes " was also published at this time.

During the last 13 years he has worked with equal success on microwave semiconductor devices and, particularly, very low noise parametric amplifiers and field effect transistors — his great contributions in this field earned him individual merit promotion to D.C.S.O. in 1971.

Professor George Wesley Austin, O.B.E.—Obituary

It is with sorrow that we record the death on 5th March of Professor George Wesley Austin, O.B.E.—an old and valued colleague and latterly a true and valued friend.

Born on 1st March, 1881, Professor Austin graduated in Metallurgy at the University of Birmingham and despite a multitude of other interests, metallurgy remained his consuming technical passion throughout his life. The first world war found him fulfilling post graduate research at the Aachen Technical High School in association with the German steelmaking industry and he was fortunate in leaving Germany just in time to avoid internment. On his return to the United Kingdom, he joined the Admiralty where his metallurgical expertise was put to good use, both in Sheffield and in Barrow. Later he joined the Department of Scientific Research in the Admiralty and served for 25 years, firstly in the Torpedo Factory, Greenock, and secondly, as the development resources grew under his careful and far-sighted stewardship, in the Torpedo Experimental Establishment where he became the Chief Scientist. During his long and enthusiastic service in this capacity he was responsible for a number of important torpedo projects including the mainstays of our torpedo armament during the second world war namely, the 21 in. Mark 8 torpedo and the air launched 18 in. Mark 15 and 17 torpedoes. His enthusiasm for improving the performance and safety of existing torpedoes and for developing new ideas over a wide range of disciplines was infectious and stimulating to his staff. This resulted in important advances across a broad front in novel forms of fuel propulsion, in the introduction of electric propulsion, in preliminary design of torpedo homing systems and influence exploder mechanisms.

Towards the end of the second world war, he was given a protective commission and sent to Sicily where he followed quickly behind the advancing Allied Forces and was successful in inspecting and assessing the Italian torpedo potential from which he was able to deduce considerable valuable information about the performance of the in-service German Torpedoes. In a similar capacity he was quickly into Kiel and Eckernforde at the capitulation in Europe, where he was able to verify his earlier deductions.

In 1945 Professor Austin left his direct involvement in the Admiralty to become Goldsmiths' Professor of Metallurgy at the University of Cambridge, where he applied his usual skill and dedication in building up his Department and providing the country with some outstanding metallurgists. In addition, he was a founder Fellow of the Institution of Metallurgists and a Fellow of Trinity Hall. In parallel with these activities he found time to become a member of the Council of the British Welding Research Association and Chairman of their Research Board from 1949 - 1957.

He never however, lost his interest in, or enthusiasm for Naval matters and during his time at Cambridge and in the years after his formal retirement he continued to serve on technical committees set up by the M.O.D. (N) and to act as Consultant on a variety of subjects to both the Underwater Weapons Department and Director General Ships Departments. In these capacities he continued to encourage and advise, seeking always in a benevolent fashion to establish truth and provide help in a variety of ways varying from direct technical advice to kindly introductions to other technical specialists he had encountered during his long life.

Professor Austin was a born raconteur who could keep either a symposium or a private dinner party spellbound with his descriptions of situations both grave and amusing. In his career he met many people and all were in some part better for having known him. His friendship, enthusiasm, approachability, kindliness and above all his humility and simplicity were a lesson to all. Those of us who were associated with him and who served under his direction will miss him in our councils and deliberations and we extend our sincere sympathy to his wife and son whose loss must be even greater.

Launching of Destroyer H.M.S. *Newcastle*

H.M.S. *Newcastle*, fifth of the Royal Navy's new *Sheffield* Class Type 42 destroyers was launched at the Wallsend shipyard of Swan Hunter Shipbuilders Ltd. on Thursday, 24th April.

H.M.S. *Newcastle* has a standard displacement of about 3,500 tons, an overall length of 410 ft, and a beam measurement of 47 ft. Her armament will comprise a *Sea Dart* missile system with an automatic 4·5 in. Mk 8 gun. She will also be equipped with twin-engined WG 13 Lynx anti-submarine helicopter.

As in other ships of the Class, H.M.S. *Newcastle's* living accommodation will be of a high standard, with bunk sleeping and centralised messing. The accommodation, offices and operational spaces will be air-conditioned throughout.

This is the eighth ship of the Royal Navy to bear the name *Newcastle*. The first entered service in 1654; the last, a cruiser of the *Southampton* class completed in 1937, won battle honours for distinguished service in the Mediterranean (1940), Burma (1944 - 45) and Korea (1952 - 53) before being withdrawn from service in 1958.

Duchess of Gloucester Launches Frigate

The Duchess of Gloucester launched H.M.S. *Ardent*, the seventh of the Royal Navy's new Type 21 *Amazon* Class frigates at the Glasgow shipyard of Yarrow (Shipbuilders) Ltd. on Friday, 9th May. The religious ceremony was performed by the Rev. P. E. Bustin, Rector of Barnwell, Northants.

The 2,500 ton *Amazon* Class of ships was designed as a collaborative venture by Vosper Thornycroft Ltd. and Yarrow (Shipbuilders) Ltd.

H.M.S. *Ardent* is 384 ft long, has a beam of just over 41 ft, and will be powered by Rolls Royce Olympus and Tyne gas turbines. She will be fitted with an *Exocet* surface-to-surface missile system, a *Seacat* surface-to-air-missile system, and a 4·5 in. Mk 8 gun as well as anti-submarine homing torpedoes. She will also carry the twin-engined Lynx helicopter.

H.M.S. *Ardent* is the ninth ship to bear the name. The first, a third rater, was captured from France in 1746 and the last, a destroyer, was launched in 1929. She was sunk in action against the German battle-cruisers *Scharnhorst* and *Gneisenau* off the Norwegian coast in 1940.

Sea Risks Spotted by Naval Hydrographers

Surveys by the Hydrographer of the Navy have revealed shoals in apparently deep water near the Hebrides and off East Anglia which could have proved dangerous, the former to oil production platforms under tow from their construction sites to the North Sea, the latter to shipping in a seaway.

These are some of the facts to emerge from the Hydrographer's introduction to his 1974 annual report, just published.

He also comments on the surprising number of uncharted wrecks discovered in the English Channel last year. The 95 wrecks came to light during the survey of the Channel which will continue in 1975.

A new development in 1974 was the department's involvement in the exploration of undersea energy resources. Much valuable information has been produced, and further requirements in support of offshore industrial operations have emerged—notably the need to prove safe routes for the projected movement of enormous concrete production platforms from their deep water construction sites on the west coast of Scotland to the oil and gas fields east of the Shetland Islands.

'These and other urgent needs highlighted the shortage of hydrographic resources' says the report, 'forced a reappraisal of the original programme and caused the setting up of the Hydrographic Study Group'.

The purpose of the Study Group, involving both Government departments and outside maritime interests, was to determine the size of the hydrographic task, assess the resources required to meet it, and consider how and when they should be provided. The Group has completed its work, and its report is being considered.

Other activities have included surveys and chart modernisation covering Fiji, the New Hebrides and the Maldive Islands; similar work in the Caribbean; and exploratory and hydrographic tasks in the Falkland Islands and Antarctica.

New Regulations on Carriage of Nautical Charts

Rules requiring ships to carry appropriate nautical charts and other publications come into operation on June 1st, 1975. They apply to all ships registered in the United Kingdom, other than those less than 12 metres in length, for which compliance could be impractical, and fishing vessels for which similar rules have been made.

In future charts must be carried which are of a sufficiently large scale and sufficiently detailed as to show navigational marks, known dangers and other specified information appropriate for each part of an intended voyage. Those ships which intend to go to sea beyond five nautical miles from any coastline will in addition be required to carry directions and information appropriate to their voyage and contained in publications such as, for example, the International Code of Signals, Merchant Shipping Notices, the Mariners' Handbook and Sailing Directions. A complete list of the publications to be carried is given in the Schedule to the New Rules which are being made under Section 86 of the Merchant Shipping Act 1970.

Norwegian Trench Survey by Vickers Oceanics Sets New Standards of Accuracy

Entirely new standards of accuracy in underwater survey work have been achieved by a *Pisces* submersible — operated by Vickers Oceanics Limited — while carrying out a preliminary route survey in the Norwegian Trench on behalf of Norsk Hydro A.S.

A 20-kilometre section on the west side of the Trench was chosen for close-range investigation by submersible because of the lack of resolution in the data obtained by conventional surface methods. Vickers Oceanics collected high-resolution bathymetric, depth and side-scan sonar data relative to a very accurate acoustic long-baseline navigation co-ordinate system.

The Charter called for the Barrow-based company to provide a complete package consisting of a report and fully-referenced map of the route section, plus video recordings of selected portions and a series of core and water samples.

The chosen section was surveyed by *Pisces II* in only four days—operating in conjunction with the company's latest support ship, *Vickers Viking*—and using a newly-installed sensor suite which VOL believes to be the most advanced in commercial use anywhere in the world.

The suite includes a highly-accurate bathymetric sensor which is a high-resolution narrow-beam sonar with instrumentation to correct for vertical movement of the survey platform; and a side-scan sonar which maps the sea floor 150 metres on each side of the submersible's path. Other equipments used are a two-axis Doppler log, an accurate depth sensor, a corer, a discreet water sampler, a video camera, and an acoustic long-baseline system—*ATNAV*.

During the survey, which was carried out in the early part of February, the support ship was positioned by *SATNAV*, and the relative position of the submersible was charted using the *ATNAV*. All submersible sensor data were linked by the use of synchronous clocks and the submersible's position within the *ATNAV* array was known to within ± two metres.

The advantage of the side-scan sonar was that it enabled specific interesting features close to the route to be easily identified and subsequently charted in detail by grid searches with the submersible's contour profiling system. Three further days were spent on this work. Mounting side-scan sonar on a submersible is justified by standards of accuracy far superior to those obtained using a towed fish, the position of which can never be precisely determined.

The wealth of data obtained could only be significantly reduced by subsequent computer analysis, but during the survey data samplings were used to produce initial maps of areas of special interest. From a geologist's point of view, the survey was a breakthrough in that—for the first time—cores and water samples were obtained with 100% certainty that, within the accuracy of the acoustic transponder system, they came from specific points on features of interest. In addition, it was possible to take geologists down in the submersible to let them inspect key sites for themselves.

The survey also represents a number of significant "firsts" in the field of offshore exploration. It was the first occasion on which a manned submersible had penetrated the 200-250 metre waters of the Norwegian trench; it was the first commercial survey ever carried out with such sophisticated and comprehensive instrumentation; and — significantly — it was the first in which it was not necessary for the submersible pilot to see the strip of sea floor being surveyed. The accuracy of *Pisces II's* sensor suite enabled it to be flown on instruments at a fixed height above the bottom.

Vickers Oceanics Limited—already the world's largest and most experienced operator of work submersibles—is rapidly expanding its capabilities in the survey field, and by the end of 1975 will have five fully-equipped support ships in service with associated submersibles.

H.M.S. *Penelope* Rededicated

On Thursday, 15th May, the frigate H.M.S. *Penelope* was rededicated in the presence of Vice-Admiral P. A. Watson, M.V.O., Director General Weapons (Navy) and other distinguished guests including Mr. John Sutherland, Managing Director of Marconi Radar Systems Limited, a GEC-Marconi Electronics company. The ship has been refitted for the sea-trials of *Seawolf*/GWS25, a new self-defence system for warships of the Royal Navy.

Seawolf/GWS25 is required to meet one of the most exacting technical specifications ever laid down for naval short-range anti-missile/anti-aircraft defence of frigate class ships. The missile is required to be cheap but fast, small and light-weight with no on-board testing. Thus it is very simple and contains only a minimum of electronics; the shipborne installation contains all the necessary sophistication. Because of the concentration of electronics Marconi Radar Systems Limited is the overall systems contractor, with special responsibility for the radars, the missile being the responsibility of the British Aircraft Corporation, the launcher that of Vickers Limited.

The threat is various and covers a wide range of speeds up to and in excess of Mach 2·0 in trajectories from very low to very high altitudes. The echoing area is extremely small and must be seen in severe weather/sea conditions, even in close proxmity to land. These conditions make rapid reaction imperative. Thus there is no time for human intervention in the loop and the system is entirely automated from first detection to interception once the peace/war switch has been thrown.

The system uses line of sight guidance with radar differential tracking and a microwave radio command link. Two new radars, working as a back-to-back S and L band assembly provide all requirements for air and surface surveillance. The L band radar is a high-power self-adaptive system, with sophsticated data handling, which resolves both velocity and range ambiguities, track initiates, carries out threat evaluation, takes the engagement decision and performs attack allocation by assigning a tracker and feeding it track co-ordinates. IFF (indicate friend or foe) and system ECCM (electronic counter-counter-measures) are incorporated.

The tracker radar is also new. It operates in X band and is also a self-adaptive system. It is capable of controlling and directing a number of *Seawolf* missiles to interception, by using differential tracking and providing radio control for line-of-sight guidance.

The system also includes a differential tracking optical television system which monitors the firing and is capable of taking over missile control during flight should circumstances demand. The missile installations include a six-round launcher, which is capable of rapid rates of fire and of firing salvoes. Hand loaded, auto-controlled, with high slewing rate and pointing accuracy, the launcher also provides all-weather protection for the missiles. The missile has a solid propellant boost motor and is capable of supersonic performance.

The complete system is kept in a state of immediate notice over long periods with a very high degree of reliability by an automatic monitoring system and built-in test equipment, all under control of the system computers. There is also provision for on-board training.

Royal Navy evaluation has commenced and trials have confirmed that *Seawolf*/GWS25 meets its specification, a wide variety of fully representative missile and aircraft targets having been successfully engaged, proving the systems capability against small targets travelling in excess of Mach 2·0.

[dstl]

Information Centre
Knowledge Services
[dstl] *Porton Down,*
Salisbury
Wilts
SP4 0JQ
Tel: 01980-613753
Fax 01980-613970

Defense Technical Information Center (DTIC)
8725 John J. Kingman Road, Suit 0944
Fort Belvoir, VA 22060-6218
U.S.A.

AD#:
Date of Search: 15 February 2007

Record Summary:
Title: Journal of Naval Science
Covering dates 1975 Jul 01 - 1975 Jul 31
Availability Open Document, Open Description, Closed Until 2005
Held by The National Archives, Kew

This document is now available at the National Archives, Kew, Surrey, United Kingdom.

DTIC has checked the National Archives Catalogue website (http://www.nationalarchives.gov.uk) and found the document is available and releasable to the public.

Access to UK public records is governed by statute, namely the Public Records Act, 1958, and the Public Records Act, 1967.
The document has been released under the 30 year rule.
(The vast majority of records selected for permanent preservation are made available to the public when they are 30 years old. This is commonly referred to as the 30 year rule and was established by the Public Records Act of 1967).

This document may be treated as UNLIMITED.