AD NUMBER
AD398997
LIMITATION CHANGES
TO: **Approved for public release; distribution is unlimited.**
FROM: **Distribution authorized to DoD only; Foreign Government Information; JAN 1969. Other requests shall be referred to Controlling Organization: British Embassy, 3100 Massachusetts Avenue, NW, Washinton, DC 20008.**
AUTHORITY
DSTL ltr dtd 15 Feb 2007.

JOURNAL

OF THE

ROYAL NAVAL SCIENTIFIC SERVICE

Vol. 24 ★ JANUARY 1969 ★ No. 1

RESTRICTED

Journal of the ROYAL NAVAL SCIENTIFIC SERVICE

Vol. 24, No. 1 JANUARY, 1969

RESTRICTED

CONTENTS

AD-398997

VERY STRONG STEELS

A. Muscott, B.Sc. Tech., R.N.S.S.

Naval Ordnance Inspection Establishment

SUMMARY

Many very strong steels are now available to the designers and there is no reason to doubt that the use of these materials in ordnance will increase. The increase is expected to be most rapid in the structural members of guided weapons and will probably take place component by component.

In major gun forgings, the change will probably take the form of a steady rise in the strength level of all components. At present, the strength levels are about 25% lower than those at which conventional steels are used in aerospace applications and there is scope for advance of design in these materials. However, the time is clearly approaching when experience should be gained with newer steels which have greater long-term potential. In particular the 9 - 4 nickel cobalt steels appear to be worthy of trial.

Quality control will require a special awareness of the end product, throughout manufacture and the nomination of special tests.

Introduction

The past few years have seen the introduction, mainly into aircraft structures of steels with proof strengths in excess of 100 tonf./in.^{2}. The potential attractions in weight saving are considerable; often the final weight saved in producing a given piece of equipment or effect in the field can be much greater than the direct weight saving on the components originally selected for lightening. Also, the possibility of producing the required strength in a smaller space, offers the designer additional freedom and a challenge to make the best use of this strength. There can be no doubt that the usage of very strong steels (V.S.S.) will spread into ordnance, particularly in missiles and selected components of guns and already, a small number of such components are in service. However, it is interesting to note that the British steel manufacturers have frequently expressed disappointment at the progress in introducing V.S.S. and additionally at the tendency of designers and manufacturers to understress these materials, thereby failing to take advantage of the available mechanical strength. Often there are good reasons for this reticence, such as:—

(a) Lack of advantage. Replacement by a lighter section of higher specific strength may be impracticable when the application calls for rigidity, because there is no direct means of achieving a comparable increase in modulus. This can sometimes be overcome by modification of design, *e.g.* the use of space frames, honeycomb structure, *etc.*

(b) Lack of confidence arising from inadequate knowledge of the deficiencies of V.S.S. There is a natural lag in the accumulation of test data on mechanical properties, *etc.*

(c) Basic cost and difficulties in handling the new material *e.g.* the need to exercise extreme care at all stages of manufacture and service, *e.g.* steel manufacture, forming, machining, welding, and protection from corrosion.

(d) Testing and inspection. The conventional mechanical tests need to be supplemented with other tests, such as detailed later.

TYPES OF STEEL

Almost any fully hardenable steel can be made to withstand an 0·2% proof stress of 100 tonf./in.^{2} in small simple components. Indeed some V.S.S. are relatively simple metallurgically, and have been in service many years. Examples include music wire for springs, with 0·85% Carbon, 0·6% Manganese and UTS up to about 170 tonf./in.^{2} and ball bearing steels, 1% Carbon, 1·2% Chromium and about 750 HV. However, such steels are normally used only in small ruling sections and under well understood stress systems. When very high strength is required in bulk, it is necessary to multiply the component *e.g.* wires in suspension bridge ropes, and for bending, biaxial, or triaxial stressing, to change to materials which have a better balance of mechanical properties *e.g.* transverse strength and ductility, notch toughness and fracture toughness. The larger the component the more difficult it is to achieve uniform properties and hence many V.S.S. are of complex composition and unusual heat treatment.

In all V.S.S. more than one basic type of hardening mechanism is operative but for convenience the various groups have been classified by their main characteristic hardening process as follows:—

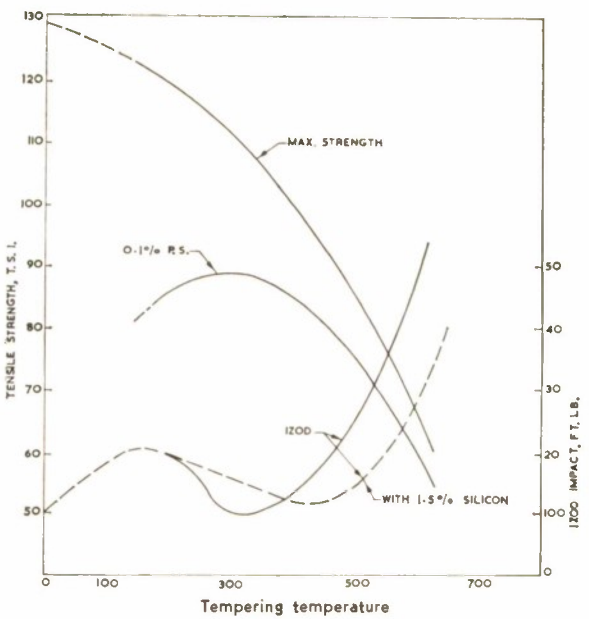


FIG. 1. Tempering diagram for EN24 steel, oil quenched.

COMPOSITIONS OF SOME MARAGING STEELS.

Alloying Element	*a.*	*b.*	*c.*	*d.*	*e.*
Carbon	0·03 max.	0·03 max.	0·03 max.	0·03 max.	0·03 max.
Nickel	25	18 - 20	17 - 19	7 - 10	4·5
Titanium	1·5	1	0·5	1	0·6.
Aluminium	0·25	0·2	0·1	–	0·1
Niobium	0·5	0·4	–	0·4	0·1
Cobalt	–	–	8	–	14·5
Molybdenum	–	–	5	–	4·5
Chromium	–	–	–	11/13	12·0
Copper	–	–	–	2	–

Type 1, Martensitic—Tempered at Low Temperature

These are usually nickel-chromium-molybdenum steels of B.S.S.95 - 99 types but with specially low contents of sulphur, phosporus and other impurities. They are hardened conventionally by quenching in oil and the high strength is retained by using abnormally low tempering temperatures and in some cases by adjustments to composition *e.g.* addition of 1 - 1·5% silicon and 0·2% vanadium which retard softening on tempering. Unfortunately, in these partially tempered ranges the toughness does not rise continuously with increasing tempering temperature. Fig. 1 shows typical embrittlement troughs at 300°C and 450°C. The modified compositions do not eliminate the trough, but tend to shift the whole diagram to higher temperatures.

Type 2, Martensitic—Secondary Hardening

These are usually 3 - 5% chromium, 1% molybdenum, 1% vanadium and 1% silicon steels. Most of this family were developed from the hot working die steels and nitriding steels, which were compounded to resist tempering and to develop secondary hardening at about 500 - 550°C. An exception is the copper-silicon-molybdenum-vanadium steel developed by R.A.R.D.E. which was originally formulated to minimize volume changes on transformation. Cleanness and freedom from impurities are again of great importance.

Types 1 and 2 have substantial carbon contentts —0·3 - 0·5%. Welding of light sections is now routine, under very closely controlled conditions but it is necessary to carry out a full heat-treatment cycle after welding. To-date, the man uses in ordnance are in airframe parts, and particularly motor bodies for G.W. When tempered to rather lower strength levels, around 70 tonf.in.^{2} minimum proof stress, these materials are suitable for gun tubes. The hot strength characteristics of the secondary hardening types are a special advantage in thin walled mortar tubes.

Type 3, Martensitic Stainless Steels

This group is based on the 12% chromium (EN.56, S.62) or 17% chromium (S.80) types. In corrosion resistance, they are inferior to austenitic stainless steel and except when tempered well outside the normal range, hardly fall into the category of V.S.S. In the lightly tempered condition the fracture toughness, stress-corrosion and fatigue behaviour are suspect and usage is normally restricted to small components *e.g.* needles, valve spindles, strikers *etc.*

A further strengthening of steels, types 1, 2 and 3, may be achieved by a thermo-mechanical process known as Ausforming. Each of the above steels depends for its hardness and tensile strength on the presence of the micro-constituent martensite. The strength (and brittleness) of these martensites is due mainly to the carbon content 0·3-0·5%, depending on the other elements present.

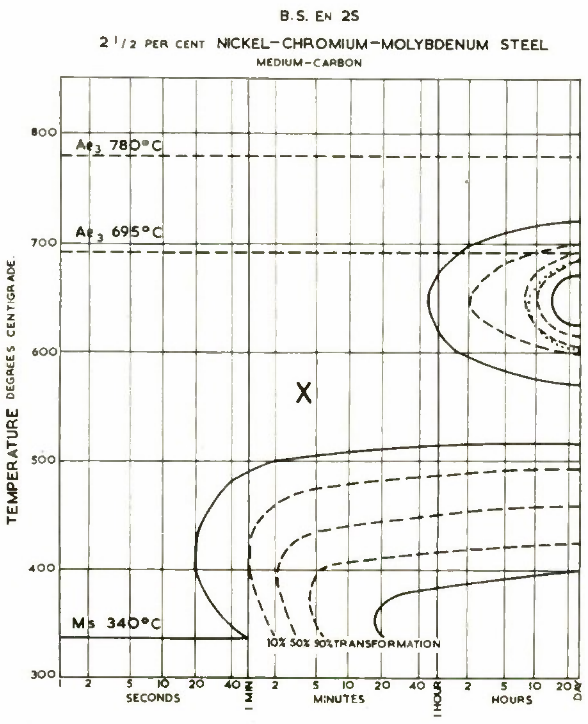


FIG. 2. Duration of isothermal temperature.

Martensite is produced by decomposition of the high temperature form of steel at temperatures below about 320°C. This is achieved by cooling the steel rapidly so as to prevent the slower growth of the softer structures which form at temperatures 700° down to 400°C. Fig. 2 is a time/temperature/transformation diagram typical of Type 1 steels and shows that there is a time lag during which the high temperature form, austenite is meta-stable, *i.e.* in the time/temperature field marked X in Fig. 2.

If steel in this condition *e.g.* quenched from 900°C into a salt bath at 550°C is mechanically worked before being allowed to transform to martensite, the final product may be as much as 25% stronger than the directly quenched and tempered sample, with little or no loss in ductility and toughness.

The process is particularly suited to rapid, reproducible working processes, *e.g.* strip rolling, high strain-rate forming, impact extrusion, *etc.*, where the timing of the interrupted quench can be linked with the forming time for each piece. Achievement of uniform properties requires large reductions of section, over 50% and uniform working temperatures. Ausforming is therefore usually reserved for simple, light sections. Attention is drawn to the impossibility of corrective work or re-heat treatment other than tempering.

Type 4, Precipitation—Hardened Martensitic Steels

This group comprises a very wide range of highly alloyed steels in which extremely tough martensites with very low carbon contents are produced by thermal or thermo-mechanical treatments. Such steels are subsequently hardened by precipitation processes, usually by re-heating for a few hours at about 450 - 550°C. (The high nickel maraging steels are a special sub-group and are discussed separately under 2·06.)

Several steels of this class were developed from the EN.56, EN.57-high chromium type of steel or from the 18/8 chromium nickel types and the chemical compositions are usually within the following ranges:—

Element	%
Carbon	0·04
Silicon	0·5
Manganese	nil - 1·0
Chromium	12 - 17
Nickel	3 - 6
Molybdenum	1 - 4
Copper	0 - 4
Titanium	0 - 1
Aluminium	0 - 1·5
Niobium	0 - 0·3

Varying contents of cobalt, nitrogen, and boron may be added.

In order to produce the martensite in the best condition for precipitation treatment it may be necessary to employ a double solution treatment and to refrigerate to −70°C before proceeding to precipitation-harden. Tensile strengths up to 120 tonf./in.^{2} have been obtained, in bar specimens of non-welding varieties, along with good stress-corrosion resistance. However, production is mainly in sheet and in the softer types, reasonable weld efficiencies are obtained in the "as welded" condition. Re-precipitation treatment, or, if practicable, full re-heat treatment is preferred for optimum weld strength.

Type 5, Cold Rolled and Precipitation Hardened Austenitic Steels

These are essentially 17 Cr. 7 Ni, 1 Mo steels *i.e.* 18/8 stainless steels with the composition slightly adjusted to assist partial breakdown of austenite to martensite. The main uses are for small/medium springs, strips and sheet. The strength is irreversibly lost by heating above about 400°C and the material is therefore unsuitable for welding.

Type 6, Maraging Steels

Maraging steels derive their high strength and ductility by combining *mar*tensite hardening of an iron-nickel alloy with *age*-hardening of the martensite. Many highly alloyed steels could be correctly described as maraging but the term is usually reserved for the high nickel types, a, b, and c, (see p.3), which generally show the best notch toughness.

Minor additions of calcium, zirconium and boron may also be made, but the normal steel making elements, carbon, silicon and manganese are regarded as impurities. In the U.K. and U.S.A. the main interest has been in type (c) which offers the best balance of mechanical properties, but types (d) and (e) have been developed to fill a need for corrosion-resistant, high strength steels. In the solution treated condition, maraging steels are relatively soft (ca. 280 HV.), ductile, readily machinable and weldable. Maraging at the relatively low temperature ca.480°C is accompanied without significant volume change and very good properties are obtained in weldments, without the need for re-solution treatment. Hence it is usually possible to apply the final precipitation (maraging) treatment as the last operation, thereby avoiding all the problems of manufacturing, machining, joining, *etc.*, in the hardened state.

The maraging temperature must be closely controlled. At 500°C and above, over-ageing occurs accompanied by partial reversion of the martensite to soft stable austenite. The strength can only be recovered by a full heat-treatment cycle.

The highest toughness levels can only be obtained with fine grain sizes, for which it is necessary to continue the hot working down to below 950°C and to avoid subsequent reheating to higher temperatures. Extremely high strength can be achieved in the 'purest' grades by cold working before maraging but this field is highly experimental and, like ausforming, applicable only to simple, light sections.

Type 7, 9% Nickel, 4% Cobalt Steels (H.P.9 - 4-X)

These steels have been recently developed by the Republic Steel Company of U.S.A. and are probably the toughest of all known steels at tensile strength levels between 95 and 130 tonf./in.^{2}. Available in two main carbon levels, 0·45% and 0·25% for optimum properties the heat-treatment of both grades consists of quenching from 900°C into a salt bath at about 400°C so that the austenite transforms isothermally at a temperature just above that at which martensite forms. The micro-structure then consists of lower bainite which is hard and strong and in these steels, much tougher than martensite. An alternative heat-treatment consists of oil quenching followed by tempering at 300°C - 500C. However, mixed micro-structures containing upper bainite or martensite have considerably lower toughness.

The 0·25% carbon grade can be welded in the quenched and tempered condition without pre-heat or postheat.

MATERIAL CONSIDERATIONS

Structure

Often the main factor preventing the use of V.S.S. is the possibility of catastrophic failure when operating at stresses substantially below the very high 0·2% proof stresses. This is particularly the case with the higher strength quenched and tempered martensitic steels, which from experience are known to be particularly notch sensitive. Although considerable effort has gone into determining the metallurgical factors which control the strengthening of steels, namely structure, grain size, particle morphology, distribution and size, the structural factors controlling the fracture toughness of steels have not been studied in such detail and are less widely understood. One reason is the complexity of the subject coupled with the fact that the assessment of toughness is usually by means of empirical tests. While extremely useful for assessing the toughness of the lower strength steels, such tests have proved to be insufficiently discriminating for low toughness, high strength steels. Two inter-relating factors, namely notch acuity and size effect greatly influence the likelihood of brittle failure in service and it is not possible to understand the effect of these by using standard acceptance type tests, *e.g.* Charpy V notch tests.

Catastrophic failures which have occurred when using V.S.S. for aerospace structures stimulated a large scale research effort to determine the factors controlling the brittle failure of such steels. This resulted in two basic aproaches, namely:—

(a) Consideration from an engineering aspect of the method of assessment of toughness and advanced design philosophy based on 'fail safe' structures.

(b) Modification of the composition and processing of existing steels and the development of new steels with increased fracture toughness.

Assessment of Fracture Toughness

The more conventional methods of fracture toughness assessment, namely the Charpy and Izod tests, are capable of comparing steels at a given section size and a particular stress concentration factor, but are not capable of extrapolation to predict the fracture characteristics of the same material in a large structure where a whole range of stress concentration factors are liable to be present. One method of assessment which has gained wide recognition (in U.S.A.) is that of linear elastic fracture mechanics. In this approach, the two most complex variables, namely notch acuity and section size are controlled. Fracture toughness parameters are determined from specimens of specific geometric shape containing a notch from which is initiated an actual crack. Both the fracture appearance and the level of fracture toughness designated G_C will depend, for a given material and test temperature, on the thickness of the material.

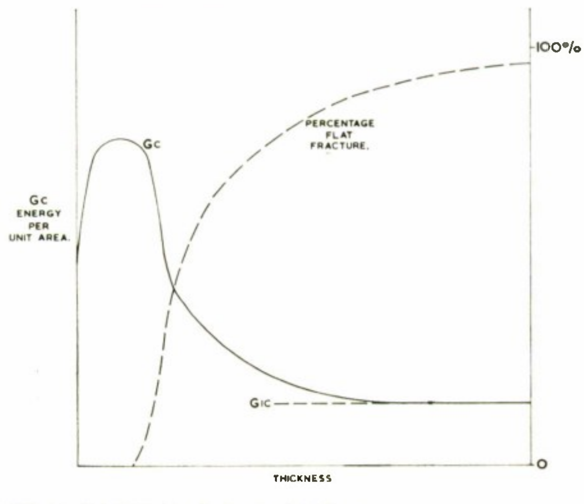


FIG. 3. Relationship between fracture toughness G_c and thickness.

The general form of the fracture toughness/thickness curve is shown in Fig. 3. The initial ascending portion of the curve at low thickness is associated with fully shear type fractures, whilst the descending portion is associated with the progressive dominance of flat fracture in the centre of the specimen. At sufficiently high thickness levels, the fracture contains only a negligible proportion of shear type fracture and the G_C value approaches a lower limiting value G_{IC}, called the plain strain fracture toughness. This characterizes crack growth in a material under the most severe conditions of constraint, and can be considered, to all intents and purposes, a material property, being independent of thickness.

The fracture mechanics approach, therefore, determines the conditions under which crack growth will take place in a given structure and realistically assumes that cracks and crack initiating defects are already present. Although every effort is made to ensure by vigorous inspection techniques that a structure is free from such imperfertions, after fabrication, cracks may also develop in service due to fatigue and/or stress corrosion. On attaining a critical crack size, subsequent crack growth can be catastrophic.

Doubts have been cast on the serviceability of maraging steels because the fatigue limit or fatigue strength at very long life is less than for other steels at similar or lower strength levels. This is probably due to the low capacity of maraging steels for work hardening as evidenced by the high yield strength/ultimate strength ratio and the abnormally large reduction of area obtained in tensile tests. However, provided the fracture toughness of the material and the operating stresses are known, the critical flaw size necessary to cause failure in a given structure can be calculated. From a knowledge of the initial flaw size in the material, usually fixed at the limits of non-destructive inspection techniques, and the sub-critical flaw growth rate, it may be possible to predict the safe life of the structure.

There are obvious limitations to the linear elastic fracture mechanics approach. As the name implies, G_{IC} measurements are only valid when failure occurs prior to general plastic yielding *i.e.* the net section stress is below the 0·2% proof stress. Consequently, the concept is at present only applicable to materials with 0·2% proof strength greater than about 90 tonf./in.^{2}, since at lower strengths with increased toughness the specimen dimensions necessary to ensure a valid measurement become prohibitive.

A further method of measurement for fracture toughness, which is at present being assessed in the U.K., utilizes a pre-cracked Charpy specimen. In this test the normal 0·010 inch radius notch (stress concentration factor 3·9), is sharpened by an actual crack introduced to the root of the notch, usually by fatigue.

A report from the U.S. Defense Information Centre at Battelle, lists no fewer than 15 tests for the fracture toughness of sheet materials alone and such tests are under constant review. For the present, however, particularly since in the absence of any extreme embrittling feature such as burning or hydrogen embrittlement, steels for ordnance are mainly employed below the strength levels, for which the fracture mechanics approach is sufficiently valid, reliance for quality control is placed on the standard Charpy test, preferably at −40°C. Increased acuity of notch, increased speed of loading and lowered temperature of tests all reduce the energy to fracture. Hence, in addition to measuring the performance of the material at −40°C, testing at this temperature with a standard test piece gives a useful indication of the likely effects of increasing the severity of the other two variables, as could occur to a component in service *e.g.* a breech ring with fine sharp fatigue cracks in the root radius.

QUALITY OF MATERIAL

From the foregoing, it will be apparent that the successful use of very strong steels requires careful attention to the features enumerated below. It may not always be practicable to check these by inspection after manufacture, or to define limits of acceptability and it is therefore essential that the Inspectorate be aware of these necessary safeguards at all stages including the preparation of specifications

(a) **Cleanness**

Any inclusion, due to oxide, carbide, silicate, sulphide, *etc.* may act as a serious stress raiser, particularly if it intersects a surface or is located at some other, unavoidable stress raiser. Processes which produce very low residual oxygen and hydrogen contents, and those which remove or disperse coarse inclusions, *e.g.* consumable electrode remelting, electro slag refining, and vacuum treating, are strongly favoured.

(b) **Purity**

Sulphur, phosphorus, antimony arsenic and tin, should be kept at the lowest feasible levels. Fig. 4 indicates the effect of sulphur and phosphorus on wide bend test performance, and work at R.A.R.D.E. has shown the drastic effect of phosphorus and tin on the low temperature impact properties of gun steels hardened and tempered to 115 tonf./in.[2] U.T.S.

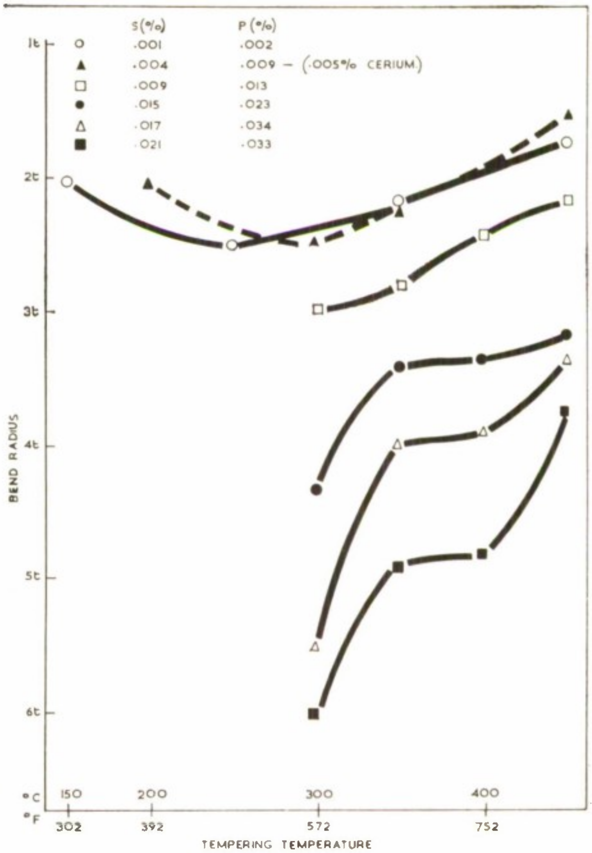


FIG. 4. Wide bend test properties, 1% chromium - molybdenum steel.

(c) **Metallurgical Structure and Grain Size**

Directionality of mechanical properties, especially toughness and ductility is never completely eliminated even when inclusions, stringers *etc.* have been reduced to very low levels. The cast, cored, dendritic structure which exhibits a range of chemical compositions in each "as cast" grain, is not completely broken up and dispersed by working. The dendrites become rotated and elongated into the direction of forging or rolling, to produce a 'fibre', which can be seen as banding of the macrostructure in sheet and bar products. In thin sheet such as is used for G.W. motor bodies, the material is very heavily worked and usually cross rolled so that the material has two 'good' directions in the plane of the sheet. The poor direction is across the thickness of the sheet and is therefore not usually stressed.

However, with smaller amounts of reductions, as in gun forgings, the material is in effect a composite containing agglomerates of varied chemical composition and therefore different hardening characteristics. With conventionally hardened steels such as types 1, where the hardenability may be marginal, the less highly alloyed material persisting from the first part of each dendrite to freeze, may in the heavier sections suffer an inadequate quench and transform to give mixed upper bainitic structures, which cannot be tempered to a satisfactory combination of strength and toughness.

It is important that this should be borne in mind when material is released on the results of test pieces cut from the ends of forgings.

Increasing the amount of hot work *e.g.* by previous upset forging and the use of water quenching, whenever this does not introduce risk of cracking, are both favourable, but tend to be under-employed in the U.K.

(d) **Avoidance of Stress Raisers**

(i) **Geometrical**

This is primarily the field of design but all involved in the use of V.S.S. should be alerted to the avoidance of these features. No sharp notch or root radius should be permitted unless it is both imperative and has been fully considered along the lines of section 3·1. Scratches or indentations, must be avoided and markings, engravings, hardness impressions should be confined to unstressed locations.

Welding reinforcement and undercut, interference fits, and local thermal treatments all act as stress raisers and must be closely guarded against.

(e) **Defects and Other Deterioration in Processing**

(i) All types of surface defect *e.g.* seams, folds, cracks should be eliminated at the earliest possible stage to prevent entrapment, unsatisfactory grain flow, *etc.* in further work.

(ii) **Decarburization**—reduces strength and particularly fatigue strength of all quenched and tempered steels. However, at very high strength levels and where fatigue strength is not required, a very small, controlled depth of partial decarburization, about 0·001 in., may be beneficial.

(iii) **Carburization** of all types but particularly types 4, 5, 6 and 7 is undesirable. Pick up of other elements in welding or heat treatment *e.g.* sulphur, oxygen, nitrogen, may also be strongly deleterious.

(iv) **Machining.** Sometimes, this may leave favourable compressive stresses in the surface layers but even these may precipitate cracking below the surface or at an adjacent point. More frequently, however, the effect of local heating is to leave residual tensile stresses as shown in Fig. 5. Except under special control, the aim should be to minimize residual stresses. Machining in the hard condition requires very sharp tools, rigid sets and usually special cutting fluids of reduced water content. Brittle untempered martensite with residual tensile stresses may be produced by severe local heating of the surface layers during poor, conventional machining. Electro-engraving and spark machining are also particularly liable to produce this defect and classical examples of failures have occurred to high carbon steel springs for Naval Service. Special machining processes which do not introduce surface stresses include chemical, electrochemical (anodic), and ultrasonic machining.

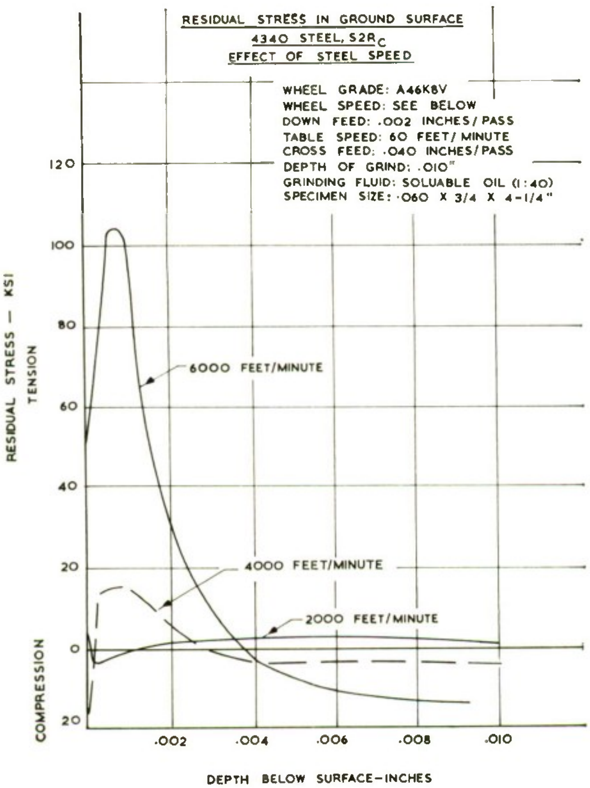


FIG. 5. Effect of wheel spin on residual stress in surface grinding 4340 steel, 52 R$_c$.

(v) **Corrosion and Stress-Corrosion.** Stress-corrosion is one mechanism by which mechanical failure occurs at a static stress level much lower than the material would have sustained in the absence of any corrosive media. The amount of actual corrosion may be extremely small but is often concentrated at the grain boundaries, similarly to the season cracking of brass. Failure may be relatively rapid or long delayed, and differs from hydrogen embrittlement by being more progressive in character. Although the two processes may occur simultaneously it is not now believed that stress-corrosion is due to progressive, corrosion induced hydrogen embrittlement. Obviously, stress-corrosion cracking is greatly affected by (and in extreme cases may be entirely due to) residual stresses but in general, members subjected to sustained loads are most vulnerable. It has been known for an aircraft undercarriage axle pin to withstand many landings but to fail during prolonged parking. Likely corrosive service environments include:—

salt spray, funnel gases, especially sulphur dioxide, de-icing fluids, and rocket effluents.

All chloride bearing materials should be especially suspected.

All steels, including so-called corrosion resistant steels, have some liability to corrosion and therefore to stress-corrosion. Where it is not possible to maintain non-corrosive conditions it is essential at least for low chromium steels, to apply a protective treatment *e.g.* electro-plating or phosphating. These are not without their problems (see vi below). Tests for stress-corrosion cracking are normally used to select materials, and stress relieving treatments and to a lesser extent, protection from corrosion. The tests are basically simple *e.g.* cantilever loaded or jig bent specimens in exposure racks, but may be prolonged and are not suitable for routine quality control.

(vi) **Hydrogen Embrittlement.** This is defined in the Defence Specification DEF-162 (replacement for D.T.D.934) as "loss of ductility, or failure by brittle or relatively brittle fracture, under a static sustained load less than the breaking stress as normally determined, caused by hydrogen in the steel". The severity of the embrittlement generally increases with the amount of hydrogen, concentration of inclusions and impurities, residual stresses, stress concentrating features and actual hardness level. Most evidence suggests that in low-alloy high strength steels of normal purity about 0·03 - 0·05 p.p.m. of hydrogen which can be extracted at 200 - 300°C, is the threshold level for significant embrittlement. For high purity steels, the threshold may be up to 10 times as high. Hydrogen at these levels is able to diffuse into and out of unplated steel and embrittlement may take place at any stage of manufacture or service when the steel is exposed to nascent (*i.e.* atomic) hydrogen, as in pickling, electroplating, phosphating, or subsequent exposure to moisture and corrosion. The embrittlement may be removed by baking at 200°C for a few hours and much more slowly at lower temperatures, by reduction of hydrogen content, or possibly by merely reducing the worst concentrations.

Most electroplates are impervious or nearly so to hydrogen and complete removal of the gas may be very lengthy.

At least one modification of cadmium electroplate has been developed which causes little embrittlement, but this process is difficult to operate and control. Increasing attention is now being given given to vacuum deposited coatings, particularly to improve adhesion. Several tests have been devised to monitor the amount of hydrogen absorbed by the steel and the rate of its release during various treatments. Such tests are not suitable for direct application to individual components.

Other forms of test *e.g.* sustained load tests, often accelerated by refrigeration, may be possible on simple components *e.g.* high-tensile bolts and pins.

NON-DESTRUCTIVE TESTS

Non-destructive tests must be used to certify freedom from manufacturing defects. Less reliance may be placed upon destructive testing of small samples. Generally, the manufacturer is concerned with economy and will test at the earliest relevant stage. The Inspection Authority however is mainly concerned with safety and performance and must therefore consider the possible introduction of new defects in further processing and service. Certain types of test may need to be repeated at several stages, with different sensitiveness, *e.g.* ultrasonic testing of forgings at:—

(a) Bloom stage for large internal defects.

(b) Finished forging with cleaned up surfaces for folds, hairline cracks, inclusions.

(c) Throughout service, for fatigue or stress corrosion cracks.

Generally as the strength is raised, the critical defect size is reduced and therefore a relative increase in the number of non-destructive tests, and a re-scaling of acceptance levels must be expected.

Many types of N.D.T. will be involved—visual, magnetic, eddy current, ultrasonic and radiographic being the most relevant and it may often be necessary to supplement one method by another. The main problems will be in the provision of standards—both for the simple purpose of measurement and for the more complex acceptance issues. It may well be necessary for designers to give guidance on this on the store drawing, but this stage may well lag behind production and only be achieved by pressure from the Inspection Authority.

Engineering, Materials and Design Exhibition

Admiralty Materials Laboratory, Admiralty Compass Observatory and Central Dockyard Laboratory, participated in the above exhibition which was held at Olympia from 2nd to 6th of December. The R.N.S.S. was also represented at the conference held in conjunction with the exhibition when several papers were read by members of the staff. The photograph shows Sir Barnes Wallis, C.B.E. discussing gas bearing development with Mr. A. G. Patterson of A.C.O.

NAVAL SATELLITE COMMUNICATIONS

G. Harries, Ph.D., F.Inst.P., R.N.S.S.
Admiralty Surface Weapons Establishment

and

K. Milne, Ph.D., A.M.I.E.E.
Plessey Radar Ltd.

SUMMARY

The important parameters of shipborne satellite terminals are reviewed with special reference to the 6 ft dish experimental system developed for operation in the first defence satellite network and installed for trials in H.M.S. Wakeful. A brief description is given of the operational system designed for fitting in H.M.S. Intrepid to operate within a British communications system.

Introduction

The operation of a satellite communication link terminal on board ship presents a number of problems which are not found in ground-based terminals. The size and weight restrictions mean that only terminals of modest size can be employed and hence their information capacity is restricted. The angular error signals derived by shipborne terminals are weaker but the angular velocities (required to compensate for ship's motion) are much higher than the corresponding values in a ground-based terminal tracking a synchronous or high-altitude satellite. The shipborne terminals must also operate in a semi-hostile environment caused by the proximity of radars and high-power communication systems and by shocks due to the ship's motion and gun-firing. Furthermore, there are problems of radiation hazards and on-mounting of high power sources on the aerial assembly.

Despite these difficulties, successful operation of a shipborne terminal has been demonstrated, and the present paper presents a brief review of the more important facets of such terminals. The data on which the paper is based are largely a result of experience gained in designing the Naval Experimental Satellite Terminal (NEST) and the SKYNET Type V terminal.

NEST in H.M.S. *Wakeful*

The NEST system was a feasibility study aimed at resolving several important factors:—

(a) The basic design techniques.
(b) The study of problems created when two or more stations seek access to the satellite simultaneously.
(c) The communications traffic capacity of a small terminal in a strategic network.
(d) The techniques necessary for acquiring the satellite and communicating in a ship environment.

The programme was part of the DCSP (Defence Communications Satellite Programme) Phase I in which the U.S. Department of Defence provided the " space segment " which at present comprises 25 satellites slowly orbiting the earth at about 20,000 miles over the equator at a rate of 1° per hour. The main parameters of the satellites are:—

Type	Active Repeater
Power Output	38 dBm
Frequency	Input 8 GHz; Output 7 GHz
Bandwidth	25 MHz
Weight	50 kg

The design aim in NEST was to produce a system with capacity for a voice channel (or several telegraph channels) in a ship-to-ship mode. This mode during the trials was simulated by transmitting signals via the satellite back to the ship.

The photograph (Fig. 1) shows the installation in H.M.S. *Wakeful.*

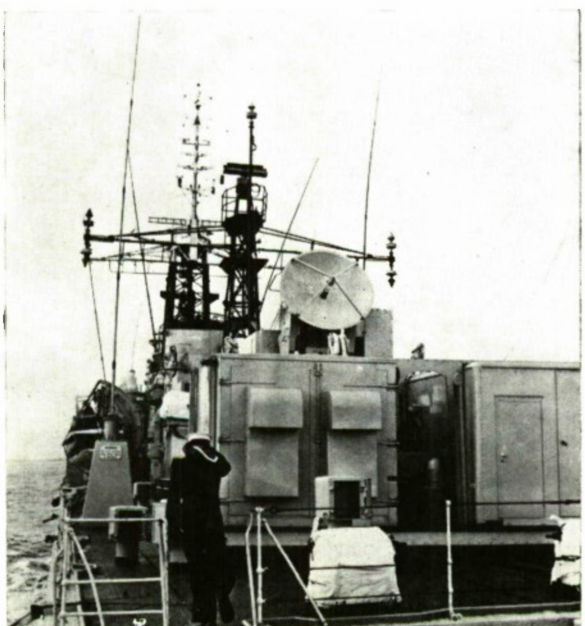


FIG. 1. NEST in H.M.S. Wakeful.

The system parameters are:—

Dish size	6 ft. (parabolic)
Beam Width	1·5°
Tx Output	10 kW (nominal)
Aerial Mounting	2-axis (Az-El)
Stabilisation	Off-mounting computer
Preamplifier	Parametric (cooled to 20°K) Noise Temperature 110°K
System Noise	250°K
Autotracking	Static split system
Baseband frequency	0-4 kHz
Modulation	FM/FDM

The power output is comparatively large and was designed generously to explore the multiple-access problems in company with large stations. The two-axis mounting was acceptable for operation in home waters where zenithal aerial positions are not encountered. The stabilisation system was a conventional gunnery system.

H.M.S. *Wakeful* was the first European ship to be fitted for satellite communications and the results of the trials afloat, and subsequently ashore, can be summed-up as follows:—

(a) Exact confirmation was obtained for the power budgets.
(b) The ease of satellite acquisition was established. Acquisition can be established in less than one minute.
(c) Successful voice and telegraph trials were conducted with a large station in U.K. and a small (6 ft. dish) station in the U.S.A.

In the following chapters the general problems of satellite communications will be considered and the desired parameters of a ship system derived.

Terminal Equipment

Fig. 2 illustrates the essentials of a shipborne terminal.

Information signals to be transmitted by the terminal to the satellite (and hence to another terminal) are passed through the baseband equipment to the exciter which produces a low-power microwave carrier modulated by the baseband signal. The microwave signal is amplified to a level of a few kilowatts and fed to the waveguide system and the aerial.

The waveguide system contains the necessary filters and diplexers for separating received from transmitted signals. It also contains a monopulse comparator for deriving angular misalignment signals which are fed to the tracking receiver. Received communications signals are first amplified in a low noise preamplifier. After down conversion to a suitable intermediate frequency and filtering to select the wanted carriers, they are then demodulated in the receivers. The latter feed the baseband equipment with outputs to the user equipment (telephone, telegraph, etc.).

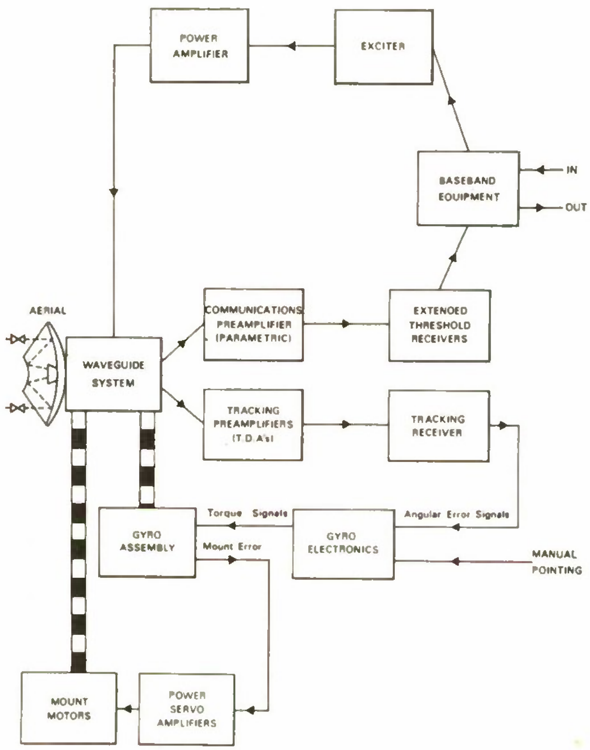


FIG. 2. A shipborne terminal.

Stabilisation of the aerial beam is by reference to the gyro assembly which provides an inertial reference pointing angle. Any misalignment of the beam with respect to this reference is sensed by the gyro pick-offs and corrections applied via the power servo and drive motos. The inertial reference pointing angle is up-dated by signals derived from the tracking receiver so that the beam always points at the satellite. Manual pointing data for initially acquiring the satellite may also be fed to the gyro assembly.

Aerial and Preamplifier

The essential requirement of the aerial and preamplifier system is that the figure of merit G/T (ratio of aerial gain to overall system noise temperature) should be as high as possible.

The system noise temperature comprises three main parts:—

(a) Waveguide and filter loss (about 50°K).
(b) Aerial noise temperature (20°K to 50°K).
(c) Preamplifier noise temperature (about 150°K for ambient temperature parametric amplifiers).

The overall system noise temperature is thus around 250°K. It should be noted that attempts to improve this figure may not always achieve the desired end, since the terminal will become more susceptible to interference from other equipments as the noise temperature is reduced.

The use of Cassegrain geometry for the aerial system enables the bulky waveguide system to be accommodated behind the reflector, and gives a degree of freedom to the feed design since the system magnification may be varied to suit the beamwidth of the feed.

Both aerial noise temperature and gain may be improved by reducing spillover, *i.e.* energy from the feed which escapes beyond the edge of the subreflector. Reduction of spillover presents some design difficulties when the feed has to provide error channel signals in addition to the sum signal.

The preamplifier noise is an important component of the system noise temperature. Very low noise temperature can be achieved using a parametric amplifier by cooling the amplifier to, say, liquid nitrogen or helium temperatures. However, this is not desirable in a shipboard environment and the development of an ambient temperature parametric amplifier has been part of this satellite programme. The resultant amplifier with a noise temperature less than 150°K will be considered later.

Down-Link Power Budget

The limitation in performance of any satellite communication system is mainly determined by the down link. Table I illustrates a typical down link power budget for a 6 ft. diameter terminal operating at 8 GHz. The ratio of received carrier power to noise power density is 56 dB Hz, if, for example, the satellite power is assumed to be 44·5 dBm.

TABLE I

Satellite Power	+ 44·5 dBm
Free-space Attenuation, $\left(\frac{4\pi R}{\lambda}\right)^2$	— 201 dB
Receiver Aerial Gain (1·5° beamwidth)	+ 40·5 dB
Miscellaneous Losses	— 2 dB
Received Carrier Power	— 118 dBm
Receiver Noise Power Density (300°K)	— 174 dBm/Hz
Carrier-to-Noise Power Density	+ 56 dB.Hz

Table II illustrates carrier-to-noise densities required for various types of traffic using frequency modulation and comparison with Table I enables the system capacity to be obtained. Thus, if the whole of the satellite's power is made available for one 6 ft. terminal, it can support one CCIR quality speech channel or three tactical speech channels or 80 telegraph channels.

TABLE II

Service	*dB.Hz.*
Television	83
Telephony—CCIR Quality (50dB S/N)	56
Telephony—Tactical Quality (25dB S/N)	51
F.S.K. Telegraph (100 bauds)	37

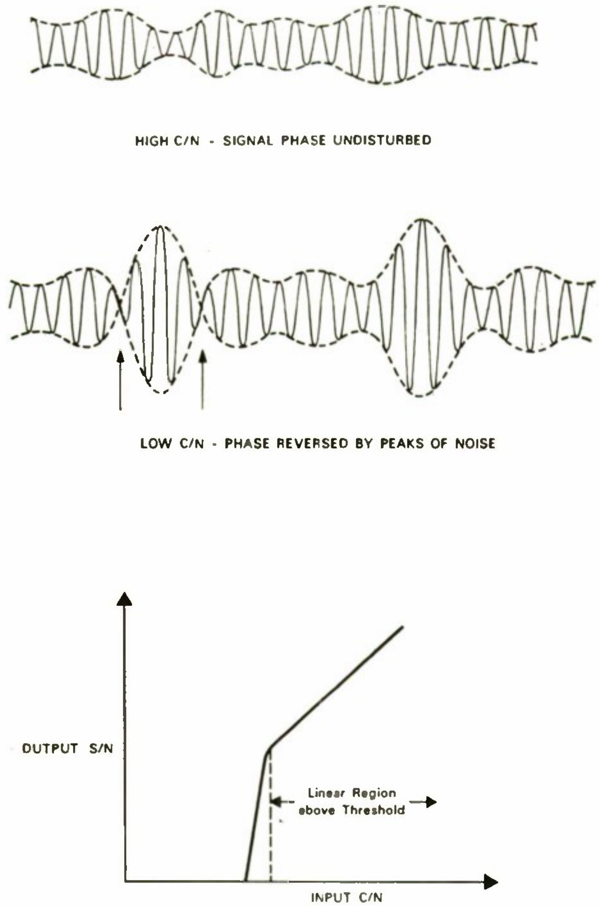


FIG. 3. Threshold in F.M. receivers.

Extended Threshold Receivers

Extended threshold receivers are necessary to achieve the performance given in Table II, and a brief description of threshold and means of extending it follows.

The basic characteristic of f.m. systems is that when the modulation index (ratio of frequency deviation to modulating frequency) is unity, the demodulated signal-to-noise ratio (S/N) is the same as that of an a.m. system. If the modulation index is increased the S/N ratio increases proportional to m^2. Larger r.f. bandwidths are required as m is increased.

The improvement of S/N with increasing m continues until the carrier-to-noise C/N (ratio of carrier power to noise in the r.f. bandwidth) is so low that noise peaks reverse the phase of the carrier, as illustrated in Figure 3. The demodulated output then contains sharp spikes and the S/N deteriorates rapidly with further increases in m. The same phenomenon is observed if the r.f. bandwidth and modulation index are fixed and the input C/N ratio is varied (see Figure 3). Threshold occurs at a C/N of about 10 dB for conventional f.m. demodulators.

Threshold extension may be obtained by employing frequency feedback techniques to reduce the deviation before the demodulator so that a narrower pre-detector bandwidth may be employed. Figure 4 illustrates the essentials of a phase-lock demodulator (P.L.D.).

The voltage controlled oscillator is locked to the instantaneous frequency of the incoming carrier by the feedback applied from the phase sensitive detector output. Threshold occurs when the peak error at the phase sensitive detector exceeds $\pi/2$ radians: the r.m.s. error is then about 0·4 radians. This corresponds to a carrier-to-noise ratio of about 10 dB in the closed loop bandwidth. The latter is controlled by design of the shaping network and generally approximates to the geometric mean of the modulation bandwidth and the signal bandwidth.

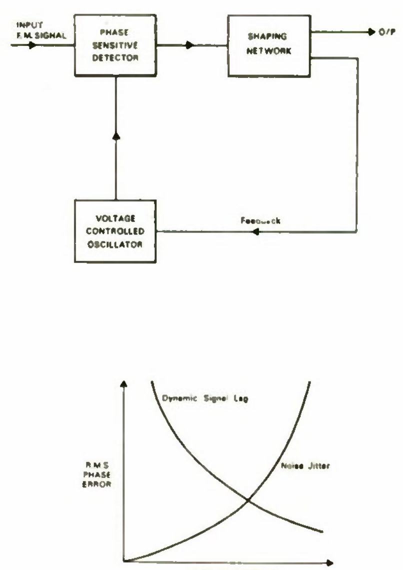


FIG. 4. Essentials of P.L.D.

Up-Link Power Budget

The main requirement for the transmitter at the earth terminal is that the power reaching the satellite shall be sufficiently large with respect to the satellite's internal noise to ensure that the signal is not suppressed by noise and that the satellite's transmitter is fully saturated. A carrier-to-noise ratio of about 10 dB is sufficient.

TABLE III

Transmitter Power	+ 67 dBm
Transmitter Aerial Gain	+ 41 dB
Free Space Attenuation	− 202 dB
Satellite Aerial Gain	+ 14 dB
Miscellaneous Losses	− 2 dB
Received Carrier Power	− 82 dBm
Satellite Noise Power (3000°K, 2 MHz)	− 101 dBm
Carrier-to-Noise Power Ratio	19 dB

Table III illustrates a typical up-link power budget for a terminal employing a 6 ft. diameter aerial with a 5 kW transmitter operating in the 8 GHz band. The C/N ratio is 19 dB in a bandwidth of 2 MHz: the terminal could therefore be used for accessing similar satellites with bandwidth up to 24 MHz.

Power Sharing

When a satellite is simultaneously serving a number of different size terminals the transmitter power of each terminal must be controlled to ensure a balance of power. The problem is best illustrated by an example and Fig. 5 shows a situation in which a satellite is serving four terminals with five separate carriers passing through the satellite repeater.

The transmission routes are as shown. The solid lines indicate the critical paths which determine the power sharing arrangements.

Table IV, which is based on Tables I and II, shows the minimum satellite power (EIRP) required for each carrier. The highest power required is that for carrier 1 to the 6 ft. terminals. Under these conditions the total satellite power required is 42 dBm. Table IV also shows the transmitter powers required to achieve correct power sharing.

The limiter in the satellite suppresses the weaker carriers and the corresponding terminal transmitter powers must be increased by about 6 dB to compensate for this effect.

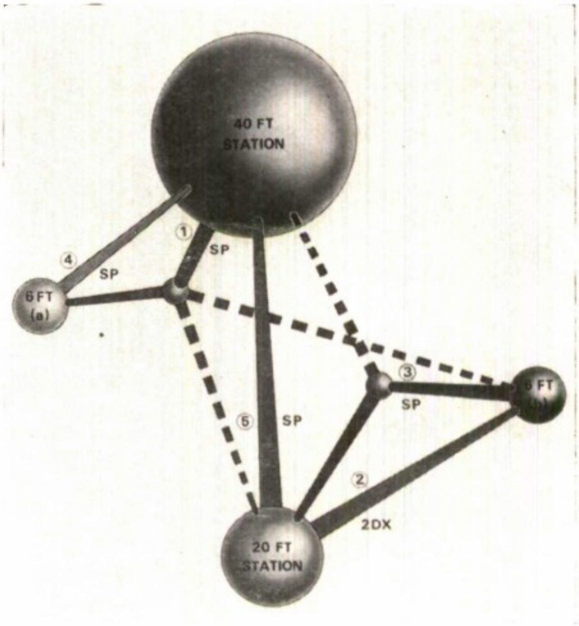


FIG. 5. Typical routes in mixed station environment.

TABLE IV

Satellite Carrier	1	2	3	4	5
Traffic Channels	1SP	2DX	1SP	1SP	1SP
Receiving Terminal	6a (+6b+20)	6b	20 (+40)	40	40
Required C/N density at receiver (dB.Hz)	51	40	51	51	51
Required Satellite EIRP* (dBm)	41	30	30·5	24·5	24·5
Transmitting Terminal (dBm)	40	20	6b	6a	20
Transmitter Power (dBm)	55	56†	67†	61†	50·5†

* Includes 1·5 dB margin for intermodulation
† Includes 6 dB margin for suppression of weaker carriers in satellite transponder.

Aerial Mount and Stabilization

For a terminal located on a ship, a two-axis mount is inadequate for tracking the satellite through regions requiring high mount velocities and accelerations. The three-axis mount shown in Fig. 6 is suitable.

For low elevation satellites the training and level axes may be used to provide the equivalent of an az-el mount whilst for high elevation satellites the cross-level and level axes form an X-Y

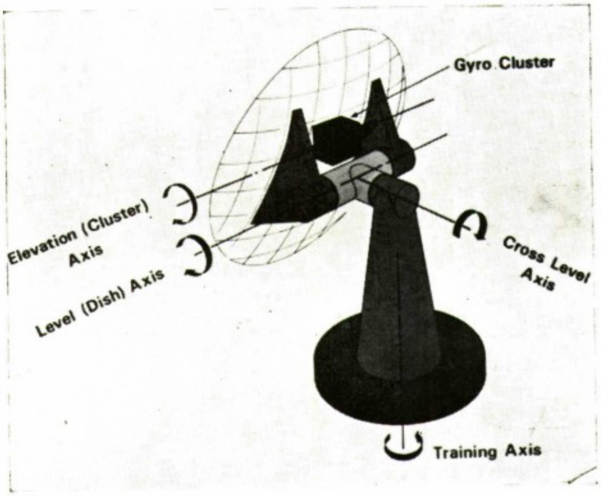


FIG. 6. On-mounting stabiliser.

type mount. In both cases, the third axis may be used to counter ship's motions and hence limit the velocity demands on the other axes.

The angular error signals produced by the monopulse system could be used directly to control the mount via a servo system of sufficient bandwidth to counter the effect of ship's motion, but the problem of acquiring the satellite (or of re-acquiring if the beam is "wooded" by the ship's superstructure) requires some form of stabilisation.

If the ship carries its own inertial reference the three-axis mount may be slaved to this data but a more attractive solution, especially for smaller ships, is to use an on-mounting stabiliser so that the only external input required is the ship's compass. This is a novel feature of the shipborne system considered here.

The on-mounting stabiliser illustrated in Fig. 6 takes the form of a cluster of three gyros and two accelerometers which are located on an axis parallel to the level axis. The cluster provides a vertical and azimuthal reference, senses deviations of the mount from this reference and applies signals to the power servos so that the mount deviation is corrected and the gyro cluster maintained horizontal at all times.

A simplified diagram of the stabilisation and control system is shown in Fig. 7 for one axis only.

This indicates that there are three main servo loops, viz. power servo, accelerometer and auto-tracking.

The gyros are of the rate-integrating type in which the pick-off output is directly proportional to the rotation of the gyro case, thus any mount movement induced by ship's motion produces an

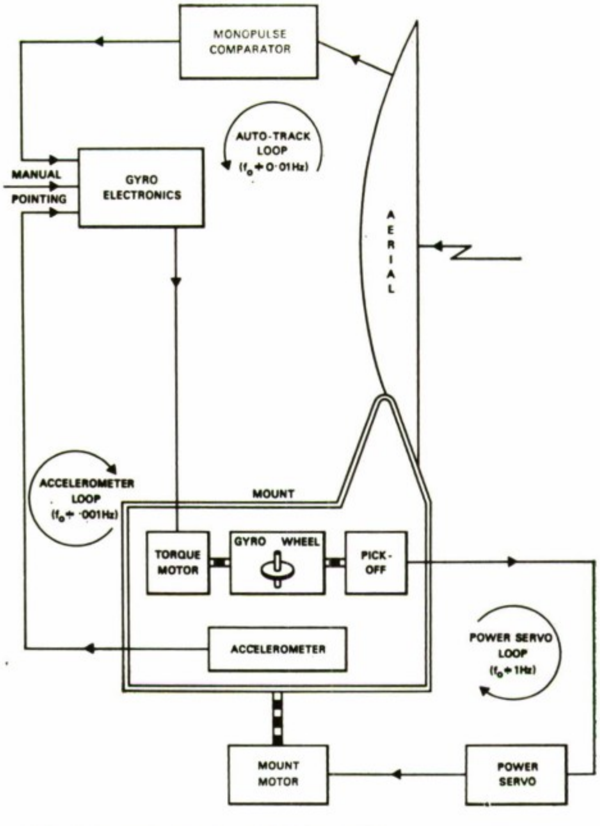


FIG. 7. On-mounting stabilisation system.

output which is applied to the power servo which re-aligns the mount. Since the power servo has to cope with ship's roll having accelerations of up to 10°/sec², its bandwidth must be fairly wide. A type 2 servo is used with a natural frequency of about 1 Hz.

Accelerometer loops are needed to correct for gyro drift and to maintain the vertical reference pointing at the earth's centre as the ship traverses the globe. A similar loop using ship's compass data provides a long term azimuthal memory. Very low natural frequencies (about 0·001 Hz) are adequate, due to the low rates involved.

The auto-tracking loop has to cope only with the satellite drift rate which may be one or two degrees per hour.

SKYNET

SKYNET is a British military satellite system. The satellite, in this case, will be stationary and positioned in an equatorial orbit at about 50° East longitude. This will provide a coverage as given in Fig. 8. The satellite will be manufactured and launched in the U.S.A. It will have a higher power than the DCSP satellites and will have two

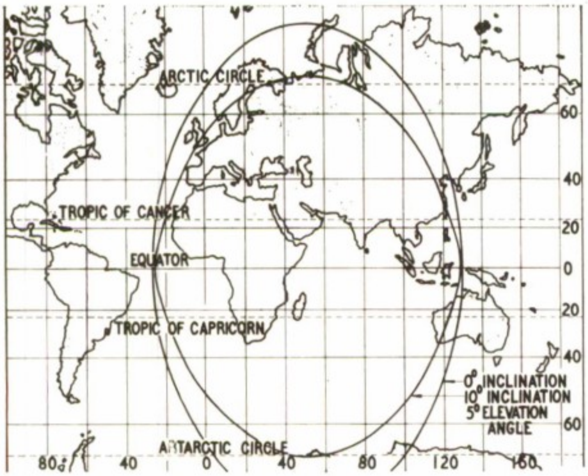


FIG. 8. Satellite coverage.

separately powered amplifier channels. One of these will be used by tactical users (such as ships) and will therefore be free from multi-access problems caused by high power stations on the strategic net. This is a very important facility which makes satellite communications for small stations a viable proposition.

In the present programme a 6 ft. terminal known as SKYNET Type V is scheduled to be fitted in H.M.S. *Intrepid* and H.M.S. *Fearless*. (Fig. 9)

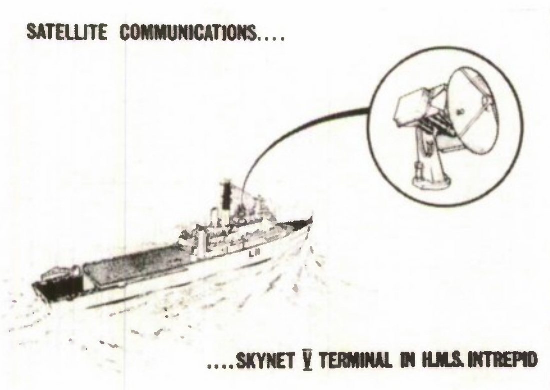


FIG. 9.

SKYNET V

Unlike NEST, the SKYNET V type is designed as an operational system, although both systems have portable features for retrospective fitting. The essential differences between the two systems are considered:—

(1) *Preamplifier*

For operational purposes an uncooled amplifier is necessary without appreciable reduction in performance. To this end an amplifier, VX1747, has been developed with the following characteristics:—

Type	Parametric (two stages)
Gain	30 dB
Noise Temp.	130°K at 20°C
Weight	20 kg

The paramp is illustrated in Fig. 10.

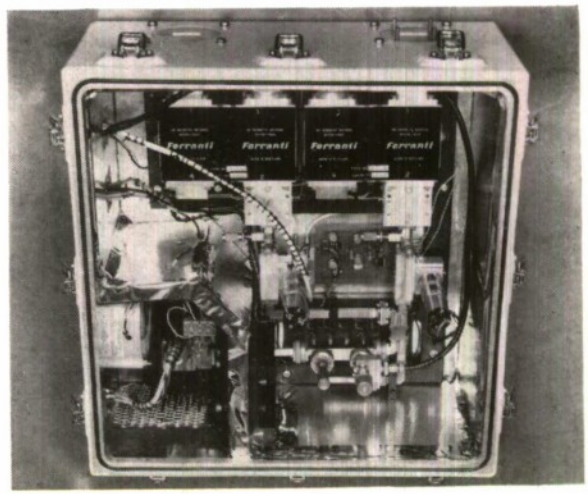


FIG. 10. Parametric amplifier.

(2) *Aerial Noise*

The aerial noise temperature will be improved by reducing spill-over, the energy from the feed which escapes beyond the edge of the Cassegrain sub-reflector. A corrugated waveguide feed is expected to reduce this spillover from 20 to 2%.

(3) *Transmitter Power*

Since the multiple-access through the satellite is less of a problem in SKYNET the transmitter power can be reduced to that which will sustain a voice channel and a few telegraph channels. At 5 kW transmitter power several problems are relaxed: the radiation hazard, the cooling of the transmitter tube and the power supply requirements. The cooling and power supply requirements are readily appreciated when one considers that the klystron tubes necessary for the transmitter are about 30% efficient. A large amount of energy has to be generated and dissipated in confined areas.

(4) *Mounting and Stabilisation*

This has been considered in a previous chapter. In SKYNET V an on-mounting gyro-accelerometer assembly will be used. This will dispense with the need for a reference from a master stable element in the ship and the system will require compass (or heading) information only from the ship. This makes the system attractive for fitting in any ship. It will also be able to operate freely at the zenith positions experienced in equatorial areas since the three-axis stabiliser is designed integrally with a mounting with three corresponding axes. This type of mounting is being used for the first time in the Royal Navy. The pointing accuracy of the dish will be better than 0·25° r.m.s. under ship motion conditions, which is adequate for a dish which has a beamwidth of about 1·5°.

Conclusion

The programme outlined above has been intensively pursued in order to provide the Royal Navy with reliable communications independent of atmospheric conditions which limit the effectiveness of HF systems. With the disappearance of overseas bases, which are expensive to build and to run and whose tenure becomes uncertain, satellite communications should provide a viable replacement. Satellite communications, in the near future, should be a cost effective method of providing long range Naval communications.

Acknowledgement

The authors wish to thank the teams in ASWE and the Space Department of Plessey Radar Ltd. for their industry and enthusiasm in these co-operative projects.

MODERN CATHODIC PROTECTION PRACTICE

R. Holland, Ph.D.
Central Dockyard Laboratory

The economic consequences of corrosion in terms of replacement costs, component failure and the cost of preventive measures are so great—a recent National Bureau of Standards report estimates the cost in the U.S. at $10 billion per annum[1]—that any method which can claim to arrest corrosion completely must command respect. The effectiveness of cathodic protection in justifying this claim has been conclusively demonstrated in numerous large scale engineering applications as well as in the laboratory and it is now established as an essential technique in many aspects of modern engineering.

Though paints and similar coatings will continue to be the first line of defence against corrosion, their application requires direct access to all parts of the structure both initially and for later renewal as even the most modern high duty coatings are impermanent. Furthermore imperfections in application are almost inevitable and damage may be expected during erection and service of such structures as pipelines, jetties and ships. Corrosion is then concentrated at such imperfections and may lead to deeper penetration than might be expected on uncoated steel. Cathodic protection supplements coatings by preventing corrosion of bare areas of buried or immersed steelwork. It should therefore be regarded, not as an alternative to paint, but as complementary to it. It is effective while the structure is no longer accessible for coating maintenance and it was probably this factor which mainly stimulated the interest of American oil companies in the first large scale applications of cathodic protection to buried pipelines before the Second World War. The successes achieved in this field led after the war to more widespread application to other buried or immersed structures where the consequences of corrosion were somewhat less disastrous but where its prevention was still very desirable. The extension of the range of applications has been accompanied by the development of new materials and operating techniques, so that whereas in 1950 the designer of a cathodic protection scheme had a simple choice between fitting magnesium anodes or using a manually controlled impressed current system with iron or graphite anodes he may now choose from a dozen anode materials and may specify quite complex servo-controlled systems.

The cost of alternative systems for a given structure may vary widely; capital costs or running costs may be dominant and either may be affected by the choice of coating. Quantitative comparisons are beyond the scope of this article but an attempt is made to summarize the options available in the design of modern cathodic protection systems.

GENERAL CONSIDERATIONS

The cathodic protection principle is basically simple: the tendency for metal ions to enter solution is opposed by a net flow of current to the surface from anodes placed in the surrounding soil or water so that the whole surface becomes a cathode. Current flows either because of the natural e.m.f. of the cell so formed, depending on the choice of anode material, or by the imposition of an external source of e.m.f. The now universally accepted criterion of complete protection is the attainment of a specified potential difference between the metal and a reference electrode in the adjacent electrolyte rather than a particular current density. Experience has shown that steel is protected at a potential of $-0{\cdot}85V$ relative to a $Cu/CuSO_4$ electrode in oxygenated soil or water or at $-0{\cdot}95V$ in anaerobic conditions where sulphate reducing bacteria may be active. The $Ag/AgCl$ electrode is widely used in sea-water and steel is protected at $-0{\cdot}8V$ relative this electrode.

Current Requirements and Distribution

In aerated water the most important cathodic reaction is the reduction of dissolved oxygen—the liberation of hydrogen becomes a significant factor only at potentials more negative than $-1.0V$ relative to $Cu/CuSO_4$.

At $-0{\cdot}85V$ oxygen is consumed to form hydroxide ions as quickly as it arrives at the surface and consequently the current density for protection varies considerably with water speed. Provision must therefore be made to meet the varying requirements of such structures as ships' hulls and water-cooling plant. In static or slowly moving water the current initially required to protect bare steel is of the order of 10 to 20 mA/sq. ft. Alkali formation at the protected surface, however, deposits chalk from sea-water or other hard waters and this may reduce the current demand to less than half the original value—a desirable feature which is often encouraged by deliberate use of a high initial current density.

Most structures are coated and the integrity of the coating is the main factor determining current demand. In sea-water 1 to 2 mA/sq. ft. is typical rising to 3 to 5 mA/sq. ft. in rapidly moving water. The current demand of buried structures tends to be less, depending on the moisture content of the soil, and the high resistance of modern coatings may reduce current requirements to very low levels. The potential of a protected structure, measured against a portable reference electrode, is rarely uniform because of potential gradients in the electrolyte. Potential gradients within the structure itself should be minimised by bonding. Areas close to the anodes are polarised to a greater extent than those which are more remote. Overprotection may lead to blistering or softening of coatings, especially in marine environments, and wasteful consumption of current. Siting anodes at greater distance from the structure leads to more uniform current and potential distribution. A single anode sited at a distance comparable to the largest dimension of the structure can often provide a good distribution but it is usually necessary to distribute a number of anodes to achieve the best results. Reduction of current demand by the provision of high duty coatings reduces the potential gradients in the electrolyte and thus contributes towards a more uniform current distribution. The best economic compromise requires an assessment of the cost of surface preparation and coating which must be comparable with possible saving in the cost of multiple anode installations, long cable runs and reduced running costs.

Power requirements depend not only on the current but also total voltage required. The greatest potential gradients, *i.e.* the major component of electrolyte resistance, occur in the vicinity of the anodes where current lines crowd together. The size of the anodes is therefore the major factor, other than electrolyte resistivity, determining the total circuit resistance. The effective resistance of buried anodes is often reduced by surrounding them with a bulky conducting backfill of coke-breeze or suitable salts. The spread of current is analogous to the spread of lines of force in electrostatics. The crowding of current lines due to anodes being sited close together increases the circuit resistance whereas the distance of the anodes from the structure has little effect.

Interaction Corrosion. If a foreign structure lies in the electric field of a cathodic protection system the potential gradient across it is often sufficient for current to enter it at one point and leave it at another, closer to the protected structure. Where current leaves the foreign structure corrosion occurs, presenting practical and legal problems. The effect is reduced by factors which reduce the potential gradients in the soil or water and by coating the foreign structure. It may be overcome by partial cathodic protection, *e.g.* by attachment of galvanic anodes or more usually by bonding the affected structure to the protected one. The usual criterion of the necessity for remedial action is that the positive shift of potential of the foreign structure when the cathodic protection system is swiched on should not exceed 20 mV.

ANODE MATERIALS

Galvanic Anodes. Cathodic protection can be obtained by directly coupling the structure to anodes of a material such as magnesium, zinc or aluminium which is sufficiently reactive to provide protective current by its own dissolution in the electrolyte. These materials are known as galvanic, reactive or sacrificial anodes and are used in the form of special alloys fabricated to ensure as far as possible continuous uniform activity. The anodes usually contain steel inserts which provide continued electrical connection until the reactive material is exhausted. Typical examples are shown in Fig. 1 (a) and (b).

The oldest and most versatile galvanic anode material is magnesium, usually alloyed with 6% zinc. Until 10 years ago it was the only such material available and although the development of zinc alloys has reduced its use on ships it still finds extensive applications. Its principal characteristic is a highly negative potential, $-1{\cdot}55$V to $Cu/CuSO_4$ giving a driving voltage of 0·7V against protected steel. The electrical output is some 50% of the theoretical value of 1000 A.h./lb., the difference being accounted for by hydrogen evolution from the anode. When used in moist soil magnesium is surrounded by a "back-fill" of powdered gypsum and bentonite, usually supplied prepacked around the anode. Magnesium is widely used in fresh and salt waters. Its high driving voltage may cause damage to marine paints, so that anodes used to protect ships' hulls must be either surrounded by an insulating shield or, for laid-up vessels, sited some 10 ft. from the hull. Control of current output is sometimes effected by variable resistance in the anode lead, by shrouds over the anodes or by varying the number of anodes in circuit.

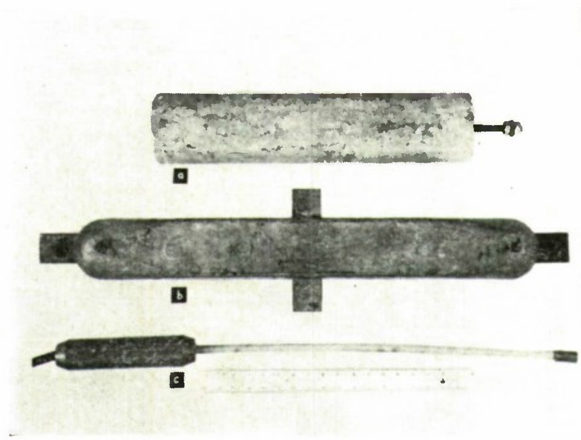


FIG. 1. Typical anodes:
(a) magnesium; (b) zinc for ships' hulls;
(c) platinized titanium.

The development of zinc anodes was a major advance particularly for the protection of ships. The low driving voltage of 0·25V between zinc and protected steel is an asset for this application since it is insufficient to cause much paint damage and the current output varies to some extent with the requirements of the structure. This is because changes in the potential of the structure cause a larger percentage change in the driving voltage between the zinc and the structure, and hence in the current. Ordinary zinc is unsuitable as the presence of more than 0·0015% iron causes the formation of a stifling layer of adherent product. The presence of 0·5% aluminium reduces the deleterious effect of iron and Crennell and Wheeler[2] showed that the presence of silicon in addition to aluminium in the melt enables the iron to be removed with the dross. The further addition of 0·1% cadmium makes the corrosion product softer and less adherent.

In recent years several aluminium anodes have been marketed for marine use. Their potential is similar to that of zinc but outputs ranging from 600 to 1300 A.h./lb. are obtained compared with 500 and 350 A.h./lb. available from magnesium and zinc respectively. Experience with these materials is limited and results sometimes fall short of expectations. The aluminium-5% zinc alloy has a potential some 50 mV more noble than that of zinc and its output is low. Addition of up to 0·1% tin to this alloy in conjunction with heat treatment can markedly improve both its potential and yield. Recently developed aluminium-zinc-indium and aluminium-zinc-mercury alloys are also capable of high output.

Impressed Current Anodes. Where scrap steel is readily available and periodic anode replacement is not regarded as a problem it forms a useful anode material provided the anodes are properly designed to avoid premature disconnection. Iron and steel are consumed in sea-water at the rate of 20 lb. per Ampere-year. Attack proceeds from the edges, so the connection should be made to the middle of the array and be well insulated, *e.g.* by potting in pitch. A regular replacement schedule may be based on the weight of material installed. Anodes weighing about a ton laid on the sea-bed have been widely used in Naval Dockyards for the protection of large fixed structures.

More durable anodes, at which the primary anodic reaction is the evolution of chlorine or oxygen are widely preferred. High silicon iron, containing 15% silicon, can be used in sea-water or in the ground. It is slowly consumed, however, and is brittle. Similar considerations apply to graphite in marine conditions. The principal use of graphite is for buried anodes surrounded by carbonaceous back-fill. Such anodes can have very long life and are used to protect the majority of buried structures.

The most durable anodes for use in sea-water are lead-silver alloy and platinised titanium. The lead alloy contains 1 to 2% silver and usually 6% antimony. It develops a conducting, self-renewing, chocolate-brown layer of lead dioxide which protects the underlying lead and acts as a site for chlorine evolution. At current densities greater than 10 A/sq. ft., blisters containing white non-conducting lead chloride may develop (Fig. 2) but at lower current densities a life of over 10 years may be anticipated. Platinum spines inserted into the lead surface enable much higher current densities to be used. Lead anodes deteriorate rapidly in mud and therefore cannot be used on the sea-bed.

Platinised titanium is the most promising anode material for use in both fresh and salt waters. It consists of a thin layer of platinum, 0·0001 to 0·00025 inch thick, plated on to titanium, often in the form of rods as in Fig. 1(c). Titanium is virtually inert at potentials less positive than +8V but is itself incapable of action as an anode. Platinised titanium can carry very high current densities, 100 A/sq. ft. being common, and its life is reputed to be much longer than that of any of the materials mentioned above. Failures

FIG. 2. Blisters formed on lead-silver alloy anode at 15 A/sq.ft.

have been experienced however, often due to bad design which allows the voltage on the titanium to exceed 8V but it has also been suspected that electrical ripple may cause dissolution of the platinum. The evidence for this effect is inconclusive and it is maintained that it does not occur if full-wave rectification is used. Platinised tantalum has also been used on a limited scale as it can tolerate very high voltages without breakdown.

APPLICATIONS

Buried Structures. Of the 250,000 miles of oil pipelines and 600,000 miles of natural gas lines in the world, the majority is buried and a high proportion is cathodically protected. Modern coatings reduce the required current density to very low levels so that in undeveloped parts of the world anodes may be sited great distances apart. In more developed areas however the possibility of interaction corrosion of other structures greatly limits the disposition of ground beds so that multiple low current anodes must be sited close to the structure. Alternatively the 'deep well' technique may be used: an anode is sited in an augured hole 100 ft. or more below the surface to give an excellent current spread but this does not avoid interaction between long crossing lines. In soils of 3,000 ohm-cm. resistivity or less, magnesium anodes may be used but most underground structures are protected by graphite or silicon iron anodes. Current is adjusted manually at intervals of 3 months or more.

Potential measurements require that electrical contact must be made to the structure and provision for this must be made at the design stage. All parts of a structure must be bonded, usually by welded straps.

Marine Structures. The natural life of jetties and sheet piling is liable to be 30 to 60 years even without protection but though corrosion prevention is not a matter of great urgency the cost of such structures and consequences of possible collapse make cathodic protection a worthwhile proposition, especially as coating renewal is not usually practicable. Massive iron anodes on the sea-bed have been widely used where there is no risk of interaction corrosion of ships moored alongside. Suspended lead or silicon iron anodes distributed among jetty piles confine the electrical gradients within the area of the structure. The modern method is to use platinised titanium rods mounted on brackets which can be secured to the piles above the water-line. Such anodes should be sited where they will not be damaged by shipping or used as moorings for small boats.

Isolated marine structures such as dolphins and drilling rigs are often protected by galvanic anodes. The current demand is fairly constant except in very shallow tidal water so that magnesium anodes are suitable for this purpose. As small anodes would require frequent replacement it is better to place a few large anodes on the sea-bed and connect them by cables to the structure. Zinc or aluminium can also be used but because of the lower driving voltage greater numbers are required. Very large anodes have been designed with a life of up to 10 years—a danger with these is that the requirement to replace them eventually may be forgotten!

Ships' Hulls. Corrosion, like fouling, progressively increases the skin friction and consequently the fuel consumption of an active ship. Cathodic protection not only reduces running costs but may significantly reduce maintenance in dry dock. Two factors which have a major influence on cathodic protection design are the variation of current requirement with water speed and possible damage to outer bottom paint. Consequently zinc is preferred to magnesium for galvanic protection and continuous control must be exercised over the output of impressed current systems either by manual adjustment or automatic control. Installation costs of an impressed current system are considerably greater than for a galvanic system but running costs are of minor importance as there is usually ample power available. Consequently impressed current systems are preferred for new construction and vessels of over 50,000 sq. ft. wetted area but zinc anodes may be more economical on older and smaller ships.

As impressed current anodes must be mounted on the hull, except in the case of the little-used trailing wire anode, steps must be taken to avoid paint damage due to the high potential gradients in their vicinity. They are therefore surrounded by a shield either as a sheet of insulating material (Fig. 3) or a high duty coating such as solventless epoxy resin. Conventional marine paints are damaged by overprotection but are often satisfactory at the general hull potential of −0·8 to −0·85V relative to Ag/AgCl. Coal-tar epoxy or vinyl paints however, which are more resistant to cathodic protection, are often used overall.

A manually controlled system requires adjustment whenever there is a large change of ship's speed so that the potential measured against hull mounted reference electrodes is kept within the required range. This is burdensome to ships' staff and subject to human error so that automatic control is generally preferred. Silicon controlled rectifiers are probably more stable than the magnetic amplifiers which have been widely used hitherto. The potential difference between the hull and fixed sensing electrodes is continuously compared with a preset potential difference and the anode currents are automatically increased or decreased accordingly. Zinc sensing electrodes are often used for this purpose—the hull is maintained at +200 to 250 mV relative to these electrodes—but Ag/AgCl is more stable.

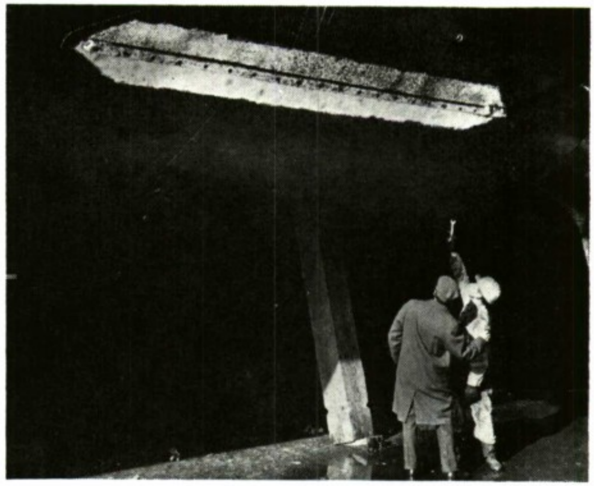


FIG. 3. Hull anode with plastics shield.

Because of high turbulence and possible galvanic coupling of the propellers to the hull, the greatest current requirement is in the stern area. Propellers are often bonded to the hull by a slip-ring and brush arrangement because normal contact is intermittent. This enables the propellers to share in the protection but can lead to severe corrosion of the stern area if the protection system is not working.

Ships' Tanks. The greatest use of galvanic anodes is for the protection of cargo compartments of oil tankers which contain alternately oil cargo and sea-water ballast. Chlorine evolution from impressed current anodes is unacceptable. Magnesium anodes were formerly used because of the weight factor and because they maintain a clean surface. The sparking hazard when steel falls on the anodes or when the anodes fall on the deck is now unacceptable. Zinc is therefore the primary material for this purpose despite its weight and the fact that crude oil may cause some clogging of the surface. Aluminium is increasingly used though there are restrictions on its disposition because of the slight possibility of incendive sparking on impact.

Plant. The internal protection of bare pipes is restricted by the limited spread of current so that it is only applied to pipes of 2 ft. diameter or larger. Anodes are sited some 5 to 10 diameters apart. Condenser water-boxes in power stations are now protected by platinised titanium anodes either as rods inserted through the wall or in the form of the so-called " coaxial anode "[3] in which long runs of platinised titanium surrounded by perforated plastics tube is secured to the interior of the structure close to the non-ferrous tube plate. Automatic control of these anodes to cope with varying water-speeds has reached a high degree of sophistication in the C.E.G.B. Small, cheap control units are now available which allow automatic control to be applied economically to quite small coolers.

FUTURE DEVELOPMENT

Cathodic protection has now reached the stage of technical maturity but there is still scope for further developments both in materials and techniques. New aluminium alloy anodes will probably be developed and rarer noble metals may prove more durable than platinum when plated on titanium or tantalum. Marine coatings are constantly being improved and high duty coatings less dependent on surface preparation would be widely welcomed. Cheap automatic controllers are already available and may well replace manual control of the protection of buried and marine structures. Standardisation of systems for typical applications will become more general as experience indicates the best economic compromises of capital and running costs for various combinations of coating and cathodic protection systems.

REFERENCES

(1) Corrosion Prevention and Control, **14**, (8), September 1967, p. 7.

(2) Crennell, J. T. and Wheeler, W. G., *J. Appl. Chem.*, **8**, 1958, p. 571.

(3) Matthewman, W., Corrosion Prevention and Control, **9**, (10), October 1962, p. 5.

THE HIGH TEMPERATURE SOLID OXIDE FUEL CELL—
A Literature Survey

D. R. Morris
University of New Brunswick

SUMMARY

The literature of the high-temperature solid oxide fuel cell is reviewed. The probable electrolyte compositions and operating temperatures of operable cell systems are defined, and the difficulties of fabrication of such systems are discussed, the problem areas of operation are isolated and the areas in which research work is required, before the successful construction of a solid oxide fuel cell can be undertaken, are outlined.

Introduction

Solid oxide electrolytes of the stabilized zirconia type have unique properties which make them attractive for use in fuel cells. In such a cell a mechanically stable electrode-electrolyte interface is possible with an all solid construction, and kinetic problems frequently encountered in fuel cells are reduced or eliminated at the operating temperatures (700° to 1000°C). Temperatures of this magnitude are necessary in order to achieve adequate electrical conductivity in the solid electrolyte.

The technology of fuel cells using solid electrolytes has reached a relatively advanced stage; in the U.S.A. a design of a coal gasification solid oxide electrolyte fuel cell system has been published[1]. A cost estimate of a "filter press" assembly has been made, whether a cell assembly could be competitive in capital and running costs with a diesel-electric generator would depend upon achieving good cell life and low cell manufacturing costs.

Dr. D. R. Morris is Associate Professor, Dept. of Chemical Engineering, University of New Brunswick, Canada. This work was carried out whilst he was engaged as a vacation consultant at A.M.L.

It is the purpose of this paper to review the technical literature upon fuel cells using stabilized zirconia as the electrolyte.

Fundamental Studies of Solid Electrolytes

Mechanism of Electrical Conductivity in Solid Electrolytes

Wagner[2] has pointed out the essential difference between an ionic and an electronic conductor of electricity. The passage of an electric current between a metal involving electronic conduction and an ionic conductor necessarily leads to an electro-chemical reaction at the boundary. At the cathode metal is deposited or non-metal dissolved, and at the anode, non-metal is liberated or metal is dissolved. It does not necessarily follow that both cations and anions are mobile. Tubandt[3] proved the validity of Faraday's law with an accuracy better than 0·1% for the silver halides, and showed that practically only the cations are mobile; while practically only the anions are mobile in the chlorides and bromides of lead and barium. However, in respect of the alkali halides and lead iodide, both cations and anions were found to be mobile.

In contrast to the behaviour of metal halides, many oxides and sulphides show little or no electrochemical reaction at the boundary metal/metal oxide or sulphide[4]. In these substances electrical conduction is due largely to the migration of electrons. It has also been shown that small deviations from the ideal stoichiometric ratio of metal to non-metal may greatly affect the mode of conductivity.

In respect of oxide systems, Kingery[5] has stated that ionic conduction is not expected to be appreciable until high temperatures, because of the high energy required to form lattice defects, necessary for ion movement, and the large activation energy of such movement. The presence of impurities or intrinsic semi-conducting properties at particular oxygen pressures normally give rise to sufficient electronic conduction to mask any ionic contribution. This is particularly true since the mobility of electrons is much greater than that of ions and even a small concentration of electrons can overwhelm ionic behaviour. Exceptions to this are compositions of zirconium oxide or similar materials, in which a lower valency ion such as calcium or yttrium is in solution. The zirconia then becomes a comparatively good conductor due to the mobility of oxygen ion vacancies, formed by the replacement of a tetravalent zirconium ion with, for example, a trivalent yttrium ion, which requires less oxygen ions than the zirconium, leaving vacancies in the lattice. The vacancies will be statistically distributed over the lattice, and the conductivity corresponds to the mobility of oxygen ion vacancies[6]. The electrical conductivity of a ZrO_2-CaO solid solution is linear with the inverse of the temperature over a wide temperature range. From this data and use of the Nernst-Einstein equation, the diffusion coefficient of the oxygen ion can be calculated. The results of these calculations agree well with experimental data.

The theory of electrical conductivity in such oxide systems has been presented by Kiukkola and Wagner[7] and by Kingery *et al.*[8]. The large concentration of oxygen ion vacancies is fixed by composition and is independent of oxygen pressure so that the ionic contribution will not be pressure dependent. However, as the oxygen pressure is changed there may be a change in the electronic contribution to the total conductivity. Thus, one may write for the formation of conduction electrons:*

$$O^{2-}_{lattice} = \square O^{2-} + \tfrac{1}{2} O_{2(g)} + 2\ominus \quad \ldots (1)$$

or for the formation of electron holes:

$$\tfrac{1}{2} O_{2(g)} + \square O^{2-} = O^{2-}_{lattice} + 2\oplus \quad \ldots (2)$$

Clearly the concentrations of the excess electrons or electron holes depend on the oxygen pressure. In view of the low total conductivity, consequent upon the low concentrations considered, association or interaction effects may be assumed to be small. Then applying the mass action law to the above equations:

$$K_1 = \frac{C^2_{\ominus} \cdot P_{O_2}^{\frac{1}{2}} \cdot C_{\square O^{2-}}}{C_{O^{2-}_{lattice}}} \quad \ldots (3)$$

$$\text{and } K_2 = \frac{C^2_{\oplus} \cdot C_{O^{2-}_{lattice}}}{C_{\square O^{2-}} \cdot P^{\frac{1}{2}}_{O_2}} \quad \ldots (4)$$

Since $C_{O^{2-}_{lattice}}$ and $C_{\square O^{2-}}$ are fixed by the composition, the concentrations of electrons and of the electron holes are given by:

$$C_{\ominus} = k_1 \; P_{O_2}^{-\frac{1}{4}} \quad \ldots (5)$$

$$C_{\oplus} = k_2 \; P_{O_2}^{+\frac{1}{4}} \quad \ldots (6)$$

Even if the concentration of electrons or electron holes is small they make a substantial contribution to the electrical conductivity since the mobility is high. The total conductivity is given by:

$$\sigma = \sigma_{ion} + F \mu_{\ominus} C_{\ominus} + F \mu_{\oplus} . C_{\oplus} \quad \ldots (7)$$

$$\sigma = \sigma_{ion} + k_3 \; P_{O_2}^{-\frac{1}{4}} + k_4 \; P_{O_2}^{+\frac{1}{4}} \quad \ldots (8)$$

Equation 8 predicts a pressure dependence of the total conductivity if there is any appreciable contribution due to electrons or electron holes.

Table 1 summarizes the experimental results of A.C. conductivity measurements as a function of the oxygen pressure. It was found that the electrical conductivity of $(ZrO_2)_{0.85}(CaO)_{0.15}$ was independent of oxygen pressure. In the case of $ThO_2-La_2O_3$ and ThO_2-CaO solid solutions however, a measure of electronic conductivity was found. This may have been due to the presence of impurities[7].

Studies of the type of conductivity of solid oxides have also been made by measuring the electro-motive force of the cells: [7, 9, 10]

---

* For nomenclature see page 34.

Me^1, Me^10, Pt	Solid electrolyte	Pt, $Me^{11}0, Me^{11}$	I
$0_{2(P1)}$,Pt	Solid electrolyte	Pt, $0_{2(P2)}$	II

These cells are concentration cells; the e.m.f. is given by:

$$E= \frac{RT}{4F} \cdot \ell n \left(\frac{P_2}{P_1}\right)_{0_2} \quad \ldots (9)$$

For an electrolyte which is behaving as an ionic conductor the measured e.m.f. will equal E, the e.m.f. calculated from thermodynamic relations.

The qualitative results of these studies are included in Table 1.

The above studies have shown that the transfer number for the oxygen ion in, for example, $(Zr0_2)_{0.85}$ $(Ca0)_{0.15}$ and $(Zr0_2)_{0.9}$ $(Y_20_3)_{0.1}$ is near unity. Since the prime conducting species in these oxides is the oxygen ion, an estimate of the oxygen ion transfer number may be obtained by measuring the decrease in d.c. conductivity of a specimen as the surrounding atmosphere is depleted of oxygen. The removal of oxygen reduces the anionic conductivity by elimination of the conducting species. By allowing for the variation of electron conductivity due to the partial pressure of oxygen, the residual conductivity can then be attributed to electronic and/or cationic conduction. Tests based on this principle were conducted by Strickler and Carlson(11) upon solid solutions of $Y_20_3-Zr0_2$. They concluded that the transfer number of the oxygen ion in such solutions was greater than 0·99. From the above studies it may be concluded that solid oxide solutions of the stabilized zirconia type are ionic conductors of electricity, further, that the mobile species is the oxide ion with a transport number near unity.

Thermodynamic Studies using Solid Oxide Electrolytes

The establishment of specific solid solutions of oxides as anionic conductors of electricity has enabled their use for the determination of thermodynamic data on metal/oxide systems by the measurement of the e.m.f. of galvanic cells of Type I. As an example of this technique, reference should be made to the study of the lead-oxygen system by Alcock and Belford(12). These authors used as electrolytes, solid solutions of either 14 mol% Mg0 in $Zr0_2$ obtained from Degussa, Frankfurt/Main, or 15 mol% Y_20_3 in $Th0_2$ supplied by Thermal Syndicate, Wallsend. A predominance of anionic conductivity had been demonstrated in both electrolytes by showing that the conductivity was independent of partial oxygen pressure in the range 10^{-11} to 10^{-16} atm. The conductivities at 800°C were found to be sensibly constant at $5{\cdot}2 \times 10^{-4}$ and $4{\cdot}6 \times 10^{-5}$ $(ohm.cm)^{-1}$ respectively, that is considerably lower than those measured in the $Zr0_2/Ca0$, $Zr0_2/Y_20_3$ systems as shown in Fig. 1.

Physical Properties of Solid Oxide Electrolytes

Phase Relationships

The system $Zr0_2-Ca0$ exhibits a limited solid solution as shown in the phase diagram given by Kingery(5). Pure $Zr0_2$ exhibits a monoclinic to tetragonal phase transition at 1000°C which involves a large volume change and makes the use of pure $Zr0_2$ impracticable as a ceramic material. Addition of Ca0 to form the cubic solid solution, which has no phase transition, is the basis of "stabilized zirconia". Duwez *et al.*(13) have also studied the system $Zr0_2-Y_20_3$ and concluded that no solid state transformation occurs when the yttria concentration is above approximately 6 mol per cent. They also report that oxides similar to yttria, namely scandia and the rare earth oxides from elements number 62 to 71 should exhibit the same properties when added to zirconia.

These systems have also been studied by Strickler and Carlson(11). In the $Zr0_2-Y_20_3$ system, the low-yttria cubic solid solution boundary was found to be between 9 and 9·5 mol % Y_20_3. The high cubic solid solution phase boundary fell between 45 and 50 mol% Y_20_3, in agreement with the results of Duwez *et al.*(13) In the $Zr0_2-Ca0$ system the solid solution low and high limits were placed at 12 to 13 and 20 to 21 mol% Ca0 respectively, in agreement with the findings of Tien and Subbarao(14). Some studies were also conducted upon the ternary system $Zr0_2-Y_20_3-Ca0$.

Phase diagrams for the systems $Zr0_2-Mg0$, $Zr0_2-Ca0$, $Zr0_2-Ca0$, $Zr0_2-La_2\,0_3$, $Zr0_2-Y_20_3$, and $Zr0_2-T_i0_2$ have been presented by Ryshkewtich(6) based upon the work of Duwez *et al.*

Transport Properties

Of primary importance in any attempt to exploit the solid oxides as electrolytes in a

TABLE 1. Type of Conductivity of Solid Oxides.

System Composition (Mol %)	Temperature °C	Range of Oxygen Pressure Atm.	Type of Electro-chemical Cell	Electrical Cond. ohm^{-1} cm^{-1}	Type of Cond.	Ref.
85% ZrO_2, 15% CaO	870	10° to 10^{-22}	–	$1{\cdot}6 \times 10^{-14}$ constant	ionic	7
ditto	1427, 1627	10° to 10^{-10}	–	0·27,0·68 constant	ionic	8
85% ThO_2, 15% La_2O_3	870	10° to 10^{-22}	–	variable	mixed	7
75% ThO_2, 25% LaO_3	870	ditto	–	variable	mixed	7
85% ThO_2, 15% CaO	870	ditto	–	variable	mixed	7
100% ZrO_2	600 – 1000	–	I	–	mixed	9
2% CaO, 98% ZrO_2	800 – 1000	–	I	–		
10% CaO, 90% ZrO_2	600 – 1000	–	I	–	ionic	7, 9
36% CaO, 64% ZrO_2	800 – 1000	–	I	–		
40% CaO, 60% ZrO_2	750 – 1100	–	I	–		
70% CaO, 30% ZrO_2	600 – 1000	–	I	–	mixed	9
100% CaO	800 – 1000	–	I	–	mixed	9
100% ThO_2	700 – 1000	–	I	–	mixed	9
15% CaO, 85% ThO_2	600 – 1000	–	I & II	–	mixed	9
50% CaO, 50% ThO_2	800 – 1000	–	I	–	mixed	9
70% CaO, 30% ThO_2	600 – 1000	–	I	–	mixed	9
15% MgO, 85% CeO_2	600 – 1000	–	I & II	–	mixed	9
15% CaO, 25% CeO_2, 60% ZrO_2	600 – 1000	–	I	–	mixed	9
100% TiO_2	600 – 1000	–	I	–	electronic	9
91% ZrO_2, 9% Y_2O_3	1000	Anion Transfer Number Determination		–	ionic	11
15% CaO, 85% ZrO_2	600 - 1100	–	I & II	–	ionic	7, 9

fuel cell system are the electrical conductivity and the oxygen ion diffusion coefficient. It is generally accepted that the conductivity of stabilized zirconia is essentially ionic over a wide temperature range and that the mobile ions are the anions.

In an ionic conductor the electrical conductivity and the transfer number (fraction of the total current carried by each charged particle) are related to the diffusion coefficient and absolute mobility by the Nernst-Einstein equations:—

$$D_i = B_i\ k\ T \qquad \ldots (10)$$

$$\sigma_i = n_i\ z_i^2\ e^2\ B_i = \sigma t_i \qquad \ldots (11)$$

$$D_i = \frac{\sigma t_i\ k\ T}{n_i\ z_i^2 e^2} \qquad \ldots (12)$$

These relationships have been verified for a number of solid electrolytes[5].

The magnitude and temperature dependence of conductivity depends on the concentration of vacancies or interstitial ions that are capable of contributing to the conduction process and on the energy required for moving these defects through the structure. Under conditions in which the concentration of lattice defects is fixed, the temperature dependence of conductivity is of the form:

$$\sigma = \sigma_o \exp.(-u/kT) \qquad \ldots (13)$$

Experimental Data on Electrical Conductivity of Ionic Oxides

Typical results of a number of experimental investigations in the ZrO_2-CaO, $ZrO_2-Y_2O_3$ and $ZrO_2-Yb_2O_3$ systems are given in Fig. 1. It may be seen that the results are of the form of equation (13).

Strickler and Carlson[11], studying cubic solid solutions in the ZrO_2-CaO, $ZrO_2-Y_2O_3$ binaries and $ZrO_2-CaO-Y_2O_3$ ternary system reported the highest conductivity to lie in the $ZrO_2-Y_2O_3$ binary at 9 mol% Y_2O_3, near the low yttria cubic solid solution limit. They stated that this appeared to be due to a lower activation energy rather than to the number of oxygen vacancies dictated by composition. In the system ZrO_2-CaO, the maximum conductivity occurred at the composition 13 mol% CaO[11, 14, 15, 16].

More recently Strickler and Carlson[17] studied the electrical conductivity and phase relations of several $ZrO_2-M_2O_3$ systems, where M represents La, Sm, Y, Yb, or Sc, in order of decreasing cation and size. Generally, they found the electrical conductivity, for those systems which formed cubic solid solutions, to increase as the size of the substituted cation decreased. The electrical conductivity for each binary system attained a maximum at about 10 to 12 mol% M_2O_3. They also reported the region of the cubic solid solution to be narrow, at ambient temperature, in the case of the $ZrO_2-Sc_2O_3$ system (6 to 8 mol% Sc_2O_3). The data for the $ZrO_2-Yb_2O_3$ system is included in Fig. 1.

In all these systems, the electrical conductivity decreases with further addition of stabilizing oxide. Since the vacancy concentration increases with such addition the electrical conductivity would be expected to increase, but this is not so[17, 18]. The reason for this is not fully understood, possible explanations have been put forward by Strickler and Carlson, who conclude that further work is necessary to clarify the mechanism responsible.

For the composition 85 mol% ZrO_2, 15 mol% CaO, over the temperature range 700°C to 1725°C, Kingery *et al.*[8] give

$$\sigma = 1{\cdot}50 \times 10^3 \exp\left(\frac{-1{\cdot}26}{kT}\right) \qquad \ldots (14)$$

The activation energy, 1·26 e.v. agrees well with that calculated from the results of Hund[19] and of Volchenkova and Palguer [15]. Tien[20] has studied the effect of micro-structure on the ionic conductivity of 86 mol% ZrO_2, 14 mol% CaO. At temperatures below 800°C specimens of smaller grain size exhibited a slightly higher electrical conductivity. Above 1000°C, differences due to micro-structure were insignificant.

The systems ZrO_2-CaO and HfO_2-CaO have been reported as having similar electrical conductivities[20a].

Experimental Data on Oxygen Ion Diffusion

Kingery *et al.*[8] determined oxygen ion diffusion coefficients directly by measuring the rate of exchange of oxygen between a gas phase enriched in O^{18} and heated spherical particles of 85 mol% 15 mol% CaO over temperature range 700°C to 1100°C. The experimental data can be represented by the relation:

$$D = 1{\cdot}0 \times 10^{-2} \exp\left(\frac{-1{\cdot}22}{kT}\right) \qquad \ldots (15)$$

One can calculate the electrical conductivity resulting from the ion mobility determined from the diffusion coefficient by means of equations 10 to 12 (see Reference [5] p.649 for the necessary conversion factors). Within experimental error, the entire electrical conductivity can be attributed to oxygen mobility, *i.e.* the transfer number for oxygen ions in this system is near unity.

Rhodes and Carter[21] have also studied ionic diffusion in this system. They showed the anion diffusion coefficients, calculated from thermal conductivity to be greater than the cation diffusion coefficients by a factor of 10^6.

Experimental Data on Electron and Cation Transport Numbers

It will be apparent that transport numbers in solid oxides are of considerable importance in their use as electrolytes in galvanic cells. If more than one charge carrier contributes to the electrical conductivity, a partial conductivity for each particle may be defined: Thus for particle i:[5]

$$\sigma_i = \mu_i(n_i z_i e) \qquad \ldots (16)$$

The total conductivity is then:

$$\sigma = \sigma_1 + \sigma_2 + \ldots\ldots + \sigma_i + \ldots\ldots \qquad \ldots (17)$$

The fraction of the total conductivity contributed by each charge carrier is:

$$t_i = \frac{\sigma_i}{\sigma} \qquad \ldots (18)$$

Clearly, the sum of the transport number must be unity:

$$t_1 + t_2 + \ldots + t_i + \ldots = 1 \qquad \ldots (19)$$

The charge carriers may be ions, electrons or electron holes.

A number of experimental studies have shown that the transport number of the oxygen ion is close to unity. Additional evidence has been given by Weissbart and Ruka[21] who found Faraday's laws to be obeyed closely for current flow through $(ZrO_2)_{0.85}(CaO)_{0.15}$ as determined by microweighing of the oxygen transferred. The electronic contribution to the total conductivity was found to be less than 0·5% near 1000°C in an oxygen atmosphere.

Transport numbers for the stabilizing cations can be determined by measuring composition changes that result from the passage of an electric current through the electrolyte. Such a study has been made by Bray and Merton[23] in respect of yttria stabilized zirconia. They found the apparent electron transport number to be $\leqslant$ 0·005 in an oxidizing atmosphere and $\leqslant$ 0·01 in a reducing atmosphere at 900° - 1100°C and the apparent transport number for yttrium ions to be $\leqslant 4 \times 10^{-9}$. They also conclude that the zirconium ions move very little with respect to one another.

As pointed out by Bray and Merton, cation transport could be a serious problem in a fuel cell. They show that for operation at about 1000°C, with an electrostatic potential difference across the electrolyte, due to ohmic losses of the order of 0·1 volt, electrolyte unmixing can be a problem of practical significance due to the possibility of the precipitation of a new phase at one of the electrodes. However, as stated earlier experimental investigation has demonstrated a very small cation mobility.

Mechanical Properties of Solid Oxide Electrolytes

Elasticity and Strength

Some of the major reasons why ceramics generally are not used more widely are the facts that they fail with brittle fracture and that their impact resistance is low. This is due to a process in which fracture occurs with little or no plastic deformation.

Kingery[5] has summarized the theory of brittle fracture. He suggests that two steps are required, first the production and then the propagation of a crack to final fracture. Either can control the overall failure process. He points out that the theoretical strengths calculated from atomic bonding forces (of the order 1 - 10 $\times$ 10^6 p.s.i.) are several orders of magnitude greater than typical observed strength values (of the order 1 - 100 $\times$ 10^3 p.s.i.). These differences are believed to be due to the presence of small cracks or flaws around which a strong stress concentration arises when the solid is stressed. This is supported by the experimental observation that freshly drawn glass rods show strengths up to 10^6. These values drop instantly to the usually observed values of about 20,000 p.s.i. when the sample is touched by a hard body. Many experiments have illustrated that the flaws leading to fracture are normally present at the surface rather than in the interior. This observation is the basis of a number of industrial processes. Thus, toughened glass is produced by quenching the surfaces from high temperatures in an air blast; when the interior cools its contraction produces a compressive stress in the already colder and rigid surface layers. This means that an additional tensile load must be applied in order to bring the surface tensile stress up to the fracture level.

Alternatively, the surface of a ceramic may be coated with a material of lower thermal expansion coefficient which will be under compressive stress upon cooling.

With regard to ceramics of the stabilized zirconia type (solid oxide electrolytes), little quantitative information is available. Some data are given by Kingery and Ryshkewitch on "sintered-stabilized zirconia" but the values will probably depend in large measure upon the method of preparation of the specimens. Some information is also available in the literature[24, 25]. The possibility of improving the mechanical characteristics if found to be unsatisfactory by surface coating, using a solid oxide electrolyte of lower thermal expansion than the substrate oxide electrolyte might be investigated. This is clearly a matter for experiment.

Thermal Properties

It is common experience that ceramic materials possess a poor resistance to thermal shock. The effect of thermal stresses on different kinds of materials depends on the stress level, stress duration and also on material characteristics such as ductility, homogeneity and porosity. Hence it is impossible to define a quantitative "thermal stress resistance factor". For a full discussion, reference should be made to Kingery[5].

Some quantitative information is available in the Russian literature[25, 26] in respect of stabilized zirconia.

Nielsen and Leipold[27] have published data on the thermal expansion of zirconia stabilized with yttria. It is in the range 9 to 13×10^{-6} °C^{-1} between 200 and 2200°C. Ryshkewitch[6] gives the thermal expansion of zirconia stabilized with calcia as 11 to 12 $\times10^{-6}$ °C^{-1}. He also describes stabilized zirconia as having a "good thermal shock resistance".

Fabrication Techniques

For the fuel cell to be a workable proposition, the electrolyte must be impermeable to the gases present on either side and techniques of fabrication as a thin film are essential. This last requirement follows from the electrical conductivity of the solid oxide electrolyte. Techniques of fabrication to produce discs and tubes of the order 0·05 cm thick have been developed by U.S. workers [1]. No information of the techniques used is available. However, there is information upon a number of possible processes. These are electrophoresis, evaporation, halide hydrolysis and spraying.

Relevant to the problem of fabrication is the subject of grain growth in the solid solutions produced. Tien and Subbarao[28] have studied isothermal grain growth in stabilized zirconia of composition 84 mol% ZrO_2, 16 mol% CaO in the temperature range 1600° to 2000°C. They found that the grain size increased as the 0·4 power of the time; the activation energy was about 80k cal/mol. It is clear from this work that in order to reduce porosity by sintering, temperatures of this order are necessary.

Electrophoresis

Collins[29] has recently reported on the production of alumina bodies by electrophoresis. The process involved deposition upon an electrode which was subsequently removed and the ceramic body consolidated by sintering. No information is available on the possibilities of using this technique for stabilized zirconia.

Evaporation

Schultz *et al.*[30] have described a technique which involves impingement of a high energy beam of electrons on a target of oxide, vaporization of the oxide and condensation of the vapours on a cool substrate. They stated that the films were porous and not sufficiently adherent. They also reported some change in composition of the stabilized zirconia during evaporation, presumably due to the different vapour pressures of the two components.

Halide Hydrolysis

This technique involves the hydrolysis of a volatile salt and subsequent deposition of the oxide on a hot substrate. Thus for example:[31]

$$2AlCl_{3(g)} + 3H_2O_{(g)} \xrightarrow{t\sim1000^\circ C} Al_2O_{3(s)} + 6HCl_{(g)} \quad \ldots (20)$$

The technique has also been used for coating with zirconia[32], but calcium oxide presents greater difficulties due to the relative non-volatility of calcium compounds.

Spraying

Basically, spraying processes involve introduction of the material to be sprayed into a high temperature gas stream and projecting the resultant molten particles at high speed at a substrate. Oxy-hydrogen or oxy-acetylene guns are most commonly used. Alternatively there is the detonation gun using oxy-acetylene and the plasma arc torch, which uses electrical energy. Flame spraying guns have maximum gas temperatures of about 3000°C, whereas plasma-arc equipment is capable of attaining 20000°C.

The techniques of flame spraying have been reported in the literature in respect of alumina[33] and stabilized zirconia[34]. Thus Bliton *et al.*[34] have described the production of stabilized zirconia films down to 0·01 cm thick by flame spraying on to a metal disc pre-coated with sodium chloride. The film was removed from the substrate by dissolution of the chloride. The films were subsequently impregnated with zirconyl nitrate solution to reduce the porosity and fired to 1500°C. Platinum electrodes were subsequently applied by use of a commercial platinum paste and firing at 870°C. In an earlier paper these authors detailed methods of measuring the physical properties of the films so produced[35].

Some general information on the use of plasma equipment is available in the literature[36] but little quantitative information on the physical properties of the products. It is claimed to produce a denser deposit than the flame-spraying process.

The Solid Oxide Electrolyte Fuel Cell

Before proceeding to an account of the attempts to assemble a solid oxide fuel cell some mention should be made of the supposed advantages of this type. The practical operating temperature of a fuel cell is dictated by kinetic factors rather than thermodynamic. The fuel of major interest is methane, or hydrocarbons generally, direct electrochemical oxidation of which is not practicable at temperature below about 400°C due to low reactivity. Hence the interest in high temperature cells in which hydrocarbons may be reformed with steam in the cell or external to the cell to produce hydrogen and carbon monoxide which are subsequently oxidized electrochemically:

$$CH_{4(g)} + H_2O_{(g)} \rightleftharpoons CO_{(g)} + 3H_{2(g)} \quad \ldots (21)$$

$$CH_{4(g)} + 2H_2O_{(g)} \rightleftharpoons CO_{2(g)} + 4H_{2(g)} \quad \ldots (22)$$

The standard free energy changes of these reactions at 500°K and 1000°K are respectively +26·1k cal and −6·4k cal for the reaction 21 and + 18·2k cal and −7·2k cal for the reaction 22. The technology of these reactions is well established.

Reactions which may prove troublesome when using hydrocarbons as the fuel, due to carbon deposition are:—

$$2CO_{(g)} \rightleftharpoons CO_{2(g)} + C_{(s)} \quad \ldots (23)$$

$$CH_{4(g)} \rightleftharpoons 2H_{2(g)} + C_{(s)} \quad \ldots (24)$$

Reaction 24 is a typical thermal decomposition reaction. The mechanism of these reactions is complicated as intermediate hydrocarbons of high molecular weight are formed. The standard free energy changes for these reactions at 500°K and 1000°K are respectively −20·2k cal and +1·1k cal for reaction 23 and +7·8k cal and −4·6k cal for reaction 24. Reaction 23 can be particularly troublesome due to its ability to transport carbon (in effect) from regions of high temperatures to regions of lower temperature.

The two principal schemes of interest in the field of high temperature fuel cells are the carbonate electrolyte cell and the solid oxide electrolyte cell. Shultz *et al.*[30] have discussed the relative merits of the two schemes. They list the disadvantages of the carbonate cell as follows:

1. Molten carbonate require a finite partial pressure of carbon dioxide at both anode and cathode, to prevent decomposition.
2. The carbonate electrolyte is corrosive.
3. The carbonate electrolyte presents retention difficulties.

The ideal properties of a solid oxide electrolyte are listed by them as follows:

1. Oxide ion transport number of unity.
2. Rapid migration of oxide ions: i.e. high electrical conductivity.
3. Chemical stability in the presence of reducing gases or oxygen at temperatures up to 1000°C.
4. Physical stability.
5. Impermeability to gases.

The extent to which these ideal properties are attainable will become apparent. It is Item 2 which is the principal factor in determining the operating temperature.

The first attempt to construct a solid oxide electrolyte fuel cell was made by Baur and Preis[37] in 1937. The electrolyte was in the form of a ceramic tube of $(ZrO_2)_{0.85}(CaO)_{0.15}$, a material which had been shown by Nernst[38] to possess a relatively high ionic conductivity. However, the performance was poor probably due to excessive electrolyte thickness, and it was concluded that sufficient cation conductivity was occurring to damage the cell.

However, a fuel cell is merely a special type of galvanic cell, in which the free energy change of oxidation of a fuel is converted directly to electrical energy, thereby avoiding the limitations of the Carnot cycle imposed whenever intermediate conversion of the chemical energy into heat is practised. A number of recent researches have shown that oxide solid solutions of the stabilized zirconia type are ionic conductors of electricity and further, that the transport number of the oxygen ion is near unity. Thus, there is no reason in principle why a cell of Type II (solid oxide electrolyte) should not be utilized as a fuel cell, in which the oxygen partial pressure at one electrode (the anode) is maintained at a low level by reaction with a fuel, for example, hydrogen.

Such an attempt was made by Weissbart and Ruka[22] in 1962. They had previously constructed vacuum tight cells of Type II:[39]

$$O_{2(P1)},Pt \mid (ZrO_2)_{0.85}(CaO)_{0.15} \mid Pt, O_{2(P2)}$$

The voltages obtained were in good agreement with the theoretical e.m.f. calculated from the thermodynamic relationship for reversible transfer of oxygen for cathode to anode. (Equation 9).

$$E = \frac{RT}{4F} \ell n \left(\frac{P_2}{P_1}\right)_{O_2} \quad \ldots (9)$$

This showed that the cells act as oxygen concentration cells. Such cells are useful for measuring oxygen partial pressures in, for example, furnace gases[39, 40] at temperatures above 600°C.

In the fuel cell variant of the above, it was arranged to flow pure oxygen at atmospheric pressure past the cathode and various water/hydrogen or water/methane mixtures through the anode chamber. The cell may be written as (using water/hydrogen as the example):

$$\begin{array}{l|c|l} H_2,H_2O.Pt & & Pt.O_2 \\ & (ZrO_2)_{0.85}(CaO)_{0.15} & \\ \text{Anode } (P_1)_{O_2} & & \text{Cathode } (P_2)_{O_2} \end{array}$$

The overall cell reaction is:

$$2H_{2(g)} + O_{2(g)} \rightleftharpoons 2H_2O_{(g)} \quad \ldots (25)$$

The equilibrium constant is related to the free energy change by: (for cell operation at pressures near atmospheric)

$$\ln K = \ln \left[\frac{P^2_{H_2O}}{P_{1_{O_2}} \times P^2_{H_2}}\right] = -\frac{2\Delta G^\circ}{RT} \quad \ldots (26)$$

Substituting in (9) for $(P_1)_{O_2}$ from (26):

$$E = -\frac{RT}{4F}\left[\frac{2\Delta G^\circ}{RT} + 2\,\ell n\left(\frac{P_{H2O}}{P_{H2}}\right) - \ell n(P_2)_{O_2}\right] \quad \ldots (27)$$

Weissbart and Ruka found agreement to within 5 mv of their measured e.m.f. with that predicted from Equation 27.

For a fuel cell to be useful as an energy conversion device, it must be capable of maintaining the e.m.f. when current is drawn from the cell. Under these conditions, the e.m.f. may be lower than the open circuit value, as calculated by the above procedures due to:

(a) activation polarization
(b) concentration polarization
(c) ohmic resistance of the cell.

The drop due to each of these is a function of the current density:

(a) depends on the particular rate determining chemical reactions at the electrodes,
(b) involves mass transfer of reactants and products and
(c) depends largely on electrolyte resistance.

Tests conducted by Weissbart and Ruka indicated that the cell had an output for the hydrogen/oxygen reaction essentially resistance limited and thus should be capable of higher output by decreasing the thickness of the electrolyte.

Tests with methane and steam as the fuel gases indicated that the major electrochemical reactions probably involved the reformed gases CO and H_2 rather than methane.

This cell has been further developed by Archer *et al.*[1] using $(ZrO_2)_{0.9}(Y_2O_3)_{0.1}$ as the electrolyte. A typical cell was constructed using a disc of the zirconia, calcia electrolyte, 5 cm in diameter and 0·04 cm thick; the resistance was about 0·1 ohm at 1000°C; this agrees with a calculation based on the data of Fig. 1. Porous platinum electrodes were applied to each side of the disc. Air was fed to the cathode chamber (partial pressure of oxygen equal to about 0·2 atm) and fuel fed to the anode chamber thereby reducing the oxygen partial pressure to about 10^{-16} atm and giving rise to a terminal voltage of about 1·0 volt. Archer *et al.* give the working characteristics of the device as follows:

Open circuit voltage (with hydrogen)	1·15 volts
Operating temperature	1010°C
Maximum power	0·85 watts
Voltage at maximum power	0·55 volts
Current density at maximum power	160 ma/cm²

A further development involved the construction of short, cylindrical electrolyte segments shaped so that they could be fitted one into the other and connected into a long tube by bell and spigot joints. Electrodes were applied to the inside and outside of these segments which then had an overall resistance of about 0·2 to 0·3 ohm. The inner electrode of one segment is attached to the outer electrode of the electrode of the adjacent segment, in this way connecting the individual segments electrically in series. The cell segment characteristics were as follows:—

Overall length	1·11 cm
Mean diameter	1·07 cm
Electrolyte material: $(ZrO_2)_{0.9}(Y_2O_3)_{0.1}$	
Electrode material: Porous platinum	
Segment weight	2 gm
Segment volume	2 cc
Electrolyte thickness	0·05 cm

Archer *et al.* give performance characteristics of a three cell assembly with hydrogen fuel fed to the inside of the tube and air or oxygen to the outside; the following data pertain to the tests using air:

Open circuit voltage	2·6 volts
Operating temperature	1000°C
Maximum power	2·1 watts
Voltage at maximum power	1·2 volts
Current density at max. power	450 ma/cm^2

Performance was slightly improved using oxygen instead of air as would be expected on theoretical grounds.

Binder *et al.*[42] have also described a fuel cell of this type using a disc of $(ZrO_2)_{0.85}$ $(CaO)_{0.15}$ as the electrolyte and porous platinum as the electrodes. The cell was operated with hydrogen, carbon monoxide, reformed propane and reformed hexane. They stated that reforming of the hydrocarbons was necessary to avoid thermal cracking reactions. The reforming reaction was carried out in a converter prior to the electrochemical reaction. The measured open-circuit e.m.f. agreed with the calculated e.m.f. to within 1%. Measurements of the cell voltage as a function of the current density up to 100 ma/sq cm revealed a minor polarization at small current densities in addition to the voltage drop due to ohmic resistance of the electrolyte. The electrolyte conductivity computed from the slope of the current-voltage plots was the same as that determined by a.c. measurements.

The majority of studies have been conducted using platinum electrodes. The essential properties of the electrodes may be listed as follows:

1. They must be electronic conductors of electricity.
2. The materials must have a low vapour pressure to minimize loss by vaporization.
3. The materials must be thermodynamically stable in the respective environments, *i.e.* the anode must be stable in a reducing atmosphere and the cathodes in an oxidizing atmosphere.
4. The bond between the porous electrode and the electrolyte must be mechanically and chemically stable.

A number of patents have been taken out in respect of these fuel cells.[42, 43, 44] For example, Tragert has taken out patents in respect of a design using a carbon fuel electrode (anode), a molten silver oxygen electrode (cathode) and a solid oxide electrolyte. Mobious has suggested the use of electrodes of the same structural type as the electrolyte but containing cations of variable valency resulting in electronic conductivity. The materials of the electrodes and electrolyte may then be sintered together. The advantages of such an electrode material are claimed to be mechanical stability and the avoidance of losses due to vaporization, inherent in the use of metallic electrodes.

It is apparent from the foregoing discussion that the technology of the solid oxide fuel cell is sufficiently advanced to enable construction of such a cell; the major factor limiting the output is the electrical resistance of the electrolyte; the major difficulty in fabrication of the cell is the manufacture of the electrolyte to close tolerances.

Conclusion

It has been shown that solid oxide electrolytes of the stabilized zirconia type in which a zirconium ion is substituted by a cation of lower valency (*e.g.* Ca^{2+}), are ionic conductors of electricity. The transport number of the oxygen ion is near unity.

In order to achieve ionic conductivity in these oxide systems, operating temperatures of at least 600°C are necessary. Realistic operating temperatures are in the range 800°C to 1000°C to achieve high conductivity. The most highly conductive materials are $(ZrO_2)_{0.9}$ $(Y_2O_3)_{0.1}$. A final choice must be based upon economic and technological considerations. Materials utilizing other rare earth oxides to stabilize the zirconia and impart ionic conductivity would probably be too expensive.

A fuel cell based upon such an electrolyte has been shown to be a workable proposition using hydrogen or reformed hydrocarbons, with none of the polarization problems associated with low temperature fuel cells. The output of the cell has been shown to be limited almost entirely by the resistance of the electrolyte.

It is apparent that the success of the cell depends largely upon the ability to fabricate the ceramic oxide electrolyte in thin, impermeable sections and to close tolerances, particularly if ceramics of tubular form are desired.

Little information is available upon the mechanical and thermal properties of these ceramic materials. It is recommended that, should it be decided to commence research aimed at building a high temperature fuel cell, attention should be devoted to fabrication problems, and the mechanical and thermal properties of the materials produced.

In order to achieve maximum density in the ceramic, a vaporization process would appear to be most advantageous as involving particles of molecular size. However, such a process is likely to be slow and expensive, in view of the high temperatures necessary. Of the alternative techniques, plasma flame spraying of the ceramic on a heated substrate would appear the most promising.

If possible it would be desirable to maintain the substrate at a temperature of about 1500°C to aid sintering of the deposited material.

The most commonly used electrode material has been porous platinum. This would probably be too expensive for general use and search should be made for alternatives. A limiting factor in the use of other metals is the high operating temperature of the cell resulting in loss of material by evaporation.

Nomenclature

B absolute mobility (drift velocity/unit applied force).

C concentration (carriers/cm^3)

D diffusion coefficient (cm^2/sec)

e electronic charge ($1{\cdot}6 \times 10^{-19}$ coulomb)

E Thermodynamic e.m.f. (volts)

F Faraday constant (96,514 coulomb/voltequiv.)

ΔG° Standard free energy change of reaction (cals/mol)

i (subscript) pertaining to carrier i

K equilibrium constant of reaction

$K_1, K_2, k_1, k_2, k_3, k_4$ constants

k Boltzmann constant ($1{\cdot}38 \times 10^{-16}$ erg/°K)

n number of ions per unit volume (ions/cm^3)

P pressure (atm)

R gas constant (1·986 cal/°K mol)

t transport number

T absolute temperature (°K)

u activation energy for ion mobility (electron volts)

z valency

σ electrical conductivity (ohm cm)$^{-1}$

μ mobility (cm sec^{-1}) (volt cm^{-1})

$0^{2-}_{lattice}$ ion on a normal lattice site

$\square_{o^{2-}}$ oxygen ion vacancy

$\ominus$ excess (conduction) electron

$\oplus$ electron hole

For conversion factors of practical and e.s.u. units, see Kingery[5] p.649.

References

(1) Archer, D. H. *et al.*, Proc. 16th Annual Power Conf. (1962) 34 - 39; Proc. 17th Annual Power Conf. (1963) 99 - 100; Proc. 18th Annual Power Conf. (1964) 36 - 40; *Advances in Chem.* Series No. 47 (1965) 332 - 342, *ibid* 343.

(2) Wagner, C., *J. Electrochem. Soc.*, **99** (1952) 346c-354c.

(3) Tubandt, C. *et. al.*, *Z. physik. Chem.*, **87** (1914) 543; *Z. anorg. u. allgem. Chem.*, **110** (1920) 196; *ibid.* **115** (1921) 105; *Z. Elektrochem.*, **29** (1923) 313; *Z. anorg. u. allgem. Chem.*, **197** (1931) 225.

(4) Wagner, C. *et al.*, *Z physik, Chem.*, **B22** (1933) 212; *ibid.* **B37** (1937) 155; *ibid.* **B21** (1933) 42; *ibid.* **B23** (1934) 469.

(5) Kingery, W. D., "Introduction to Ceramics" John Wiley & Sons Inc., New York, London 1960.

(6) Ryshkewitch, E., "Oxide Ceramics", Academic Press, New York and London 1960.

(7) Kiukkola, K. and Wagner, C., *J. Electrochem. Soc.*, **104** (1957) 379 - 87.

(8) Kingery, W. D. *et al.*, *J. Am. Ceram. Soc.*, **42** (1959) 393 - 398.

(9) Palguev, S. F. and Neuimin, A. D., *Trans. Inst. of Electrochem.* (1) "Electrochemistry of Molten and Solid Electrolytes" Authorized Translation from Russian, Consultants Bureau, New York 1961, pp.90-96.

(10) Bray, D. T. reported by Dixon, J. M. *et al.*, *J. Electrochem. Soc.*, **110** (1963) 276 - 280.

(11) Strickler, D. W. and Carlson, W. G., *J. Am. Ceram. Soc.*, **47** (1964) 122 - 127.

(12) Alcock, C. B. and Belford, T. N., *Trans. Far. Soc.*, **60** (1964) 822 - 835.

(13) Duwez, P. *et al.*, *J. Electrochem. Soc.*, **98** (1951) 356 - 362.

(14) Tien, T. Y. and Subbarao, E. C., *J. Chem. Phys.*, **39** (1963) 1041 - 7.

(15) Volchenkova, Z. S. and Palguev, S. F., *Trans. of Inst. Electrochem.*, (1) "Electrochemistry of Molten and Solid Electrolytes" Consultants Bureau, New York 1961, pp. 97 - 103.

(16) Dixon, J. M. *et al. J. Electrochem. Soc.*, **110** (1963) 276 - 80.

(17) Strickler, D. W. and Carlson, W. G., *J. Am. Ceram. Soc.*, **48** (1965) 286 - 288.

(18) Cocco, A. and Barbariol, I., *Ric. Sci. Rend. Sez.*, **A2(3)** (1962) 296; *Chem. Abs.*, **58,** 7462a.

(19) Hund, F., *Z. physik. Chem.*, **199** (1952) 142 - 51.

(20) Tien, T. Y., *J. App. Phys.*, **35** (1964) 122 - 24; *J. Am. Ceram. Soc.*, **47** (1964) 430 - 3.

(20a) Johansen, H. A. and Cleary, J. G., *J. Electrochem. Soc.*, **111** (1) (1964), 100 - 103.

(21) Rhodes, W. H. and Carter, R. E., 65th Annual Meeting Am. Ceram. Soc., New York, April 30th, 1962, *Am. Ceram. Soc. Bull.*, **41** (1962) 283 (abstract).

(22) Weissbart, J. and Ruka, R., *J. Electrochem. Soc.*, **109** (1962) 723 - 726.

(23) Bray, D. T. and Merten, U., *J. Electrochem Soc.*, **111** (1964) 447 - 452.

(24) Uei, I., *et al.*, *Mem. Fac. Ind. Arts. Kyoto Tech. Univ. Sci. Technol.*, **8** (1959) 63 - 78; *Chem. Abs.*, **55,** 927c.

(25) Lukin, E. S. and Poluboyarinov, *Ogneupory*, **28,** 7 (1963); 318 - 23; *Chem. Abs.*, **59,** 11081 f.

(26) Fridman, Y. B. *et al.*, *At. Energy (U.S.S.R.)*, **10** (1961); 606 - 19 *Chem. Abs.*, **56,** 1114 i.

(27) Nielsen, T. H. and Leipold, M. H., *J. Am. Ceram. Soc.*, **47,** 3 (1964) 155.

(28) Tien, T. Y. and Subbarao, E. C., *J. Am. Ceram. Soc.*, **46** (1963) 498 - 92.

(29) Collins, A. M., Conference on New Materials and Processes, Eastbourne, England, May, 1965. Scientific Instrument Research Assoc., South Hill, Chislehurst, Kent.
(30) Shultz, E. B. *et al.*, *Am. Chem. Soc. Div. Petrol Chem. Preprints*, **6**, No. 4B (1961) 9 - 19.
(31) Blocher, J. M. and Oxley, J. H., *Am. Ceram. Soc. Bull.*, **41**, (1962) 81.
(32) Batelle Memorial Institute Reports Nos. 1468, 1471.
(33) Huffadine, J. B. and Thomas, A. G., *Powder Metallurgy*, **7** (1964) 290 - 299.
(34) Bliton, J. L. *et al.*, *Am. Ceram. Soc. Bull.*, **42** (1963) 6 - 9; *Chem. Abs.*, **58**, 8751c.
(35) Bliton, J. L. *et al.*, *Am. Ceram. Soc. Bull.*, **40** (1961) 683 - 8.
(36) Jones, C. E. C. and Griffiths, H., *Brit. Welding J.*, **10** (1963) 546 - 551; Orbach, H. K., *Ceram. Ind.*, **79**, 5 (1962) 72 - 5; Davis, L. W., *Metal Progr.*, **83**, 3 (1963) 105 - 8.
(37) Baur, E. and Preis, H., *Z. Elektrochem.*, **43** (1937) 727.
(38) Nernst., W., *Z. Elektrochem.*, **6** (1900) 41.
(39) Weissbart, J. and Ruka, R., *Rev. Sci. Instr.*, **32** (1961) 593.
(40) Schmalzreid, H., *Z. Elektrochem.*, **66** (1962) 572 - 6.
(41) Binder, H. *et al.*, *Electrochem. Acta*, **8** (1963) 781 - 93.
(42) Tragert, W. E., to General Electric Co. (U.S.A.) U.S. 3,138,487; 3,138,488; 3,138,490 (1964); *Chem. Abs.*, **61**, 5206f, 7951f, 9174g.
(43) Mobius, H. H., German Patent (East) 22030 (1961); *Chem. Abs.*, **58** P243a.
(44) General Dynamics Corps., Belg. Pat. 634, 204 (1963); *Chem. Abs.*, **60**, P15443c.

RETIREMENT

J. R. HARDWICK

Mr. John R. Hardwick retired in August 1968 after 31 years service with the Admiralty. John Hardwick was one of several engineers who were recruited to the then Mine Design Department in H.M.S. *Vernon* to assist with the re-armament programme then under way. He came from Standard Telephones & Cables, for whom he had spent many years abroad, notably in Burma. He was allocated to the Controlled Mining Section under Mr. N. E. Noble, later transferring to the operating mechanism of aircraft laid mines.

When U.C.W.E. was absorbed into U.W.E. Mr. Hardwick took over control of workshops and drawing offices at Southwell. Subsequently, he relinquished this to take charge of R and D finance, a field in which he became very knowledgeable in respect of the procedures best designed to achieve the desired end.

John was a quiet, almost reserved, but well respected member of A.U.W.E. On his retirement he was presented with a barometer with the good wishes of his colleagues for his continued enjoyment, with his wife, of health and leisure.

A RE-APPRAISAL OF BOSE'S THEORY OF NONLINEAR SYSTEMS

By R. W. K. Moilliet, B.Sc., R.N.S.S.
Admiralty Surface Weapons Establishment

SUMMARY

Little practical use seems to have been made of the ideas of Amar G. Bose since he first published them in 1956. The main reason for this is probably that in most cases of practical interest, they are likely to lead to immensely complicated solutions to design problems for which far simpler designs will suffice.

In this article, after a re-statement in condensed form of the more important of Bose's ideas (in which most of the mathematics are sidestepped), some means of implementing these ideas are suggested which take full advantage of modern data handling equipment. In the light of the availability of such equipment, the methods proposed can no longer be said to be impracticable, and might in certain cases be found to be worth while.

Introduction

Twelve years ago, Amar G. Bose investigated, amongst other things, the problem of inventing a procedure for designing the optimum filter for any particular application. This work was contained in his paper "A Theory of Nonlinear Systems "[1]. Not that the resulting design of filter would necessarily be nonlinear; but the methods he derived were in no way *limited* to linear applications. (The linear device is, after all, but a special case of the nonlinear one.) This ability to step outside the linear regime is important. It can be shown that the best possible *linear* filter for a particular application is defined only by the autocorrelation function of the input, and the cross-correlation of the input with the required output[2]. Auto-correlation and cross-correlation functions are 'first order' statistics, relatively simple to derive, and (correspondingly) not very informative. All sorts of very different signals can have the same auto-correlation function. Thus the best possible linear filter may not be very specifically matched to the job it has to perform. One would expect that in all applications where the input and desired output can at all closely be defined statistically, a nonlinear filter could be found which would do the job better.

Before describing Bose's ideas, it is best to make a few general points about filters. The idea of a filter is that it should pass on its input in modified form. This means that its output must *depend* on its input. One can at once invent two classes of filter—those whose output depends only upon the *simultaneous* value of the input (*e.g.* a diode "function generator" network used at low frequencies), and those whose output depends upon not only simultaneous but also *past* values of the input (*e.g.* a simple lag network). The latter class can be said to have *memory,* which the former lack. Certain types of circuit can even be said to have infinite memory: thus a bistable will " remember " being switched into one state or the other indefinitely, until a further input modifies its 'response'. In a great many practical cases, however, the system's memory is of quite limited extent. Thus in the case of the lag circuit, features of the input signal occurring more than 4T seconds ago, where T is the time constant, will have very little effect on the output.

Bose's method is relatively unrestricted. It is theoretically applicable to any filter, linear or nonlinear, with or without memory, provided only that the memory is not infinite (practical considerations about how long a memory it is feasible to treat, will be dealt with later). It rests upon the building-up of a mathematical model of what a general nonlinear system is and does. We shall now proceed to build up such a model.

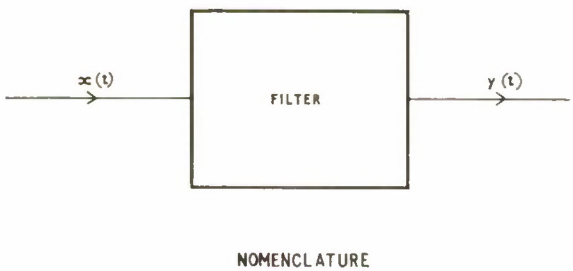


FIG. 1.

The Memory Faculty

We consider first the case of a filter with memory. Let it receive input x(t) and yield a response y(t) (Fig. 1). Since the output is to be constructed from the current value of the input and its past, we must find some way of expressing the recent past of the input. If we drew a graph of the last T seconds (say) of input signal against time, and fitted it with a Fourier series, then the set of Fourier coefficients could be said to *characterize* the recent past of the input, and the required instantaneous output must be some function of this set of numbers. Any other series representation of the recent past of x(t) would yield a set of coefficients from which y(t) could equally-well be calculated. Bose suggests the use of the set of Laguerre functions in τ for fitting $x(t-\tau)$ for all $\tau>0$, and takes the set of coefficients which yield a least-mean-square error fit as being a suitable set of numbers to characterize the past of x(t). However, the method of characterizing the past of x is in no way central to his theory, and we will here take a more direct approach and suppose that we divide up the recent past of x(t) into short intervals, and characterize it by the series of numbers representing its average value in each interval. Thus as the present instant (t) advances along the t-axis, a constantly-changing array of numbers appears (Fig. 2) and these collectively represent the material which will define y(t). Let them be labelled u_1, u_2, . . . u_s, u_1 being taken (quite arbitrarily) to be the most recent " sample ", u_s the most remote.

We will not discuss, for the time being, how the functions u_1, u_2, . . . u_s, may be produced (though the reader will doubtless already have some ideas about this in his mind): we merely assume that a circuit can be built into which we feed x(t), at one input terminal, and from which outputs $u_1(t)$ $u_2(t)$ . . . $u_s(t)$ come from s output terminals, continuously generated.

Function Space

To shed light upon the next problem—how, mathematically, to model the construction of y(t) from the set of u's at any instant—Bose introduces the concept of " function space ". Imagine an s-

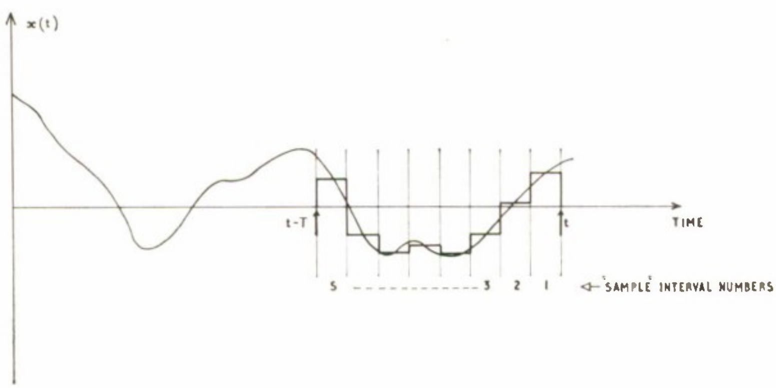


FIG. 2. One way of representing the past of x(t).

dimensional vector-space. Any vector in this space can be regarded as being made up of the sum of its constituent projections along each of the s axes. If the axes were three in number, and labelled x, y and z, then one might say that the vector was made up of so much x-ness, so much y-ness and so much z-ness compounded together. Since the present and past of x(t) at any instant is considered as being made up of so much in interval 1, so much in interval 2, . . . so much in interval s, it is clear that the present and past of x(t) can be represented as a vector in a " function space " whose axes represent the interval-one-ness, interval-two-ness, . . . interval-s-ness of x. Thus the components of the vector representing x(t) and its past along each of the s axes are $u_1, u_2, \ldots u_s$. These, of course, are all functions of time, so one must envisage the resultant vector as darting about all over function space as time advances, its position at any instant representing the input and its past, at that instant. If we know just where in function space* this input vector is at any moment, then we know everything that goes to define the filter's output at that moment.

Gate Functions

At this point it is convenient to introduce gate functions. The gate function of any argument a is $\phi(a)$ and is a special function devised by Bose. It will here be defined slightly differently and rather less strictly, because Bose's definition is unnecessarily restricted for our purposes. Broadly, if any variable a varies over some fixed range, from a=p to a=q, let this range be divided up into n intervals labelled 1 to n, starting at p and finishing at q. The intervals need not be equal in width; but they must be non-overlapping and contiguous—*i.e.* there are no gaps. Then $\phi_j(a)$ is defined by:—

$$\phi_j(a) = \begin{cases} 1 \text{ if a is in } j^{th} \text{ interval.} \\ \text{o if a has any other value.} \end{cases}$$

A graph of $\phi_j(a)$ against a is shown in Fig. 3.

As a matter of fact, we have already been using gate functions without saying so, in our representation of the past of the input.

For consider the expansion.

$$x(t-\tau) \doteq u_1\phi_1(\tau)+u_2\phi_2(\tau) + \ldots +u_s\phi_s(\tau)$$
$$= \sum_{r=1}^{s} u_r\phi_r(\tau) .$$

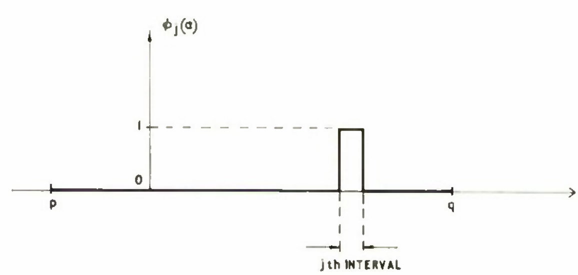
* The reason for christening the space *function* space derives from Bose's original suggestion that $x(t-\tau)$ should be characterized in terms a the set of Laguerre functions in (τ). The basis vectors of the space are then the Laguerre *functions* themselves, and the u's are the coefficients attached to them.

—where the set $\{\phi(\tau)\}$ is defined over the interval τ=o to τ=T. Each term in the series is zero except when τ is within its own associated interval: and when τ does lie within this interval, the term has the value of the associated coefficient u. Thus we actually have been using a gate-function expansion of x to characterize x and its recent past; the basis vectors of our function space are the set of S gate functions, and the u's are gate function coefficients.

The Model

This however, is not what Bose designed and used gate functions for. Bose was concerned with " quantising " the u's, so that function space is divided into cells. If each u is considered as the argument of a set of gate functions, then if the set is defined over a range of the argument sufficient to cover all possible values of u, the effect is to partition each axis of function space into intervals. Since each u is a single-valued time function, only one such interval can be occupied at any instant, which means that only one gate function of each u is non-zero. In Fig. 4, the generation of such gate functions is depicted, and it is assumed that there are the same number (n) of gate functions of each u, and that these appear as electrical signals emanating from a total of ns terminals.

The partitions between the intervals along each axis are in fact planes perpendicular to the axis, and they jointly partition the whole of function



THE jth GATE FUNCTION OF ARGUMENT a

FIG. 3.

space into cells. The compilation of a list of non-zero gate functions of all the u's at any instant amounts to an identification of the single cell in function space in which the tip of the vector representing the input and its past resides at that instant. This cell we shall refer to as the " occupied cell ".

In Fig. 4 is indicated an array of multipliers. Each of these has s inputs, one of which comes from each ϕ generator block. Thus the first multiplier's inputs might be the ϕ_1 function of each of the u's, the second multiplier's inputs might be ϕ_1 of all

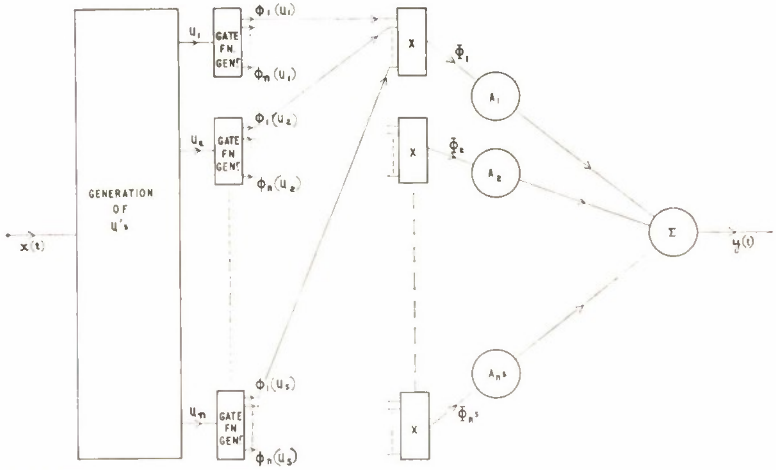


FIG. 4. Bose's model of a general nonlinear system.

but u_s, together with ϕ_2 of u_s, and so on. Enough multipliers are provided to cover all possible such selections of ϕ_j (u_k). It is easy to see how many this comes to: the first of the s input terminals on any one multiplier can be connected to any of the n terminals on u_1's gate function generator (n possibilities); the second input terminal can be connected to any of the n terminals on u_2's gate function generator (thus a total now of n^2 possibilities); and so on, down to the last of the multiplier's input terminals, whose n possible connections raises the final total to n^s possibilities. We thus need n^s multipliers, to cover them all.

This of course is exactly the number of cells into which function space has been divided. A moment's reflection will show that since the output from a multiplier is non-zero only if all its inputs are unity (when its output is, of course 1), there can at each instant be only one multiplier whose output is non-zero. Each multiplier in fact represents a cell in function space, and that which has a non-zero output is the one and only occupied cell.

To complete Bose's mathematical model of a general nonlinear system, it only remains to pass the output Φ of each multiplier to a "gain device", which multiplies it by a constant coefficient, and then to sum the resulting signals to produce an output. The output is completely defined by which cell in function space is occupied; its size at any instant is governed by the magnitude of the corresponding coefficient, all other inputs to the summer being zero. The coefficients are simply chosen to give the required output—or rather, the nearest possible approximation to it, given the limitations imposed by the selected number of dimensions (s) in function space, and the fineness of the cell structure (n) into which it is divided. We shall refer to the set of coefficients as $\{A_\alpha\}$, $(\alpha=1, 2, \ldots n^s)$, and shall call them the Bose coefficients for the system being modelled.

Discussion

The material so far presented represents merely a mathematical model of a nonlinear or linear system with memory. It gives a rather fascinating insight into what a nonlinear system can be regarded as being and doing, but it may already have been dismissed by the reader as being too complicated to be realizable. I hope to show later that this need not necessarily be the case. Meanwhile, however, it is profitable to discuss the model further.

Firstly, its application to systems other than the single input, single output, memory-type system is fairly obvious. For a single input, single output system *without* memory, the 'memory' section of the model—that which generates the u's from x(t)—is dispensed with. The function space in this case is one-dimensional, so all the u's except one vanish, and that one (say u_1) becomes x(t) itself. There are no multipliers, therefore. The model simply looks to see what the current value of x(t) is, and outputs an appropriate value of y(t).

For a multi-input, non-memory system, the set of multiple inputs becomes the set of u's, and the function space has as many dimensions as there are inputs. As before, a single vector in function space represents the combined instantaneous input, and the rest of the model identifies the occupied cell and outputs the appropriate y(t). For a system with some inputs to which it responds 'memoriously' and others to which it responds 'simultaneously', the former inputs each have u-generators, and contribute S_1, S_2 *etc.* dimensions to function space, while the latter inputs contribute one dimension each. For systems with multiple outputs all of which are independent functions of all the inputs, each output is the result of a separate summation of gain-factored multiplier outputs, the sets $\{A_1\alpha\}$, $\{A_2\alpha\}$ *etc.* of Bose coefficients being chosen to implement the various output functional relationships. Thus the basic model is very general in its application.

Secondly, it is clear that since the ϕ's are all either 0 or 1, the process from the generation of the ϕ's onwards is basically a digital one. The multipliers are really gates, and the model becomes a particular variety of digital computer. This makes it interesting to remark that all filters are computers in fact, though the more usual filters are analogue rather than digital ones: they receive an analogue (*e.g.* voltage) input, compute from this (by analogue methods) the appropriate output, and feed this out again as an analogue quantity. To do the same thing digitally, one has to use analogue-to-digital convertors (ADC's) at the inputs, and digital-to-analogue convertors (DAC's) at the outputs of one's digital computer. It is then noteworthy that digital processing generally can be tackled in two ways; by actually computing every function from scratch, as it is needed, or by storing in the computer memory a table of values, and looking them up when required. The processing which occurs in the right-hand half of the Bose model is really just a table-entry type data processing procedure which, given a particular configuration of non-zero ϕ_j (u_k), looks up the appropriate y(t) and outputs it. The gate function generators are a form of ADC, and the $\{A\alpha\}$ embody a form of DAC. These remarks will be carried to a logical conclusion later in this article. They lead to a simple but quite radical "re-casting" of the Bose model.

Identification of an Optimum Filter

It was indicated in the introduction that if the improvement to be expected from the use of a nonlinear filter in place of a linear one was to be realizable, higher-than-first-order statistics of the input and desired output signals must be available. The assumption that this requirement is met implies that extensive lengths of typical signals are available for analysis, since this sort of information is hardly the kind of thing one can guess at from general considerations. However, even given such samples, the actual evaluation of the higher order statistical functions is laborious to the point of unpracticality.[3] Hence Bose chose to devise a method of identifying the optimum filter for a particular job which did not use the higher-order statistics as such, but instead made direct use of the signals themselves, which of course contain all their own statistics. By "identifying" is meant finding the set of Bose coefficients: for this set completely defines the filter mathematically (given the configuration of the model used—*i.e.* values of n, s, *etc.*), and can be inserted into a model in order to make it behave like the system identified by the coefficients. So what Bose in effect says is "Give me a sample of your input x(t) and your desired output z(t), and I will define for you the filter you want". (See Fig. 5).

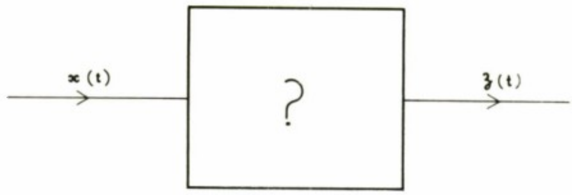


FIG. 5. The problem Bose sets out to solve.

Bose's identification equipment is indicated in Fig. 6. A single input, single output memory-type system is assumed. The inputs to the multiplier are connected successively in the n^s possible configurations and for each configuration, the whole of the x(t) sample and the whole of the z(t) sample are played in concurrently, so that the identification equipment is "told" exactly what output z(t) is required at each juncture in response to the then-obtaining input and its past. At the end of each of these n^s runs, the number of times the α^{th} cell was entered and the sum of all the instantaneous values of z(t) given that the α^{th} cell was occupied, appear as two signals at the output. Obviously, dividing the latter by the former yields the average value of z(t) given that the α^{th} cell was occupied, which is the required best estimate of $A\alpha$. It can be shown that the use in a corresponding Bose model of this value of $A\alpha$ for all α results in a minimized mean-square error between y(t), the actual output of the model in response to x(t), and z(t), the desired output. Fig. 6 is somewhat changed from Bose's original scheme, consequent upon the omission of a weighting term G(t) whereby a *weighted* mean-square error could, if desired, be minimized. This facility is here omitted for brevity, and to make the mathematics self-evident. In passing, we may also note that if enough multipliers *etc.* were provided, one could identify the whole set of $\{A\alpha\}$ in one run.

It is worth remarking that in general, the desired output z(t) may well turn out to be a physically impossible response to the prescribed input x(t). This makes no difference: the set of $\{A\alpha\}$ obtained will be that yielding the least-mean-square approximation to the impossibility asked for. But of course, this may well be a very bad approximation.

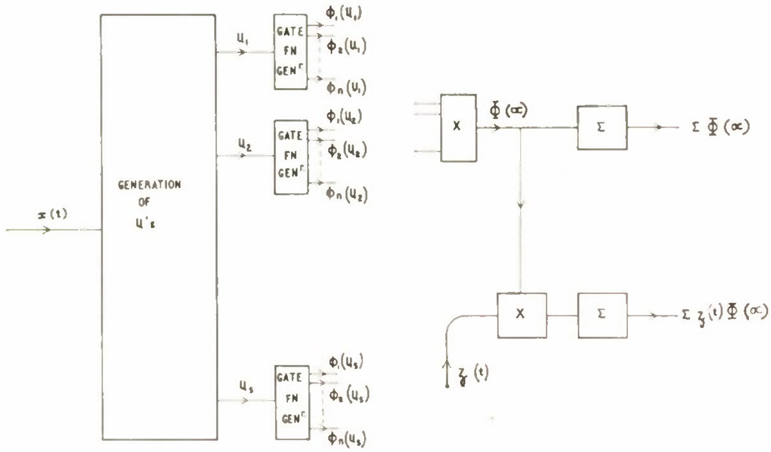


FIG. 6. Bose specification of optimum filter.

Black Box Identification

Clearly the problem of identifying (in Bose terms) the system contained in a given black box, is simply a sub-problem of the optimum filter identification problem. For if x(t) were fed in parallel to both the black box and the identifying equipment, and the output of the black box were then fed into the identifying equipment in place of z(t), then the resulting $\{A\alpha\}$ would be those defining the system in the black box. This is mentioned here for two reasons: firstly, the best way to try out any practical implementation of Bose's theory is to take a known nonlinear system, identify it (*i.e.* find its Bose coefficients), and so construct a Bose model of it. The outputs of the model and the original system, when fed in parallel from any input, can then be compared. Secondly, the identification of plant whose parameters are unavoidably varying with time is a necessary step in one method of optimally controlling such plant. Methods of identifying (and of modelling) " black boxes" are therefore of interest to adaptive control engineers.

Accuracy of Identification

The residual mean-square error between z(t) and the output of a Bose model using the set of coefficients $\{A\alpha\}$ identified as above, has been calculated by Bose, and turns out to be:

$$E \text{ min.} = \overline{z^2(t)} - \sum_{\alpha} A\alpha \quad \overline{\Phi(\alpha) \,.\, z(t)}$$

—where the bars indicate time-averages taken over the whole duration of the identification run. We shall later see how the evaluation of this expression can be used to help choose the right configuration (*i.e.* value of s, n *etc.*) for a Bose model of the system.

Adaptation to Modern Methods

The above concludes this rather hackneyed presentation of the gist of Bose's ideas. We shall now explore the way in which these can best be bent to enable us to take full advantage of modern data-processing equipment.

Firstly, consider the process of generation of the $\Phi(\alpha)$'s, (the functions representing the contents of the cells of function space.) All one is after, here, is identifying the occupied cell, the non-zero$\Phi(\alpha)$. The zero Φ's contribute nothing to the output of the model (Fig. 4); neither do they feature in the answers produced by the identification equipment (Fig. 6).

Since at any instant, all but one of the Φ's are zero, the net result is a vast preponderance of idle equipment.

Instead of generating gate functions, and then combining these to form Φ functions, let each u signal be passed to a crude sort of ADC. By crude is meant that the accuracy required is only, say, 3 binary digits. The voltage representing the u is first sealed so that the expected range of u fills the range of the ADC. Then the 3 binary bits plus their sign will divide the given range of u into 16 sections (remembering that +0 is not the same thing as −0 in this context), and the section containing the current value of u will be represented by the 4-bit output. If all u's are similarly treated, then every time the ADC's receive a " convert " command, a 4S-bit word becomes available, formed by putting the 4 bits from each of the S ADC's side by side in a single register, in the order $u_1\ u_2\ u_3\ \ldots\ u_{s_s}$. This 4S-bit word directly references the occupied cell in function space.

This word, which we may call $\alpha_0(t)$, since it references the occupied cell, and varies with time), can

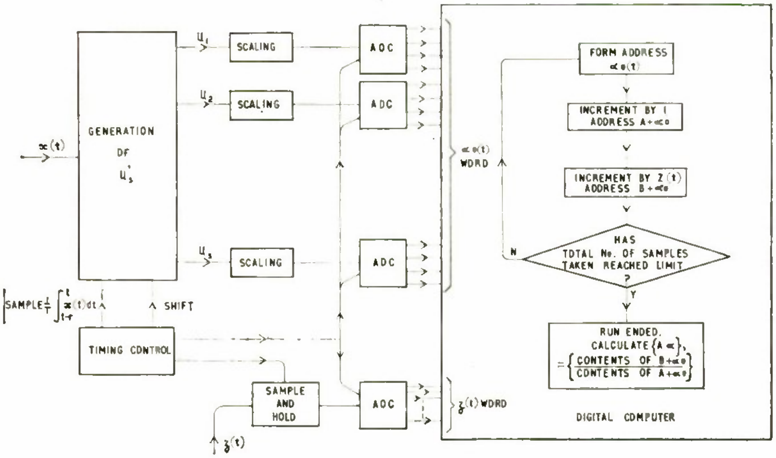


FIG. 7. Digital implementation of Bose identification.

be used to address two store locations in different parts of the computer's memory (See Fig. 7), and cause them to be incremented by one and z(t) respectively. From the totals in all the store locations at the end of a run, the complete set of $\{A\alpha\}$ can be found by simple division.

In Fig. 7 it is implied that the u's are generated in analogue form. by an integrator and a chain of track/store units (the analogue equivalent of a shift register). One could of course use a delay line, though it would be advisable to precede it with a suitable low-pass filter. Alternatively, the whole process could be done digitally, though this would be a relatively slow process if Laguerre coefficients were to be used as the u's. The advantage of a hybrid approach is that it allows a real physical system to be used in black box experiments, and makes it easy to compare results obtained using Laguerre or other orthonormal functions to represent the past of the input with those obtained using samples of the input's past. In Fig. 7, again, *each* ADC is shown as producing a 4-bit output, as indeed is assumed, in the accompanying discussion. There is, of course, no reason why this should be so. Indeed it seems rational to divide into fewer intervals those u's which represent the more remote past of the input, on the assumption that the exact values of these will have less affect on the filter's output than do the values of the 'more recent' u's. It is however convenient, and makes the best use of computer memory capacity, to use numbers of intervals that are powers of two, so that there are no impossible bit-configurations in the output of any ADC, and hence all bit-configurations in the $\alpha_0(t)$ word have meaning as addresses.

For synthesis (*i.e.* modelling) of an already-identified system, the same system of generating $\alpha_0(t)$ can be used, but having generated this, all that is required is to get the computer to look up the corresponding value of $A\alpha$ (stored, conveniently, in a memory location whose reference is $\alpha_0(t)$), and output it, *via* a DAC. One is thus not using the computer as a computer at all, but simply as a quick-reference table. So the computer can be dispensed with, and, say a magnetic disc store with digital addressing facility used instead. The result is still a very expensive and complex filter, but it much cheaper than using a computer as a filter, and might, in certain circumstances, be a sensible proposition.
An alternative method of implementing a Bose filter might be found in the use of optical processing as a table-reference device. If the $\alpha_0(t)$ word were split into two words (*e.g.* by simply chopping it in two), and each word was applied *via* a DAC to one pair of deflection plates in a cathode ray tube, the resulting spot position on the screen would be a unique representation of $\alpha_0(t)$. In fact, one would have mapped function space on to a plane. The resulting "picture", focused on to a photographic plate whose density at all points represented the Bose coefficient there stored, could thence pass to a photomultiplier whose output would define the model's output. If some 330 spot positions in each direction could be distinguished, then 10^5 coefficients could thus be stored. Such a system would be faster in operation than currently available disc stores (which have an access time of 5 ms or more) but it might pose several development problems (not the least being d.c. restoration of the filter's output), whereas the disc store is available off-the-shelf.

How Many Coefficients?

The only established fact about the identification method indicated above is that it yields a least-mean-square-error result. But how small is "least"? Clearly this depends upon (a) whether the number of u's used is sufficient—(*i.e.* is the past of x(t) adequately represented?)—and (b) whether the number n of intervals into which the u's are quantized is adequate.

In other words, does our function space have enough dimensions, and is its cell-structure fine enough? Are we, in fact, catering for enough coefficients?

An expression for calculating the size of the mean-square error to be expected from a given set of coefficients has already been given. Consider first a case of black box identification. This expression can easily be computed by the digital computer at the time of identification, and will indicate, for the chosen values of S and n, whether the identification is good enough.

This is valid, because one knows in this case that the demanded output (z(t) in Fig. 6) is a physically-possible response to the input specified by x(t); so the errors remaining must all be due to deficiencies in the structure allotted to function space. Now an unacceptably large error due to the use of too course a cell structure should be evidenced by the nature of the Bose coefficients obtained. The way in which the occupied cell moves about in function space is subject to certain rules. In particular, if simple consecutive 'samples' are being used to represent the past of the input, and if n is the number of gate functions used to quantise the most recent sample, then there are only n possible successors to any one occupied cell. A computer programme which compares the co-efficient associated with each function space cell with those associated with its n possible succeeding cells, can thus give warning of a preponderance of large jumps in the model's output, which might betoken too course a cell structure or too few dimensions in the chosen function space. Similarly, consistently tiny jumps, which might indicate a wastefully over-fine structure, should be detectable by the same means.

The best approach initially will probably be to program the computer to prepare a histogram showing the frequency of occurrence of jumps of different sizes. Experience with identifying systems should eventually establish some rules about the most desirable shape for this distribution. The computer could then be made to list the cells contributing to any histogram ordinates considered to be inappropriate. Armed with this, one should be able to devise a modification to the chosen function space which would improve matters.

In applying the same technique to the case of the identification of an *optimum* filter, we must recognize first that the identification error may be large because we are asking for the impossible. In this case, no amount of de-designing of function space can help much. However, the successful development of a method such as the above for checking and improving one's function space design, would make it possible to ensure that the final result was indeed somewhere near the best obtainable.

The question remains, however, as to how one gets a rough idea, before embarking on a Bose filter project, whether such an approach might conceivably be worth while, in a particular case. What *order* of complexity of equipment is one likely to want? The answer is that, until some more work has been done on "identifying" simple filters, and then synthesizing them and comparing the model output with that of the original filter nobody knows. Intuitively, one might guess that, say, a simple lag circuit $(1/(1+ST))$ could hardly be successfully synthesized with less than five samples of the past of the input, and less than 10 intervals into which to quantize the u's. Yet even this apparently crude representation would require the evaluation of 10^5 coefficients.

The sheer length of a system's memory is not by itself a limitation; for if its frequency response is correspondingly limited, the fine structure of the input signal becomes irrelevant, and thus a very coarse representation of the past of the input will be sufficient. If 'samples' are used, each sample could be an average over a considerable stretch, and the total number of samples used need still not be very large. It is when a long memory is *combined* with a good response to high frequencies, that one may expect to have to use a large number of u's, and in this case, the number of coefficients required could become astronomical. Thus, for example, lightly-damped systems with a high natural frequency might prove difficult to model. The *complexity* of a system need not, *per se* affect the number of coefficients. Hence experiments on quite simple systems should suffice to provide quite useful experience.

Prognosis

One might conclude that Bose's work has remained unapplied all these years with good reason! It is true that of the two *suggested* applications known to the author, both were restricted to a consideration of plant which could receive only 2-level inputs. The past of the input could then be represented as a series of 0's and 1's in a shift register, and n becomes equal to 2. Thus a stretch of input-past 10 samples long could be accommodated in a function space of only 1024 cells.

These studies, by Roy and De Russo[4, 5] and by Harris and Lapidus[6], are of interest to control engineers, as 2-level input devices (such as relay-controlled motors, or processes controlled by on/off thermostats) are quite widely used. Roy and De Russo additionally make the interesting suggestion that a vastly-smaller store of coefficients would be adequate for modelling a system, if only an appropriate cross-referencing system could be devised, to refer the model to the stored value nearest to the value of the coefficient actually indicated. This would involve doing arithmetic on the addresses, and would necessitate the use of a computer when modelling (synthesizing) the system. They suggest that the computer might be made self-teaching in respect of this arithmetic processing.

Mention should also be made of the *predictive* possibilities of the Bose model. Consider a black box identification. If, in Fig. 6, $z(t+T_1)$ is fed in in place of $z(t)$, then assuming that $x(t)$ is a valid sample statistically of the expected input, the array of Bose coefficients obtained will define a system which would predict, with least mean-square error, the probable output of the black box in T_1 seconds time.

One has thus made a faster-than-real life model of the system. Such a model has application in adaptive control,[7] and in the case of adaptive control of a large plant, such a facility might be worth paying for dearly in terms of complex on-line equipment.

Finally, in data-processing equipment we have a field in which development is still rapid. What is just too complex to contemplate today may not be so tomorrow. Some hundreds of thousands of coefficients could quite happily be accommodated in a disc store today; we can expect even larger numbers to be accommodatable, perhaps quite cheaply, in the future*

* Magnetic tape storage, which allows the use of many millions of coefficients, may well prove most valuable in some process control applications; however, it is likely to remain too slow for most other requirements.

Such considerations suggest that it is high time that someone obtained some actual results on the identification and synthesis of simple systems by these (modified Bose) techniques, so that some factual information becomes available on the number of coefficients actually involved, and the quality of the results. At the time of writing, the author is actively pursuing this investigation: however, a good deal of work needs yet to be done before the area of profitable application of these ideas can be at all clearly defined.

Acknowledgements

The author is indebted to Dr. N. G. Meadows of the University of Sussex for the suggestion that Bose's theory was worthy of further investigation and evaluation with particular reference to the use of hybrid computation, and for arranging the necessary facilities for carrying out this work. Also to various colleagues and members of staff in the School of Applied Sciences at the University of Sussex for many helpful discussions; and particularly to Mr. G. L. Page for valuable suggestions about the application of some standard items of data processing equipment.

References

(1) M.I.T. Research Lab. of Electronics Tech. Report, 309, May 1956.

(2) The well-known Wiener-Hopf Relationship. See, *e.g.*, Gibson, J. E., Nonlinear Automatic Control, McGraw-Hill 1962.

(3) West, J. C. "Analytical Techniques for Nonlinear Control Systems." E.U.P. 1960. Section 12.2.3.

(4) "A Digital Orthogonal Model for Nonlinear Processes with Two-Level Inputs." IRE Transacttons on Automatic Control, Vol. 7. 1962.

(5) "An Adaptive-Predictive Model for Nonlinear Processes with Two-Level Inputs." IEEE Trans. on Applications and Industry, May 1964.

(6) "The Identification of Nonlinear Systems." Industrial and Engineering Chemistry, Vol. 59, No. 6, June 1967.

(7) Kelley, C. R. "Closing the Loop with Predictive Controllers." Control Engineering, Vol. 15, No. 5. May 1968.

THOUGHTS – WITH EXAMPLES

L. Banks, B.Sc.(Tech.), C.Eng., M.I.E.E., A.M.B.I.M.
R.N. Armament Depot, Singapore

Introduction

Mr. R. D. Woods, in his article in the January 1967 issue of this Journal, suggests that the prerogative of the creative designer is to be able to stand on his head and see if things look better that way up. A book more recently published suggests that idea production is a process of lateral thought. Notwithstanding what creativity may be or how it should be described, a few steps backwards and a fresh look at one's own job from a new angle, either upside down or laterally, can be recommended as a most interesting experience.

This article is a further plea for creative thought, and not just by designers. My interests lie in the production and testing of missiles, and, as in most other defence engineering projects, there is scope for outrageous thought. In fact, in any development, be it cracking nuts or busting dams, one must first of all have impossible ideas and then somehow try to make these more acceptable until a technique can be found which makes the impossible possible. This is the way that technology progresses at the rate demanded of it, there is nothing natural about this, but if one is searching for laws this may be a good starting point.

As the boundaries of knowledge spread the scope for further improvement broadens, but at an even faster rate. We can almost regard ourselves as sitting on the island of knowledge casting dredges into the unknown sea to consolidate and spread our perimeter. The island gets bigger giving us more scope for increasing the number of dredges we use, but at the same time the surrounding sea becomes deeper demanding an ever increasing effort.

This analogy ignores the artificial limitations set up on all types of knowledge acquisition. In the commercial field, if a business can get by with producing what it produced 20 years ago there is no incentive to produce something new. Failing to recognise and exploit the potential growth rate of an organisation by innovation, (not "by Parkinson"), leads to stagnation, and our national situation at the present time has all the symptoms of this sad condition. With weapon systems one would expect the possibility of an unlimited increase in sophistication; in fact very severe artificial limitations are applied. These depend upon many things but perhaps the most significant factors prove to be:

(a) the expected life span of the in-service weapon systems in a peacetime environment,

(b) changes in the possible threat by a potential enemy,

but of prime importance is

(c) the amount of money that the Government can spare for research into new systems, coupled with the efficiency with which the research and production are carried out.

Thought costs nothing so we can be as extravagant as we wish with ideas. Many ideas will prove to be outlandish and no one is more aware of this or more careful to avoid ridicule than the Professional Engineer. Nevertheless progressive ideas border on the ridiculous so that the discussion of ridiculous ideas, without bias, sets in motion thought processes which can give some positive future advance. This is the basis of various brainstorming techniques and to get results demands an atmosphere free from any form of prejudgement, criticism, ridicule or evaluation. Could some long-distance brainstorming technique be developed by readers of this Journal ? A panel from the research organisations could provide the censoring element required and perhaps suggest subjects for discussion.

The Ministry of Defence has vast Professional resources scattered over large areas. Many are out of touch with their fellows and the need for security presents a gigantic barrier to communication. The combined think power of this pool of brains is phenomenal. Unfortunately each individual brain can only get a certain distance into pure speculation by itself before entering the realms of fantasy. It is conceivable, however, that a large enough group, properly led, could think their way through generations of weapon systems without spending a brass farthing on hardware. It is probable that inverted and lateral thought on the subject of more fully utilising the brain resources available to us would pay off better dividends than any single technical breakthrough. The development of creative trends must be a conscious effort by each individual. What is known as creative thought is probably just the useful knack of being able to de-think established ideas and re-think them again, but using the advantages of a much later jumping off point.

Organisations of any size could do much more than they do to encourage creativity. Most ideas conceived by individuals probably have very little to do with furthering the aims of their organisation. Nevertheless the useful encouragement of conscious thinking habits may result from setting up a small section whose objectives would be to evaluate and provide a clearing house for any ideas, on any subject, from any member of the organisation.

Many good ideas on politics, economics and public administration are probably thrown away because there is simply no outlet for them, yet someone somewhere may be looking for just such ideas to become the next Mr. Churchill.

The same applies to ideas coming under the category of invention, the Patent process is really so sloppy and expensive as to be a downright barrier to creativity. There are very few inventions which can be positively protected and a central clearing house within the organisation may prove to be an alternative means of encouraging the exchange of ideas. The few truly patentable ones are easily recognised by anyone having patent experience. The remainder could be circulated to groups of qualified members of the organisation for comment and expansion, very much on the same basis that portfolios of prints are circulated within photographic societies. In this way a continuous flow and feedback of information circulates, upon which the thinking processes of a large number of individuals could be nurtured.

This may appear to be more of a social service to members than a furtherance of the objectives of the organisation. But, when one considers that creative thought provides the means by which the Nation earns its living and defends itself, and from which the individuals engaged on constructive tasks gain the incentives and satisfaction of their chosen career, it does seem that a lot more thought has to go into this essential management technique.

One of the phenomenons of our age is the amount of creative thought conceived by German Scientists at Peenemunde. This pool of information, generated very rapidly under conditions of limited resources, has only recently dried up. Before the end of World War II multi stage I.C.B.M's had been designed even though there was then no known warhead to make the project an economic proposition. The "Polaris" type operations had been foreseen and a practical design for a space vehicle launcher of higher thrust than "Atlas" had been designed. Inertial navigation principles had been developed for the V2 and beam riding anti aircraft missiles were at the point of production and were rapidly adopted by the U.S.A. as the Nike-Zeus. The ratio of original thought to expenditure on hardware fell drastically after the war simply because there was a more vital need for practical development.

Britain is now entering a new age in which it seems vital to rethink all our established precepts and institutions. The basic problem is—how can a highly populated country with less natural resources and available reserves than ever before continue to increase its standards of life ? Fortunately that is no more my own problem than it is of all my neighbours.

Quite naturally the defence field has its own responsibilities under this changing situation and it may be that if properly exploited by a lot of thought and very little hardware, the present conditions of financial stringency could very well prove to be the nursery bed of the next generation's defence system.

Of most direct interest in my own field are, the overall weapon system, the terminal weapon—missile or torpedo, production buildings and logistic facilities, and systems test equipment.

By way of example some lateral and inverted thought on these subjects is given below. It may seem that I have trampled over the hallowed

ground of the specialist, but it is very likely that the same ideas have been well thought over and already rejected by those most directly concerned. It cannot be too strongly stressed however, that the one most closely concerned with a job does not always have the best perspective, and these examples may help to drum home the vital need and the opportunity for a different angle of view on all the practical things which we now accept as familiar and unchangeable.
A broader review of all the other factors affecting our way of life may show even better opportunities.

Weapon Systems

There has always been a requirement for a level flying projectile preferably skimming the earth or ocean surface for long distances at a height of 6 ft. or less. The impossibility of such a proposal has repeatedly tended to push the demand out of the way and it is now probably disregarded, but is it really so impossible in the missile age?
There must be a particular need for effective defence against ship launched missile carriers and perhaps an attack missile with a terminal stage of level flight could do this job best of all.
The requirements for such a system would include long range, supersonic speed, a warhead adequate to disable or sink the target and the line of attack to be in the least defensible zone.
There are now several defensive systems capable of dealing with attack missiles on guided ballistic trajectories so perhaps the low, level-flying missile could best fulfil the latter requirement.
A missile skimming the ocean at high speeds has no scope whatsoever for manoeuvring in pitch by any nomal instrument technique, but could consistent low level flying be achieved by exploiting the ground effect by purely aerodynamic means? I understand "ground effect" to be the extra lift found in the landing of aircraft, probably as a result of the compression of air between the aerofoil section and the ground suface. I may be wrong in this interpretation in which case the idea comes into the ridiculous category. Nevertheless these are the avenues which have to be explored in gaining a better understanding.
The launching of such a missile would be from an aircraft or on a ballistic trajectory from a surface vessel. In order to level off the missile at a suitable position for its final low level attack, it may be best to consider the use of a horizon seeking phase of flight. It is conceivable that a set of three horizon seekers (optical, infra red, radar or other types) looking aft and 60° each side of the line of flight could gauge both altitude and attitude. The attitude of the missile would then be programmed by its altitude and speed so that it would level off at a point from which ground effect takes over.
The horizon seeking phase would commence when the missile, in its initial launched trajectory, crosses a predetermined tangent to the earth's surface, see Fig. 1.

FIG. 1. Possible trajectory of ship to ship guided weapon.

Some active homing device brought into operation at a set distance from the target would make course corrections and, as height and attitude are automatically controlled throughout flight, this correction would be in the azimuth plane only.
It may be that a waterline contact with the target is desirable and this can be arranged by terminating the flight with a dive activated by comparing range with missile height.
An attack, using ballistic decoys to keep the enemy defensive system busy, while sneaking in a few active attack missiles might be the best tactic.

Missiles

A significant advance in long or medium range shiplaunched surface to air weapons would be derived from semi-automatically mating the explosives warhead and the more sensitive propellants to the flight body at the launcher. When this is possible the hazardous components can be stored in the ships' magazine while the flight hardware, representing no hazard to the crew, would be stored and maintained close to the launcher. The overall effect would be to cut down valuable ship space and to make handling and testing arrangements both at the depot and the ship very much easier. Some idea of the reductions in cost and inconvenience at depots can be gained from the next section of the paper dealing with missile assembly.
The link between the boost and the missile is a major structural and flight feature, but the problem is not a particularly complex one as all the mechanical limitations are very well known. If a solid fuel sustainer is necessary then this could probably be in tandem with the boost providing a compound attachable component.

The warhead offers some problems but why not use a " halo " or " collar " type? This would slip over the missile to complete the ballistic casing profile. The locating and locking and the transfer of essential connections can be by a relatively simple device. Perhaps the aircraft brake shoe expanding bag could be of use here.

The " halo " type warhead might well be a useful configuration ! Perhaps a three part expanding rod could be fitted into the larger surface area of the halo so that when it detonates it would produce a clover leaf pattern of steelhoops covering a greater frontal area than the single rod, whose present size is limited by the chance of an aircraft passing through it untouched. See Figs. 2 and 3.

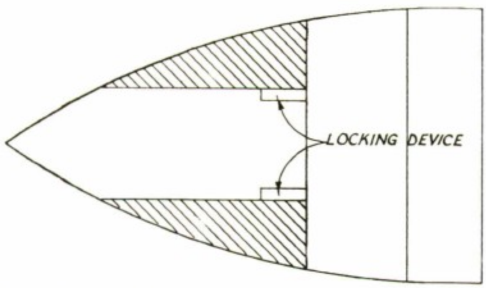


FIG. 2. Collar or Halo type warhead

Anti Submarine Torpedoes

Modern Naval tactics involve more action underneath the ocean than on the surface. Improvements in sonar could very well outstrip the range of operation of anti submarine torpedoes and to get over this difficulty " IKARA " has already been developed. This is a guided missile flight body which carries anti submarine torpedoes from a surface ship to a computed interception point. The torpedo is then released and continues its journey by parachute and once in the water under its own power to the target using active homing guidance. In service already are large quantities of semi-diesel torpedoes of superb quality and likely to be redundant only when surface shipping finally becomes obsolete. If a number of these were modified to carry concentrically either the smaller air dropped torpedo, or a HE warhead of similar proportions, they could then assume a dual purpose. They could be used, with the modified warhead, to attack surface shipping as before, or to give a high speed piggyback ride to the smaller air launched torpedo to a point of interception from which the latter would be launched for the final submarine attack. It could be arranged that, on discharge of its anti submarine torpedo, the spent delivery torpedo would remain floating on the surface, so that it would not interfere with the attack, and then sink after a period of time, or be recovered if conditions were favourable.

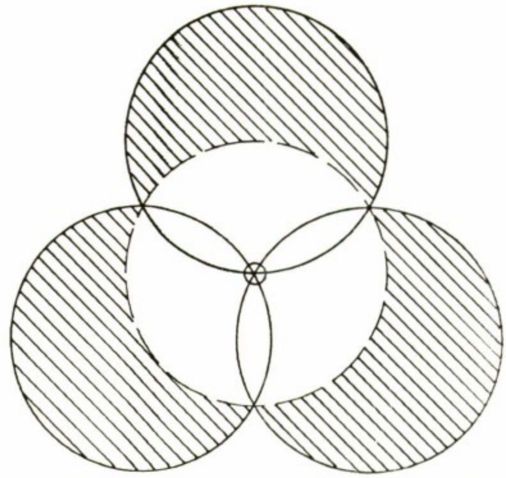


FIG. 3. Extra area covered by 3 loop warhead configuration.

This deployment would not be as efficient as IKARA on the grounds of range, speed and noise, but it would mean an extension of the reach of our well proved weapons systems. It could be launched from submarines or any other vessel capable of launching standard torpedoes. It would simplify the supply situation and would not require more than a gentle tug on the Treasury purse strings.

Missile Assembly Facilities

Until we can successfully extricate the missile electronics and flight systems from the explosives, we must regard such missiles as a hazard and plan our logistic support accordingly. There is no doubt that, as discussed later in this paper, we are not too far away from having a universal test equipment capable of testing any specialised system with the minimum of specialised test black boxes. Ideally this should permit the assembly of several different missiles at one time, arranging their completion so that the test equipment will be kept fully occupied and so justify the large capital outlay. In anticipation of this happy day, we can consider the hypothetical layout of buildings that would meet this need; once again at no extra expense to the taxpayer.

In cases where explosives and propellants are indivisible from the missile during system tests, then the full requirements of explosives storage and handling regulations have to be met. Explosives and electronics certainly do not mix; the solution is the best compromise of the two conflicting issues and this has proved to be quite acceptable. Confidence in the safety of the system is such that the explosive element of the missile is more of a psychological than a real hazard.

Due respect is enforced by standing instructions, which include a comprehensive safety routine; this is also coupled with the knowledge that a one in seven million chance sometimes turns up. Testing has to be done remotely through extension leads, and this all adds considerably to the length of the testing time and increases the chances of inaccuracies.

Explosives handling philosophy is such that the missile under test is housed in an explosives building, while the test equipment and personnel are housed in an adjacent strengthened building capable of preventing damage to its contents if the explosives element of the missile should detonate at the test position. The explosives building is generally a light-weight box structure surrounded by an earth bund or mound so that blast, if an accident should occur, is directed upwards. The building itself offers no tamping effect to the explosion and disintegrates into small enough fragments so that the kinetic energy of these is largely absorbed in flight.

When thinking of buildings nowadays I am much attracted to the cylindrical or igloo shape. These shapes enclose space more economically than squares or rectangles. A round house would probably have many advantages over a square one (above a certain floor area) particularly in reducing circulating space and in having the windows facing in more convenient directions. A round building lends itself to such techniques as automatic bricklaying, automatic plastering, automatic painting and automatic cleaning and its construction requires only one reference point. The brick arch and the brick dome structure have been in use since the Greek and Roman pre-biblical days. It is the strongest and most natural structure and requires no shuttering or reinforcement or the other paraphernalia of the modern building industry. The Pantheon in Rome is a composite brick arch/concrete dome spanning 142 feet and is now more than 1,800 years old. This must have been constructed with the aid of the arch centreing only.

I am not suggesting anything quite so exotic as the Pantheon—a missile assembly building greater than 46 ft. diameter could be an embarrassment particularly as successive missiles will reduce in size for the same performance, but let us consider the hypothetical layout shown in Fig. 4. The central king post in this case provides the member for the swivelling hydraulic king post crane which covers the entire floor area of the building. The king post is also used to centre the lighting fitting which provides both direct and diffused light throughout the building; in the latter case by using the reflective properties of the dome. This same fitting could also be used to contain the television camera and its floodlights. Fixed test equipment connections and other essential services are all kept below the 6 ft. level so that maximum clear dome space is left to enjoy the benefits of automatic maintenance, which is possible by fitting a temporary trammel arm to the king post at such times.

To give a certain amount of daylight, the dome is truncated and provided with a second smaller dome or head which is cut away at an angle and weatherproofed with translucent fibre glass or plastic sheeting material. This also provides the pressure release point in the event of an explosion.

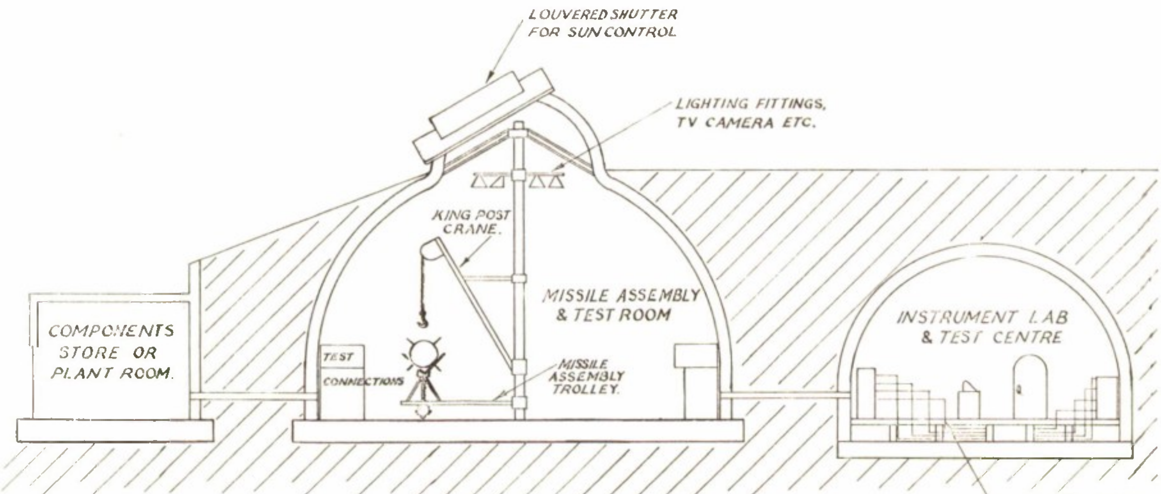


FIG. 4. Cross section of missile assembly/test facilities.

The dome is waterproofed with a two layer bitumastic membrane protected by a brick or edge skin and then the mound is built up around the building itself; as is quite common practice. Waterproof ducts in the mounds convey cables from one room to the other. Air conditioning or plenum heating is supplied tangentially to the upper dome and the exhaust is returned via the entrance corridors which gives reasonable air flow arrangements.

The assembly cradle itself is mounted on a two wheel trolley which is pivoted around the king post and can be securely locked to the floor at any point in its turning circle. The body of the missile is delivered to the building in the usual type of transit trolley where it is picked up by the king post crane and centred on the pitch circle of the assembly cradle. The transit trolley is removed from the building, and the assembly cradle is brought around the king post beneath the body which is then dropped onto it and cradle and load can be turned to the assembly position. It may be convenient to have a common assembly point for all types of missile, say at 12 o'clock, and then to turn the particular assembly through 90° to connect it to the test position pick up points. The test points for two other types of missiles may be possible at 180° and 270° respectively so that, without too much inconvenience, the same building may be used for three different types of missile.

There could be as many as six assembly and test buildings served by the same test room but probably the bottleneck would then be in the length

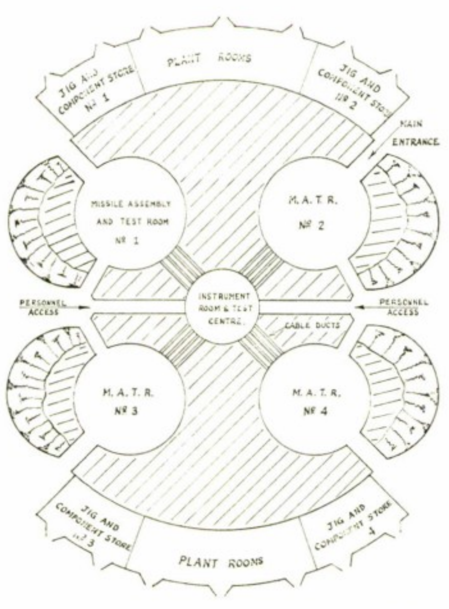


FIG. 5. Layout of missile assembly and test buildings.

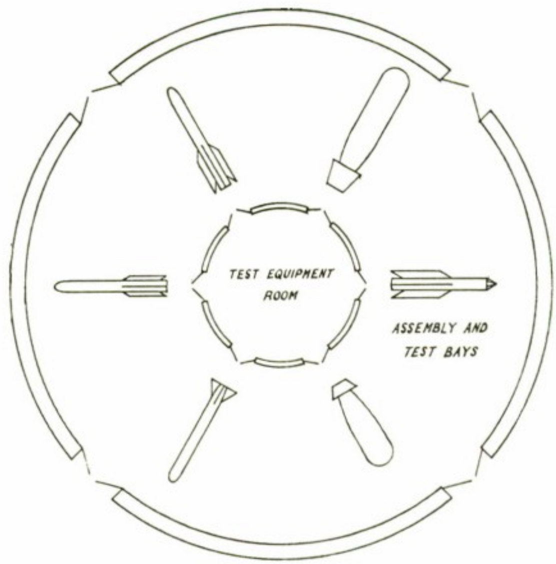


FIG. 6. Round philosophy used with assemblies having no explosive content.

of tests. The plan in Fig. 5 shows an arrangement of four assembly buildings served by a single test equipment room with stores and plant rooms.

For the test equipment room, a complete igloo without the king post seems to offer the requirements of strength and safety. The false floor of the room would be suspended about 18 in. above the structural floor, as in computer installations, allowing free runs of cables between various equipment racks and so avoiding the usual interconnecting test rooms clutter on cable racks. The test desk would be situated centrally in the room while the standard power supplies, patch boards and test equipment together with special to type test equipment would be located in standard racking around the wall of the room.

The arrangement of interconnecting corridors between the test equipment room and the assembly room can be designed to meet the safety requirements demanded, and avoids excessive walking excursions which are a feature of all remote test sequences.

The circular building also has applications even where explosives are not involved in missiles, or where other complicated electro-mechanical system testing is required on units such as military or civil aircraft or vehicles. In this case the universal test equipment would be set up in a central round room which would be surrounded by an annular space formed by a much larger round building in which the particular system would be coupled to the test equipment. See Fig. 6.

Bricklaying in the Round

The sketches Figs. 7 to 9 show how the construction of round buildings could be undertaken. The basic requirements are a guyed king post tube, drilled with peg holes every six courses. A plumb bob is suspended centrally down the tube and two sets of peepholes are bored through the tube so that the plumb line can be centralised by means of an adjustment of the guy bottle screws. Collars are provided on the tube which are pegged at the required height, one to be used to support the trammel arm, and one for the pivot of the working platform.

The trammel arm is a tubular member pivoted around the axis of a collar on the king post tube which is in turn supported by the stationary collar pegged to the king post. The arm is counterbalanced so that it can be raised and lowered very easily by the bricklayer at the working line. The working end of the trammel arm has a vertical socket through which passes a smaller diameter tube which can be pegged to the trammel at any one of six brick courses. This arrangement means that the height of the trammel pivot of the king post need only be adjusted every sixth course while the individual courses in between are taken care of by a simple adjustment at the operator's end. Attached to the pivot of the trammel is a chamber containing mercury and protected from the elements with a transparent cover. On the mercury floats a pair of prisms cemented together so that a horizontal beam of light is reflected back along the same axis. A very simple auto-collimator with a large aperture and a reflecting swivelling eyepiece would allow the bricklayer to place either the inner wall or outer wall course of bricks to the same level throughout. The bricklayer's end of the trammel would comprise the tube of adjustable height, mentioned earlier, at the bottom of which would be a three fingered hand which would grip the brick so that the face was truly tangential to the circle and the brick itself was level. A vibrator or electric hammer operated by a trigger would take the place of the edge of the trowel in bedding the brick into its mortar. Mortar could be pumped as a bedding layer and also to fill in the spaces normally done with trowel work. A second man would lay the other bricks in the most expeditious manner using the course already laid as the reference.

The workers concerned with this operation would be accommodated on a moving platform pivoted around the king post on a two wheel undercarriage and powered by an electric motor so that a range of speeds could be selected depending upon the speed of the individuals. The height of the platform would be adjustable course by course by a hand hydraulic pump and materials could be conveyed by a belt conveyor extending from a position near the king post. If necessary, the platform could be extended over the wall forming a saddle for one of the bricklayers to use and supporting an outside platform which could be used every other course to do the necessary pointing and

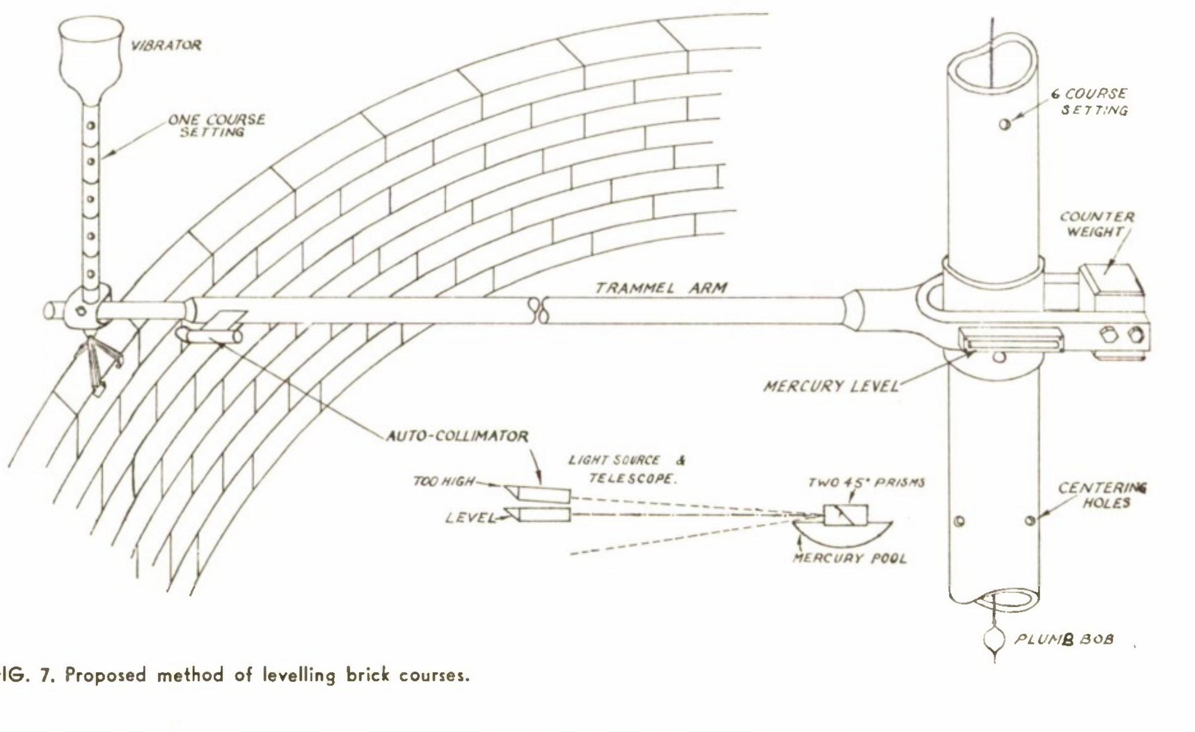


FIG. 7. Proposed method of levelling brick courses.

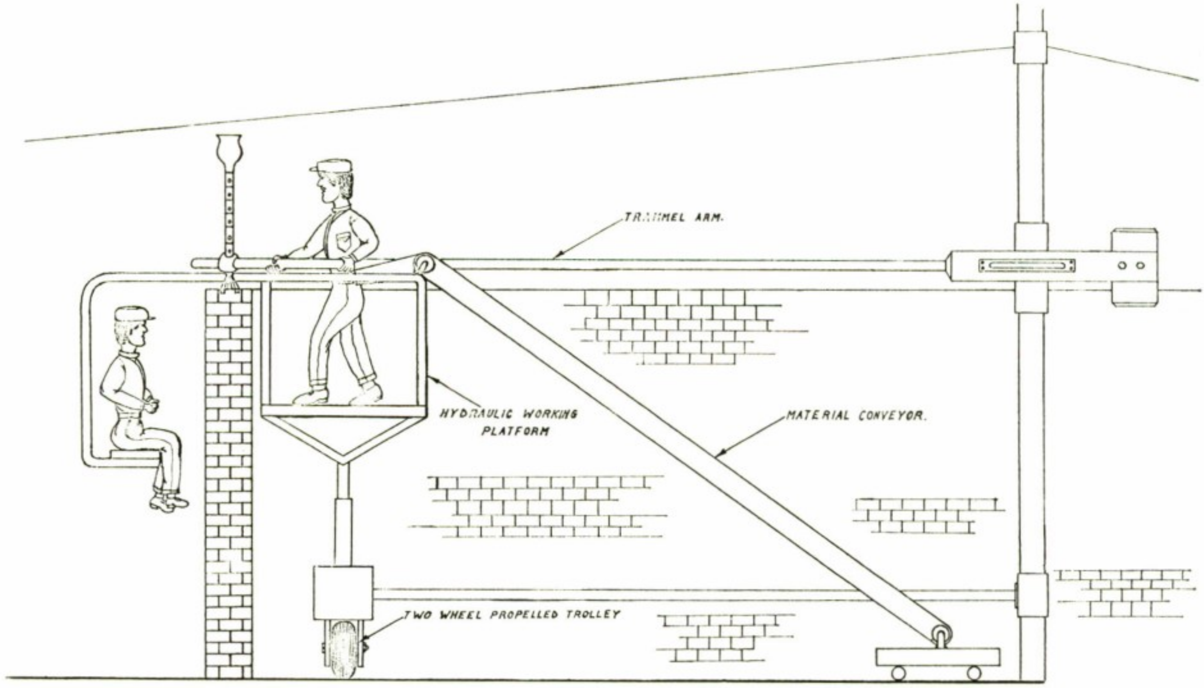


FIG. 8. Round building construction requirements,

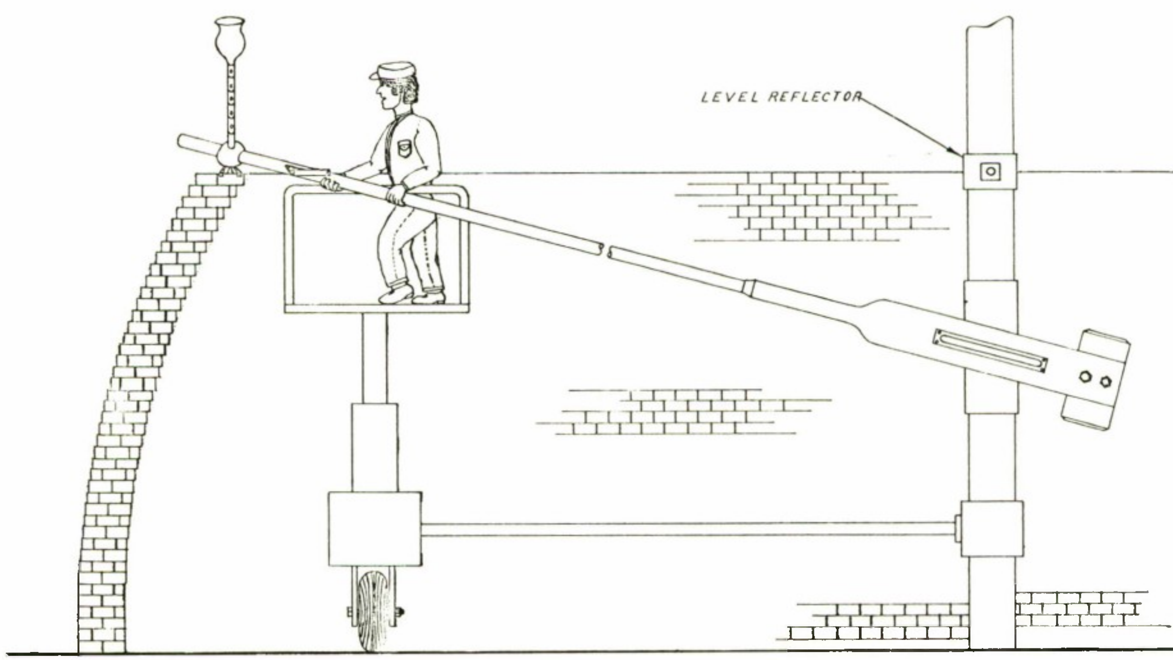


FIG. 9. Construction method for domes.

cleaning. The platform itself could be covered with an awning to keep out the weather and in a severe climate could be enclosed and heated. Alternatively on calm days the king post and guys could be used to support a canvas or polythene roof to cover the whole site. Figure 9 shows the arrangement for building a dome.

Systems Test Equipment

Systems test equipment normally provides all the electrical hydraulic and pneumatic supplies which would be available under the actual systems operation conditions. This avoids the starting of engines and the utilisation of other internal supply sources which, in the case of missiles, are invariably one shot devices of one form or another. Transducers are fitted to pick up various responses and all vital circuits are monitored.

The measurements we are interested in include AC or DC voltages, current, resistance, reactance, power, phase, mechanical movement, time, frequency and event counting. Stimuli includes AM and FM signals, infra red and pulsed inputs. The foregoing can be in any range or combination and it is little wonder that switching is the real art in test equipment design. One missile has a series of no less than 370 different tests which are manually switched and involve up to eight hours hard work. The tests themselves usually involve the frequent reference to calibration charts and tuning operations.

The length of time a missile is on test reduces the life expectancy of the systems, so perhaps this is strong argument for fully automatic sequence switched test gear. Automatic test gear tends to be too specialised and inflexible and lacks the diagnostic capacity which is required. A range of tests can be done on a single proprietary instrument, and instruments are readily available to cover any single function of all systems. It is the number and combination of these requirements alone, that has led to the manufacture of specialised test equipment by the missile contractor. This saves space, but is very expensive, more difficult to keep in tune and has no applications outside its single function so that, on the expiry of missile service life, it has only scrap value.

Automatic "Universal" test gear is already available in many different forms. Punched tape is a popular medium, the programme being written to cover switching sequences, warming up time delays, measurement of the signal, comparison with built in tolerances, go and no go indication and automatic shut down on receipt of the No Go signal; a teleprinter prints out the results. This system is expensive and the writing of the programme is no easy matter, the diagnostic capabilities are also limited.

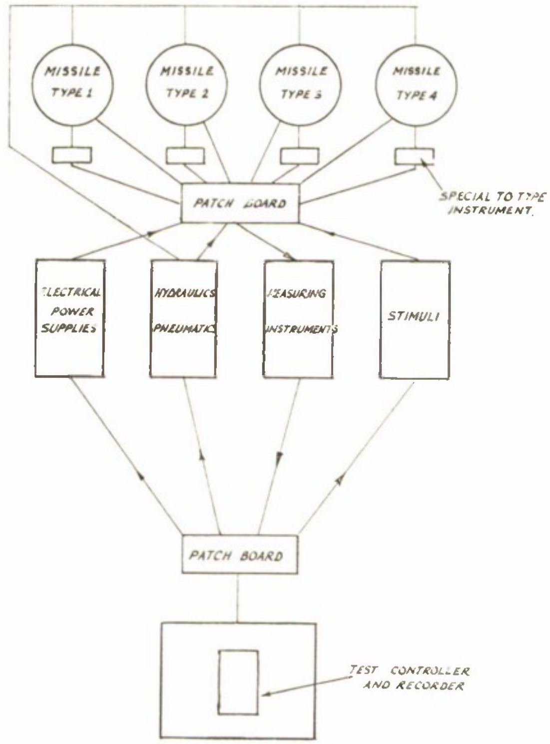


FIG. 10. Block diagram of test system elements and patch boards.

Something between the fully automatic and manual system may be acceptable and should allow tuning and diagnostic operations to proceed at the time of the test. One possibility is shown in Figures 10 and 11.

This is an attempt to combine the programme, the operator's instructions, the calibration data and a record of test results on a single chart. The test equipment set-up would possibly comprise all the elements of a Standards laboratory plus a series of "patch boards", by means of which the appropriate power supplies and test points could be switched to the missile, to the instrument bank and to the central control console. It is desirable that all switching sequences be controllable either manually or automatically from the central testing console but any deviation found to be expedient can easily be marked into the chart as an instruction to the operator. The system thus has considerable flexibility.

A typical element of the chart is shown in Figure 11. The four high box of squares at the bottom of the presentation indicate the switching function controlled (according to the column) and any one

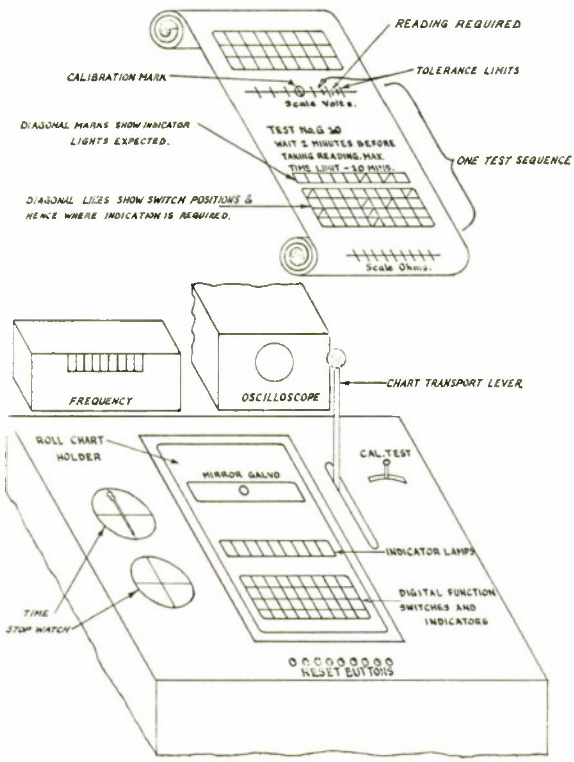


FIG. 11. Possible layout of control console and specimen chart.

of sixteen positions of the function switch in digital form. The chart would register over a box of similar squares on the instrument. In an automatic version these boxes could contain both the automatic switching elements, which would be operated by interaction with the chart markings, and a light indication which would be controlled by the switching element. In a manually switched arrangement only the lights would operate in response to the switched positions. The operator would simply have to check that a light appeared behind each of the squares marked with a diagonal line. To carry this principle further each of the corresponding squares in the instrument could incorporate a push switch so that by pressing each of the squares marked the complete switching function could be achieved, and achievement of the function would again be fed back as a light signal.

The line of squares above the bottom box on the chart would register over a line of indicator lamps. The achievement of a test is sometimes shown simply by the operation of an indicator lamp so that if the particular test was in this form one would expect to get the appropriate indicator response from the square marked with a diagonal line.

Space is provided on the chart above this line of squares, where the test concerned is numbered and any notes or instructions are given to the operator.

At the top of the presentation is a straight line scale. This is marked up in accordance with the requirements of the test and permanently indicates (a) the precise reading required, (b) the tolerances of acceptability either side of the reading, and also (c) a check calibration mark. The indicating element could be a mirror galvanometer or the Y scale of a C.R. tube. Before making a reading, the instrument can be checked for calibration by means of a key switch. After the first complete test, the results of the previous test would be still marked on the scale so that any tendency to drift would be immediately obvious to the operator. Any tuning required can be undertaken at this stage and then the scale marked at the point indicated. This is possible by a wide variety of devices, the simplest of which is a fine "China-graph" pencil.

This arrangement seems to fulfil the requirement of a diagnostic instrument while providing an element of semi-automatic operation to reduce the testing time to a minimum. Some people will regard the reversion to a galvanometer from digital instruments as a backward step because of the less degree of accuracy. It is however, well within the limits required, it is very much cheaper and also responds to transient movements which a digital meter could not detect. In any case, the present generation of Engineers are not digitally orientated and the effect of interposing unnecessary analogue/digital converters in the chain only leads to frustration.

In tests involving event counting or frequency measurement, where the digital instrument is the most natural choice, then the reading of the digital meter would be recorded in the place normally occupied by the scale. In a similar way pulse shapes read off an oscilloscope could be copied or indicated as correct.

The test chart of a particular missile would be a series of these standard sections printed in a roll form according to the number of tests required; probably on a transparent base. At the time of writing I envisage that the transport of the roll from one test sequence to the next, would be carried out in the instrument by means of a lever, similar to a fruit machine. The programme for one missile would be worked out from the basic requirements and the chart made up by hand. After proving, the copy charts would be printed from the master and would accompany each missile together with the log book through its life. Copies of test results could be printed from the missile test sheet if this were ever required.

The feasibility of this type of tester would have to be studied in relation to a large cross section of missiles already in service. It would require a much harder and more concentrated effort than I have had the time to apply, but whoever did this would emerge as a more useful individual without unduly increasing defence expenditure.

A new missile or similar system could be introduced to the test system and would require a new set of patchboards, a new programme, possibly a special to type element such as a microwave unit and perhaps an extension to some of the standard instruments.

Each of the instruments used would be capable of calibration individually against built in standards. Each could be replaced individually if a better class of instrument became available, and when replaced the surplus instrument would still have many further uses as a bench instrument elsewhere.

The cost of the first installation would probably be within the present day cost of one of the larger special to type missile testers. Subsequent missiles would be capable of being included at a fraction of the usual cost.

The presence of depot installations of this type would allow a more logical and practical approach to ships test equipment. This must of necessity be rugged, small and compact and must show on a "go" or "no go" basis whether the system or element of a system will perform its function under operational conditions. There is no need for diagnostic equipment because system elements can be exchanged on a unit basis. Experience of the depot equipment should show the very minimum of tests needed to fill this bill and then the practical test equipment could be simply designed in the most compact form.

Other Test Recording Methods

The recent introduction of the video tape recorder, using a helical scanning arrangement, opens up many new fields in the permanent recording and testing of the more simple electro-mechanical systems.

A homing torpedo reacts to external stimuli such as water pressure, angle of heel and pitch, direction of signals received, etc. The correctness of operation is usually a measure of the change of angle of the control fins. To check all aspects of a test one usually requires a multichannel recorder and associated transducers. The interpretation of results then becomes a rather specialised affair and requires a good deal of time.

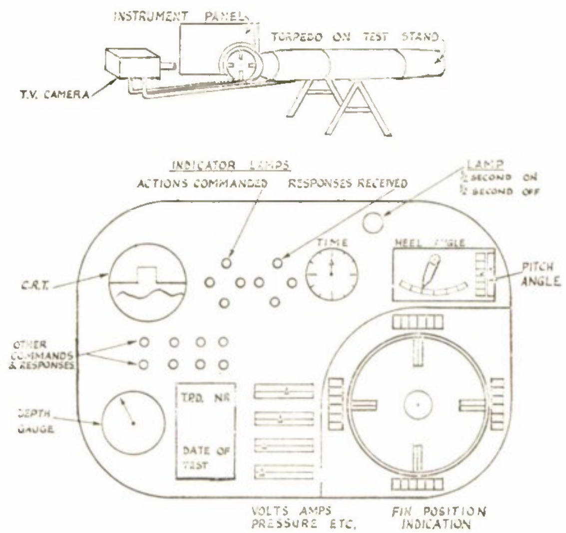


FIG. 12. Possible set up for recording test on homing torpedo.

By mounting a TV camera so that the control fins are in its field of view, together with an instrument panel, every aspect of the test can be recorded. See Figure 12. By quite simple means, it is possible to produce an almost infinite number of information channels on a single TV monitor and then one has available the sound channel in which even more information channels can be crowded. The cost of a TV tape recorder is much cheaper than a 12-channel U.V. recorder. The information obtained is immediately available for interpretation by our normal senses. The tape can be stopped at any frame, played backwards or forwards and can be played on any similar machine, so that it forms the ideal permanent recording medium. False tests can be wiped out and the tape re-used so that we have an economical automatic test recorder with a character just as universal as the ingenuity of the user.

Conclusion

Writing down one's ideas could be a useful step to consciously developing creative habits. This being so, perhaps in future issues of the *Journal* we may have more readers' viewpoints on the controversial subjects which arise from a fresh look at our familiar surroundings.

Notes and News

A.M.L.'s 21st Anniversary

A.M.L. celebrated its 21st Anniversary by holding a week of "Open Days" at the beginning of October, 1968. This event is mentioned in the next column. In the following week, before the key exhibits had been dismantled, the opportunity was taken of arranging tours for science students from several local schools.

Dr. R. Holland, P.S.O., joined the Chemical Engineering Division at A.M.L. on transfer from C.D.L. on 3rd September, 1968.

Recently developed methods of crack detection, the AMLEC Crack Detector, the Magneprint and Magnestamp techniques, have had further publicity through a demonstration by Raymond Baxter on "To-morrow's World", screened on 5th June, 1968. They were also displayed as part of the R.N.S.S. Exhibition at Portsmouth Navy Days. The latter two techniques were the subject of a paper entitled "Some Recent Developments in Magnetic Crack Detection" by D. Birchon, R. H. Warren and P. M. Wingfield, which was presented at the Annual Conference of the Non-Destructive Testing Society of Great Britain, at Warwick University on 13th September, 1968.

Dr. D. K. Ross and Mr. A. A. Law presented a paper entitled "The Development of Fuel Cell Ancillary Systems and Components" at the Power Sources Symposium at Brighton on 24th-26th September, 1968.

The following papers have also been published:—

"Mechanism of Energy Dissipation in Manganese-Copper Alloys" by D. Birchon, D. E. Bromley and D. Healey, in *J. Metal Science*, Vol. 2 (1968) p. 41.

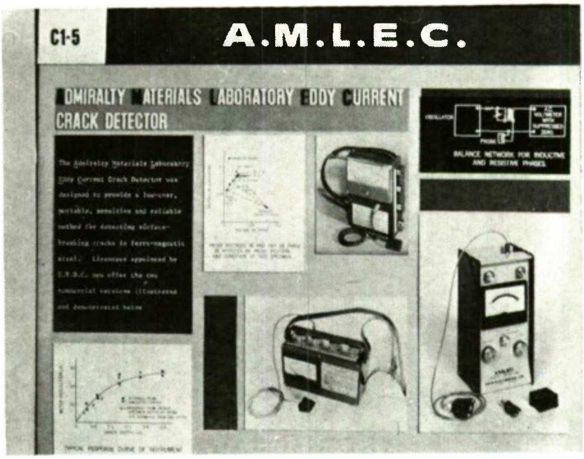
"Engineering Design and Non-Destructive Testing (The LEO Technique)" by D. Birchon, in *The Engineer*, Vol. 226 (1968) p.478.

"The Use of Ceramics in High Temperature Engineering" by D. J. Godfrey, in *Metals and Materials*, Vol. 2 (1968) p. 305.

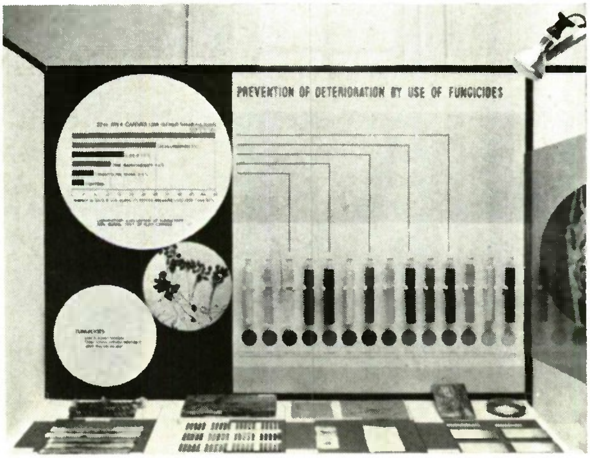
OPEN DAYS AT A.M.L.

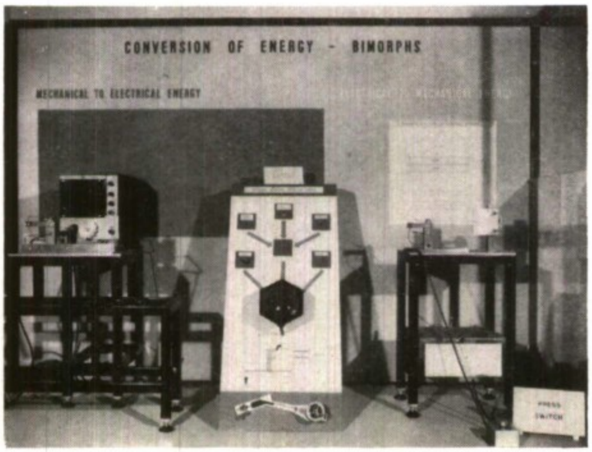
The Admiralty Materials Laboratory was founded on 1st May, 1947. In celebration of the completion of its first 21 years the establishment held a series of Open Days on 1st - 5th October, 1968, attended by about 2,000 visitors. The first two days were restricted to Defence personnel and representatives from Commonwealth and N.A.T.O. countries. The third day was reserved for the Press. The fourth day was attended by visitors from Industry, Research Associations, Professional Institutions, universities and other educational organisations, and local dignitaries. On the Saturday afternoon, 5th October, the establishment was open for the benefit of wives, families and friends of staff.

This was the first time that an event on such a scale had been arranged at A.M.L., but the enthusiasm of the staff and the willing help from others, together with the co-operation of the Clerk of the Weather, all added up to a most successful week. In particular, the interest shown by the visitors from industry, which has led to a constant stream of further enquiries, indicates that the potential of the civil fall-out from the work at A.M.L. is being increasingly appreciated.



There were fifty main exhibits covering the very broad field of interests of the establishment. Some of these are exemplified by the accompanying photographs. All these displays were produced by the staff of MTIP(N) assisted by the Workshop and Drawing Office staff at A.M.L. The final standard of presentation played a considerable part in the success of the Open Days.





Experimental Trials Ship Ordered for Royal Navy

A substantial order for a new experimental trials vessel has been placed with Messrs. Scotts Shipbuilding and Engineering Company Limited of Greenock. This was announced on the 3rd of December, 1968 by Dr. David Owen, M.P., the Parliamentary Under-Secretary of State for Defence for the Royal Navy. This contract will provide work for two years.

It is estimated that the ship and her equipment will cost between £2½ and £3 million. Construction will be to Lloyds' and Board of Trade standards, with accommodation for a civilian crew of 42 officers and men and up to 21 scientists and laboratory staff, as required.

Living spaces, sick bay, offices and other working spaces will be air-conditioned to tropical standards and winterized.

It is planned to have the new ship ready for service in 1971, when the existing Trials Ship *Sarepta I*, a vessel with limited capabilities, is expected to come to the end of her useful life.

Colleges of Technology Project Competition

This year, for the first time, Portsmouth, Brighton and Kingston Colleges of Technology jointly instituted a Project Competition. For this competition each college put forward two students who were adjudged to have carried out the best project in the Honours course for a C.N.A.A. degree in 1968. The competition took the form of each of the six competitors delivering a lecture before a panel of adjudicators on the subject of his project.

The panel awarded the first prize to Mr. S. J. Stubbing for his paper entitled " The Theoretical and Experimental Analysis of an Aerial Remote Position Control System ". Mr. Stubbing is an ex-craft apprentice at A.U.W.E. who was sponsored by S.S.P.(N) for a sandwich degree course at Portsmouth College of Technology where he obtained an upper second class honours degree in Mechanical Engineering. Since obtaining his degree, Mr. Stubbing has been appointed as a Scientific Officer at A.U.W.E.

The prize awarded by the College takes the form of a silver cup to be retained by Portsmouth College of Technology for a year, and a gold plaque for retention by the prize-winner, together with £5 in cash.

Planar Motion Mechanisms

A three day course on the use of Planar Motion Mechanisms (PMM) for investigations of the stability and control characteristics of ships will be held jointly by the Admiralty Experiment Works and University College, London on 21st, 22nd and 23rd April, 1969, at A.E.W., Haslar.

The course will mark the inauguration at A.E.W. of a PMM operating in the vertical plane and will include lectures and discussions on linear and non-linear equations of motion, and vertical and horizontal plane PMMs.

Further details and booking forms may be obtained from M. S. Chislett, Admiralty Experiment Works, Haslar, Gosport, Hampshire, PO12 2AG.

Surfacing Division Welding Institute

A Surfacing Division of The Welding Institute has been formed. It is based on the membership (about 100) of the Metal Spraying and Plastic Coating Division of the former Institute of Welding, but will cover a wider field and give a greater service to industry. The new division will cover all aspects of spraying, surfacing and cladding with metals, alloys and ceramics. Its aims are to promote research, hold conferences, issue technical publications and provide information for members and industry. Membership of the division is available either through membership of the Welding Institute, or as divisional associates or industrial sponsors. Further information is available from the Hon. Secretary, Surfacing Division, The Welding Institute, 54 Princes Gate, London, S.W.7.

Director, A.S.W.E. Appointed D.C.S.(R.N.)

Mr. D. S. Watson, C.B., O.B.E., B.Sc.(Eng.), C.Eng., F.I.E.E., relinquished his post as Director A.S.W.E. at the end of August, consequent on his appointment as D.C.S.(N) in the re-organised headquarters department of CS(RN). A Dinner given by a large number of his colleagues was held in honour of Mr. and Mrs. Watson before they left, and they were presented with a decorative plaque commemorating Mr. Watson's service as the first Director A.S.W.E.

Following the publication of an article in the *Journal* in November, 1966, describing an "Arc Generating Attachment for Milling Machines and Lathes" by Mr. J. C. Burnett, Deputy Workshop Manager, A.S.W.E., further applications were made for patent coverage in U.S.A., France and Germany and the rights of exploitation assigned to the National Research Development Corporation who are now negotiating with a company considering its manufacture. Interest has also been shown by the British Scientific Instrument Research Association who have seen the prototype used in A.S.W.E. and consider it would make a useful addition to their workshop machinery.

⚓ ⚓ ⚓

H.M.S. *Renown* Commissions

H.M.S. *Renown* (Commander Kenneth Mills, starboard crew, and Commander Robin Heath, port crew). the Royal Navy's third Polaris nuclear powered submarine, was commissioned at Birkenhead on Friday 15th November, 1968. Built by Cammell Laird and Co., she will "work up" in the Clyde then proceed next year to the United States to test fire her Polaris Missiles, after which she will join her sister ships H.M.S. *Resolution* and H.M.S. *Repulse* in the operational Polaris force. She will be armed with sixteen Polaris A.3 missiles and six 21-inch torpedo tubes, backed by the most modern and sophisticated underwater detection and navigation equipment, a far cry from the first *Renown*, a 10-gun fire ship which was captured from the Dutch in 1652.

⚓ ⚓ ⚓

Central Dockyard Laboratory

Mr. D. R. Houghton attended the Second International Congress on Marine Corrosion and Fouling held in Athens from the 20th - 21st September, 1968. He was a member of the scientific committee and joint chairman of session II on the 21st September.

Mr. J. R. Saroyan, Head of the Paint Laboratory, San Francisco Bay Naval Shipyard, California, visited the Central Dockyard Laboratory on 11th October, 1968, and met many of the members of staff during his tour of Sections in the Main Laboratory and the Exposure Trials Station at Eastney.

The first International Bio-Deterioration Symposium was held at Southampton from 9th - 14th September, 1968, and was attended by Mr. D. R. Houghton and Mrs. A. M. Mortlock, both of whom presented papers. Mr. Houghton's paper was entitled "Mechanisms of Marine Fouling" and Mrs. Mortlock's "Factors Controlling Metamorphosis in Barnacles". In addition Mr. Houghton was on the executive committee and chaired the special session on Marine Fouling.

Dr. E. N. Dodd and Mr. D. R. Houghton attended the meeting of the fourth General Assembly of the Permanent International Committee for the Preservation of Materials in the Marine Environment which was held in Athens on the 18th and 19th September, 1968. Dr. Dodd is chairman of the sub-group on Underwater Protective Coatings and Mr. Houghton is one of the vice-chairmen of the Committee and chairman of the sub-group of biologists.

⚓ ⚓ ⚓

Admiralty Underwater Weapons Establishment

The Director General Weapons held one of his periodic management team meetings at A.U.W.E. on 28th November. On the following day the management teams of D.G.W. and D.G. Ships held a joint meeting at which presentations on items of common interest were given by A.U.W.E. staff.

Mr. Peter Wallis has been promoted to Chief Scientific Officer and left A.U.W.E. in September 1968 to take up the post of Assistant Chief Scientific Adviser (Research) in the Ministry of Defence.

Mr. Wallis graduated from the City and Guilds College, London in 1944 and joined the Admiralty Signal Establishment (now A.S.W.E.). Promoted S.P.S.O. (Special Merit) in 1958, he moved to Portland in 1959 to lead a joint U.D.E.-U.W.E. Assessment Group which later became the Assessment Division of A.U.W.E. As Head of this Division Mr. Wallis acquired an international reputation for his work in undersea warfare in all its facets, being at various times Chairman of the N.A.T.O. Mine Counter-measures Working Party Technical Panel, Project Director for N.A.T.O. long-term scientific studies in A.S.W. and a leading figure in the Undersea Warfare Study Panel of Sub-Group G. His abilities were recognized with a Special Merit promotion to D.C.S.O. in 1965.

A five-day seminar on "U.D.C. in mechanical information systems" was held in Copenhagen from 2nd to 5th September, and was attended by about 60 participants from 14 countries. Brief papers were presented by Mr. Corbett of A.E.A. and Mr. H. J. Norris of A.U.W.E. on the use of U.D.C. in their organizations. Participants were given practical work in handling and logical structure of queries and U.D.C. number encoding. An account of the seminar is to be issued in the FID/CR report series.

During the autumn, staff of the Information Centre attended courses on mechanization of library routines at ASLIB.

The Director, A.U.W.E., Dr. Ralph Benjamin, and Mr. S. E. Shapcott, Acting Director A.S.W.E., recently spent a week in H.M.S. *Hampshire,* flagship of the Flag Officer (Flotillas) Western Fleet, to observe weapon problems from the user's angle in the Fleet autumn exercises. During the exercises they visited many of the ships and Dr. Benjamin became the first man to be winched down on to H.M.S. *Valiant* from a helicopter.

Mr. R. W. Willmer, D.C.S.O. joined A.U.W.E. on 2nd December, 1968 to become Head of the Weapons Department. Mr. Willmer is 52, and prior to joining the Government Scientific Service in 1947, he served in the R.A.F.V.R. and was involved in work on Gunnery Radar for bomber aircraft. Immediately after the war he joined Phillips Transmission Ltd. as a Communications Engineer.

Since 1947 he has served at R.R.E. Malvern; at M.O.A. Headquarters, and at R.A.E. Farnborough. Whilst at M.O.A. he became Assistant Director of Scientific Research (Electronics) and was concerned with satellite communications. During the past four years, Mr. Willmer has been Head of Radio Department of R.A.E. and has had contact with A.U.W.E. through his responsibilities for development work on sonobuoys.

The Director presented prizes and indentures to apprentices on Friday, 20th December. Prize winners were as follows:—

Best overall apprentice 1st year—T. Broad.

Best craft apprentice 1st year—A. Pearce.

Best craft apprentice 2nd year (RNSS prize)— R. Jackson.

Best craft apprentice 3rd year (RNSS prize)— J. Wilkinson.

Best craft apprentice 4th year (RNSS prize)—J. Collins.

Best craft apprentice 5th year (RNSS prize)— W. Spalding.

Most marked craft advance—A. Henderson.

Best overall apprentice (mechanical)—F. Tough.

Best overall apprentice (electrical)—A. Palmer.

Best O.N.C. result—J. Collins.

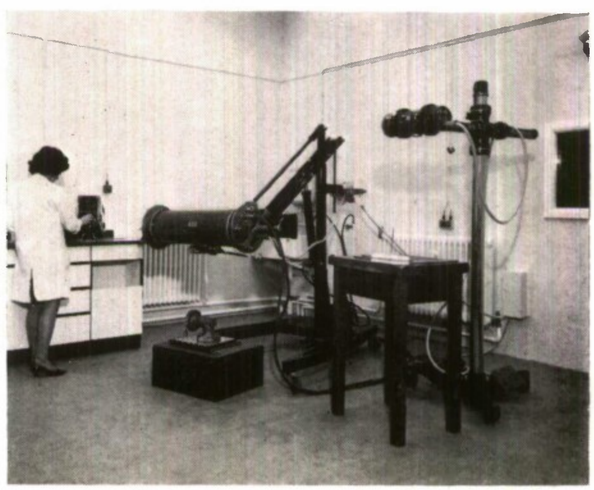
Mr. W. A. Spence, an A.E.O. at A.U.W.E. has been awarded a scholarship at the Royal Military College of Science at Shrivenham. He will study for an Honours degree in Engineering. This is the first occasion on which an officer from A.U.W.E. has been accepted for Shrivenham. The opportunity for an R.N.S.S. officer to take this course is regarded as of some importance as it will be done in a military environment and with service officers who will have been selected for their powers of command and leadership as well as for their technical ability.

At the invitation of the Turkish Navy, the annual meeting of the Mine Warfare Groups of the N.A.T.O. Military Agency for standardisation was held for the first time in Istanbul during the latter half of November 1968. The overall meeting was comprised of a series of meetings of the Mine Laying Working Party, the Degaussing Panel, the Scientific and Technical Panel of the Mine Countermeasures Working Party, and the Mine Countermeasures Working Party itself, the prime activity of the meeting being devoted towards the development of standardised techniques in M.C.M. throughout the N.A.T.O. Navies. The U.K. was represented at this meeting by Dr. J. Wood, Mr. H. A. Hudson, Mr. A. B. Cotton and Mr. D. F. Walker of A.U.W.E., together with Naval representatives from the Naval Staff, H.M.S. *Vernon* and D.G.W./D.W.E.U.(N).

Apart from the official business of the meeting, the Turkish Navy went to considerable lengths to ensure that the visit was both interesting and informative. Official entertainment included a sea trip on the Bosphorous, a tour of Istanbul's famous mosques, palaces and museums, and an evening visit to the European cup football match between Netherlands and Turkish teams.

The completion of a new floor in the southwest wing of the main building at Southwell has provided an opportunity for the drawing office staff to concentrate, while the space they have vacated will be converted into offices for staff being transferred from Bath.

Also at Southwell, the radiography facility has been removed from the main building and a new laboratory, fully protected and complete with control room and dark room, has been constructed.



A corner of the new radiographic facility at A.U.W.E.

Admiralty Oil Laboratory

The laboratory is now installed at its new address—The Admiralty Oil Laboratory, Fairmile, Cobham, Surrey. Tel. No. Cobham (Surrey) 4331.

Visitors will find it on the Portsmouth Road, A.3 just on the London side of Cobham, standing well back from the road in its own grounds opposite the conspicuous Fairmile Hotel. Recent visitors have included Captain Bodnaruk, U.S.N., from the Naval Ship Research and Development Centre, Washington D.C. and Cdr. John R. Eshmen, U.S.N., on 19th November, 1968.

Mr. R. P. Langsdon, Superintendent, Admiralty Oil Laboratory, was in Bangkok from 9th to 13th December, 1968. The Planning Officer at S.E.A.T.O. Headquarters, Bangkok, called a meeting to consider the degree of interchangeability and related problems arising from the use of POL (Petroleum Fuels Lubricants and Associated Materials) by the Armed Forces of the S.E.A.T.O. countries. Mr. Langston represented POL specialist branches of all three U.K. services at this meeting and from experience of N.A.T.O. matters was able to ensure that S.E.A.T.O. commitments in the POL field were compatible with our agreements with our N.A.T.O. allies.

POL materials constitute considerably more than half the total supplies sent to battle areas in the present type of warfare and with the need to satisfy the requirements of equipment from several nations, the need for detailed planning to prevent disastrous logistic problems, is obvious.

⚓ ⚓ ⚓

Press Visit to A.U.W.E.

On the 25th July 1968 a number of Press representatives visited the Admiralty Underwater Weapons Establishment, Portland. The visitors were welcomed by the Director, Dr. Ralph Benjamin who outlined the functions of the Establishment and with members of his staff accompanied the visitors on a tour of the various laboratories and workshops. The morning session opened with an introduction by Mr. F. Burt, Deputy Director, on torpedo development and was followed by a demonstration on the Weapon Test Table where a Mk. 24 torpedo was subjected to every manoeuvre and hazard likely to be met in its natural environment, torpedo recorders and attitude sensing units were also displayed and lively discussions ensued which terminated on departure to the Degaussing Range situated at Portland Bill, where Dr. J. E. Wood demonstrated in his " Magnetically Quiet " laboratory the techniques and methods by which ships are protected from mines. Mr. S. Williams rounded off the morning with a lecture on the aspects of Deep Diving, during which he compared past, present and a very sophisticated suit which the Navy will use in the near future.

After lunch the visitors assembled in the cinema and were shown a film of the 3D Range, then transported to the Propagation Laboratory for an Introduction to Sonar and a Presentation on Underwater Sound by Mr. W. K. Grimley and Dr. V. Flint respectively, followed by a lecture and demonstration by Dr. Morris on the Construction and Uses of Transducers. Finally Dr. Flint gave a presentation on underwater photography and projected a film in which some spectacular effects could be clearly observed. The members of the Press expressed their appreciation for an enjoyable and fruitful day and this was subsequently borne out by the copious write-ups which followed. All in all then, a very successful visit and the staff of A.U.W.E. were well recompensed for their efforts which made it so.

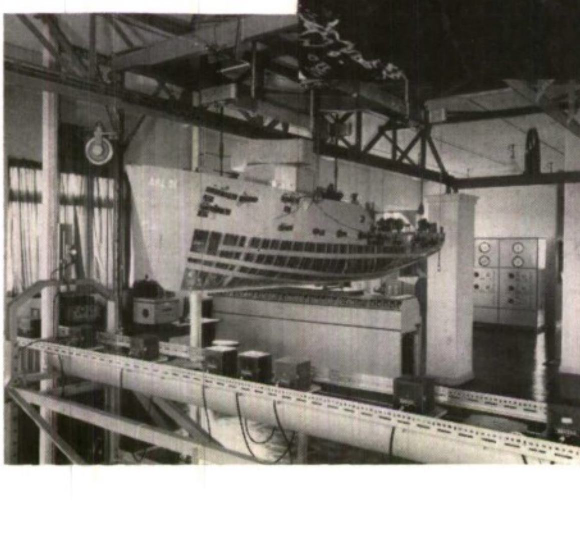
The photographs on the facing page show some of the practical demonstrations seen by the Press representatives during their visit to A.U.W.E.

Weapon test table.

Buoyant body.

Transducer demonstration.

Degaussing laboratory.



" Polyox " demonstration

Retirements

E. N. LEE, R.N.S.S.

Mr. E. N. Lee, Senior Experimental Officer of the Installation Division at A.S.W.E., retired on the 4th September 1968 after 51 years continuous Admiralty service. The last 43 years were with A.S.W.E. (and its variants since the days it was known as H.M. Signal School) during which time he had the distinction of serving under seven Experimental Commanders, 11 Captain Superintendents, six Chief Scientists and one Director. He started his career as an apprentice in the Electrical Department of H.M. Dockyard, Portsmouth towards the end of World War I. He joined the Drawing Office of H.M. Signal School in 1925 and was concerned with Ship Installation work; he later joined the Shore Station Section, increasing its staff to two, in order to handle all the specifications, drawings, preparation of 'E' Lists and Instructions to Dockyard Departments for Shore Stations. In 1933 he was concerned with the planning and erection of a number of 150 ft steel towers to support large directional aerial arrays on Horsea Island. With the outbreak of World War II, Mr. Lee was evacuated to Haslemere as an Experimental Officer to take charge of a rapidly expanding Royal Naval Air Station Section. This work included major modernization programmes for about 32 airfields and later a large UHF conversion programme. He remained in this Section and was associated with Naval Air Stations until shortly before he retired. Mr. Lee served for many years as a member of the I.P.C.S. Committee and devoted considerable time to the sporting activities of the Establishment, being a founder member of the Tennis, Rifle and Films Sections. He has always been very popular and will always be remembered by most of the Establishment as the Films Officer responsible for arranging weekly film shows, a task he carried out most satisfactorily for over 13 years. Mr. Lee was presented with a portable transistor radio, a pair of binoculars and a cheque by the Director, Mr. D. Stewart Watson, in the presence of a large gathering of his colleagues. His many friends and colleagues wish Mr. and Mrs. Lee a long and happy retirement in their home in Southsea.

W. FEARON, B.Sc., A.C.G.I.

Mr. William Fearon, Senior Principal Scientific Officer and Deputy Superintendent of the Admiralty Engineering Laboratory retired on the 3rd January, 1969 after nearly 30 years in Admiralty Service.

"Bill" Fearon took his B.Sc. Hons. II degree in the Department of Mechanical and Civil Engineering of City and Guilds Institute in 1926. Following this he served a three year pupilage at Messrs. W. H. Allen and Sons, Bedford, and subsequently became an experimental assistant in the diesel engine department of that firm—an experience which held him in good stead later in his career. In 1931 he joined Messrs. Lancashire Dynamo and Crypto Ltd. as a development engineer, and in 1934 moved to Messrs. Ilford Ltd. as assistant to the Chief Engineer in the Central Engineering department.

He entered A.R.L. as a Technical Officer in March 1939, where he was concerned with the design of mathematical linkage mechanisms and other fire control equipment for A.A. predictors and gunnery systems. When A.G.E. came into being at Teddington in 1943 he was transferred to that establishment. In 1947 he joined the headquarters department of D.A.E.R. and was seconded from there to the British Joint Staff Mission in Washington for three years, and later to the Naval Catapult section of R.A.E. Farnborough, being concerned with the development of steam catapults and arrester gears.

He returned to A.R.L. Teddington in January 1957, to become Group Leader of the new Engineering Group, a post he filled until his appointment as Chief Scientist A.E.L. West Drayton in May 1966. During this period at A.R.L. he tackled the problem of waterside attack in diesel engines, and the resulting investigations led to the development of the A.R.L. oil cushioned piston. Through this work he became well known to and highly respected by practically every major engine manufacturer in this country and in many countries abroad.

He and his wife will take with them the good wishes of his many friends and colleagues throughout the Service for a long and happy retirement.

Book Reviews

The Corrosion of Light Alloys. By Hugh P. Godard, W. B. Jepson, M. R. Bothwell and Robert L. Kane. Pp. x + 360. New York; John Wiley and Sons, Inc., 1968. Price 112s.

Two thirds of the book is devoted to the corrosion of aluminium alloys whilst the remainder of the volume contains sections on beryllium, magnesium and titanium alloys. The general approach in each section is to present a sufficient amount of basic information followed by an account of specific forms of corrosion which the materials can suffer and finally to cover the behaviour of the alloys in the environments they are likely to encounter in service.

Despite the considerable amount of work which has been carried out concerning the corrosion of aluminium alloys Dr. Godard's contribution is the first collective account of any significance. Failure of aluminium alloys in service can usually be associated with the use of the wrong alloy or the misuse of the correct alloy, a view which the author appears to share. The detailed account of all the corrosion troubles one is likely to experience with aluminium alloys is augmented by the author's access to a vast amount of hitherto unpublished information regarding the behaviour of aluminium alloys in known environments. This data will be appreciated by the user as no attempt has been made to disguise the unsuitability of some alloys for certain applications. The American Aluminium Association system of wrought alloy identification is used in the book and U.K. readers will find the given brief outline of this system helpful.

The section concerned with beryllium will certainly be of interest to the specialized users in the nuclear or aero space fields although it is concise enough to give the general reader a good basic understanding of the properties and behaviour of beryllium-base materials.

Magnesium alloys have never received the same attention from research workers that aluminium alloys attract despite the use of magnesium, particularly in the aircraft industry, for many years. Dr. Bothwell gives a concise account of the corrosion behaviour of magnesium alloys, the section dealing with magnesium anodes is of particular interest.

Titanium alloys are becoming widely recognized as practical alternatives to the more conventional light alloys. Higher material and fabrication costs are partly compensated by the smaller section thicknesses made possible by the high strength to weight ratio. The author discusses various types of corrosion and the effects of specific media including a tabulated account of a wide range of components in chloride environments.

The book is comprehensive in content and well produced. Each section contains a list of references. An author index and a subject index are at the end of the volume. It would be a worthwhile acquisition and is particularly recommended to those concerned with aluminium alloys.

B. N. Hall

The Admiralty Chart. By Rear Admiral G. S. Ritchie. Pp. xi + 388. London; The Bodley Head. Hollis & Carter Imprint, 1967. Price 84s.

The author of this book, Admiral G. S. Ritchie, went to sea in 1932 and joined the Hydrographic Service in 1937. Having commanded four surveying ships before being appointed Hydrographer of the Navy, when he writes a book touching on the history of the Admiralty Chart there is no doubt he knows his subject. But not many people could have brought such an able and charming talent to the telling of the story.

One of the main difficulties may have lain in maintaining some sort of continuity but, with so many surveys in hand simultaneously in so many quarters of the globe, continuity is hard to achieve. Sufficient that herein all the great marine surveyors, many of whose memorials can be found in the names of geographical features, surveying ships and instruments in common use today, are duly recalled and given their places in history.

A wonderful fund of information, this book, among a great many other things, gives a picture of Beaufort (who retired at 81) under whose guidance came "the high noon" of hydrography. The author places Cook as probably the greatest surveyor of all time, but there are so many other people and things of which one would have liked to have heard more—if only what Matthew Flinders said when his brother Samuel failed on three occasions to wind the chronometers (which is pleasantly described as a "horological hat-trick"). If one lists Flinders' 6½ years' captivity under De Caen on the Isle de France, Dalrymple's reluctance to make use of D'Enrecasteaux captured charts, the excessive expenditure of candles under Hurd (the latter possibly the father of the Surveying Service, who produced the first Catalogue of Admiralty Charts), Parry's introduction of Sailing Directions or the activities of Belcher (that unpopular man) at Hong Kong, Spratt in the Crimean War or Denham at Liverpool, one is merely touching the fringe of the fascinating material which Admiral Ritchie so skilfully brings to light.

Not that, as the author admits, the mystery of the surveying service should go entirely unchallenged. He himself quotes considerable extracts from that remarkable publication "The Bogus Surveyor", and your Reviewer remembers various occasions when it became necessary to tow a surveying ship, captained by a future Hydrographer, off the coral in the Persian Gulf because, apparently, surveyors rarely looked where they were going. There is, too, the classic case of the Fleet Flagship hitting an uncharted pinnacle rock at Skiathos when the name of the officer who conducted the survey, as given on the chart, was Lieutenant XYZ who by that time was Hydrographer of the Navy.

The bestowal by surveyors of place names could in itself be the subject of an interesting treatise. How easy it was in the early days to name a mountain after one's ship or oneself or one's girl friend, yet an acquaintance of your Reviewer, doing a running survey inside the Farazan Islands, marked the only outstanding feature—a camel thorn with an arab and two goats beneath it—as "Shepherd's Bush", only to have the name disallowed by the Hydrographer of the day.

The author, for full measure, deals with the laying of ocean cables, the Compass Department of the Admiralty (that most distinguished organization which, though at one time under the Hydrographer, later escaped from this yoke), the Challenger voyages and that curious and vaguely unsatisfactory occupation called "Vigia hunting", which implies clearing the charts of various alleged hazards sometimes, as has been said elsewhere, observed and reported by Masters of passing vessels when appearing between two pink elephants against the setting sun.

The production of the book is good and there appear to be only two misprints. If, on page 329, the author writes of the swinging of the compass, when he means the swinging of the ship to adjust the compass, one must remember that even Homer nodded. There is an extensive bibliography and an adequate index.

Your Reviewer enjoyed the book immensely, and it should give great pleasure, not only to that distinguished company, the marine surveyors, and to the navigators (who, lacking the surveyors, might well be on national assistance), but also to anyone whose interests lie in the comings and goings of such as pass on the seas upon their lawful occasions.

A. V. Thomas

Basic Ideas in Neurophysiology. By T. D. M. Roberts. Pp. xi + 108. London; Butterworths, 1966. Price 25s.

This book surveys a growing field of general interest. The titles of the five chapters indicate its scope. Basic concepts, mechanical properties of skeletal muscle, electrical activity of nerve and muscle, how nerve impulses convey information, the muscle spindle and the stretch reflex. Through it all runs the idea that we are dealing with a complex of interacting mechanisms and it is only in exceptional circumstances that we can look at them separately. Hence the emphasis on terms like reflex and feedback. Free use is made of flow diagrams to illustrate various control mechanisms. A very useful feature of the book is that each chapter apart from the first is provided with a concise summary. The list of sources, bibliography and index all look adequate. The style is good, the sentences being clear and concise.

H. N. V. Temperley

On Documentation of Scientific Literature. By Th. P. Loosjes. Pp. ix + 165. London; Butterworths, 1967. Price 42s.

This book is a " review of the state of the art " which would be useful to the serious student who is prepared to go to the original literature using the book as a director and also as a reference source to the general practitioner. The field is one in which there is a notable shortage of reliable text-book; it is one in which the Dutch excel and the author is an eminent Dutch librarian who published the book in German some years ago and who has added to the text for this English translation.

As is to be expected in a book of 165 pages of text, it consists largely of summaries of other publications, though it is meticulous in acknowledging the original authors and in giving bibliographies at the end of each chapter. The 31 line diagrams are also copied or adapted from other publications, though one of them, devised by Mr. Loosjes himself, appears no less than seven times in different variations: it depicts aspects of the flow of documentation and to the reader who merely wants to find references to particular topics it is apt to be not particularly helpful. It should not be supposed that the author presses his own ideas or opinions unduly, and although his name appears in several of the chapter bibliographies it does not occur in the index at all.

The text is highly condensed and there are of course some omissions and awkwardnesses. Thus Boolean algebra gets a bare mention although there is an overlapping diagram from a publication in 1962, and a diagram of Vickery's is reproduced on page 145 with a small difference in layout and an incorrect ascription: the diagram is referred to as very clear, but the change in layout makes the description confusing to anyone not already familiar with the matter and in any case the principle is more simply explained in words than by the diagram. The second half of the book deals with retrieval systems and covers the important points, though sometimes in a rather unusual way; thus the author sees the use of a List of Chosen Descriptors as a vital distinction, which brings UDC and feature cards together in one group with Uniterms and automatic indexing in the other: the distinction between item and term indexing is in fact adequately dealt with later in the book.

The author's summaries have a natural tendency to restrict themselves to the main points of the original publications and thus to omit any reference to their special circumstances, which are sometimes vital, but the book will be invaluable to users who need a guide to the original literature and have not maintained their own indexes. The translation, by A. J. Dickson, is so good that it shows very little sign of being a translation at all; the layout is clear, printer's errors are few, the index is adequate and the price is reasonable.

A. Holloway

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 the Royal Naval Scientific Service
Covering dates 1969 Jan
Availability Open Document, Open Description, Normal Closure before FOI
Act: 30 years
Note Vol 24 No 1
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.

NOTICE TO READERS

The Editor extends to all readers of the JOURNAL a cordial invitation to contribute articles of R.N.S.S., 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 material within its pages does not constitute open publication and does not prejudice the subsequent use of the material 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 3,000 and 5,000 words, 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.

Views and opinions expressed in the JOURNAL are not necessarily endorsed either by the R.N.S.S. or by the Editor.

All communications should be addressed to:—

The Editor,

Journal of the Royal Naval Scientific Service,

Ministry of Defence,

Station Square House,

St. Mary Cray, Orpington, Kent.

Telephone Orpington 32111 Ext 345