

Founded 1925

Incorporated  
by Royal Charter 1961

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

VOLUME 41 No. 12

DECEMBER 1971

# THE RADIO AND ELECTRONIC ENGINEER

The Journal of the Institution of Electronic and Radio Engineers

## Changes in Government Science

OUT of a total estimated R and D expenditure in Great Britain for 1971-2 of £645M, the Research Councils of the Department of Education and Science will account for about £109M. This is the financial background to a recent Government 'Green Paper' on the long-term strategy for an important area of scientific research and development in this country, some of which affects directly or indirectly the electronics industry and the community of electronic and radio engineers. The Head of the Central Policy Review Staff, Lord Rothschild, F.R.S., has contributed to the Green Paper a report on 'The Organisation and Management of Government R & D', while Sir Frederick Dainton, F.R.S. as Chairman of a Group appointed by the Council for Scientific Policy, has presented a report on 'The Future of the Research Council System'—both reports have appeared together under the title 'A Framework for Government Research and Development'.\*

The accompanying Government statement seems at first sight to indicate a wide measure of agreement with the Rothschild thesis, noting merely that the C.S.P. report 'is not at variance with' the 'customer/contractor' principle advocated by Rothschild. But, the C.S.P.'s proposed change in the Research Council system of setting up a Board of the Research Councils is hardly compatible with the Rothschild proposals that—with the exception of the Science Research Council which is to remain with the D.E.S.—the Agricultural, Medical and Natural Environment Research Councils should lose some autonomy by being brought much closer to appropriate Government Departments from whom they would derive most of their financial support. The intention is that 'applied R & D' should be linked more closely to the Departments and distinguished from 'basic research' which would continue to come under the D.E.S. through the S.R.C.

The Dainton report sees things differently, wishing to retain independent Councils which will work in closer co-operation through the new Board, and maintaining the view of the past few years that 'basic' and 'applied' research benefit by not being arbitrarily divided. The good working relationship between the universities and the Research Councils is considered to have inherent advantages, not to be obtained if there were close links with Government Departments.

As part of the Rothschild plan to make the R & D functions of the Departments more effective, it is envisaged that there should be a Controller R & D who would be the Chief Executive and responsible for the 'contractor' aspects. The 'customer' view of the department is to be associated with a Chief Scientist. This concept of the 'customer-contractor' relationship *within* a department is an interesting one which is intended to ensure that a programme of work is not embarked upon without clearly establishing responsibilities for expenditure. Nevertheless, provision is to be made for some research outside the directly contracted work by the device of a 'general research surcharge' of about 10% of sanctioned expenditure.

While the prospect of financial considerations being more intimately brought into scientific research is likely to worry the traditionalists, the encouragement which Lord Rothschild gives to the proposals of the Fulton Committee on the Civil Service for enabling scientists to gain administrative experience will be more widely approved. However, his corollary that a non-scientific senior civil servant might become Director of a large Government Laboratory by virtue of his administrative ability may be less easy for some to accept.

Apart from its endorsement of the 'customer/contractor' principle, the Government has indicated in the Green Paper that 'it would be right to preserve the Research Councils under the sponsorship of the Department of Education and Science.' Again, it is the Government's view that a body of authoritative advice should be available in the allocation of the D.E.S.'s Science Budget. Thus, in so far as the Rothschild and Dainton reports may be considered to be conflicting, there seem to be hints that a compromise will be evident when the final White Paper appears in the Spring.

\*Cmnd. 4814, H.M.S.O., price 52½p.

## Contributors to this issue



**Mr. I. White** studied electrical engineering at the Royal Radar Establishment College of Electronics, Malvern, and after obtaining the Higher National Diploma in 1964 worked for the Plessey Company on research and development of communications systems. In 1967 he went to the University of Birmingham to study for the degrees of M.Sc. and Ph.D.; his thesis dealt with various aspects of target recognition. He

is currently with the Plessey Company, engaged on research in the field of pattern recognition.



**Mr. Ralph E. G. Keon** (Member 1963; Graduate 1961; Student 1950) studied electronic engineering in France, Belgium and England and obtained engineering qualifications in all three countries. After an apprenticeship in the microwave laboratory of C.S.F. in Paris, he joined the radio valve division of S.T.C. at Paignton where he worked for seven years on development and design of many types of vacuum

and gas-discharge valves. From 1957 to 1959 he was engaged on development of high power transmitting valves at the M.O. Valve Company and subsequently on research on gas-discharge devices. In 1963 he moved to the electronic instrumentation field and to marketing, becoming export sales manager of Airmec Limited. Since March 1970 Mr. Keon has been a director of Culton Control Systems Limited.



**Mr. Harry Bisby** graduated with first class honours in physics from King's College, University of London, in 1938. From 1940 to 1947 he was with the Ministry of Supply, specializing in mobile ground radar. During the war years he served in the Reserve of Air Force Officers, with the rank of Flight Lieutenant, as radar adviser to H.Q. Middle East and to Supreme H.Q. Allied Expeditionary Force. In 1947 he trans-

ferred to the Atomic Energy Research Establishment, Harwell, and he is now group leader for data acquisition and control systems in the Electronics and Applied Physics Division. Mr. Bisby has played a leading part in the development and standardization of instrumentation systems. He is the U.K. Laboratory Delegate to Executive Group of the ESONE Committee, the U.K.A.E.A. Representative to SIRA Institute, and a member of B.S.I. Committee NCE/6 and he has written numerous papers and articles on nucleonic instrumentation.



**Mr. J. F. McIlvenna** (left) is an electronic engineer and **Mr. C. J. Drane** (right) a physicist at the U.S. Air Force Cambridge Research Laboratory where they have been working together in the microwave physics laboratory for several years on aerial array theory and design. Their earlier paper on this subject has recently received the Brabazon Award as the outstanding contribution on radar published in the Institution's *Journal* last year. Fuller details of their careers appeared in *The Radio and Electronic Engineer* in January 1970.



**Mr. B. Martin** graduated in electronic engineering at the University of Sheffield in 1968. Since then he has been working on device aspects of the Gunn effect at the University of Sheffield with the support of a Ministry of Defence (Procurement Executive) contract.



**Dr. G. S. Hobson** has been working in the Department of Electronic and Electrical Engineering since 1968 and earlier this year he gained his doctorate for a thesis on 'Some fluctuation and modulation properties of transferred electron oscillators'. He has recently spent three months in the School of Electrical Engineering of Cornell University, New York.

A fuller note on his career was published in June 1970.



**Mr. R. P. Owens** received the degree of B.Sc. in physics from the University of Bristol in 1962, and in the same year joined EMI Electronics Ltd as a graduate apprentice. From 1964 to 1967 he worked on the design of aerial systems and microwave components in the Radiation Laboratory at EMI and during this period he studied part-time for the University of London M.Sc. degree in microwave physics, which was

awarded in 1968. In 1967 he became a research assistant in the Electronics Branch of the Royal Military College of Science, where his work has been concerned with the general characteristics and equivalent circuits of Gunn diodes.

(Mr Owen's paper appeared in the November Journal)

# The CAMAC Interface and some Applications

By

**H. BISBY,**

B.Sc., A.K.C., C. Eng., M.I.E.E.†

*Presented at a meeting of the Instrumentation and Control Group in London on 29th April 1971.*

CAMAC is a definitive style for implementing the interface conditions which exist when many channels of input/output information share a common data-processor/controller. The features of CAMAC are described to indicate the applicability of CAMAC-compatible equipment and programming to real-time situations. Some of these are illustrated by typical systems which use either a computer or a simpler device as the central processor/controller.

## Glossary of Special Terms

**Interface** The boundary between two different systems, e.g. operator/equipment interface.

**System** An assembly of equipment to perform a given task.

**Dataway** A multi-wire highway for the transfer of data.

**Real-time** Applies to systems which operate at sufficient speed to analyse or control external events happening concurrently.

## 1. Introduction

Real-time situations are becoming more common in the progressive areas of electronic applications, where control and data-processing by computers plays a major part. The centre of interest in these situations may be a process plant, a section of a transport network, a generator of electric power, a hospital patient or the behaviour of a nuclear particle. These systems combine measurement, control and display and require a 'processor/programme/controller' as the system-component that is shared by all these functions. For this reason all other equipment is, by convention, considered peripheral to the central system-component. For example, printers, punches, plotters, displays and disks are the recognized peripheral equipment to a computer and for the purpose of this paper, any source or acceptor of information (e.g. any sensor or actuator), will be treated as peripheral equipment.

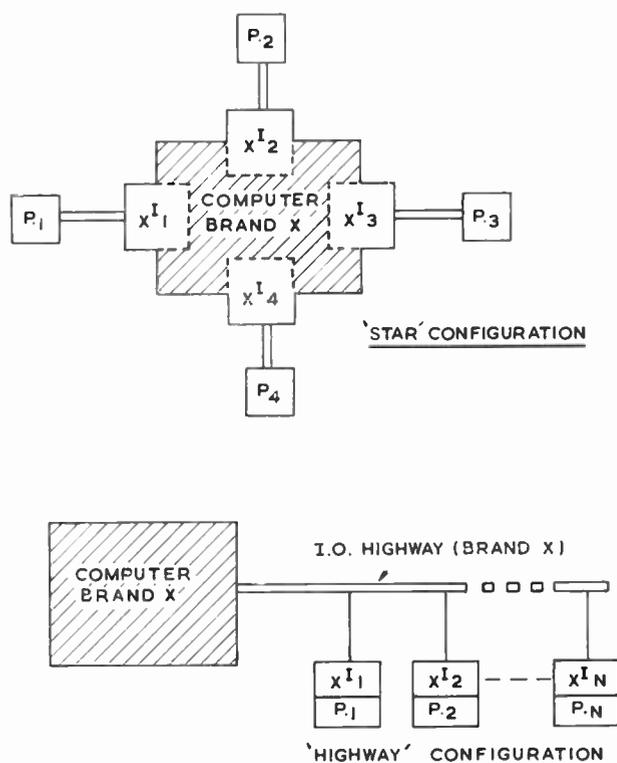
'Peripherals' may be connected to the computer in 'star' or 'highway' configurations (Fig. 1). Since both the discrete input/output channels and the highway are normally characterized by features of the computer, each connexion requires some interfacing electronics to handle mis-match between the peripheral and the computer, e.g. differences in signals, signal conventions, unit of data-transfer, speed of operation. In organizations which employ many different types ( $C$ ) of computer and many different, but commonly used, peripherals ( $N$ ), there will be  $(N \times C)$  types of interfacing device. The introduction of a peripheral, with new and desirable features such that it is required on all these computers, will generate  $C$  new interface devices.

If, however, all computers were equipped with the same standard input/output channel or standard highway, the number of interface devices required would be

† Electronics and Applied Physics Division, Atomic Energy Research Establishment, Harwell, Berkshire.

reduced to  $(N+C)$ . The British Standard Interface BS4421, inspired by the National Physical Laboratory, does this for 'star' configurations. It defines standard conventions for transferring data in parallel 8-bit bytes from a source to a receiver. It is available on some peripherals and, as an option, on some U.K.-based computers. BS4421 is not, however, applicable to a highway configuration and therefore not of use in situations where the number of peripherals is both large and variable such as exist with a large body of computer users in nuclear laboratories. These laboratories have therefore defined CAMAC as a highway interface which is unique and independent of both the type of computer and the type of input/output peripheral. It has been defined by, and agreed between, 50 major nuclear laboratories in 14 countries of Europe, Canada and U.S.A.

In practice, because of the defined analogue and digital signal standards of CAMAC, it is possible for one



**LEGEND**  $P_N \equiv$  PERIPHERAL  
 $x^I_N \equiv$  PERIPHERAL INTERFACE

Fig. 1. 'Star' and 'highway' arrangements.

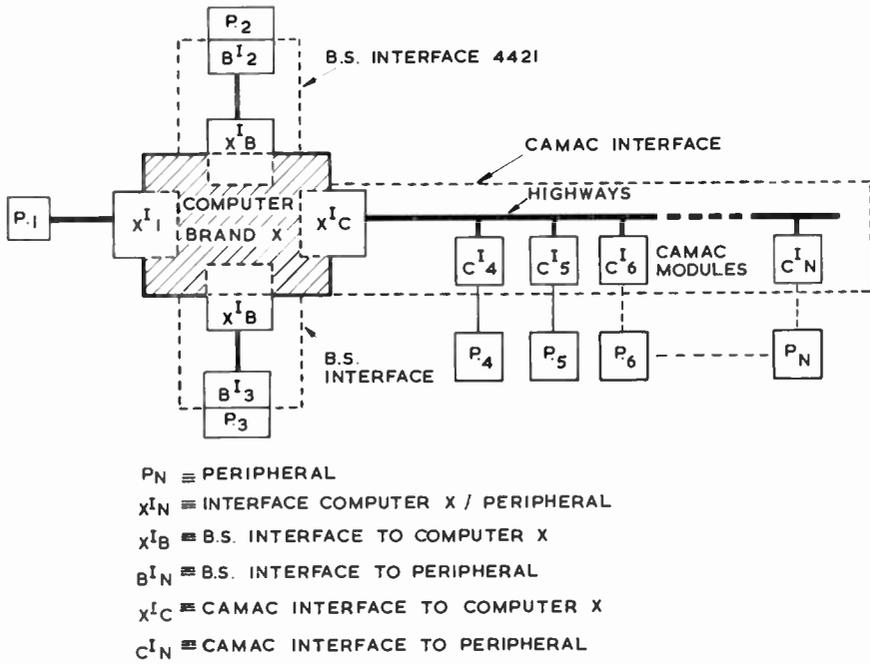


Fig. 2. Configuration with BS4421 and CAMAC Interfaces.

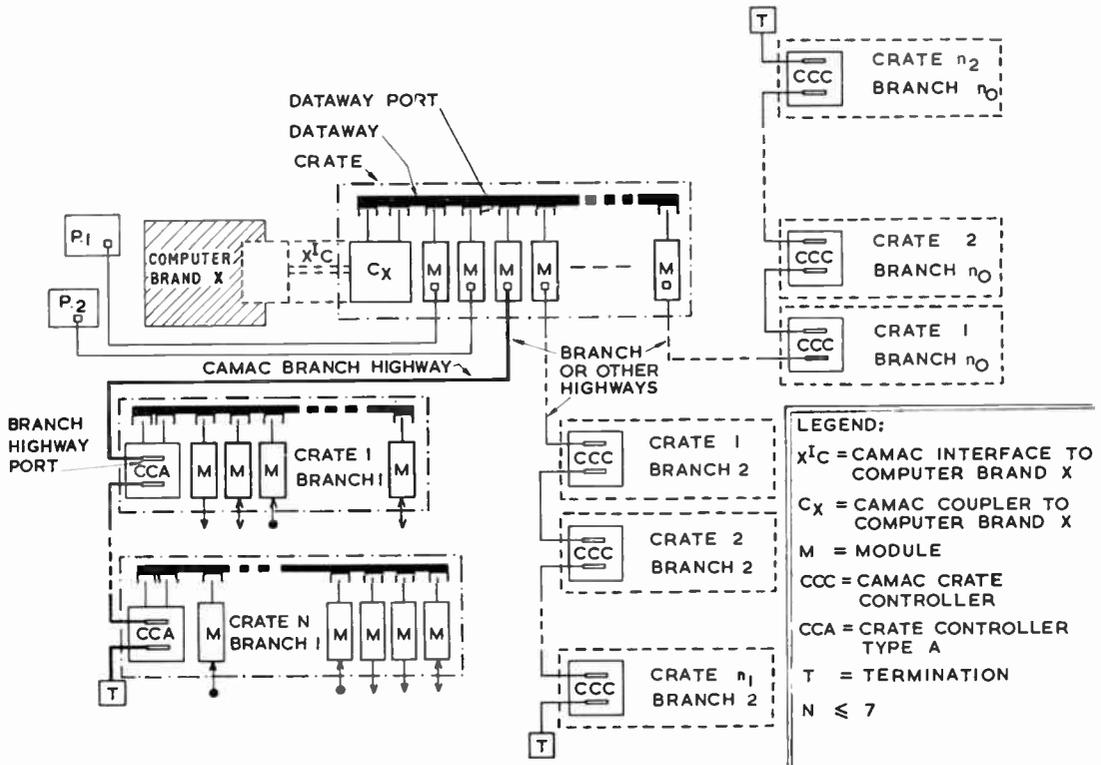


Fig. 3. The highways of CAMAC.

CAMAC unit to satisfy the interfacing conditions of many similar peripherals. Furthermore, these CAMAC units can be designed and made available without any knowledge of which peripheral or computer they will subsequently be used with. Thereby the 'library' of CAMAC units needed in large organizations can be substantially reduced and the equipment situation rationalized.

As illustrated in Fig. 2, the use of CAMAC in no way precludes the use of a computer's specialized input/output channels or BS4421, where it is more appropriate or economical to do so. In fact, many of the present systems which use CAMAC have 'star' configurations for some standard peripherals and a CAMAC highway to expand them, in the most economical and flexible manner, to more peripherals.

The 'hardware' implications of CAMAC offer great economies in effort, time and money but even greater simplification and therefore economies are offered in the field of program 'software' due to CAMAC's defined structure and modes of data transfer.

## 2. Description of CAMAC

CAMAC defines a design style applicable to situations where a computer, or similar device, is to communicate, along a highway, with a few or very many peripheral devices. The highway configuration can be complex or simple. Because of its unique definition, the highway can be interfaced at suitable points to any computer via a coupler or controller, or to any peripheral, via a module, these two interfacing problems being entirely independent of each other.

CAMAC therefore must specify the methods of communication but, in addition, it defines the physical, electrical and data transfer characteristics of the component parts, yet only so far as necessary to ensure compatibility between them. The specifications for CAMAC are contained in references 1, 2 and 3. References 4, 5 and 6 provide background information.

### 2.1. The CAMAC Highways

The CAMAC multi-wire highway can transfer 24-bits of information, as a parallel word, in both the 'read' and 'write' modes. A word can enter or exit this highway via electrical 'ports' which are, normally, T-junctions onto the highway but may be end-junctions in particular circumstances. The highway can be one of two types—the 'Dataway' and the 'Branch Highway'. These have a general similarity but also important differences, in detail, because of their different purposes. The dataway is always used in any CAMAC system and the branch highway only when either the number of interface units is large and/or their geographical disposition is such that the dataway needs extension (see Fig. 3).

The dataway has 25 ports and a fixed physical size within a framework, or 'crate', which is a standard 19-in. rack-mounting unit (Fig. 4). The ports are 86-way edge-connector sockets whose pins are separately attached to the wires of the dataway. Each port is an integral part of a 'station', which is of specified dimensions. Equipment units can fit into the stations and make electrical contact with the ports via an edge-connector

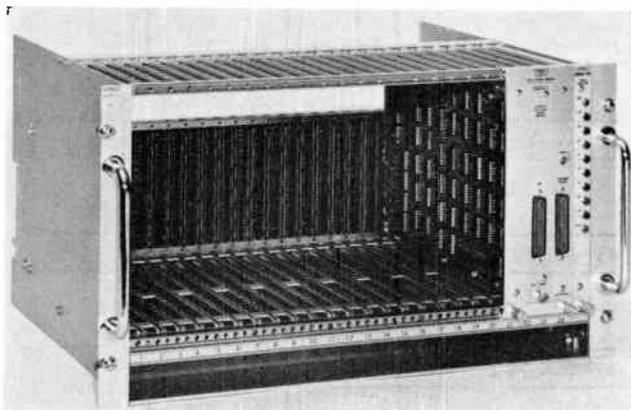


Fig. 4. CAMAC crate and dataway.

plug on the unit. The stations are physically identical so that CAMAC-compatible units are 'modular' and typical ones are shown in Fig. 5.

Some of these units provide the interface conditions for transfer of information between a peripheral device and the dataway. They may also provide other signal or data processing functions which, in some circumstances, may be controlled by information supplied to them via the dataway. These units, by CAMAC conventions, are called 'modules'. Other units provide the interface and control conditions for transfer of information between the dataway and, say, the computer directly. These units are called 'controllers', to distinguish them from 'modules', and also may have functions over and above those mentioned. Each crate, equipped with modules, therefore has a controller of some kind to organize the sequence of operations whereby information is transferred on the dataway between a specific module and the source/acceptor of this information, be this the controller itself or some other device acting through the controller. This 'other device' could be a computer, connected directly to the controller, or a branch highway which in turn is connected directly to a computer.

A specific kind of controller (crate controller type A—see ref. 2) is defined for connexion to one of up to seven defined ports along the length of the branch highway. Thereby seven crates, each fitted with modules and a crate controller type A, may be connected via this highway to a common unit called the 'branch driver'. These ports are physically different from the dataway ports because the geographical disposition of crates can be variable. For this reason also, a data transfer operation between a

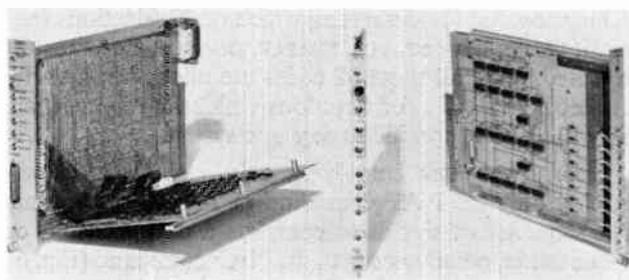


Fig. 5. CAMAC modules.

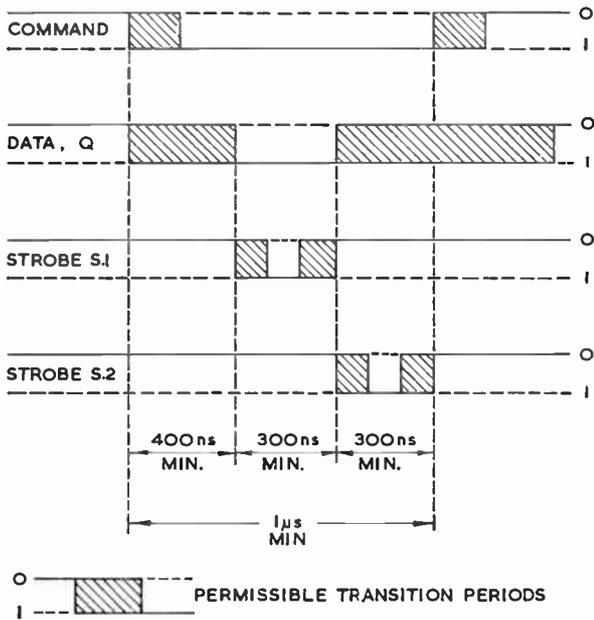


Fig. 6. The dataway operation.

module and the branch driver has to be self-adapting to the length of the branch highway involved.

2.2. Methods of Communication

A bi-directional communication system, as described so far, should be able to transfer a parallel word in either direction and in addition permit this action to be instigated by the central processor or by any module/peripheral within the system. In addition, if the central processor is to be thoroughly current in its knowledge of the system, it must be able to establish the status of the component parts.

To determine which component part is involved, either for data transfer or status operations, CAMAC provides an address coding system which can uniquely identify a crate (C) in a branch, a station (N) in that crate and therefore the module that is plugged in that station. A CAMAC module is allowed to have up to 16 sub-stations (A) and anyone of these can be identified. The full address code (CNA) requires 12-bits (C ≡ 3 bits for 7 crates, N ≡ 5 bits for 23 station addresses and A ≡ 4 bits for 16 sub-addresses) giving a total location of 2576 variables. The redundant codes of N are utilized for multiple-station addressing and so on. A 5-bit code (F) is provided for defining which of 32 functions (e.g. READ, WRITE, TEST STATUS, ENABLE, DISABLE, etc.) must be executed. Not all these 32 codes are allocated; a useful sub-set is available for functions which may be particularly useful and specific to any one system.

Clearly a system employing only a single crate will employ only the NAF command of 14-bits (maximum). The sequence of signals to complete a data transfer, on the dataway, commence with the NAF command (Fig. 6). After a short time-lapse in which these signals have settled to their 'l' state and been interpreted at the

module, the data or status signals are established on the dataway lines. At this point in the sequence, the controller generates a strobe signal (S1) to open appropriate gates either in the controller (for read and status operations) or the module (for write operations). Another strobe (S2) follows S1 and can be used for operations which can be auxiliary to read/write, status testing and so on. After S2, the command can be removed to prepare for the next operation. This may well be instigated by the computer as the next step of a main program sequence or a sub-routine brought into action by a program interrupt being received from a module via a look-at-me (L) signal. Depending on the type of controller and computer, this L signal may instigate a direct transfer of data into or out of computer store without intervention by the c.p.u. of the computer (i.e. autonomous transfer).

On the branch highway, the dataway signals are basically relayed by the controller; however, the method of establishing the conditions required for a transfer are different. The branch driver requires a positive response from the addressed crate before the transfer sequence is commenced. In addition, it is not feasible to relay onto the branch highway all the possible L signals which might be generated from even one crate, so that a selection principle is invoked if polling of all possible sources of L is to be avoided. This is done by the 'graded L' mode of operation on the branch highway, whereby the pattern of a 24-bit word set up on the read/write lines of the branch highway gives the status of 24 sources of L signals selected from all possible sources of L signals in the total branch.

In addition to the facilities for reading and writing data from and into modules, respectively, testing status and allowing interrupts, other modes of communication give facilities for implementing on all units commands like C (CLEAR all data and status signals) I (INHIBIT all data taking) and Z (INITIALIZE all registers to a pre-set condition).

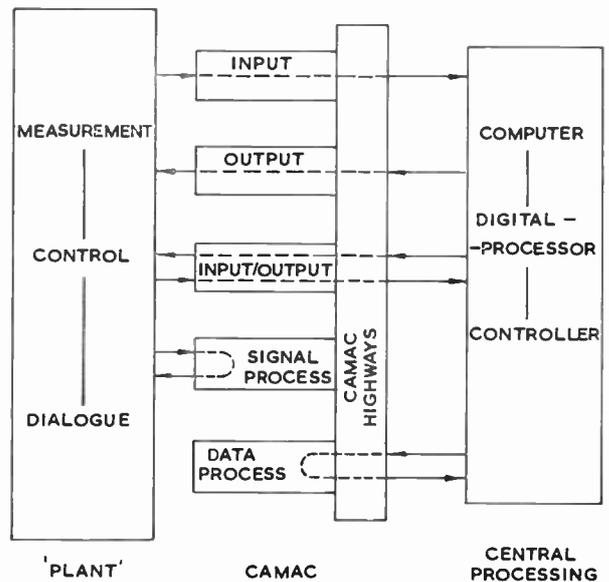


Fig. 7. Information flows for module classification.

### 3. Applications of CAMAC Facilities to Modules

The facilities offered by CAMAC can be used to generate modular units which can perform tasks which may be highly specific to a particular peripheral or may be of more general application to a variety of peripheral duties.

Employing the CAMAC convention, a typical module would contain circuit elements whose functions would be to decode the command signals (NAF), to indicate status signals, to respond to the common commands (CIZ) and to input or output data. It is in this latter function where the variability arises in the type of tasks which CAMAC-compatible units can perform. These tasks may be classified according to the flow of information indicated in Fig. 7.

#### Input Modules

Information, contained in signal inputs, is processed and converted to the CAMAC dataway standards for further transfer on the highways.

#### Output Modules

Information, contained in digital signals on the dataway, is processed and transferred to external peripheral devices.

#### Input/Output Modules

Information can be handled in both the read and write direction, typically for external peripherals which have read/write properties.

#### Signal Processing Modules

Signals received from peripheral devices can be processed, under the control of a register in the module, for re-transmission.

#### Data Processing Modules

These are the hardware equivalents to software programs for processing data within a system. They accept data, via the dataway, perform an algorithm and re-transmit the subsequent data onto the dataway.

Many of the CAMAC-compatible modules currently available<sup>7</sup> can now be described in fairly brief terms by giving examples in the above five classes.

#### 3.1. Input Modules

These would accept current or voltage analogue signals, representing external conditions such as temperature, pressure, or strain, and provide the necessary signal conditioning and/or analogue-to-digital conversion to a data word which can be read. Alternatively an external multiplexing a.d.c. may present a data word for the input module to gate onto the highway when requested to do so. The data word could be from an angle encoder, a BS 4421 driver (source) or it could represent the on/off state of switch contacts, perhaps the limit switches of a device under control or switches set by an operator. A CAMAC system may require to be instructed by paper-tape, so that an interface module for a paper-tape reader would be typical. One common task uses a register in the module to integrate the number of serial input pulses from flow meters, radiation detectors and regular-pulse generators. Such modules may control the duration of an operation by indicating to the computer when a preset number has been reached. Alternatively the number in

the register after a preset time will give pulse-rate. Voltage to frequency converters or Wilkinson-type pulse analysers, external to CAMAC, can be interfaced to the dataway by serial-to-parallel converter modules.

#### 3.2. Output Modules

Typically, output modules can generate signals which are the analogue equivalent of the data word fed to the module from the highway. These can be used to control solenoids and voltage sensitive devices, for motor speed control and driving meter or c.r.t. displays. Outputs may be in parallel or in serial form for driving printers, punched paper-tape, X-Y incremental plotters, for switching lamps or other external devices and for digital displays. A module in this class may control an external relay—switching multiplexer handling many channels of analogue information. A BS4421 (source) output module, combined with the (acceptor) input module referred to earlier, can provide in CAMAC a two-way multiplexing system for peripheral devices fitted with BS4421.

#### 3.3. Input/Output Modules

Some typical tasks for these modules may be to couple standard peripheral devices to the CAMAC dataway, such as a magnetic tape or disk, a teletypewriter, visual display units or a core-store. Each module, generally, will be specific to the type of peripheral and can provide an alternative method of adding proprietary peripherals from different manufacturers to the same computer. A module, coupling to a P.O. modem, offers an opportunity of a serial data link for connecting into distant sources/acceptors of data. Many input/output modules have the ability to store data on registers and transfer this to the dataway or a peripheral when instructed to do so either under instruction from the computer or on receipt of a signal from the peripheral.

#### 3.4. Signal Processing Modules

Typical units could condition external analogue signals, or perform high-speed multiplexing, using f.e.t. switches, or could select output routes for input pulses or perform any of a range of logic functions on combinations of input signals. Any of these actions could be directly controlled by program.

#### 3.5. Data Processing Modules

The type of typical arithmetic or algorithmic task would be binary to b.c.d. conversion, multiply/divide, floating-point calculations and such modules provide these options on any computer which is interfaced to CAMAC. They may be used to provide an add-on facility for any algorithm which is frequently required and which otherwise would occupy excessive computer c.p.u. time. A binary-to-b.c.d. converter has been found most useful for non-computer systems in which binary input information on the dataway is to be printed or displayed in decimal format.<sup>8</sup>

### 4. Applications of CAMAC to Systems

Another view of CAMAC is that its two types of highway provide two levels of multiplexing capability. The

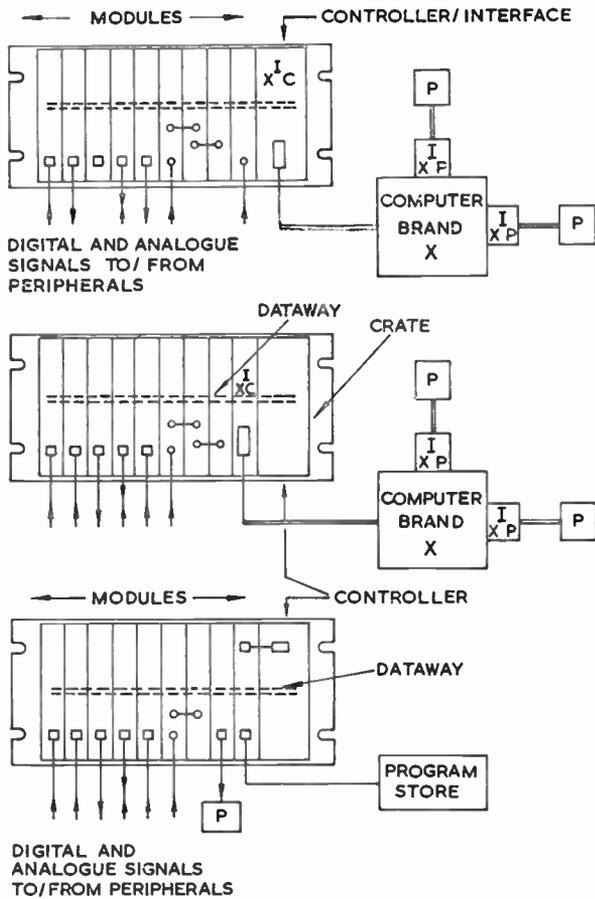


Fig. 8. Examples of single crate systems.

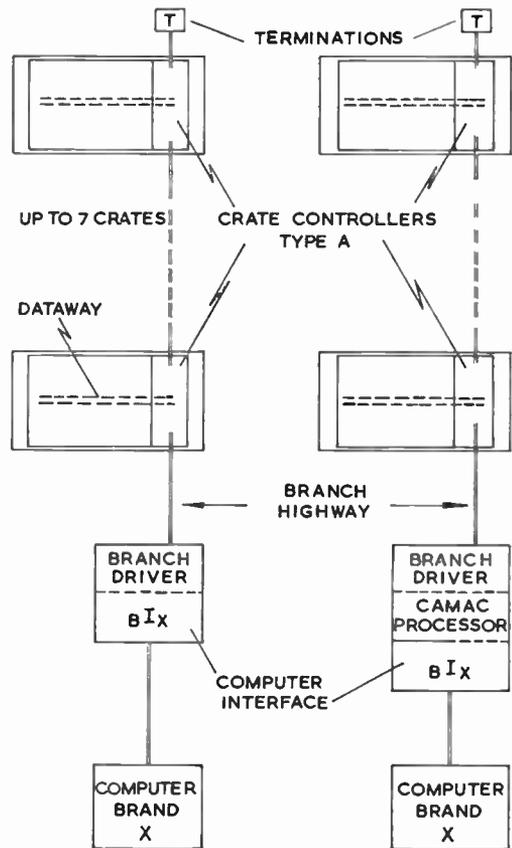


Fig. 9. Multi-crate systems.

dataway enables data transfers in either direction between the controller and either any one module or any one sub-address of any module within a single-crate configuration. Equally, the branch highway enables data transfers, in either direction, between the branch driver and any one crate and thereby any module or part of a module in that crate.

There are many minor variations on several basic system arrangements. Three variants of single crate systems, shown in Fig. 8, are typical and self-explanatory. In the first example, note that the crate-controller is also the interface to the computer and will be specific therefore to that type of computer. The modules however will be independent of the type of computer and this applies to the second example with the exception that one of the modules is the computer interface. The controller now is not directly related to the features of the computer used and is therefore a unit of more general applicability. The third example illustrates how the controller itself is a computer-like device that can be programmed from an external store.<sup>9</sup> This store may be a plug-board, matrix or other type of read-only-memory. Alternatively, if it is a corestore then it may be used for program or data storage.<sup>10</sup>

There are similar variants in systems which cannot be accommodated in a single crate and therefore need the higher level of multiplexing. There are, clearly a number of ways of arranging multi-crate systems in 'star' or

'highway' configurations. For such systems requiring fast, parallel-data transmission, CAMAC has specified the branch highway and recommends the use of a crate controller (Type A) unless there are over-riding considerations to do otherwise. Figures 9 and 10 illustrate basic configurations using one or more branches, according to requirements. In Fig. 9, the CAMAC coupler can interface the branch driver and the input/output channel of the computer. Everything from some point inside the coupler is independent of the standards of the computer and the coupler forms a convenient 'neutral' component between the computer and the rest of the system. A somewhat different approach could combine the branch driver, the computer interface and a processor for CAMAC transfers which directly access the memory of the computer without occupying time of the computer's c.p.u.<sup>11</sup> The branch driver is shown in Fig. 10 as a module in a 'master crate' under instruction from a 'system controller'. This crate could contain more than one branch driver module, to give a 'star' arrangement for several branch highways, and also accommodate interface equipment for peripherals that are common to the whole system; Fig. 10 shows an alternative in which the system controller is independent of the computer and modules interface to more than one computer. The branch highway, in common with any parallel highway, has limitations in length either for technical reasons or due to cost considerations. Furthermore, in many systems the high-speed transfer rate of which it is capable is unnecessary.

In these cases a serial-by-bit link may be suitable and could be inserted in the branch or treated as any other peripheral would be (Fig. 3).

Overall control of a total system, which may or may not include a computer, therefore presents a range of complexity according to the type of system, the multiplicity of tasks and the operational flexibility demanded. This range may be covered economically if executive action in parts of a complex system can be devolved onto local controllers. CAMAC, with its different levels of multiplexing capability, offers an opportunity to do this and a hierarchy of controllers can be envisaged. In each level of hierarchy there is scope to design controllers for restricted or general application tasks.

A very simple crate controller exists<sup>12</sup> and although it is hardly likely to find application in any system, it has proved most useful for testing individual crates and modules and also in training personnel to understand CAMAC. This controller has a rotary switch to select the addressed station, toggle switches to set the bits defining sub-address and function, and press-buttons to either step through the phases of a dataway operation or cycle through two dataway operations with independently selected functions but the same selected address. Write data can be set up on 24 switches and read data is indicated on 24 lamps.

At the other end of the range are controllers that directly interface CAMAC to a small computer. Each will be specific to a particular computer or family of computers and their complexity will depend on the facilities they provide, for example, whether they enable only programmed transfers or allow transfers involving direct access to the computer store. The most common requirement is for a controller which executes a dataway operation under computer program control. Typically, a CNAF code is put into the accumulator of the computer and transferred to the controller. The controller executes the operation, say read data from module, and the data are put into the accumulator and then into the computer's store as required. Program transfers of this kind are

relatively slow, because they employ several instructions to the c.p.u. of the computer. If greater speed is required, the controller has to transfer data between CAMAC and computer store autonomously, e.g. by cycle-stealing. With some recent computer designs it is possible to provide software-controlled switching between programmed transfers and autonomous transfers in the same CAMAC controller. Such a controller can only organize autonomous transfers from an address specified by program, or from a sequence of related addresses starting from a specific address. A more complicated controller is required for independent control of autonomous transfers in that the controller has to provide, additionally, the CNAF command to be established in the CAMAC crate and the address of the data location in the computer's memory. Such a controller<sup>13</sup> provides a very flexible and powerful facility and may itself be modular so that the capability of a basic controller can be added to with additional modules.

### 5. Programming with CAMAC Systems

In any application of CAMAC, the modules provide the required interaction with the peripheral devices, such as transducers/actuators, input/output machines and control/display facilities. The prime task of controllers is to enable data transfers to be made on the highways in an ordered fashion. All these operations must therefore either be instructed or supervised by a program. This program can be contained in a plug-board diode matrix or other 'hardware' routine device, or be resident within the memory of a computer and thereby classified as 'software'.

In a CAMAC system, no matter what the nature of the program, each program step will result in an executable code defining the crate(s) (C-code), the module(s) (N-code), the subsection(s) of the module (A-code), what function is to be carried out (F-code) and what type of information is to be transferred. Such a program can be generated by *ad-hoc* programming in machine code or in the assembly language of the computer used. It might also be generated by modification of existing device

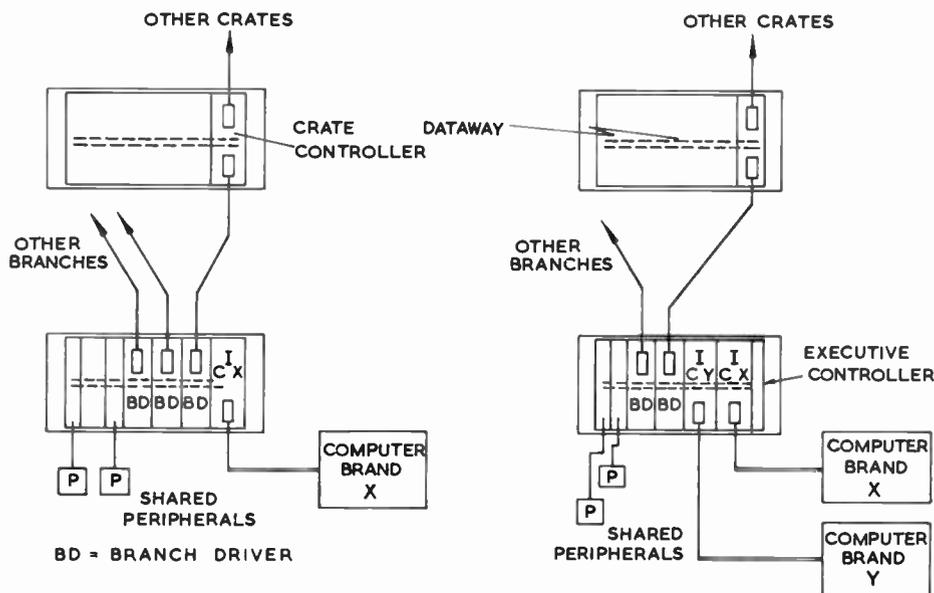


Fig. 10. Multi-crate, multi-computer systems.

handlers, provided with the computer, to adapt these to the CAMAC system's configuration and facilities.

These methods have been used with existing CAMAC systems and can be expensive in programming effort, and inefficient in the use of core-store space and computer time. The ESONE Committee therefore has a Working Group studying alternative and more efficient programming techniques for CAMAC systems. The detailed recommendations from this study, in terms of possibly a high level language, are not predictable as yet, however an intermediate stage is being actively pursued and recommendations are expected towards the end of 1971. This intermediate stage will define a syntax for CAMAC statements which can be pre-translated into an existing base language (e.g. *Fortran*, *Algol*, *Pearl*). Thereby a CAMAC user will be able to write a source program using symbolic names.

In general statements in the source program will contain three fields. The first is an *action* field which specifies what is to be done (e.g. READ/WRITE, TEST, STATUS), the second is an *external reference* field giving the locations of CAMAC devices and the third is an *internal reference* field pointing to the computer core locations or registers involved. The latter two fields may be multi-valued for the purposes of applying the same action to several CAMAC devices simultaneously and/or defining an array involving some sequence of operations.

An interesting possibility for a CAMAC system is that in which the controller has a processor section to deal with CAMAC transfers and working in a parallel-processing mode with the c.p.u. of the computer.<sup>11</sup> In this way an in-line program which is c.p.u.-controlled could branch into a parallel program, which is controlled by the CAMAC c.p.u., on receipt of a CAMAC demand. These parallel programs would be segments resident in core, identified by pointers and taken out of store without computer intervention. The in-line program would continue when a terminate signal from the parallel program segment is received in response to a call from the computer c.p.u. The object code of these segments would be in CAMAC-controller-code rather than computer machine-code and thereby the statements in the user source program would be independent of the instruction set of the computer.

### 6. A Selection of CAMAC Systems

It is not clear where the first true CAMAC system appeared; one seen at the Reaktor Centrum Nederland (Petten) in October 1969 must be among the first to have come into operation. The system controlled the positioning of nuclear fuel samples in a neutron irradiation facility and acquired data related to the determination of absorption cross-sections of fission products. The computer was a Honeywell DDP-516 and complete interfacing employed a range of CAMAC modules and a DDP-516/CAMAC interface controller providing program and direct access input/output transfers. Since then many systems have become operational in nuclear laboratories as far apart as Serperkov in Russia and Berkeley in U.S.A., with the largest concentration in Europe. An estimate of close on 100 systems employing ~£1M of CAMAC

equipment is not unrealistic at this time. The majority of these are in 'nuclear' laboratories where CAMAC is best known and understood. It is not exceptional, however, to find some of these systems divorced from the traditionally nuclear-electronic problems and used with n.m.r., X- and mass-spectrometers, chromatographs and X-diffractometers. CAMAC systems may also be found in research projects supporting astronomy, electric power generation and distribution, hospital patient monitoring, atmospheric pollution, industrial plant control, marine technology and computer-aided design.

Because of the rapid spread of CAMAC applications there is considerable difficulty in knowing all that is going on and, in any case, it is an unenviable task to select examples of typical systems. Therefore, no apology is made for taking examples close to the author's area of activity, being guided by topicality and/or the existence of documentation for further detailed reference.

## 7. Non-Computer Systems

### 7.1. Control of an X-Ray Fluorescence Goniometer

The X-ray intensity, at specified angles on an arc relative to the primary beam and sample orientation, has to be measured on a routine basis and the program repeated at each sample orientation. This total operation is programmed on a paper-tape fed to the teletypewriter (Fig. 11) and at each setting of the X-ray detector, pulses from the detector are counted within real-time or total-count boundary conditions which, when reached,

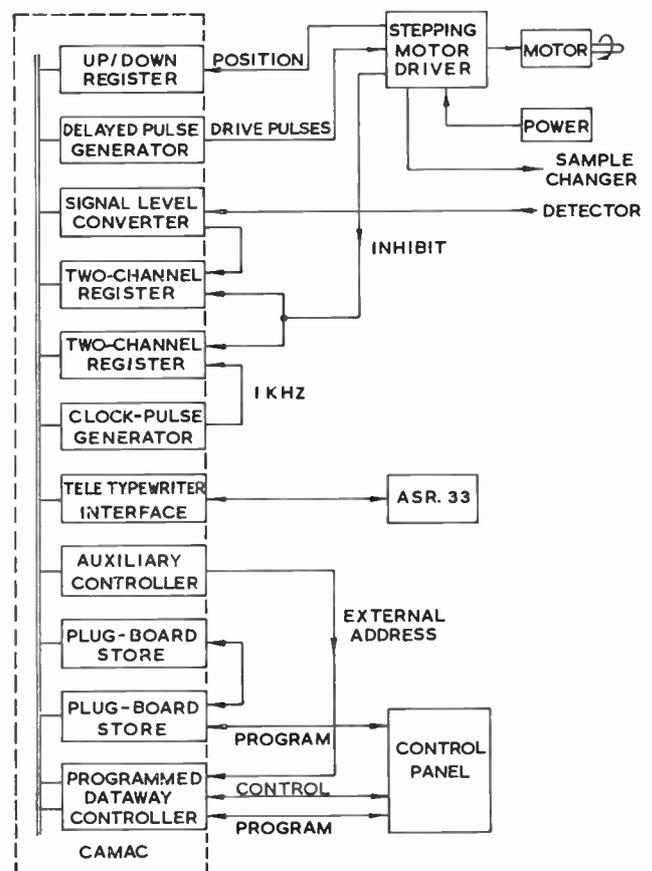


Fig. 11. Program control and data acquisition system.

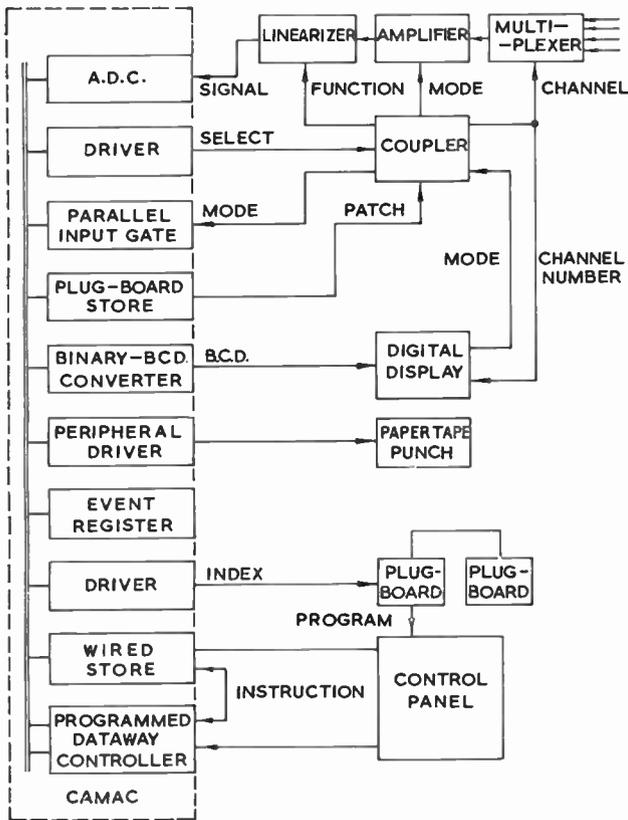


Fig. 12. Programmed data-logging system.

initiates a stepping motor to move the detector to the next specified angular position and so on. All data are punched out on the teletypewriter tape together with any identification or specification header included in the input tape. This particular system† is notable in that it is built up entirely from the library of standard CAMAC modules/controllers in the Harwell 7000 Series and therefore minimal system design, programming effort and time lapse to attain an operational system is required.

7.2. CAMAC Controlled Data Logging System

CAMAC enables an initial small non-computer system to be expanded as the requirements develop for increasing input/output channels, or as the data processing capability develops and justifies the use of a small on-line computer. An example of this is a non-computer data logging system<sup>14</sup> demonstrated at the Physics Exhibition 1971 and now marketed commercially. Figure 12 gives a schematic of the system. It employs a controller which can be programmed to perform all dataway operations, and such arithmetical manipulation of the data as required for tolerancing the data channel information. The logging program is set up on a diode matrix plug-board and channel data can be selected for decimal display from 256 data channels. Analogue signal channels are multiplexed by relay or m.o.s.t. switching and low impedance differential or high impedance non-inverting amplifiers are followed by a module giving a choice of two linearizing functions selected by program.

† Ward, L. D., Private communication.

A choice of a.d.c. is possible giving conversion times of 20 ms or 50 μs. A second plug-board store is used to select the operating conditions and limit tolerances relevant to each channel:

- upper and lower limits.
- amplifier gain and mode.
- linearizing function.

The output devices may be selected from a teletypewriter (printer and punch), magnetic tape, X-Y plotter, c.r.t. display, each interfaced by a suitable CAMAC module. To expand into a computer system, the programmed controller may be connected, via a module in CAMAC, to the teletypewriter input/output channel of the computer or the controller may be replaced by a specific computer interface controller, such as are available for DEC and Honeywell computers.

8. Computer-based Systems

8.1. The Analysis of Analogue Data on Magnetic Tape

The Marine Laboratory, Aberdeen, is engaged on a research project to evaluate the effect of acoustic noise produced by fishing gear on the efficiency of conventional fish-capture techniques. To do this, signals from a seabed hydrophone array are recorded on magnetic tape as a fishing gear equipped with a pulser (pinger) is towed over the array. The noise and pinger pulse signals are separated by filters, and then interfaced using the CAMAC system shown in Fig. 13, for spectral and correlation analysis by computer program.<sup>15</sup>

The analogue signals from one or more tape recorder tracks are digitized by a separate 8-bit a.d.c. for each track, so avoiding the need for multiplexing. Pinger pulse events are timed relative to a count of reference clock pulses held in a scaler. These data are accessed under program control to a Marconi-Elliott 920C computer,

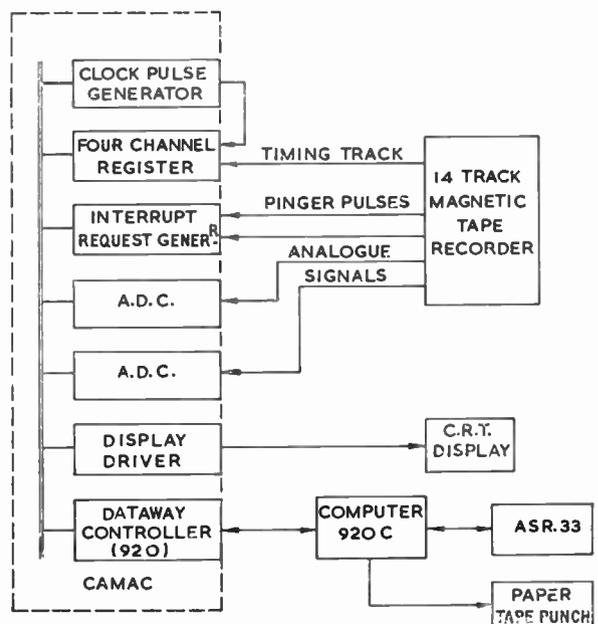


Fig. 13. System for acoustic noise evaluation of fishing gear.

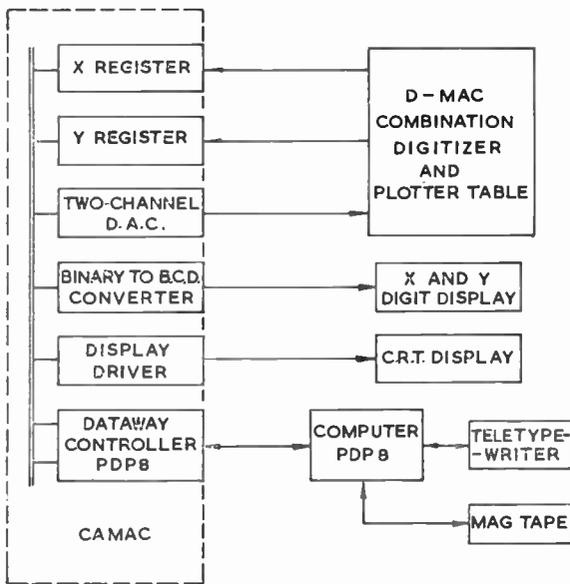


Fig. 14. Computer-aided design system.

via a controller designed and constructed by the Marine Laboratory. The computer program then carries out the analysis and necessary housekeeping on the data to produce power spectra, auto-or cross-correlation functions as required. The CAMAC system is again used for the display of these results. A real-time display is provided on an oscilloscope driven by a d.a.c. For permanent records, however, precise numerical results are output on a fast paper-tape punch.

For this application, it has been necessary to design an instrumentation system incorporating a considerable degree of flexibility to cater for a series of experiments with varying objectives, and consequently different requirements for data analysis and presentation. The use of a CAMAC compatible system has provided the necessary flexibility, while minimizing the engineering effort to implement hardware for specific experiments. The instrumentation system described here will be used for the analysis of data produced by the first of these experiments, which started in June 1971.

### 8.2. Computer-Aided Design

Within a wide variety of fields, such as architecture and building, civil engineering, pipe layout in industrial process plants, town planning and electronic circuit board design, there is a need for low cost automation design aids which combine immediate access to an up-to-date visual display of either the current state of overall design or the effects of amendments to localized features. At the Imperial College of Science and Technology, CAMAC equipment is being used to interface multiple and fully interactive, digitizing/draughting tables to a PDP-8E computer configuration. The system is known by the name CADMAC and is marketed commercially.<sup>16</sup> A schematic of the system is given in Fig. 14 and demonstrates the CAMAC advantage of being able to expand a single table system to a multiple table capability, as demanded and at marginal additional cost.

### 8.3. A Multi-Task System

Techniques previously used to measure radioactive fall-out resulting from nuclear weapon testing are being extended to the measurement of non-radioactive constituents of air, rain and dust samples collected during environmental studies.<sup>17</sup> The routine analysis of  $\gamma$ -ray spectra of thousands of such samples, after neutron-activation, demands a high degree of automation. Normally the accumulation of a statistically valid spectrum on a sample may take a very long time because the  $\gamma$ -ray intensity is very low. Several independent measurement systems, each with a wired-program multi-channel analyser, could increase the throughput of samples. However, the same situation can be obtained using a small computer with the advantages of greater flexibility and a computational ability. A system configuration (Fig. 15) used at Harwell provides up to eight channels of sample analysis, each with an appropriate a.d.c. and live-time controller. The a.d.c.s are interfaced to the CAMAC dataway by CAMAC-compatible modules (H.7000 Series) and multiplexed by a controller/interface into the DDP-516 computer. An X-Y plotter and a c.r.t. raster display are driven via the dataway, but other standard peripherals are directly interfaced to the computer although some of these could now be interfaced via CAMAC. The CAMAC controller allows programmed input/output transfers or block transfers via the direct memory access channel to the computer store. This latter facility allows the c.p.u. freedom to execute other programs during block transfers.

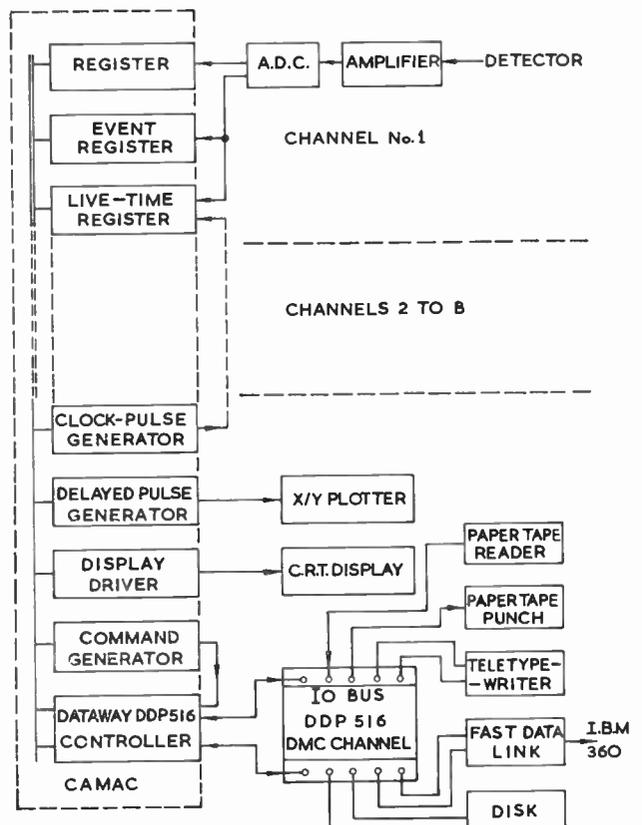


Fig. 15. A multi-task computer system.

The supervisory program allows up to eight data-partitions of variable size to be created and data can be accumulated, automatically controlled and transferred in all partitions simultaneously. The system therefore can be used for multiple, simultaneous tasks each in its own independent time-scale. Additionally, computational programs can be invoked by the supervisor program on request.

## 9. Conclusions

Many systems are implemented already with CAMAC-compatible equipment throughout Europe and North America. They use markedly different transducers/actuators and computers. The systems are front-ended with existing equipment, compatible with earlier modular unit schemes, and then interfaced for data processing purposes with CAMAC equipment. Within the foreseeable future, many of these 'front end' facilities will be found in CAMAC-compatible equipment, as medium-scale integration develops and as operational practices tend more and more towards computer control of system parameters and configuration.

Currently, also, the programming benefits offered by CAMAC are not being realized, but this situation will show a marked change now that equipment facilities are available on a massive scale and the incentive for software rationalization becomes more urgent.

If the full exploitation of the small digital computer is not to be impeded by effort and cash considerations, then a major rationalization of hardware and software is vital. CAMAC is already making its contribution to this rationalization process in nuclear laboratories, and elsewhere, by becoming a basic commodity in many diverse applications. Because it is non-proprietary, has international origins and is concerned primarily with the transfer of information, CAMAC may well provide the basis for an expanding rationalization process in a significant proportion of the applications of electronics to measurements and control.

## 10. Acknowledgments

This paper is the author's tribute to many colleagues, both in the United Kingdom, Europe and North America, without whose enthusiasm and professional ability CAMAC could not have materialized. In particular, the author wishes to thank Messrs. R. C. M. Barnes, J. M.

Richards, I. N. Hooton, A. Lewis, N. P. Whitehead and L. D. Ward, from whose writings liberal extractions have been made, and all those who have allowed their CAMAC systems to be described in this paper.

## 11. References

1. 'CAMAC—A Modular Instrumentation System for Data Handling', EURATOM Report EUR 4100e, March 1969.
2. 'CAMAC—Organisation of Multi-Crate Systems', EURATOM Report 4600e (Preliminary Issue), November 1970.
3. 'CAMAC—Amplitude Analogue Signals', EURATOM Report EUR 5100e (Preliminary Issue), March 1971.
4. 'CAMAC—a European standard for modular interface units between experiments and data processing equipment', *Nuclear Engineering International*, 14, pp. 345–47, April 1969.
5. Bisby, H., Barnes, R. C. M., Richards, J. M. and Lewis, A. 'The CAMAC Scheme', Harwell Report AERE-R6713, 1971.
6. 'Standards for nuclear electronic equipment', *The Radio and Electronic Engineer*, 37, No. 4, p. 212, April 1969.
7. CAMAC *Bulletin*, Issue no. 2, October 1971.
8. Diary of Developments, February 1971, pp. 6–7. Nuclear Enterprises Ltd., Beenham, Reading, Berkshire.
9. Ward, L. D., Mitchell, G. S. L. and Richards, J. M., 'A Programmed Controller in the CAMAC System', Harwell Report AERE-R6334, 1970.
10. Richards, J. M., 'Harwell 7000 Series CAMAC Controllers', Harwell Report AERE-R6723.
11. Hooton, I. N., Lewis, A. and Whitehead, N. P. 'Implementing CAMAC-Computer Systems', Harwell Report AERE-R6664, 1970.
12. Manual Dataway Controller Type 7024. List 7000, Issue 3 April 1970. Harwell.
13. Whitehead, N. P., 'CAMAC Dataway Control for the PDP-8 Computer Family', Harwell Report AERE-R6673, 1970.
14. Nuclear Enterprises Ltd., 'CAMAC Data Logging System'. Physics Exhibition Handbook, Exhibit B9, p. 145, 1971.
15. MacLennan, D. M. and Hawkins, A. D. 'The Investigation of Noise Produced by Fishing Gear', Spring Meeting, British Acoustical Society, April 1971.
16. D-MAC Ltd., 'CADMAC Computerized Graphics System'. Presentation at Imperial College of Science and Technology, March 1971.
17. Salmon, L. and Greevy, M. G., 'An On-Line Computer System for Instrumental Activation Analysis of Air, Water and Soil', Harwell Report AERE-R6524, 1970.

*Manuscript first received by the Institution on 21st April 1971 and in final form on 12th October 1971. (Paper No. 1420/IC 54)*

# A Technique for Standing Wave Measurements of Pulse-biased Semiconductor Devices

By

B. MARTIN, B.Sc.(Eng.)†

and

G. S. HOBSON, Ph.D., B.A.,  
A.Inst.P., C.Eng., M.I.E.E.†

A coherent sampling technique is described by which the v.s.w.r. of a pulse-biased semiconductor device mounted in a transmission line may be measured with a sensitivity comparable to that obtainable in c.w. measurements.

## 1. Introduction

The microwave properties of semiconductor devices are often determined by standard transmission line techniques involving a slotted line to monitor the standing wave ratio.<sup>1,2</sup> If the device is operated continuously the r.f. supply is usually amplitude modulated at a low frequency (1 to 3 kHz) and the signal demodulated by the slotted line detector is amplified by a narrow band a.c. amplifier (approximately 1% bandwidth is usual). Typically, such a system has a noise limitation of 0.1 to 1  $\mu$ V referred to the input.

Some devices (for example, Gunn diode oscillators) operate with a high thermal dissipation so that their temperature is appreciably above ambient and is also dependent on bias. This can be an undesirable effect for device characterization and can only be overcome by pulse biasing at low pulse repetition frequency and short pulse length. The technique to be described was designed to enable measurements to be made on a pulse-biased device with a sensitivity as good as that obtainable with c.w. operation. The equipment was more complicated and expensive than that used for c.w. measurements but the individual units will usually be found in a semiconductor microwave laboratory.

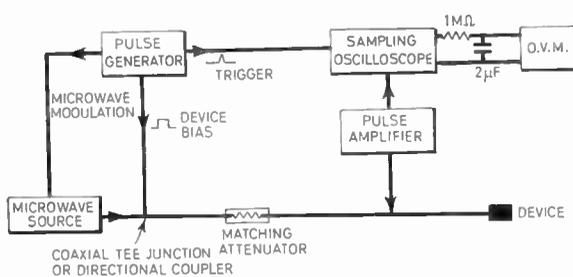


Fig. 1. The block diagram of the measuring equipment.

## 2. Experimental Techniques

A block diagram of the measuring equipment is shown in Fig. 1. The pulse generator provided the bias voltage for the device under investigation, modulated the microwave power source and provided a trigger pulse for the sampling oscilloscope. The biasing and modulating outputs were synchronized to the trigger pulse but were independently variable in height, length and delay. The sequence of pulses incident on the devices is shown in Fig. 2 together with the output waveform from the stand-

ing wave detector diode. The bias pulse width in our experiments was less than 1  $\mu$ s and the duty cycle was less than 10%. The microwave power source had a rise-time of greater than 1  $\mu$ s so it was switched on and allowed to achieve a steady output before the pulse bias

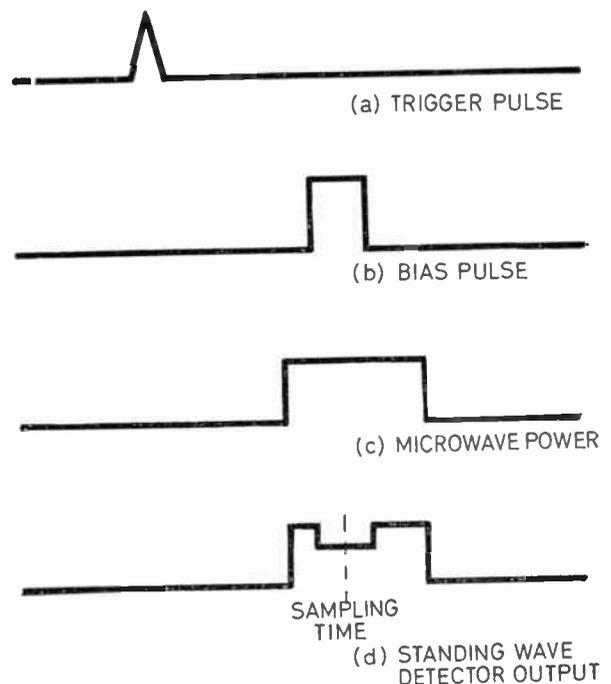


Fig. 2. The bias and microwave pulse synchronization and the output of the standing wave detector.

was applied to the device. A coaxial tee-junction or directional coupler was used to combine the microwave power and pulse bias. Commercially-available bias tee-junctions were not suitable because they degraded the bias pulse shape. The combined signals passed through an attenuator in order to match the input to the slotted line and were incident on the device mounted against a short circuit at the end of the coaxial line. Simple and obvious modifications would have to be made for measurements in a waveguide system.

The pulse amplifier was a simple video amplifier with a gain of up to 1000 and a rise-time of 0.1  $\mu$ s or better. It was a.c. coupled at its output so that no d.c. voltages were applied to the sampling oscilloscope (to avoid damage) and drift problems were avoided during measurement. Further comment on the low-frequency cut-off will be made later. The amplifier input impedance was 1 k $\Omega$ . This was low enough to give a rise-time of

† Department of Electronic and Electrical Engineering, The University of Sheffield, Mappin Street, Sheffield, S1 3JD.

better than 0.1 μs when fed by a coaxial cable with a capacitance of several tens of picofarads (typical of a short coaxial cable). It was also approximately equal to the diode output impedance so that there was not a large voltage attenuation. The gain of 1000 was necessary to amplify the smallest input signals (less than 1 μV) to a level where drift problems within the sampling oscilloscope were not troublesome.

The sampling oscilloscope was used on manual scan, set at a time such as that shown by the dashed line in Fig. 2(d), so that only the desired operating conditions of the device were monitored in the measurements. However, the noise voltage of the pulse amplifier was about 20 μV referred to its input so the desired sensitivity of less than 1 μV was not achieved. It was necessary to average many samples so that the ratio of coherent signal to noise was increased. This was done by feeding the vertical output of the sampling oscilloscope to a series RC integrator with a time-constant of 2 seconds. In our measurements it was possible to operate at the maximum sampling rate of the oscilloscope at about 50 kHz. In this way approximately 10<sup>5</sup> samples were averaged and the desired signal was unchanged but the random noise level averaged over *n* samples was reduced by a factor of √*n* (or approximately 300 in this case) in the usual way for random quantities<sup>3</sup>. The effective noise level referred to the pulse amplifier input was reduced from 20 μV to less than 0.1 μV as was confirmed experimentally.

The output voltage of the integrator was measured with a digital voltmeter and standing wave measurements were taken in the usual way.<sup>1,2</sup> The zero level was measured by switching the sampling time to the inter-pulse period when no r.f. or diode bias was present.

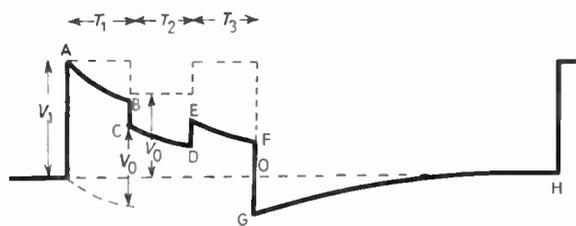


Fig. 3. Distortion of the detected pulse shape by the low frequency response of the pulse amplifier.

**3. Accuracy**

An exaggerated illustration of the pulse waveform received by the sampling oscilloscope is shown in Fig. 3 as the solid line and the dotted line is the ideal response. The desired voltage was measured at an unknown time between C and D and the zero level was measured somewhere between G and the next pulse. The difference of these voltages from the ideal due to the low-frequency cut-off of the pulse amplifier determine the maximum possible accuracy of the measuring system. This can be expressed as the maximum standing wave ratio that it is possible to measure reliably because in these conditions measurements are made when the voltage between C and D has its minimum value. A worst case estimate is now made of the error when the device conditions are such

that the voltages in the times between A and B, and E and F have values corresponding to a maximum of the standing wave pattern so that the voltage OG has its maximum value. It is further assumed that the decay of all voltages is described by a simple exponential with a time-constant, *T*, corresponding to the lower cut-off frequency of the pulse amplifier and all other level changes occur instantaneously. While the last assumption is open to question it allowed quantitative estimates to be made and these were checked experimentally. The apparent minimum voltage, *V*<sub>min</sub>, measured during *T*<sub>2</sub> will lie between the voltage difference at times D and H and that at times C and G.

$$\text{i.e. } V_0 + \left( \frac{V_1 T_3 + V_0 T_2}{T} \right) > V_{\min} > V_0 - \left( \frac{V_1 T_1 + V_0 T_2}{T} \right)$$

It has been assumed that *T*<sub>1</sub>, *T*<sub>2</sub>, *T*<sub>3</sub> ≪ *T*.

If *T*<sub>1</sub> = *T*<sub>3</sub>, as it usually did in these measurements, then

$$V_{\min} = V_0 \pm \left( \frac{V_1 T_1 + V_0 T_2}{T} \right) \approx V_0 \pm V_1 \frac{T_1}{T} \dots\dots(1)$$

if *T*<sub>1</sub> ≈ *T*<sub>2</sub> ≈ *T*<sub>3</sub> and *V*<sub>0</sub> ≪ *V*<sub>1</sub>.

In the worst case which is under consideration *V*<sub>1</sub> is the voltage corresponding to a maximum of the standing wave pattern so that the measured standing wave ratio *s*<sub>meas</sub>, is given by

$$\frac{1}{s_{\text{meas}}^2} = \frac{V_{\min}}{V_{\max}} = \frac{V_0}{V_1} \pm \frac{T_1}{T}$$

for a square-law detector,

$$\text{i.e. } \frac{1}{s_{\text{meas}}^2} = \frac{1}{s^2} \pm \frac{T_1}{T} \dots\dots(2)$$

where *s* is the correct standing wave ratio.

A realistic specification allowed only a 10% possible error in *s*<sub>meas</sub> for a standing wave ratio of 50 so that *T*<sub>1</sub>/*T* < 0.8 × 10<sup>-4</sup>.

In practice the time-constant *T* was set by the 50 Ω input impedance of the sampling oscilloscope (for low amplifier output impedance) and the capacitor in the output of the pulse amplifier. With *T*<sub>1</sub> = 10<sup>-6</sup> s the output capacitor had to be greater than approximately 200 μF. The maximum error in equation (2) was verified within a factor of 2 by connecting smaller capacitors in the pulse amplifier output and choosing the appropriate sampling times. This agreement was considered satisfactory owing to the finite rise-time of the microwave pulse which was comparable to *T*<sub>1</sub>, *T*<sub>2</sub> and *T*<sub>3</sub>. It is emphasized again that the error in equation (2) is the worst case and the accuracy will usually be better than this. Typical standing wave measurements taken with the present technique and with a conventional c.w. technique are compared in Table 1.

**Table 1. Standing wave ratio**

	Matched load		Short circuit	
	750 MHz	1 GHz	750 MHz	1 GHz
Pulse Technique	1.09	1.15	28.6	27.8
C.W. Measurement	1.08	1.12	29.1	27.6

**4. Conclusions**

A coherent sampling method of v.s.w.r. measurement for pulse-biased devices has been described. It has a sensitivity comparable to that of conventional c.w. systems and an accuracy of about 10% for a standing wave ratio of 50. This limit is also close to that set by drift in the sampling oscilloscope when incident microwave power levels between 100 μW and 1 mW were used.

**5. Acknowledgments**

This work was supported by a Ministry of Technology contract.

**6. References**

1. Barlow, H. M. and Cullen, A. L., 'Microwave Measurements' (Constable, London, 1966).
2. Ginzton, E. L., 'Microwave Measurements' (McGraw-Hill, New York, 1957).
3. Bell, D. A., 'Electrical Noise' (Van Nostrand, New York, 1960).

*Manuscript first received by the Institution on 5th November 1970, in revised form on 11th June 1971, and in final form on 20th August 1971. (Short Contribution No.149/CC 112.)*

© The Institution of Electronic and Radio Engineers, 1971

**STANDARD FREQUENCY TRANSMISSIONS—November 1971**

(Communication from the National Physical Laboratory)

Nov 1971	Deviation from nominal frequency in parts in 10 <sup>10</sup> (24-hour mean centred on 0300 UT)			Relative phase readings in microseconds N.P.L.—Station (Readings at 1500 UT)		Nov 1971	Deviation from nominal frequency in parts in 10 <sup>10</sup> (24-hour mean centred on 0300 UT)			Relative phase readings in microseconds N.P.L.—Station (Readings at 1500 UT)	
	GBR 16 kHz	MSF 60 kHz	Droitwich 200 kHz	*GBR 16 kHz	†MSF 60 kHz		GBR 16 kHz	MSF 60 kHz	Droitwich 200 kHz	*GBR 16 kHz	†MSF 60 kHz
1	-300.2	-0.1	-0.1	530	563.4	16	-300.1	0	-0.2	543	580.7
2	-300.1	-0.1	-0.2	531	567.5	17	-300.2	-0.1	-0.1	545	581.7
3	-300.0	0	-0.1	531	567.8	18	-300.0	0.1	-0.2	545	582.6
4	-299.9	-0.1	-0.2	530	572.1	19	-300.1	-0.1	-0.2	546	583.9
5	-300.1	0	-0.2	531	572.4	20	-300.2	-0.1	-0.1	547	584.9
6	-300.2	-0.1	-0.2	533	573.3	21	-300.2	-0.2	-0.2	549	586.6
7	-299.9	0	-0.2	532	573.5	22	-300.0	-0.2	-0.2	549	588.6
8	-300.2	-0.2	-0.2	534	572.4	23	-300.2	-0.1	-0.2	551	590.5
9	-300.1	-0.1	-0.2	535	574.5	24	-300.2	-0.2	-0.2	553	592.0
10	-300.2	0	-0.2	537	574.8	25	-300.1	-0.1	-0.1	554	593.1
11	-300.1	-0.2	-0.2	538	576.8	26	-300.0	-0.1	-0.2	554	594.2
12	-299.9	-0.1	-0.2	537	577.8	27	-300.1	-0.1	-0.2	555	595.4
13	-300.2	-0.1	-0.2	539	578.8	28	-300.1	-0.2	-0.2	556	596.9
14	-300.1	-0.1	-0.2	540	579.6	29	-300.2	-0.2	-0.2	558	598.5
15	-300.2	-0.1	-0.1	542	580.5	30	-300.2	-0.2	-0.2	560	600.4

\* Relative to UTC Scale; (UTC<sub>NPL</sub> - Station) = + 500 at 1500 UT 31st December 1968.

† Relative to AT Scale; (AT<sub>NPL</sub> - Station) = + 468.6 at 1500 UT 31st December 1968.

# Amplitude and Phase Modulated Pulse Trains for Radar

By

M. H. ACKROYD,  
B.Sc., Ph.D.†

*Based on a paper presented at a meeting of the Aerospace, Maritime and Military Systems Group in London on 16th December 1970*

The paper reviews the theory by which amplitude and phase modulated (a.m.ph.m.) pulse trains can be represented by complex number sequences and the means by which they can be generated and processed. The synthesis of pulse trains to have a desired autocorrelation function is described, some useful short Huffman sequences derived from Barker sequences are presented. Finally, the z-transform is applied to the study of combination codes formed by combining shorter sequences.

## 1. Introduction

Pulse compression is a well-known technique for increasing the range of a radar for a given peak transmitted power without loss in range resolution. Because the range resolution of a radar depends on the *bandwidth* of the transmitted signal it is possible to obtain good range resolution even with signals of long duration. All that is necessary is that the transmitted signal should be modulated in amplitude, phase and/or frequency to give it appropriate spectral properties and that the received echoes should be 'compressed' by proper processing at the receiver. The latter is usually done by a matched filter—a filter whose impulse response is a time-reversed version of the transmitted signal.

The most commonly used pulse compression signal is probably the 'linear f.m.' waveform,<sup>1</sup> where the transmitted pulse remains constant in amplitude but the instantaneous frequency increases or decreases linearly with time for the duration of the pulse. One reason for the ubiquity of the linear f.m. signal is the simplicity with which it can be generated and the ease with which matched filters can be constructed using dispersive delay devices.

The use of digital microcircuits continues to be more attractive in signal processing applications as their compactness and cheapness increase. An attraction of digital signal processing methods is their flexibility; the radar system designer is no longer restricted to the use of a few signals which can conveniently be generated.

One class of signals which is particularly suited to digital implementation is the class of amplitude and phase modulated (a.m.ph.m.) pulse trains. This paper outlines some methods by which systems using these signals can be implemented, with emphasis being placed on schemes using digital circuit devices. The problems of designing a.m.ph.m. pulse trains to have specified properties are approached. In particular, attention is paid to the problem of designing a pulse train to have a specified autocorrelation function.

The z-transform, familiar in the theory of sampled-data control systems, proves to play an important part in the study of a.m.ph.m. pulse trains. The later sections of the paper, devoted to problems of the design of a.m.ph.m. pulse trains, rely heavily on the use of the z-transform.

† Department of Electronic and Electrical Engineering, University of Technology, Loughborough, Leicestershire LE11 3TU.

## 2. Representation of A.M.Ph.M. Pulse Trains

A typical pulse train of the type considered in this paper is represented in Fig. 1. Such a signal consists of a finite number of contiguous pulses, each of which is a section of a sinusoid whose frequency,  $f_c$ , is fixed but whose amplitude and phase may be chosen arbitrarily. In the type of pulse train considered here, all the pulses are of equal duration,  $T$ .

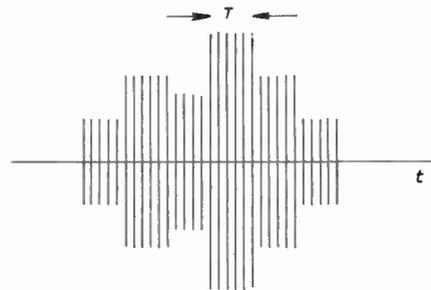


Fig. 1. A.m.ph.m. pulse train.

The pulse train can be represented as an amplitude and phase modulated waveform which can be expressed in the form

$$s(t) = e(t) \cos [2\pi f_c t + \phi(t)]. \quad \dots(1)$$

The amplitude or envelope modulation function,  $e(t)$ , and the phase modulation function,  $\phi(t)$ , may only change in jumps at the start and end of each individual pulse. For the duration of each pulse the envelope and phase modulating functions remain constant.

The theory of a.m.ph.m. pulse trains is simplified by the use of the complex envelope representation of bandpass signals.<sup>1,2</sup> Equation (1) which represents the waveform of the a.m.ph.m. pulse train can be rewritten

$$s(t) = \text{Re} [\sigma(t) \exp j2\pi f_c t] \quad \dots(2)$$

where  $\sigma(t)$  is the complex envelope of  $s(t)$  and is given by

$$\sigma(t) = e(t) \exp j\phi(t). \quad \dots(3)$$

Instead of performing operations such as convolution (linear filtering) on the bandpass waveform itself, the equivalent operation can be performed on the complex envelope. Simplification results because the carrier term  $\exp j2\pi f_c t$  no longer need be included at each stage. At the end of the calculations if it is required to have an

expression for the bandpass waveform it is merely necessary to multiply the complex envelope that is obtained by  $\exp j2\pi f_c t$  and then to take the real part of the result.

The complex envelope of an a.m.p.h.m. pulse train which contains  $M + 1$  pulses can be compactly expressed as the sum of  $M + 1$  rectangular functions each of which is of width  $T$ :

$$a(t) = a_0 \text{rect}(t/T) + a_1 \text{rect}(t/T - 1) + \dots + a_M \text{rect}(t/T - M) \dots(4)$$

Woodward's rect function is defined as shown in Fig. 2. The coefficient  $a_n$  which multiplies the rectangular function centred at time  $nT$  gives the value of the complex envelope over the range of time between  $(n - \frac{1}{2})T$  and  $(n + \frac{1}{2})T$ . The magnitude of the complex number  $a_n$  thus specifies the amplitude of the corresponding pulse while the angle of  $a_n$  specifies its phase.

The Fourier transform of equation (4) can be found by using the following standard rules of transform theory:

$$x(t - T) \leftrightarrow \exp(-j2\pi fT)X(f)$$

$$\text{rect}(t/T) \leftrightarrow T \text{sinc} fT.$$

The result is

$$\mathcal{S}(f) = [a_0 + a_1 \exp(-j2\pi fT) + \dots + a_M \exp(-j2\pi MfT)]T \text{sinc} fT \dots(5)$$

An important operation to which a.m.p.h.m. pulse trains are subjected is linear filtering. In particular, it is important to study the result of applying an a.m.p.h.m. pulse train to a filter whose impulse response is itself an a.m.p.h.m. pulse train with the same carrier frequency and the same individual pulse duration. If the amplitudes and phases of the  $N + 1$  pulses which constitute the impulse response are represented in order by the magnitudes and angles of a sequence of complex numbers  $(b_0, b_1, \dots, b_N)$  then the Fourier transform of the complex envelope of the impulse response is

$$\mathcal{H}(f) = [b_0 + b_1 \exp(-j2\pi fT) + \dots + b_N \exp(-j2\pi NfT)]T \text{sinc} fT \dots(6)$$

It is known that bandpass operations can be represented by corresponding operations performed on complex envelopes. In the case of linear filtering, a bandpass waveform  $s(t)$  may be applied to a bandpass filter with impulse response  $h(t)$  to produce an output waveform  $y(t)$ . Then  $y(t)$ , the complex envelope of the filter output has a Fourier transform which is the product of the Fourier transforms of the complex envelopes of the input waveform and the impulse response scaled by a factor of one half. In the present case,

$$\mathcal{Y}(f) = \frac{1}{2} [a_0 + \dots + a_M \exp(-j2\pi MfT)] \times [b_0 + \dots + b_N \exp(-j2\pi NfT)] \times T^2 (\text{sinc} fT)^2$$

$$= \frac{1}{2} [c_0 + \dots + c_{M+N} \exp(-j2\pi(M+N)fT)] \times T^2 (\text{sinc} fT)^2 \dots(7)$$

The inverse Fourier transform of equation (7) is

$$y(t) = \frac{T}{2} [c_0 \text{tri}(t/T) + c_1 \text{tri}(t/T - 1) + \dots + c_{M+N} \text{tri}(t/T - (M + N))] \dots(8)$$

The function  $\text{tri}(t/T)$  is defined in Fig. 2 and has the Fourier transform  $T \text{sinc}^2 fT$ . At instants of time  $0, T, \dots, (M + N)T$  the complex envelope of the filter output has values proportional to the corresponding coefficient in equation (8). At other times which lie between these instants the complex envelope of the filter output is given by linear interpolation between the values at adjacent integer multiples of  $T$ . The foregoing shows that the problem of evaluating the output waveform from a filter whose impulse response is an a.m.p.h.m. pulse train when its input is another such pulse train reduces to a problem in polynomial multiplication.

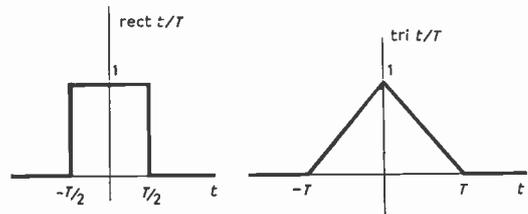


Fig. 2. rect and tri functions.

The sequence of numbers  $(c_0, \dots, c_{M+N})$  which form the coefficients of the product polynomial in  $\exp(-j2\pi fT)$  (equation (7)) specifies directly the values of the complex envelope of the output waveform at times which are integer multiples of  $T$ . The complex envelope at other times is given by linear interpolation between these values. The complex envelope in its turn specifies by its modulus and angle the envelope and phase modulation of the output signal from the filter.

A compact and convenient way of representing number sequences is by the use of the  $z$ -transform. The  $z$ -transform of a sequence  $(a_0, \dots, a_M)$  is a polynomial in powers of  $1/z$  which is given by

$$A(z) = a_0 + a_1 z^{-1} + \dots + a_M z^{-M} \dots(9)$$

To invert such a  $z$ -transform and recover the number sequence is simple; the number having index  $n$  is simply the coefficient of  $z^{-n}$  in the  $z$ -transform. For immediate purposes the use of the  $z$ -transform can be regarded as a notational convenience in which the symbol  $z$  is substituted for  $\exp(j2\pi fT)$  in expressions such as those occurring in equations (6) and (7).

In summary, the problem of calculating the output waveform from a filter whose impulse response is an a.m.p.h.m. pulse train when its input is also an a.m.p.h.m. pulse train with the same carrier frequency and pulse duration can be solved by the multiplication of two  $z$ -transforms. Two stages of abstraction are involved in representing an a.m.p.h.m. pulse train by a  $z$ -transform or by a sequence of complex numbers. In the

first stage of abstraction the real bandpass signal waveform is represented by a low-pass complex signal—its complex envelope. The second stage of abstraction is the representation of the complex envelope by a sequence of numbers representing its values at instants  $0, T, \dots$ , etc. or equivalently by the  $z$ -transform of the number sequence.

### 3. Generation of A.M.Ph.M. Pulse Trains

A potential attraction of a.m.ph.m. pulse trains is their suitability for generation and processing by digital means. This section and the following one illustrate how a.m.ph.m. pulse trains can be implemented. Digital means of implementation are emphasized here because with the development of large-scale integration techniques they can provide physically compact and relatively cheap systems. However it should not be forgotten that other means of implementation may be available and preferable in some circumstances. Surface-wave ultrasonic devices in particular may prove a good alternative to digital schemes. The arrangements presented here have been chosen primarily for illustration and so do not necessarily provide optimal engineering solutions.

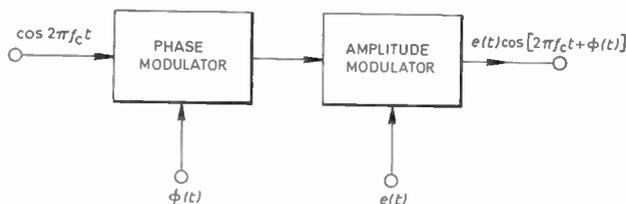


Fig. 3. Generation of pulse train by direct phase and amplitude modulation of carrier.

A.m.ph.m. pulse trains can be generated by direct amplitude and phase modulation of a carrier by the appropriate envelope and phase functions as evinced by equation (1) (Fig. 3). The envelope and phase functions, which remain constant for the duration of each pulse and change only in jumps, can be generated by the use of binary transversal filters as shown in Fig. 4(a). A binary transversal filter is a device which use a clocked shift register in place of the transversal filter delay line. A single binary 1 applied to the input of a binary transversal filter is clocked along the shift register and appears at each tap output in succession. The tappings are connected via weighting resistors to the inverting or the non-inverting inputs of a summing amplifier to produce a stepped waveform of the appropriate shape. The binary transversal filters which generate the envelope and phase functions can make use of a common shift register as shown in Fig. 4(b). Because the envelope is a non-negative function the inverting input to one summing amplifier is not needed.

A.m.ph.m. pulse trains can also be produced by direct generation of the in-phase and quadrature components of the required waveform. The right-hand side of equation (1) can be expanded to give

$$s(t) = e(t) \cos \phi(t) \cos 2\pi f_c t - e(t) \sin \phi(t) \sin 2\pi f_c t.$$

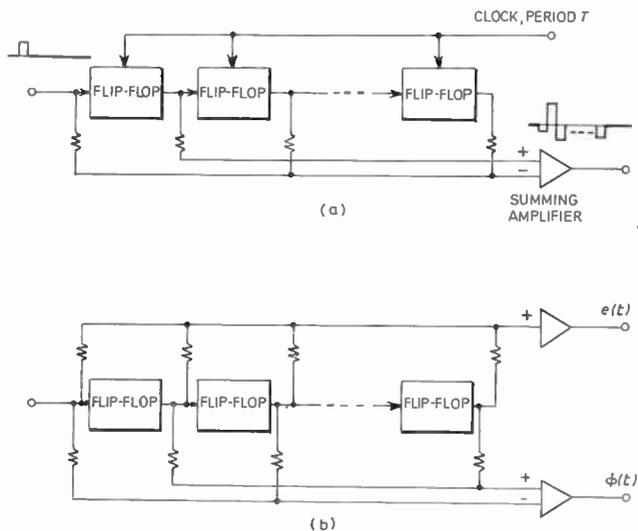


Fig. 4. Binary transversal filters.

The functions  $e(t) \cos \phi(t)$  and  $e(t) \sin \phi(t)$  can be recognized as the real and imaginary parts of the complex envelope. Each of these two functions consists of a sum of contiguous rectangular functions whose heights, in order, are given by the elements of the sequences  $(\text{Re } a_0, \dots, \text{Re } a_M)$  and  $(\text{Im } a_0, \dots, \text{Im } a_M)$  respectively. The waveforms  $e(t) \cos \phi(t)$  and  $e(t) \sin \phi(t)$  can thus be generated by a pair of binary transversal filters using a common shift register. Applied to sine-fed and cosine-fed balanced modulators these waveforms produce the required pulse train (Fig. 5).

Both of the foregoing schemes of generation require a shift register of  $M+1$  stages and, in general,  $2(M+1)$  weighting resistors in addition to two modulators. Certain special classes of a.m.ph.m. pulse train are used which are somewhat simpler to implement than the most general a.m.ph.m. pulse train. These special classes and the simplifications made possible by their use are listed below.

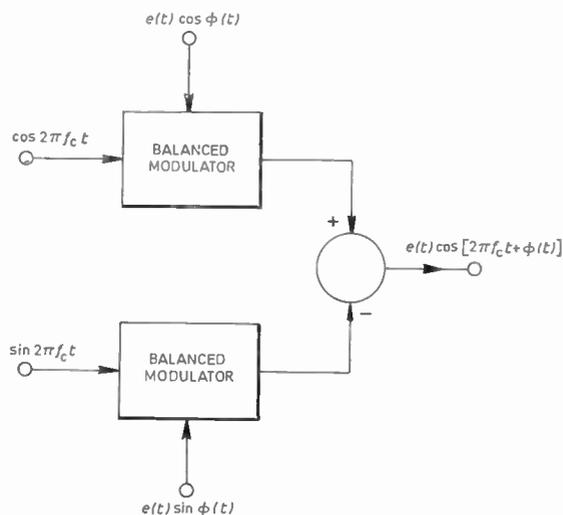


Fig. 5. Generation of pulse train by production of in-phase and quadrature components.

**Real Codes.** When the numbers  $(a_0, \dots, a_M)$  are all real but are otherwise arbitrary then the phase of each pulse is either 0 or  $\pi$ . In this case the sine-fed modulator of Fig. 5 is not needed. In the scheme of Fig. 3, the phase modulator need be simply a switching phase inverter.

**On/Off Codes.** The simplest a.m.ph.m. pulse trains to generate are the on/off codes. The numbers  $(a_0, \dots, a_M)$  have either some fixed positive value or else are zero. Thus all the non-zero pulses in the train are of equal amplitude and phase.

**Pure Ph.M. Codes.** With purely phase modulated pulse trains the amplitudes of all the pulses in a train are equal; the pulses differ only in phase. The complex numbers  $(a_0, \dots, a_M)$  are all of equal magnitude. The amplitude modulator in the scheme of Fig. 3 is not needed except that some device is needed to suppress the carrier before the start and after the end of the pulse train. Frank codes<sup>1</sup> and generalized Barker codes<sup>1</sup> are examples of pure ph.m. codes.

**Binary Codes.** Binary codes are the most widely used class of a.m.ph.m. pulse train. They present a special case of both pure ph.m. and real codes. The numbers  $(a_0, \dots, a_M)$  are all real and of equal magnitude; they differ only in sign. The phase modulation, being only by 0 or  $\pi$ , can be applied by a phase reversing switch.

On/off, pure ph.m. and binary codes have the practical advantage that the output device in the transmitter can operate under saturation conditions with consequent high efficiency.

**4. Matched Filtering of A.M.Ph.M. Pulse Trains**

Matched filter processing is very commonly used with radar signals. It provides the optimal form of processing with respect to a variety of criteria when Gaussian white noise from the receiver circuits is the dominant source of interference. Even in situations where matched filter processing is not optimal, for example when interference from clutter is significant, matched filter processing may provide a reasonable compromise between system performance and complexity. Again, matched filter processing may be used simply because information needed to design more nearly optimal schemes is unavailable.

A filter which is matched to a signal is one whose impulse response is a time reversed version of the signal, except that the amplitude scale may be different and an arbitrary shift along the time axis can be made. Thus a matched filter for a signal waveform  $x(t)$  has an impulse response  $h(t)$  which is given by

$$h(t) = kx(T_a - t) \dots\dots(10)$$

where  $k$  and  $T_a$  are arbitrary.

A matched filter for an a.m.ph.m. pulse train has an impulse response which is itself an a.m.ph.m. pulse train. If the coefficient  $k$  in equation (10) has the value unity, then the first pulse in the impulse response train will have the same amplitude as the last pulse in the signal pulse train. The second pulse in the impulse response train will have the same amplitude as the penultimate pulse in the signal train, and so on. The phase of the first pulse is the

impulse response train will be equal in magnitude but, as a result of the time reversal, will be opposite in sign to the phase of the last pulse in the signal train and so on.

A matched filter for an a.m.ph.m. pulse train can be implemented by the use of a tapped delay line as shown in Fig. 6. A filter which matches a single pulse of unit amplitude and of zero phase is followed by a tapped delay line. The taps of the delay line are connected to a summing network via phase shifters and attenuators. This system filters the pulse train directly as a bandpass signal. However it is often preferable to process signal at baseband particularly when digital realization is required.

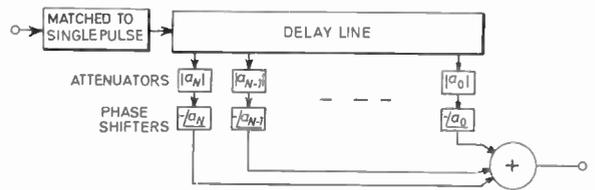


Fig. 6. Tapped delay line matched filter.

Bandpass filtering operations can be performed by first translating the signal to baseband frequencies, filtering its in-phase and quadrature components appropriately and finally translating back to carrier frequency. This process amounts to filtering the complex envelope of the signal. The physical realization of this operation is shown in Fig. 7. So far as input-output characteristics are concerned, this system is completely equivalent to the scheme shown in Fig. 6. (See, for example, reference 2.) The latter scheme uses four baseband filters. Two of these filters have impulse responses which are the real part of  $h(t)$ , the complex envelope of the impulse response of the prototype bandpass filter. The other two baseband filters have impulse responses which are the imaginary parts of the complex envelope. For an a.m.ph.m. pulse-train matched filter this scheme can be implemented as shown in Fig. 8. Although four baseband filters are required, each pair has a common input and so only two delay lines are needed for realization in transversal filter form. The addition and subtraction of the outputs of the individual baseband filters would in practice be done

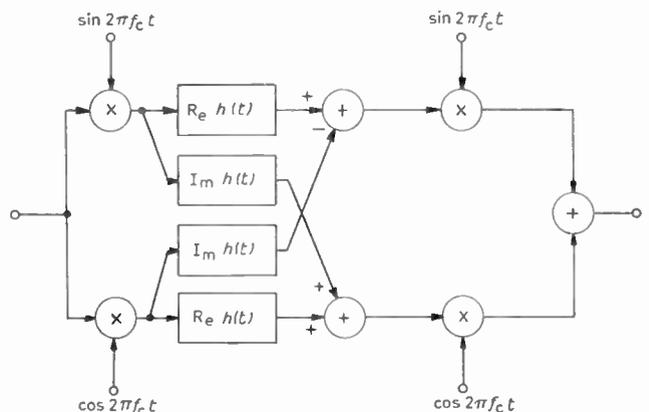
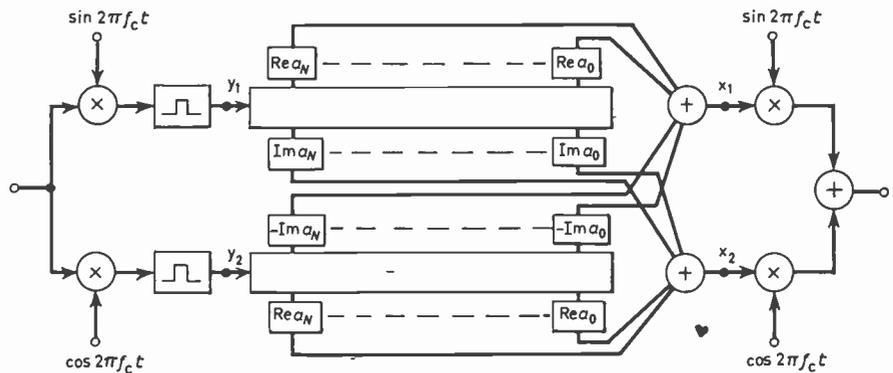


Fig. 7. Baseband matched filter.

Fig. 8. Realization of baseband matched filter.



by the summing amplifiers which are needed in any case by the transversal filters.

The output of a matched filter is normally applied to an envelope detector which may be either 'linear' or square law. If envelope detection is required there is no need for the pair of modulators at the right of Fig. 8; square-law envelope detection can be accomplished by discarding the modulators and connecting the circuit of Fig. 9 to the points marked  $x_1$  and  $x_2$  in Fig. 8.

The analogue delay lines required for the scheme of Fig. 8 are cumbersome and expensive, particularly when large delay-bandwidth products are required. The delay-bandwidth product needed is approximately equal to the number of pulses in the train and for a train of more than about forty pulses conventional artificial delay lines become impractical. The analogue delay lines can however be replaced by shift registers with consequent reduction in physical size and without limit to the total delay obtainable. To accomplish this, analogue-to-digital converters would be inserted at points  $y_1$  and  $y_2$  in Fig. 8 with their outputs feeding a number of binary transversal filters, one for each digit of the parallel-binary output of each converter. The binary transversal filter outputs would be added together with appropriate weighting as outlined by Voelcker.<sup>3</sup> The sampling period for the analogue-to-digital converters and the shift registers will depend upon the relative importance assigned to loss in resolution which occurs at low sampling rates and the cost of extra shift register stages which are needed to obtain a given delay at high sampling rates. In any case, the sampling period would not be less than the duration of the individual transmitted pulses. The number of binary digits required in the analogue-to-digital conver-

sion would depend greatly on individual circumstances such as the dynamic range required in resolving closely spaced targets. In some circumstances quantization to two levels only would be sufficient in which case the analogue-to-digital converter would take the form of a simple infinite clipper.

Some slight simplification is possible with some of the special classes of pulse train described in the previous section. With real codes the cross linking filters in Fig. 7 are not needed, which leads to a halving in the number of weighting resistors required in the scheme illustrated in Fig. 8. With binary codes all the weighting resistors are of equal value which may lead to economy in manufacture. The complexity of the matched filter system of Fig. 8 is not very significantly affected by the particular class of a.m.ph.m. pulse train used. At most the difference is in the number of resistors used and the number of resistor values required. This contrasts with the expense of providing for amplitude modulation in the generation of non-purely ph.m. pulse trains.

### 5. Desirable Properties of A.M.Ph.M. Pulse Trains

A property which is usually of prime importance in a radar signal is its ability to provide resolution between stationary targets of differing range and echoing strength. The autocorrelation function provides a measure of the ability of the signal to provide resolution between targets differing in range and strength. Loosely, when the magnitude of the autocorrelation function at a given time shift is low, good resolution is obtained between targets differing in range by a distance corresponding to that time delay. By suitable choice of amplitude and phase modulation an a.m.ph.m. pulse train can be made to have an autocorrelation function which consists of a central lobe resembling the autocorrelation function of a single strong pulse together with sidelobes of low amplitude. The level of the sidelobes sets a limit to the minimum echoing strength at which a target can be resolved when it is in the vicinity of a strong target of given strength.

A property which is desired of a radar signal is that for a given duration and peak amplitude it should carry as much energy as possible. For a given peak amplitude it is evident that the energy in a pulse train will be a maximum if all of the pulses have an amplitude which is equal to the peak. A measure of the effective energy of a

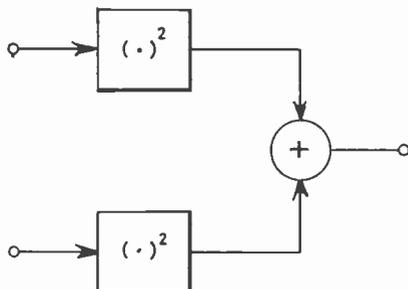


Fig. 9. Baseband square-law envelope detector.

pulse train is provided by the *energy efficiency*,  $\eta$ , which is defined as the ratio of the total energy of the pulse train to the energy that it would have if the amplitude of each pulse were made equal to the peak. The energy efficiency can be written in terms of the coefficients which define the complex envelope:

$$\eta = \frac{1}{N} \sum_{k=0}^N |a_k|^2 / \left( \max_k |a_k|^2 \right) \dots\dots(11)$$

It is evident that constant amplitude pulse trains such as binary and purely ph.m. pulse trains have an energy efficiency of unity. ▽

As mentioned in Section 3, binary codes, purely ph.m. codes and on/off codes have the advantage that, apart from some kind of on/off device, no amplitude modulator is required at the transmitter. Purely real codes, too, have an advantage in having phase modulation which takes only two values. These properties of a.m.ph.m. pulse trains namely of being either binary, purely ph.m., phase modulated by 0 and  $\pi$  only or being on/off lead to simplification in implementation and can be counted as being desirable.

### 6. Synthesis for Specified Autocorrelation Function

This Section presents a procedure by which a.m.ph.m. pulse trains can be designed so as to have a specified autocorrelation function. The autocorrelation function must of course be of a sort which can be realized by a.m.ph.m. pulse trains. The procedure was proposed by Huffman<sup>4</sup> in connexion with pulse trains and also, in different contexts, by Fejer<sup>5</sup> and by Wold.<sup>6</sup>

The output waveform from a matched filter in response to the signal that it matches has the same form as the autocorrelation function of the signal except that it may differ in scale and, with a physical filter, will be shifted along the time axis to the right by at least the duration of the signal. In Section 2 it was shown that the complex envelope of the output waveform from a filter whose impulse response is an a.m.ph.m. pulse train with an input which is another a.m.ph.m. pulse train could be calculated by the multiplication of z-transforms. It can be shown that the sequence which defines the complex envelope of the matched filter impulse response is  $(a_N^*, \dots, a_0^*)$  when the sequence which specifies the signal is  $(a_0, \dots, a_N)$ . From this, the complex envelope of the matched filter response is specified by the sequence  $(r_{-N}, \dots, r_0, \dots, r_N)$  whose z-transform is given by

$$\begin{aligned} R(z) &= r_{-N} + \dots + r_0 z^{-N} + \dots + r_N z^{-2N} \\ &= (a_N^* + \dots + a_0^* z^{-N})(a_0 + \dots + a_N z^{-N}) \\ &= z^{-N}(a_0^* + \dots + a_N^* z^N)(a_0 + \dots + a_N z^{-N}) \\ &= z^{-N} A^*(1/z^*) A(z). \dots\dots(12) \end{aligned}$$

The coefficients of  $R(z)$  are labelled with index values running from  $-N$  to  $N$  because, except for the inclusion of a factor  $z^{-N}$ ,  $R(z)$  is the z-transform of the number sequence which specifies the complex envelope of the autocorrelation function of the pulse train. The  $n$ th coefficient,  $r_n$ , gives the value of the complex envelope of the autocorrelation function at time shift  $nT$ .

The properties of  $R(z)$  are revealed by a study of its zeros in the complex z-plane. This study starts by considering the zeros of  $A(z)$ .  $A(z)$  is a polynomial of  $N$ th order in powers of  $z^{-1}$  and can be written in factored form as

$$A(z) = a_0(1 - z_1/z)(1 - z_2/z) \dots (1 - z_N/z) \dots\dots(13)$$

where  $z_1, \dots, z_N$  are the zeros of  $A(z)$ , i.e. the values of  $z$  for which  $A(z)$  is zero.  $A^*(1/z^*)$ , which appears in equation (11), can be written in factored form, too:

$$A^*(1/z^*) = a_0^*(1 - z/z_1^*)(1 - z/z_2^*) \dots (1 - z/z_N^*). \dots\dots(13)$$

An important conclusion is that if some complex number  $z_i$  is a zero of  $A(z)$  then the reciprocal of its conjugate,  $1/z_i^*$ , is a zero of  $A^*(1/z^*)$ .  $z_i$  and  $1/z_i^*$  have the same angle in the z-plane but their magnitudes are reciprocal quantities (Fig. 10). The z-transform  $R(z)$  has  $A(z)$  and  $A^*(1/z^*)$  as its factors and so has for its zeros both the zeros of  $A(z)$  and those of  $A^*(1/z^*)$ . The conclusion to be drawn is that the zeros of  $R(z)$  must all occur in reciprocal-conjugate pairs.

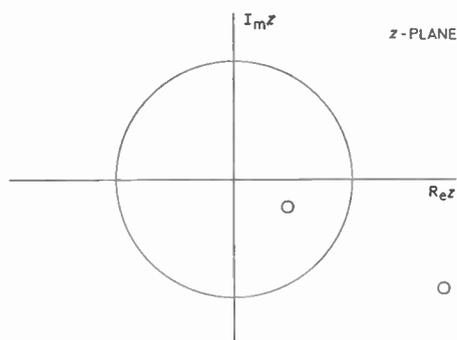


Fig. 10. Reciprocal-conjugate zeros.

From the foregoing it is inferred that a given bandpass waveform will be the autocorrelation function of an a.m.ph.m. pulse train of finite length provided that the following conditions are satisfied:

- (1) The values of the complex envelope of the tentative autocorrelation function measured at time shifts  $nT$ , where  $n = 0, \pm 1, \dots$ , should be zero for  $|n|$  greater than some value  $N$ .
- (2) The polynomial in powers of  $z^{-1}$ ,  $r_{-N}z^{-N} + \dots + r_0 z^{-N} + \dots + r_N z^{-2N}$ , should have zeros which occur only in reciprocal conjugate pairs. The coefficients of this polynomial are the sample values of the complex envelope of the autocorrelation function at times which are integer multiples of  $T$ .
- (3) The values of the autocorrelation function complex envelope for time shifts other than integer multiples of  $T$  should be given by straight line interpolation between the values at such instants.

If the tentative autocorrelation function satisfies these conditions, then at least one, and usually a whole family, of a.m.ph.m. pulse trains can be found which possess the specified autocorrelation function. The following paragraphs present a procedure for the synthesis of

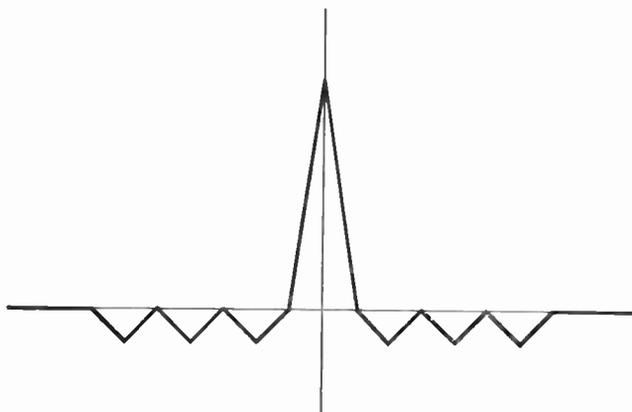


Fig. 11. Complex envelope of 7-element Barker sequence autocorrelation.

a.m.ph.m. pulse trains to have a specified autocorrelation function. The procedure is illustrated by taking as an example the autocorrelation function of a 7-element Barker sequence. Figure 11 illustrates the complex envelope of the autocorrelation function which in this case is purely real.

The first stage in the procedure is to take the values of the complex envelope of the autocorrelation function at integer multiples of  $T$  to obtain the coefficients of  $R(z)$  and then to factorize this polynomial to obtain its zeros. For the case of the autocorrelation of a 7-element Barker code  $R(z)$  is given by

$$R(z) = -1 - z^{-2} - z^{-4} + 7z^{-6} - z^{-8} - z^{-10} - z^{-12}.$$

In general a digital computer and a polynomial factoring program are necessary to effect the factorization although there is an important exception which is dealt with in the next section. Figure 12 shows the locations of the zeros of  $R(z)$  for the example. In this case, as a consequence of all the coefficients of  $R(z)$  being real numbers, the zeros which are complex occur in conjugate pairs. All the zeros occur in reciprocal conjugate pairs as they must since the autocorrelation function is indeed realizable by at least one a.m.ph.m. pulse train, namely a Barker sequence.

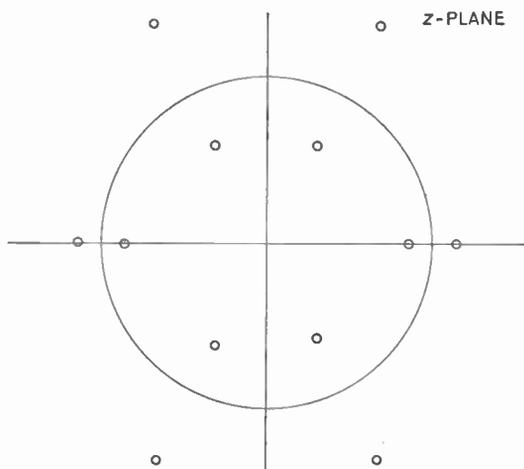


Fig. 12. Zeros of  $R(z)$  for 7-element Barker sequence.

The second stage of the synthesis procedure involves the selection of one zero from each of the reciprocal conjugate pairs. The chosen set of zeros provides a zero pattern which is assigned to  $A(z)$ ; those that are not chosen are automatically the zeros of  $A^*(1/z^*)$ . If all the zeros of  $R(z)$  are distinct then, as there are  $N$  reciprocal conjugate pairs there will be  $2^N$  different ways in which the zero pattern for  $A(z)$  could be chosen.

The final stage of the synthesis is to multiply out the right-hand side of equation (13) with the chosen zeros to obtain the coefficients of  $A(z)$ . This is a task which is best performed by a digital computer. Reference 8 presents a computational procedure which gives accurate results even when the order of  $A(z)$  is high. Figure 13 shows the complex envelopes of the pulse trains which result from four of the sixty-four possible zero patterns which could have been chosen.

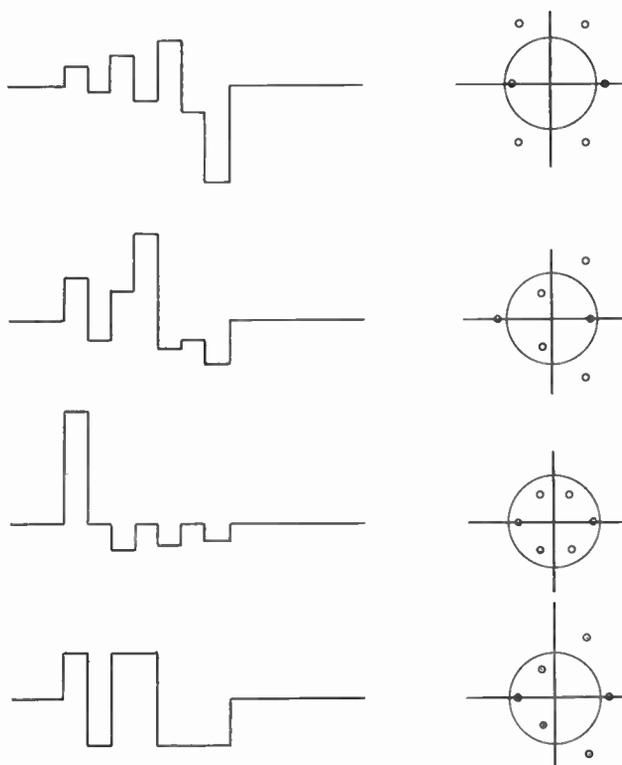


Fig. 13. Complex envelopes and zero patterns of some pulse trains with the same autocorrelation as 7-element Barker sequence.

The example illustrates several points of interest.

- (i) The complex envelopes of the synthesized pulse trains are purely real. This is a result of the complex zeros in each pattern having been selected to occur in complex conjugate pairs. Only when all the complex zeros of  $A(z)$  occur in conjugate pairs are its coefficients all real.
- (ii) When all the zeros of  $A(z)$  lie within the unit circle in the  $z$ -plane the amplitudes of the first few pulses in the train exceed those of the last few; the energy is concentrated at the front of the pulse train. Conversely, when all the zeros of  $A(z)$  lie outside the unit circle, the energy is concentrated to the rear. When

half of the zeros lie inside and half lie outside the unit circle, the energy is distributed equally between the front and rear halves of the train. This effect is discussed quantitatively in reference 9.

- (iii) In the bottom example the chosen zero pattern 'repeats', i.e. in the sector of the z-plane lying between  $\pi$  and  $2\pi$  radians the zero pattern is a repetition of the part of the zero pattern lying between angles of 0 and  $\pi$  radians. In these cases, every second coefficient of  $A(z)$  is zero. Such repetitive zero patterns occur in the design of 'combination sequences' discussed in Section 8.

This synthesis procedure presents difficulties of a computational nature. One difficulty in the procedure is the task of factoring  $R(z)$  to obtain its zeros. The factorization of high-order polynomials is a notoriously difficult computational operation. The author has found that the *Fortran* routine NEWRA,<sup>7</sup> used on an ICL 1904A computer in double precision arithmetic, works reliably with orders of  $R(z)$  up to about 50.

A second difficulty in the synthesis procedure is the choosing of a zero pattern. While any choice of zero pattern will yield a pulse train with the required autocorrelation function, some choices will result in a pulse train with a much higher energy efficiency than will others. There is no known way, short of evaluating all possible zero patterns, to find the pulse train which has the maximum possible energy efficiency for a particular autocorrelation function. For pulse trains of even moderate length the total number of zero patterns which may be as large as  $2^N$  becomes too large to permit evaluation of every possibility. It might well be assumed that pulse trains of high energy efficiency will have their energy distributed equally between their first and second halves. Even if the search for zero patterns giving high energy efficiency pulse trains is then confined to patterns in which half of the zeros lie within and half lie outside the unit circle there are still a great many possibilities.

These difficulties are not insuperable. The factorization problem is avoided in the case of Huffman codes, which are dealt with in the next Section. The problem of choosing a zero pattern may be sidestepped by simply selecting the best pulse train found from a few randomly chosen zero patterns instead of searching for the best possible.

### 7. Huffman Codes

D. A. Huffman suggested that the form of autocorrelation function shown in Fig. 14 should be used because it is as nearly as is possible the ideal autocorrelation function, namely, that of a single strong pulse. The autocorrelation function of an a.m.ph.m. pulse train of length  $(N+1)T$  is of necessity non-zero for shifts in the region of  $NT$  but otherwise this form of autocorrelation function is zero everywhere away from the central lobe.

This choice of autocorrelation function has the additional advantage that there is no difficulty in factoring the autocorrelation z-transform which is a quadratic in the variable  $z^{-N}$ :

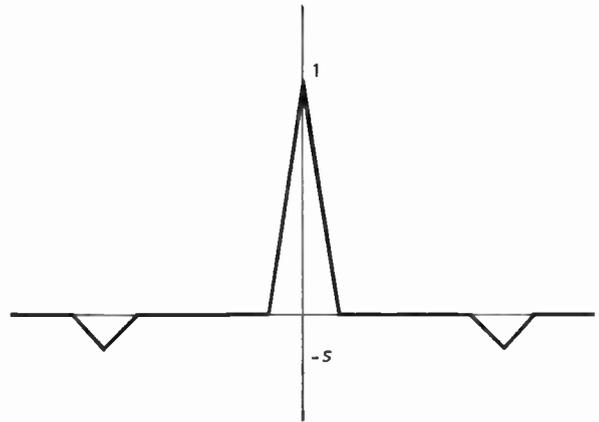


Fig. 14. Huffman sequence autocorrelation.

$$R(z) = -s + z^{-N} - sz^{-2N} = -s(X^{-N} - z^{-N})(X^N - z^{-N}) \dots\dots(14)$$

where  $X$  is given by

$$X^N = \frac{1}{2s} + \left(\frac{1}{4s^2} - 1\right)^{\frac{1}{2}} \dots\dots(15)$$

The zeros of  $R(z)$  are thus given by the  $N$  roots of  $X^N$  and  $X^{-N}$ . They lie in the z-plane at regular angular intervals of  $2\pi/N$  radians and on two circles, one of radius  $X$  and the other of radius  $1/X$  as shown in Fig. 15.

The design of a Huffman code consists of the choice of three things. These are the code length,  $N+1$ ; the circle diameter,  $X$ ; and the zero pattern. Unfortunately there is no known way short of trial-and-error for making these choices so as to produce the Huffman sequence with the maximum energy efficiency for a given length. †

The tolerable autocorrelation peak-to-sidelobe ratio, sets a limit to the minimum value of circle radius that can be used. From equation (15) the peak-to-sidelobe ratio is given by

$$s = X^{-N}/(1 + X^{-2N}).$$

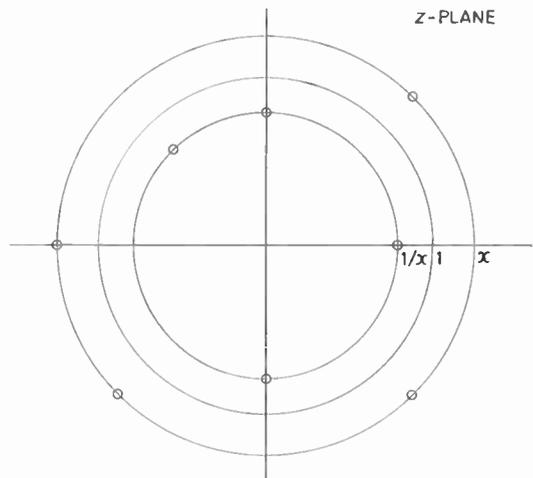


Fig. 15. Zeros of z-transform of typical 8-element Huffman sequence.

† Note added in proof: Since this paper was written a partial solution to these problems has been found.

A general trial-and-error procedure for designing a Huffman code of a given length is to choose a zero pattern, perhaps by random selection, and then to compute the sequences which result for a succession of values of  $X$ . If no value of  $X$  in the range set by the limit on permissible peak-to-sidelobe ratio yields an adequate energy efficiency, then a new zero pattern is chosen and the process is repeated. If a real code is required, the

One scheme for designing Huffman codes which has yielded some short codes with good energy efficiency is the use of a 'good' but non-Huffman code to suggest a zero pattern for the Huffman code. A 'good' code is simply one whose autocorrelation function consists of a large central spike and low sidelobes and which has a high energy efficiency. Such codes as the Barker codes are thus considered as being good codes.

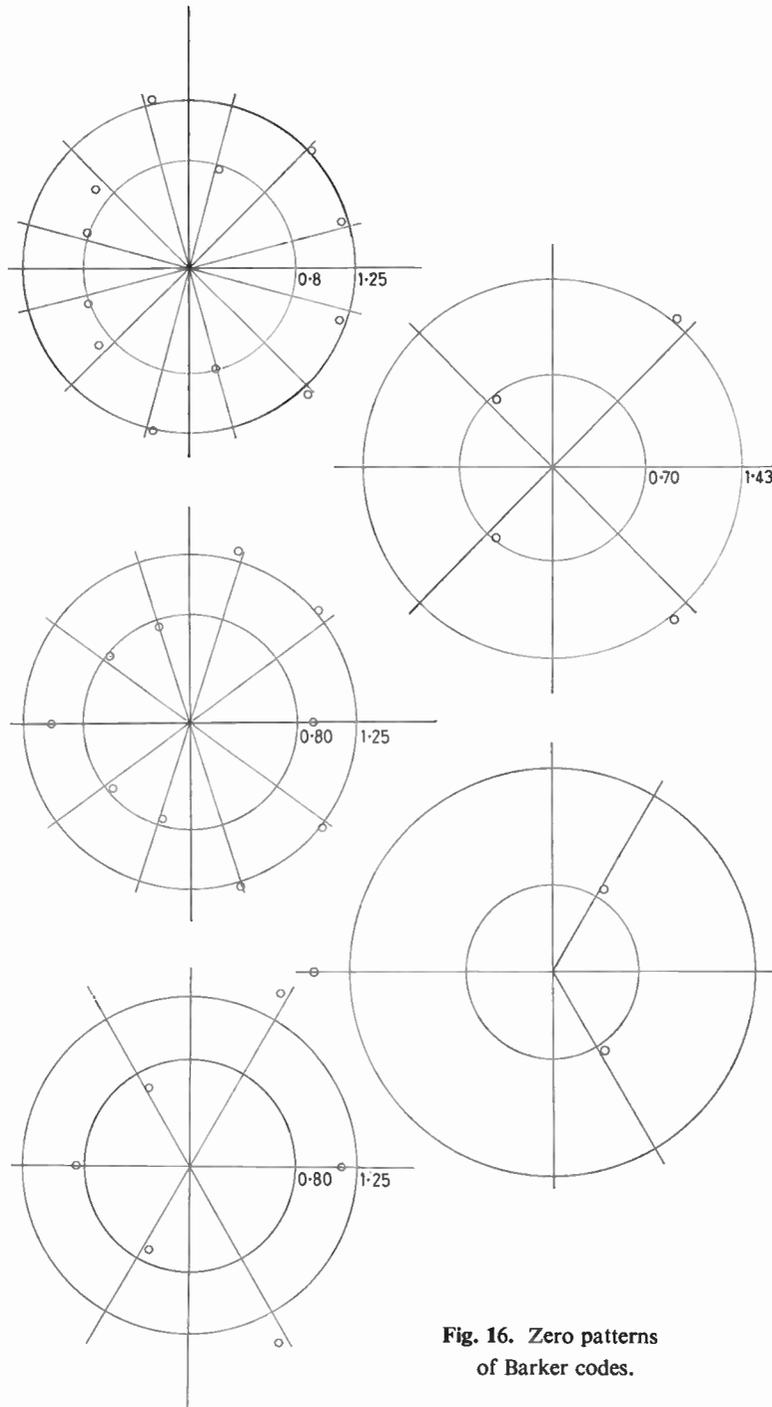


Fig. 16. Zero patterns of Barker codes.

zero pattern must be chosen so that all complex zeros occur in conjugate pairs. Experience shows that this procedure suffers from the disadvantage that randomly chosen zero patterns seldom yield codes with good energy efficiency.

Figure 16 shows the zero patterns which result from factorization of the  $z$ -transforms associated with the known Barker codes. These zero patterns are similar to Huffman code zero patterns in that only small changes in the zero positions are required to transform them into

SEQUENCE LENGTH	COMPLEX ENVELOPES		SEQUENCE	ENERGY EFFICIENCY	HUFFMAN SIDELOBE LEVEL	BARKER SIDELOBE LEVEL
	HUFFMAN	BARKER				
13			1.000 1.093 0.597 0.791 0.686 -0.905 -1.086 0.905 0.686 -0.791 0.597 -1.093 1.000	0.65	0.098	0.077
			0.353 0.336 0.161 -1.000 -0.988 -0.922 0.988 -1.000 -0.160 0.336 -0.353			
7			0.500 1.000 1.000 0.000 -1.000 1.000 -0.500	0.64	0.0556	0.143
			1.000 1.000 0.500 -1.000 1.000			
4			1.000 1.000 -0.618 0.618	0.69	0.224	0.250

Fig. 17. Complex envelopes of Huffman sequences produced by modifying Barker sequences.

true Huffman zero patterns. It might be hoped that modifying the zero pattern to a Huffman type pattern would cause little reduction in the energy efficiency of the resulting code but would eliminate the autocorrelation function sidelobes except for the inevitable pair of non-zero sidelobes.

The results of this scheme are displayed in Fig. 17 which displays the complex envelopes of the codes (which are all real), the autocorrelation sidelobe levels and their energy efficiencies. For comparison the complex envelopes and sidelobe levels of the 'parent' Barker sequences are also shown. These examples show that, contrary to a common belief, it is perfectly possible for real Huffman codes to have high energy efficiencies.

It has not yet proved possible to apply this procedure to larger codes. Figure 18 shows the zero pattern of a 49-element Frank code. It is evidently not possible to modify this zero pattern into a Huffman-type zero pattern by making only *small* changes in zero locations. A small change is one which is small by comparison with the distances between adjacent zeros.

Despite the difficulty of design, Huffman codes are a potentially useful class of a.m.p.h.m. pulse trains, particularly for applications where very low autocorrelation sidelobe levels are important.

### 8. Combination Codes

When very long a.m.p.h.m. pulse trains are required, with hundreds or thousands of pulses, two difficulties arise. One is the difficulty of designing very long codes in the face of the computational difficulties mentioned earlier. The other is the expense and complexity of the equipment needed to generate and to process very long pulse trains. This section shows how both of these difficulties can be overcome to some extent by a process of combining two good short codes to produce a longer code. Long codes produced in this way are termed in this paper *combination codes*.

The special case of binary codes produced from Barker sequences has been studied by Hollis.<sup>10</sup> Compared with other methods (c.f. references 10 and 11) the z-transform approach leads in a very direct way to the autocorrelation properties of general, not necessarily binary, combination codes.

A combination code is produced from two codes ( $a_0, a_1, \dots, a_M$ ) the *inner* code, and ( $b_0, b_1, \dots, b_N$ ) the *outer* code by the following process. The outer code is first 'stretched' by a factor  $M$  to produce a new code ( $b_0, 0, \dots, 0, b_1, 0, \dots, b_N$ ) in which every  $M$ th element is taken from the outer code and in which all other elements are zero. The stretched code is then convolved with the inner code to produce the combination code which is ( $a_0b_0, a_1b_0, \dots, a_Mb_0, a_0b_1, a_1b_1, \dots, a_Nb_1, \dots, a_0b_M, a_1b_M, \dots, a_Mb_N$ ). The combination code thus contains  $(M+1)(N+1)$  elements.

A pulse train specified by a combination code can be visualized as being produced by taking the pulse train specified by the outer code and replacing each individual pulse by a complete a.m.p.h.m. pulse train specified by the inner code. The result is a train of pulse trains.

The z-transform of the combination code,  $C(z)$ , is simply related to the z-transforms of the inner and the outer codes. The z-transform is given by

$$C(z) = b_0 A(z) + b_1 z^{-M} A(z) + \dots + b_N z^{-MN} A(z) = A(z)B(z^M), \dots(16)$$

$B(z^M)$ , which is nothing but  $B(z)$  with  $z$  replaced by  $z^M$ , is the z-transform of the stretched outer code. Equation

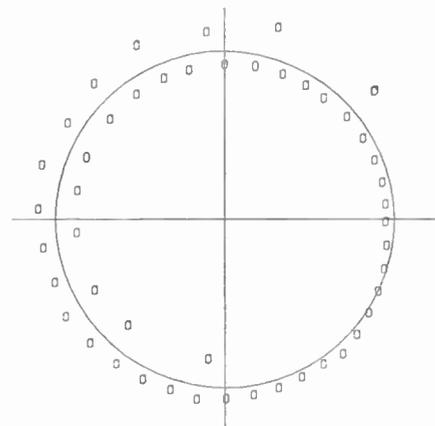


Fig. 18. Zero pattern of 49-element Frank code.

(16) shows that the zero pattern of  $C(z)$  bears a simple relation to the zero patterns of  $A(z)$  and  $B(z)$ . It is in fact the superposition of the zero patterns of  $A(z)$  and  $B(z^M)$ . The zero pattern of  $B(z^M)$  repeats every  $2\pi/M$  radians around the  $z$ -plane.

The  $z$ -transform of the autocorrelation sequence of the combination code,  $R_C(z)$ , is given by

$$R_C(z) = C(z)C^*(1/z^*) = A(z)B(z^M)A^*(1/z^*)B^*(1/z^{*M}). \dots\dots(17)$$

The  $z$ -transforms of the autocorrelation sequences of the inner and outer codes,  $R_A(z)$  and  $R_B(z)$ , are given by

$$R_A(z) = A(z)A^*(1/z^*)$$

$$R_B(z) = B(z)B^*(1/z^*).$$

Substituting these expressions in equation (17) gives the  $z$ -transform of the autocorrelation of the combination code in terms of the  $z$ -transforms of the autocorrelation of the inner and outer codes;

$$R_C(z) = R_A(z)R_B(z^M). \dots\dots(18)$$

Equation (18) shows that the autocorrelation sequence of the combination code can be calculated by stretching the autocorrelation of the outer code by a factor of  $M$  and convolving with the autocorrelation of the inner sequence. If, for example, the outer code is a Huffman code of length four and the inner code is a Barker code of length five, the combination code will have an autocorrelation sequence as is shown in Fig. 19.

It can be seen from the above process of computing the autocorrelation that if the sidelobes at the inner codes autocorrelation are of the order of magnitude  $S_A$  and the outer codes are of order of magnitude  $S_B$ , then the orders of magnitude of the sidelobes of the combination sequence autocorrelation will be as is described by Fig. 20. The central lobes are assumed to have unit magnitude here.

Combination codes can be designed by taking any two codes of suitably low sidelobe level and of appropriate length and combining them in the way described. It is worth noting that a combination code produced from two binary codes will itself be a binary code. Again, combination codes produced from purely real or purely ph.m. codes will themselves be purely real or purely ph.m., as the case may be.

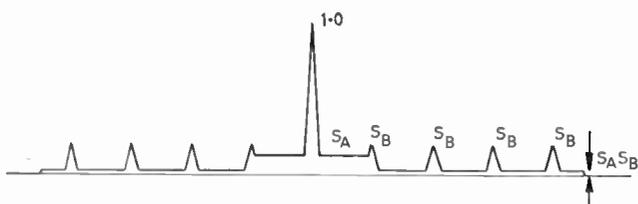


Fig. 20. Order of magnitudes of autocorrelation sidelobe levels of combination codes in terms of sidelobe levels of inner and outer codes.

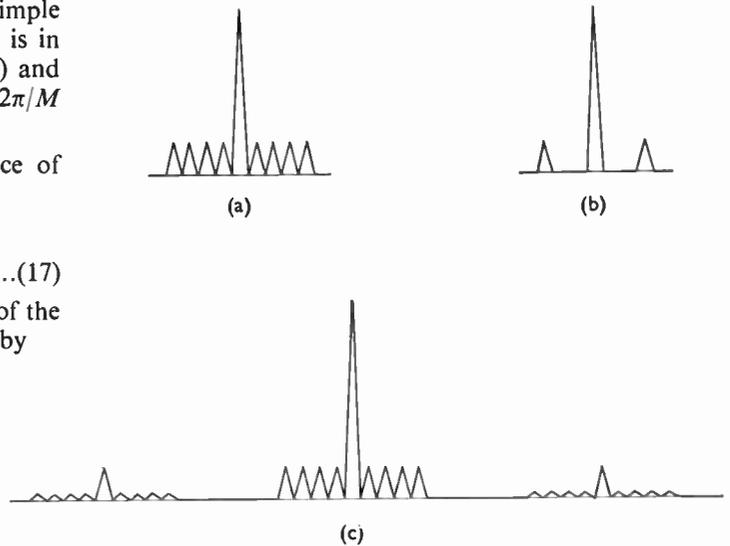


Fig. 19. Autocorrelation of (a) 5-element Barker code, (b) 4-element Huffman code and (c) 20-element combination code formed from these shorter codes.

A difficulty in implementing shift register type transversal filters for very long pulse trains is that so many tapings are needed. There are great advantages, of course, in terms of physical compactness if access is only required to every  $M$ th shift register stage when  $M$  is a large number. Figure 21 shows how the scheme of Fig. 6 can be modified for use with combination codes so as to effect economy in the number of tapping points. The transversal filter of Fig. 6 is replaced by two transversal filters, the first of which matches the stretched outer code and the second of which matches the inner code. A total of only  $M+N+2$  tapings and coefficient weights are now required in place of the  $(M+1)(N+1)$  that would be needed for the single transversal filter. This idea can be adapted to the digital scheme of Fig. 8 to enable multi-stage microcircuit shift registers to be used.

9. Conclusion

A.m.ph.m. pulse trains form a class of signals which are particularly well suited to digital methods of generation and processing. Some schemes for their digital implementation have been reviewed in Sections 3 and 4 of this paper. In some circumstances the use of pulse trains other than purely ph.m. may be precluded by the expense of providing amplitude modulation. However in other circumstances the use of a.m. may be no great disadvantage. The very low autocorrelation function sidelobe levels that can be attained with a.m.ph.m. pulse trains such as Huffman codes makes their use particularly attractive in systems requiring large dynamic range.

The later sections of the paper deal with the design of a.m.ph.m. pulse trains to have desired properties. The  $z$ -transform proves to be a central tool in this problem. The synthesis of a.m.ph.m. pulse trains for specified autocorrelation is accomplished by factorization of the autocorrelation  $z$ -transform and selection of one zero out

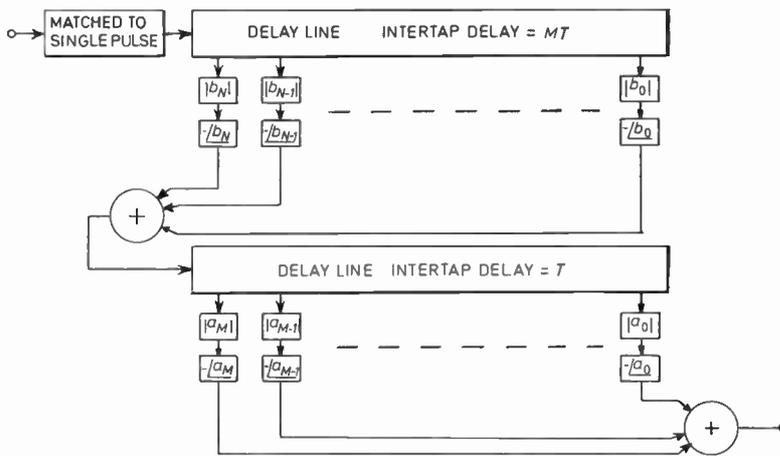


Fig. 21. Matched filter for combination code pulse train.

of each reciprocal conjugate pair. One shortcoming of the procedure is that in practice one is usually not directly concerned to find a signal with a specified autocorrelation function. Instead one more often requires a signal whose autocorrelation function satisfies certain criteria such as having sidelobes below a specified level. Further research may show how to choose an autocorrelation function which will satisfy a specification and also yield a signal with, for example, high energy efficiency when the synthesis procedure is applied.

The idea of combining good short codes to produce combination codes provides a way to sidestep some of the difficulties of designing long codes directly. Combination codes may also provide significant economies in implementation compared with other long a.m.p.h.m. codes.

10. References

1. Cook, C. E. and Bernfeld, M., 'Radar Signals' (Academic Press, New York, 1967).
2. Franks, L. E., 'Signal Theory' (Prentice Hall, Englewood Cliffs, N.J., 1969).
3. Voelcker, H. B., 'Generation of digital signaling waveforms', *I.E.E.E. Trans. on Communication Technology*, COM-16, pp. 81-93, February 1968.

4. Huffman, D. A., 'The generation of impulse-equivalent pulse trains', *I.E.E.E. Trans. on Information Theory*, IT-8, pp. S10-S16, September 1962.
5. L. Fejer, 'Uber trigonometrische polynome', *J. reine angew. Mathematik*, 146, pp. 53-82, 1915.
6. Wold, H., 'Study in the Analysis of Stationary Time Series', 2nd edition (Almqvist and Wiksell, Stockholm, 1954).
7. Schechter, H. S., SHARE Program RF 2C01.
8. Ackroyd, M. H., 'Computing the coefficients of high-order polynomials', *Electronics Letters*, 6, pp. 715-17, 29th October 1970.
9. Ackroyd, M. H., 'The design of Huffman sequences', *I.E.E.E. Trans. on Aerospace Electronic Systems*, AES-6, pp. 790-6, November 1970.
10. Hollis, E., 'Comparison of combined Barker codes for coded radar use', *I.E.E.E. Trans. on Aerospace and Electronic Systems*, AES-3, pp. 141-3, January 1967.
11. Hollis, E., 'Predicting the truncated autocorrelation functions of combined Barker sequences of any length without a computer', *I.E.E.E. Trans. on Aerospace and Electronic Systems*, AES-3, pp. 368-9, March 1967.
12. Ackroyd, M. H., 'Synthesis of efficient Huffman sequences', *I.E.E.E. Trans. on Aerospace and Electronic Systems*. (To be published).

Manuscript received by the Institution on 16th June 1971. (Paper No. 1421/AMMS43).

© The Institution of Electronic and Radio Engineers, 1971

The Author

Dr. M. H. Ackroyd was appointed in August to a Lectureship in Signal Processing in the Department of Electronic and Electrical Engineering at the University of Technology, Loughborough, and he is to be responsible for establishing a digital signal analysis centre. Since publication of a previous paper in the *Journal* in March 1970 he has obtained his doctorate for a thesis on signal theory.



# The Logarithmic Transformation of Random Processes

By

I. WHITE, M.Sc.†

This paper considers the effect of full-wave logarithmic transformation of signals with known probability density functions (p.d.f.s). A general result for the output variance is derived and specific results given for the mean and variance values of several common p.d.f.s. The output autocorrelation function for the Gaussian case is also derived.

## List of Symbols

- $A$  amplitude of sinusoidal signal
- $C(\tau)$  output autocorrelation function
- $I_0$  modified Bessel function of order zero
- $M = 1 - \rho^2(\tau)$
- $p(\cdot)$  probability density function (p.d.f.)
- $R$  sinusoid signal to Gaussian noise power ratio
- $x_1$  instantaneous input at time  $t$
- $x_2$  instantaneous input at time  $t + \tau$
- $\gamma$  Euler's constant (0.57721...)
- $\Gamma(\cdot)$  gamma function
- $\Gamma^{(1)}(\cdot)$  derivative of gamma function
- $\varepsilon = \sqrt{2} \exp(-\gamma/2)$
- $\theta[n/2] = \sum_{k=1}^m k^{-2}; m = \text{largest integer} < n/2$
- $\rho(\tau)$  normalized input autocorrelation function, denoted by  $\rho$  in text for brevity
- $\sigma_x^2$  input variance
- $\sigma^2$  output variance
- $\psi(\cdot)$  psi function
- $\psi^{(1)}(\cdot)$  derivative of the psi function
- $\chi$  instantaneous envelope of input signal

## 1. Introduction

The logarithmic amplifier finds wide application in communication and radar, usually for compressing dynamic range. The amplifier input, in many cases, can be specified by a known probability density function (p.d.f.). For example, in many communication systems the input p.d.f. is that of the envelope of sinusoid plus

Gaussian noise, whilst in neutron detectors, Gaussian and Rayleigh input p.d.f.s are common.

Although the techniques for analysing non-linear transformations of random processes are well known, the results for the logarithmic transformation involve some intricacy and do not appear to have been given previously in the literature. In this paper the constancy of the output variance is discussed and mean and variance values given for several well known p.d.f.s. The output autocorrelation function for the Gaussian input case is also derived, this being of value when the output power spectral density must be studied.

## 2. Constancy of the Variance

The relationship between instantaneous input,  $x$ , and output  $y$ , for a 'practical' full-wave logarithmic amplifier is

$$y = b \log(1 + |x/x_0|) \quad \dots\dots(1)$$

The term  $|x/x_0|$  denotes the modulus of  $x/x_0$  and represents the effect of full-wave rectification prior to logarithmic transformation (amplification).

Although equation (1) is not readily amenable to analysis it can be shown<sup>1</sup> that the mean square error in using the 'ideal' full-wave logarithmic transform

$$y = b \log(|x/x_0|) \quad \dots\dots(2)$$

is negligible if the mean square of the input  $\bar{x}^2 \gg x_0^2$ .

In the following  $b, x_0$  are taken to be unity, without loss of generality.

Since a scaling factor in a logarithmic transform is an additive constant in the output, all central moments of that output are independent of any scaling of the input.

Table 1. Mean and variance values for logarithmic transforms of common p.d.f.s.

p.d.f.	$p(x)$	range of $x$	mean output $\bar{y}$	output variance $\sigma_y^2$
rectangular	$k^{-1}$	$0 \leq x \leq k$	$\log(k) - 1$	1
sinusoid	$\frac{1}{\pi A \sqrt{1 - x^2/A^2}}$	$-A \leq x \leq +A$	$\log(A/2)$	$\frac{\pi^2}{12}$
Gaussian	$\frac{1}{\sqrt{2\pi}\sigma_x} \cdot \exp\left(\frac{-x^2}{2\sigma_x^2}\right)$	$-\infty < x < +\infty$	$\log(\varepsilon\sigma_x/2)$	$\frac{\pi^2}{8}$
Rayleigh	$\frac{x}{\sigma_x^2} \exp\left(\frac{-x^2}{2\sigma_x^2}\right)$	$0 \leq x < +\infty$	$\log(\varepsilon\sigma_x)$	$\frac{\pi^2}{24}$

† The Plessey Company Ltd., Electronics Research Laboratories, Martin Road, West Leigh, Havant, Hampshire.

Thus the output variance,  $\sigma_y^2$  is unaffected by any scaling of the input and hence is constant and independent of the input variance, a feature which is demonstrated by the values for several common p.d.f.s listed in Table 1.

The variance result for the Rayleigh distribution quoted in Table 1 is derived in ref. 1 and separately by Deutsch in ref. 2 who additionally gives an expression for the output autocorrelation function when the input is the envelope of a Gaussian signal.

### 3. Variance Result for the Rician Distribution

In the following, the p.d.f. of the envelope of sinusoidal signal plus Gaussian noise will be termed Rician after S. O. Rice,<sup>3</sup> in accordance with current usage. By the envelope of a time series is meant the formalized definition of Dugundji,<sup>2</sup> which in narrowband situations corresponds to the intuitive concept of envelope

The mean value of the logarithmic transform of the Rician distribution is

$$\bar{y} = \frac{1}{\sigma_x^2} \int_0^\infty \chi \log(\chi) \exp\left\{\frac{-\chi^2}{2\sigma_x^2} - R\right\} I_0\left(\sqrt{\frac{2R}{\sigma_x^2}} \cdot \chi\right) d\chi \quad \dots\dots(3)$$

where  $\sigma_x^2$  here denotes the variance of the Gaussian signal.

Making the substitution  $z = \chi^2/2\sigma_x^2$  and expanding the Bessel function  $I_0$  as a power series gives

$$\bar{y} = \frac{\exp(-R)}{2} \int_0^\infty \log(2\sigma_x^2 z) \exp(-z) \sum_{n=0}^\infty \frac{z^n R^n}{n!n!} dz \quad \dots\dots(4)$$

Applying the integral solution

$$\Gamma^{(l)}(n) = \int_0^\infty \log^l(z) z^{n-1} \exp(-z) dz \quad \dots\dots(5)$$

to (4) gives

$$\bar{y} = \frac{1}{2} \log(2\sigma_x^2) + \frac{\exp(-R)}{2} \sum_{n=0}^\infty \frac{\psi(n+1)R^n}{n!} \quad \dots\dots(6)$$

where

$$\psi(n+1) = \frac{\Gamma^{(1)}(n+1)}{\Gamma(n+1)}$$

In a similar manner we obtain for the mean square value

$$\bar{y}^2 = \frac{1}{4} \log^2(2\sigma_x^2) + \frac{\exp(-R)}{4} \sum_{n=0}^\infty \frac{\Gamma^{(2)}(n+1)R^n}{n!n!} + \frac{\exp(-R)}{2} \log(2\sigma_x^2) \sum_{n=0}^\infty \frac{\psi(n+1)R^n}{n!} \quad \dots\dots(7)$$

Using the following relations,<sup>4</sup>

$$\begin{aligned} \psi(n+1) &= -\gamma + \sum_{k=1}^n k^{-1} \\ \psi^{(1)}(n+1) &= \frac{\pi^2}{6} - \sum_{k=1}^n k^{-2} \end{aligned}$$

in (6) and (7) gives for the variance,

$$\begin{aligned} \sigma_y^2 &= \frac{\pi^2}{24} + \frac{\exp(-R)}{4} \sum_{n=1}^\infty \frac{R^n}{n!} \left\{ \left( \sum_{k=1}^n k^{-1} \right)^2 - \sum_{k=1}^n k^{-2} \right\} - \\ &\quad - \left\{ \frac{\exp(-R)}{2} \sum_{n=1}^\infty \frac{R^n}{n!} \sum_{k=1}^n k^{-1} \right\}^2 \quad \dots\dots(8) \end{aligned}$$

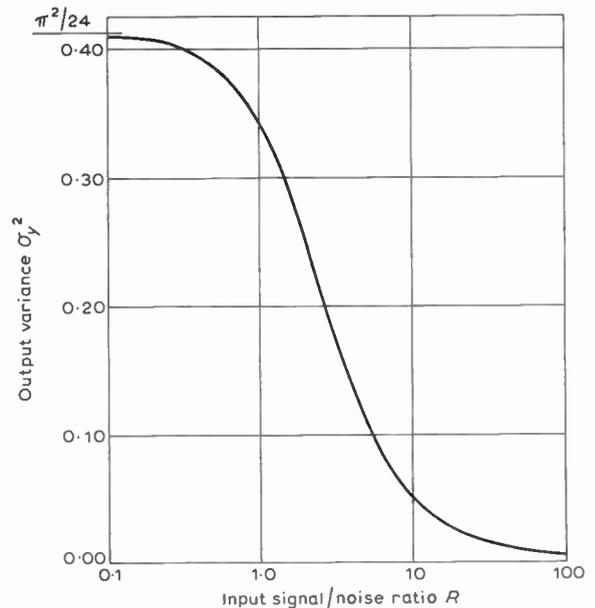


Fig. 1. Output variance for Rician distribution.

With  $R = 0$ , (8) reduces to the Rayleigh distribution result of  $\pi^2/24$ . In the Appendix equation (8) is simplified somewhat to

$$\sigma_y^2 = \frac{\pi^2}{24} - \frac{\exp(-R)}{2} \sum_{n=2}^\infty \frac{R^n}{n!} \theta[n/2] \quad \dots\dots(9)$$

It can be seen from (9) that the output variance is only a function of the signal/noise ratio and has therefore applications in signal/noise ratio measurement. A computer evaluation of equation (9) is shown in Fig. 1 from which it can be seen that useful measurement could be made over the range  $0.7 < R < 20$ .

### 4. Autocorrelation Function for the Gaussian Case

The autocorrelation function at the output of a full-wave logarithmic device with zero-mean Gaussian input is

$$\begin{aligned} C(\tau) &= \frac{1}{2\pi\sigma_x^2\sqrt{M}} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \log(|x_1|) \log(|x_2|) \times \\ &\quad \times \exp\left\{\frac{-x_1^2 - x_2^2 + 2\rho x_1 x_2}{2M\sigma_x^2}\right\} dx_1 dx_2 \quad \dots\dots(10) \end{aligned}$$

where (10) is the solution according to the Van Vleck-North method.<sup>3</sup> The solution could also be obtained using the characteristic function method.<sup>2</sup>

Making substitutions to normalize the Gaussian p.d.f. and applying Price's theorem<sup>5</sup> gives

$$\begin{aligned} \frac{\partial C(\tau)}{\partial \rho} &= \frac{1}{2\pi\sqrt{M}} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \frac{1}{x_1 x_2} \times \\ &\quad \times \exp\left\{\frac{-x_1^2 - x_2^2 + 2\rho x_1 x_2}{2M}\right\} dx_1 dx_2 \quad \dots\dots(11) \end{aligned}$$

The validity of the limits in (11) can be demonstrated by considering the four quadrants of the integral separately. If the substitutions  $z_1 = x_1/\sqrt{2M}$  and  $z_2 = -x_2/\sqrt{2M}$  are made in (11) the integral is reduced to a form whose solution is given by Rice:<sup>3</sup>

$$\frac{\partial C(\tau)}{\partial \rho} = \frac{-1}{2\pi\sqrt{M}} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \frac{1}{z_1 z_2} \times \exp\{-z_1^2 - z_2^2 - 2\rho z_1 z_2\} dz_1 dz_2$$

$$= \frac{-1}{2\pi\sqrt{M}} [-2\Gamma^2(\frac{1}{2})\sqrt{M}\rho \cdot {}_2F_1(1, 1; \frac{3}{2}; \rho^2)]$$

where  ${}_2F_1(a, b; c; \rho)$  denotes the Gaussian hypergeometric function.

$$\frac{\partial C(\tau)}{\partial \rho} = \rho \cdot {}_2F_1(1, 1; \frac{3}{2}; \rho^2)$$

$$= \frac{\arcsin(\rho)}{\sqrt{1-\rho^2}} \dots\dots(12)$$

Integrating both sides of (12) gives,

$$C(\tau) = \arcsin^2(\rho) + \text{const.}$$

With  $\tau = \infty$ ,  $\rho(\tau) = 0$ ,  $C(\tau) = 0$  neglecting the d.c. term in the output, and therefore the constant of integration is zero. With  $\tau = 0$ ,  $\rho(\tau) = 1$  and  $C(\tau) = \pi^2/8$  in agreement with Table 1.

**5. Conclusions**

The logarithmic amplifier is used extensively in signal processing. Common applications include compressing dynamic range; obtaining the logarithm of the product or ratio of two signals (by adding or subtracting respectively the appropriate outputs of two similar logarithmic amplifiers); approximating functional relationships, etc.

The variance results presented in this paper should be of particular value when estimating the signal/noise ratio at the output of a logarithmic amplifier. As is shown in Section 3, the variance result for the Rician p.d.f. can be used directly for signal/noise measurement. When the input p.d.f. is Rician, the power of the a.c. component of the output of a logarithmic amplifier has a precise monotonic relationship to input signal/noise ratio as shown in Fig. 1.

The output autocorrelation function derived for the Gaussian case is useful in studying the output power spectral density, since these two functions are a Fourier transform pair.<sup>3</sup> This particular case is encountered for example in neutron detector studies.

Collectively the results should be of some value whenever an assessment of the effect of logarithmic transformation is required.

**6. Acknowledgment**

The author would like to thank Dr. D. C. Cooper and Mr. J. M. Layton for their advice and criticism in the preparation of this paper.

Acknowledgments are also due to the Plessey Company for permission to publish this paper.

**7. References**

1. Croney, J., 'Clutter on radar displays', *Wireless Engineer*, 33, pp. 83-96, April 1956.
2. Deutsch, R., 'Non-Linear Transformation of Random Processes' (Prentice Hall, Englewood Cliffs, N.J., 1962).
3. Rice, S. O., 'Mathematical analysis of random noise', *Bell Syst. Tech. J.*, 23, pp. 282-332, July 1944; 24, pp. 46-156, January 1945.

4. Abramowitz, M. and Stegun, I. A., 'Handbook of Mathematical Functions' (National Bureau of Standards, Washington, 1964; also Dover Publications, New York, 1965).

5. Price, R., 'A useful theorem for non-linear devices having Gaussian inputs', *I.R.E. Trans. on Information Theory*, IT-4, pp. 69-72, June 1958.

**8. Appendix: Variance Solution for the Rician Distribution**

In (8) there are two infinite series, the first of which is

$$f_1(R) \triangleq \sum_{n=1}^{\infty} \frac{R^n}{n!} \sum_{k=1}^n k^{-1} \dots\dots(13)$$

We hypothesize

$$f_1(R) = b_1(R) \exp(R)$$

where  $b_1(R)$  is some unknown function. The differential equation of  $f_1(R)$  is

$$f_1^{(1)}(R) = f_1(R) + b_1^{(1)}(R) \exp(R) \dots(14)$$

Differentiating (13) gives

$$f_1^{(1)}(R) = f_1(R) + \sum_{n=0}^{\infty} \frac{R^n}{(n+1)!}$$

$$= f_1(R) + {}_1F_1(1; 2; R) \dots\dots(15)$$

where  ${}_1F_1(a; b; R)$  denotes the confluent hypergeometric function. Applying Kummer's first transform to this in (15) gives

$$f_1^{(1)}(R) = f_1(R) + \exp(R) {}_1F_1(1; 2; -R) \dots\dots(16)$$

From (14) and (16) we have

$$b_1^{(1)}(R) = {}_1F_1(1; 2; -R)$$

Integrating,

$$b_1(R) = - \sum_{n=1}^{\infty} \frac{(-R)^n}{nn!} + \text{const.} \dots\dots(17)$$

When  $R = 0$ ,  $f_1(R) = 0$  and therefore the constant of integration is zero.

A similar hypothesis can be made about the second infinite series in (8).

$$f_2(R) = \sum_{n=1}^{\infty} \frac{R^n}{n!} \left\{ \left( \sum_{k=1}^n k^{-1} \right)^2 - \sum_{k=1}^n k^{-2} \right\}$$

The differential equation in this case is

$$f_2^{(1)}(R) = f_2(R) + \frac{2}{R} \sum_{n=1}^{\infty} \frac{R^n}{n!} \left\{ \sum_{k=1}^n k^{-1} - \frac{1}{n} \right\} \dots\dots(18)$$

If we define

$$q(R) \triangleq \sum_{n=1}^{\infty} \frac{R^n}{nn!} = \int_0^R \frac{\exp(\lambda) - 1}{\lambda} .d\lambda$$

then from (14), (17) and (18)

$$b_2^{(1)}(R) = \frac{-2}{R} \{q(-R) + \exp(-R) \cdot q(R)\}$$

Integrating and applying boundary conditions gives

$$b_2(R) = -2 \int_0^R \frac{\{\exp(-\lambda)q(\lambda) + q(-\lambda)\}}{\lambda} .d\lambda$$

Applying these results to (8) for the output variance

$$\sigma_y^2 = \frac{\pi^2}{24} + \frac{1}{4} \{b_2(R) - b_1^2(R)\} \dots\dots(19)$$

Convolving the series for  $b_1(R)$  with itself gives

$$b_1^2(R) = 2 \sum_{n=2}^{\infty} \frac{(-R)^n}{n} \sum_{k=1}^n \frac{1}{k(n-k)!k!}$$

which can be identified as

$$b_1^2(R) = 2 \int_0^R \frac{\{\exp(-\lambda) - 1\}q(-\lambda)}{\lambda} .d\lambda$$

Substituting the expressions obtained for  $b_2(R)$  and  $b_1^2(R)$  into (19) gives integral solution

$$\sigma_y^2 = \frac{\pi^2}{24} - \frac{1}{2} \int_0^R \frac{\exp(-\lambda)}{\lambda} \{q(\lambda) + q(-\lambda)\} d\lambda \quad \dots\dots(20)$$

If (20) is integrated termwise the general term is

$$\begin{aligned} \mathcal{I}_n(R) &= \int_0^R \exp(-\lambda)\lambda^n .d\lambda \\ &= n! \left[ 1 - \exp(-R) \left\{ 1 + \frac{R}{1!} + \frac{R^2}{2!} \dots + \frac{R^n}{n!} \right\} \right] \end{aligned}$$

Substituted into (20) this gives

$$\sigma_y^2 = \frac{\pi^2}{24} - \frac{1}{2} \sum_{n=2}^{\infty} \frac{1 - \exp(-R) \sum_{k=0}^{2n+1} \frac{R^k}{k!}}{(n+1)^2}$$

If this series is written out and summed in columns we finally obtain

$$\sigma_y^2 = \frac{\pi^2}{24} - \frac{\exp(-R)}{2} \sum_{n=2}^{\infty} \frac{R^n}{n!} \theta[n/2] \quad \dots\dots(21)$$

where

$$\theta[n/2] = \sum_{k=1}^m k^{-2}$$

and  $m$  is the highest integer  $< n/2$ .

*Manuscript first received by the Institution on 19th April 1971 and in final form on 9th September 1971. (Paper No. 1422/CC 113)*

© The Institution of Electronic and Radio Engineers, 1971

# High-speed Measurement of Insulation Resistance of Capacitors and Capacitive Circuits

By

**RALPH E. G. KEON,**  
C.Eng., M.I.E.R.E.†

A new method is described for high-speed measurement of insulation resistance of capacitive components compatible with present day requirements for 100% insulation test of capacitors. It is suggested that such an instrument opens the way to a new approach to insulation resistance measurement which may lead to a revision of long established test specifications.

## 1. Introduction

The present methods of measurement of high insulation resistance, based upon well-founded principles, no longer fulfil many of the present requirements of the electronic industry where high speed and thorough quality control are required, especially when automated testing is involved. This requirement, placed upon component manufacturers, has led to the need for increased speed of inspection to keep up with production but, so far, only high-speed impurity test ( $\tan \delta$ , dissipation factor,  $Q$ ) has been added to the initial inspection for  $R$ ,  $C$  and  $L$  values.

High-speed measurement of insulation resistance has been neglected so far; this measurement is still usually left to the statistical uncertainty of 'batch sampling' specifications based upon sound but slow methods of measurement established long ago. These methods, ideal in the days of wax and shellac varnish insulations, are no longer compatible with the present technology based upon high-purity paper, plastics, oil and ceramic insulation materials, when the extremely high insulation resistances commonly achieved, would result in lengthy measurements. Electronic instruments employing solid state circuits often have to operate 'at once' and the insulation characteristics of a component after 60 seconds may be irrelevant since the 'switching-on' characteristic is the one that counts.

Metallized-film capacitor production engineers were the first to call for high-speed 100% insulation resistance tests to verify the effectiveness of the electrical flash-breakdown healing treatment which is part of the manufacturing process.

## 2. Present Methods

The most common method of measurement called for in specifications covering the test for insulation resistance, of capacitors and capacitive circuit elements, is based upon the measurement of the current that flows through the element when a specified test voltage is maintained across it. This arrangement suffers from several disadvantages, the major one being the long time which must elapse before very small leakage currents can be measured against the initial charging current of the capacitor, which will be many times greater than the leakage current  $i_L$  (Fig. 1). Additional delay may be needed, in some cases, to allow for extraneous currents to subside, such as charge and discharge currents due to capacitance changes

caused by electrical and mechanical stresses. Since it is difficult to establish accurately a suitable measuring time, this is usually chosen arbitrarily as 30 or 60 seconds, according to convenience.

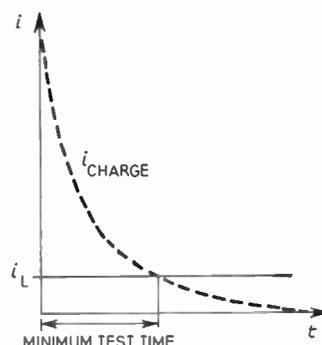


Fig. 1. Unavoidable delay with the usual methods of insulation resistance measurements.

Another disadvantage is that the supply voltage to the capacitor being tested must be known exactly if the insulation resistance is to be measured with any accuracy, and it must also be very stable, so as to avoid large charge and discharge currents which would also conceal the leakage current. Therefore this method calls for highly sensitive, stable and accurate instruments to ensure an acceptable level of measurement accuracy. The test instruments offered so far are seldom sufficiently ruggedized and simple to operate for everyday use by unskilled, production-line operators.

Another method, based upon the self-discharge of the capacitor being tested, fails to give valuable information on the behaviour of the capacitor at its peak voltage, since the steady-state electrical stresses are released as soon as the capacitor is switched off from its supply and most of the measurement is done at a much lower decaying voltage than the initial operating test voltage called for (Fig. 2).

As one might expect, the results obtained from these two totally different methods, i.e. based upon the charge or discharge of a capacitor, seldom agree since, on the one hand, the insulation is subjected to increasing stresses whilst, on the other hand, the measurement is performed with decaying electrical stresses. These methods would probably be consistent if the insulating materials followed Ohm's Law, which rarely is the case. A more recent method initiated by the French firm Lemouzy a few years ago enables the measurement to be

† Culton Control Systems Ltd., Dorking, Surrey.

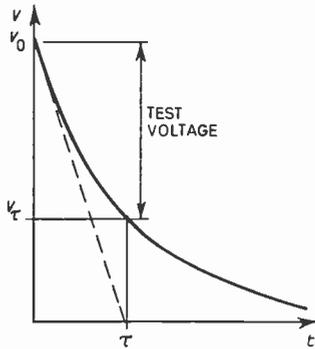


Fig. 2. Variable test voltage resulting from the method of insulation resistance measurement by self-discharge.

speeded up considerably by means of a special feedback circuit called Tripole,<sup>1</sup> which assists the charge of the capacitor so that the steady state is achieved rapidly but, even so, the measurement still takes several seconds and calls for high stability and sensitive instruments which would be difficult to incorporate in an automated capacitor test system.

Since these methods call for expensive and often fragile instruments and still remain slow and uncertain, they are often replaced by the very simple flash test method which may tell the eventual user that the component 'had not failed yet' at a particular specified voltage, but it will not tell him anything about the behaviour of the component under other prevailing operating conditions. He will only be able to guess it.

### 3. High-speed Measurements

With the recent advance in automatic classification of capacitors by value, soon followed by the classification by dissipation factor, at speeds in excess of 1800 and often 3600 components per hour, it was felt essential to design a system offering a 'valid' means of classification by insulation resistance, within predetermined selected limits, and at a speed compatible with present day sorting machines.

A preliminary survey was made<sup>2</sup> of the requirements necessary to meet the essential points of established test specifications while overcoming the disadvantages of the methods described so far. This led to the following basic requirements for the circuit to allow faster operating speeds whilst giving results compatible with established methods.

- (1) It must be virtually insensitive to the exact value of the applied voltage or of its fluctuations.
- (2) It must be virtually insensitive to the actual value of the capacitor or capacitance of the circuit being analysed, this for small values of  $C$  as indicated usually in test specifications.
- (3) The measurement must be performed, at d.c., at the peak of the applied voltage specified or as near as possible to the maximum stress point called for.
- (4) It must cover the range of insulation resistances, or equivalent time-constants, usually resulting from the present day insulating materials, or most of them in any case.

- (5) It must offer a suitable accuracy of 10% or better.
- (6) It must be simple to operate, safe and ruggedized, so that it may be used conveniently by unskilled as well as trained operators.
- (7) It must be fast and compatible with the normal speed of component classifiers. This means about 1 to 2 seconds per test in 12-channel classifiers and faster in particular set of circumstances.
- (8) It must be easily adapted for the control of automatic component sorting machines.
- (9) It must be reasonably priced.

From this analysis, point (1) leads to the choice of a discharge method of measurement but point (3) dictates that it must be done at peak voltage, therefore the capacitor must not be allowed to discharge appreciably. Point (2) calls for a comparison method whilst points (4) and (5) state that the reference component must be of the highest quality of insulation and that means must be provided to neutralize extraneous leakage currents. Points (6), (7) and (8) are practical considerations called for in the electronics industry and point (9) speaks for itself.

The aim of the instrument must be to perform the more usual measurements of up to  $10^{12} \Omega$  in 1 second or less and to extend the range up to  $10^{13} \Omega$ , and for this to be achieved in less than 5 seconds.

As a result of the speed of operation contemplated and other considerations called for in automatic classification, it was decided to dispense with a meter giving actual read-out of the value measured and to design the instrument as an insulation resistance limit detector giving the answer visually by means of ACCEPT/REJECT lamps, and electrically by means of a REJECT signal. The speed is achieved by observing the loss of charge immediately after the test voltage has been removed, and this for a short but defined time  $dt$ , rather than by measuring the current that flows after the capacitor has been charged for some time. Since usual test specifications call for d.c. leakage measurements, the capacitor is not allowed to discharge appreciably so that the voltage fluctuations are insignificant compared with the value of the applied voltage and interference from a.c. losses are avoided. The d.c. applied voltage is derived from a half-wave supply which, once the stipulated voltage is reached, maintains the charge during a short part of the cycle and allows for a substantial measuring time during the non-charging part of the cycle. Since the voltage source has a low impedance, usually increased for safety reasons, the initial charging

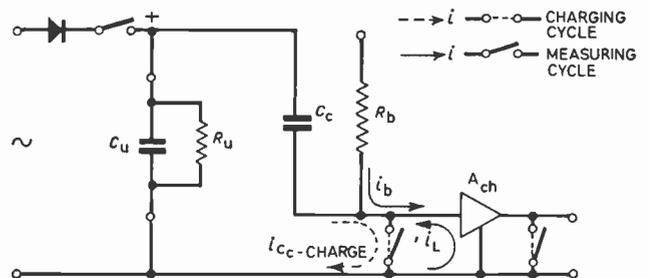


Fig. 3. Basic principle of the circuit described.

time is small, 4 to 5 cycles of a half-wave 50 Hz generator, and the specified steady d.c. voltage level is reached in some 100 ms, in the case of small-value capacitors; higher value capacitors will take a proportionally longer time to reach the stipulated voltage but, in the case of low voltage capacitors, the safety charging limiter resistor may be reduced to suit particular requirements.

The measurement of the loss of charge is achieved indirectly by providing a coupling capacitor  $C_c$  (Fig. 3), in shunt with the unknown capacitor  $C_u$ , to the input of the virtual-earth chopper amplifier  $A_{ch}$  where the leakage current  $i_L$  is opposed by a known preset bias current  $i_b$ . The input impedance of the amplifier is low with respect to the reactance of  $C_c$  at the chopper frequency of 25 Hz and both input and output of the amplifier are short-circuited during the charging part of the test cycle so as to ensure a low resistance in the charging circuit of the capacitors and also to reduce the possibility of interference in the measuring circuit. The chopper is phased, relative to the supply voltage which charges the capacitors (Fig. 4), so that the amplifier is active during a short but defined part of the period in which  $C_u$  is discharging.

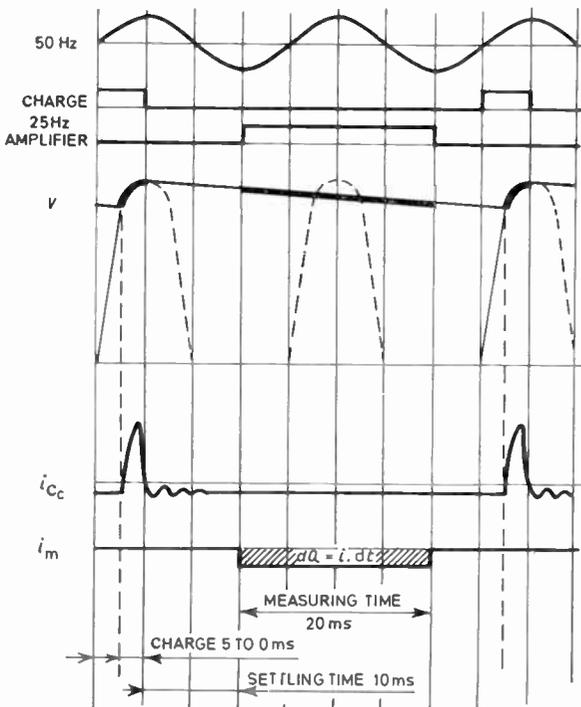


Fig. 4. Typical waveforms at various points of the circuit.

To justify the principle of the method, established in the Appendix, the d.c. source ought to present an infinite resistance during the measuring part of the cycle, in any case much greater than the maximum leakage resistance considered, and whilst the last half-wave rectifier of the multiplier or other source adopted is only subjected to a small fraction of the actual applied d.c. voltage  $V$ , its leakage current is probably excessive compared with the current to be measured. The rectifier is therefore followed

by a glass-insulated dry-reed relay which is operated at 25 Hz and phased so that the opening and closing of the contacts are made during the 'off-period' of the rectifier. (The relay, having no current to switch on or off, will have a very long life.) This relay, operating in synchronism with the choppers of the amplifier, ensures a very low series resistance during the short charging period of  $\frac{1}{4}$  cycle maximum (0 to 5 ms) and an extremely high resistance during the measuring part of the cycle.

The duration of the 'active measuring time' of the chopper amplifier is made to be exactly equal to one cycle of the line frequency and the amplifier is followed by an integrator which thereby gives rejection of interfering line signals and harmonics. A time delay of 10 ms ( $\frac{1}{4}$  cycle of 50 Hz) is allowed between the end of the charging time and the beginning of the defined time of measurement to allow for the decay of transients and avoid fluctuations during the measurement. The input chopper of the amplifier also short-circuits the charging pulse which would otherwise subtract from the measurement.

The theory developed in the Appendix substantiates the validity of the principle of the method used and explains its mode of operation as well as its peculiarities. The following refinements can be added to meet the practical requirements discussed previously.

- (1) Small value resistors may be inserted in series with either or both capacitors to limit the initial charge current or breakdown current. They will not affect the measurement of insulation resistance.
- (2) The d.c. supply voltage may be provided by a simple variable source or multiplier chain to achieve the desired voltage.
- (3) In theory, the system is suitable for measurements at any voltage level but, in practice, the highest voltage contemplated is limited by the availability of high-voltage coupling capacitors  $C_c$  with an insulation far greater than the insulation of the capacitor being tested and the ultimate sensitivity of the amplifier and detector limits the operation at low voltages.
- (4) The voltage applied to the capacitors may be measured by arranging for a second rectifier and reservoir capacitor to be connected to the same supply, the latter capacitor having a voltmeter in shunt. The series resistor  $R_b$  (Fig. 5) is used to provide the bias current  $i_b$  proportional to the

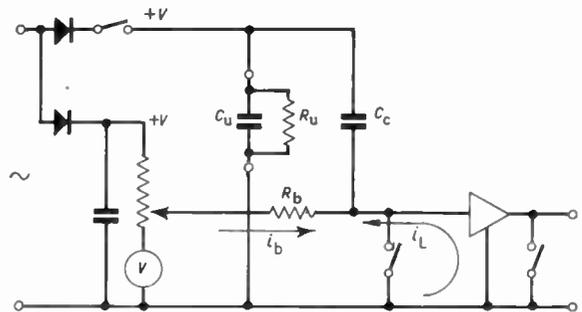


Fig. 5. Basic circuit showing the method of supply of the reference current  $i_b$ .

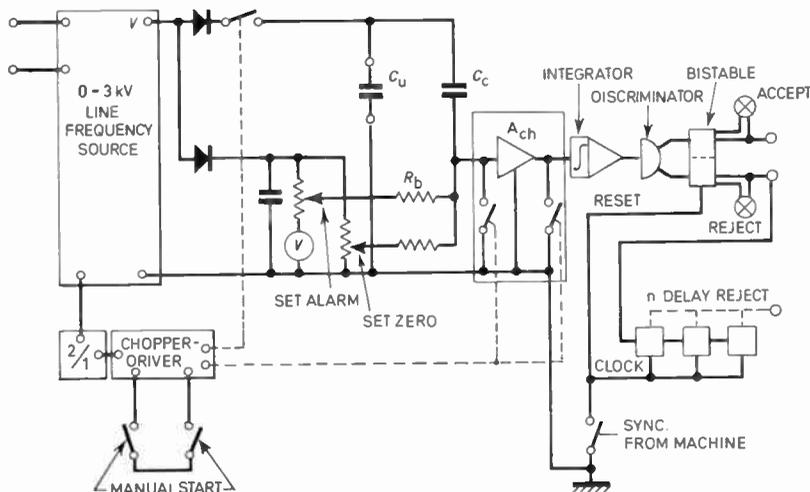


Fig. 6. Block diagram of the insulation resistance.

applied voltage  $V$ , thereby making the 'alarm' level substantially independent of the exact value of the test voltage  $V$ .

- (5) The predetermined bias current may be fed to the input or output of the amplifier, according to its relative level with respect to the unknown current, and define the level above which the alarm will operate.
- (6) The chopper amplifier may be followed by an integrator, discriminator and bistable circuit to drive a visual indication of the level of the unknown current below or above the preselected threshold.
- (7) The bistable circuit may be followed by a delayed memory circuit so that the REJECT information is available at the correct time for the classification sorting machine.

**4. Realization**

The circuit described<sup>2</sup> is embodied in an insulation resistance tester in which out-of-limit alarm level controls are derived from the decision-making dichotomiser circuit—a two-way discriminator. The latter drives a set of warning lamps as well as appropriate electro-mechanical gates when the instrument is used with a component classifier such as a Klemt Automat D12 sorting machine. When used with such a sorting system, the dichotomiser and bistable circuits are followed by a ladder of four synchronized memories which carry the REJECT signal until the end of the test sequence (Fig. 6).

A bench-top model has also been developed (Fig. 7) in which the test voltage is applied to the centrally placed component by means of lateral push-buttons. This layout has been adopted so that the operator cannot handle the component or terminals when the test voltage is applied. The component is discharged automatically upon release of the buttons.

The experience gained so far with the instrument in quality control of capacitors, and especially for metallized film and paper types has fully justified the principle of the method and its validity when compared with measurements made to established test specifications.

In the present instrument, the threshold of ' $C_u$  much greater or much smaller than  $C_c$ ' has been chosen arbitrarily as 30 nF; this value is usually accepted as a threshold in some test specifications. Higher threshold values may be used to meet particular cases.

The insulation resistance preselected levels have been chosen to follow the universally-accepted progression of resistance values, namely, 1.5, 2.2, 3.3, 4.7, 6.8, and the multiplying factor switch extends the range from 10 MΩ to 6.8 TΩ ( $6.8 \times 10^{12}$  ohms, shown as 6.8 M × MΩ on the panel of the instrument). In the case of measurement of higher value capacitors, above 30 nF, the instrument reads directly in time-constant from 10 seconds to 68 thousand seconds.

The instrument is therefore ideally suited to the measurement of insulation resistance in all cases when the actual value does not need recording.

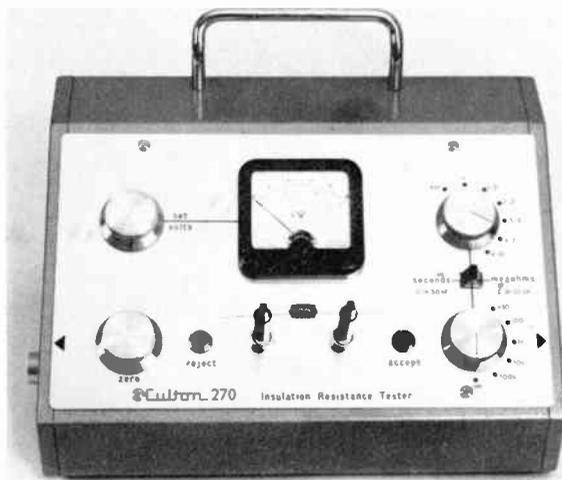


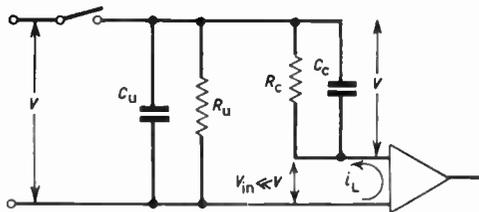
Fig. 7. The Cultron 270 insulation resistance test.

**6. Appendix**

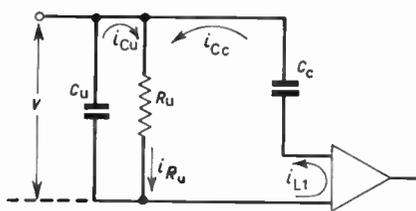
During the measuring part of the cycle the virtual-earth chopper amplifier  $A_{ch}$ , which has a very low input impedance, is connected in series with the coupling capacitor  $C_c$  and the supply voltage  $V$  is disconnected

from both the unknown capacitor  $C_u$  and the coupling capacitor  $C_c$ .

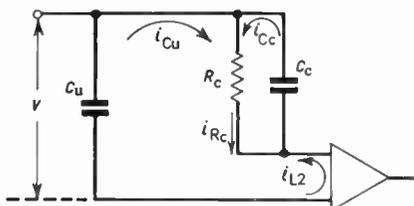
At the initial instant of discharge, the circuit is represented by Fig. 8(a) in which both capacitors behave like generators supplying currents to both leakage resistances,  $R_u$  of the unknown capacitor and  $R_c$  of the coupling capacitor; only the resulting leakage current  $i_L$  is being analysed by the amplifier and detector circuit.



(a) State of the circuit at the initial time of measurement.



(b) Influence of  $R_u$ .



(c) Influence of  $R_c$ .

Fig. 8.

By using the method of superposition of currents, one may analyse the effect of both leakage resistances, currents and charges  $t \cdot i_{L1}$  and  $t \cdot i_{L2}$ , independently as follows:

Leakage current due to  $R_u$  (see Fig. 8(b))

$$t \cdot i_{L1} = i_{C_c} \cdot t = V \cdot C_c$$

$$\frac{V}{t} = \frac{i_{L1}}{C_c} \quad \dots\dots(1)$$

and  $t \cdot i_{R_u} = \frac{V}{R_u} \cdot t = V(C_u + C_c)$

$$\frac{V}{t} = \frac{V}{R_u(C_u + C_c)} \quad \dots\dots(2)$$

Equating (1) and (2)

$$i_{L1} = \frac{V}{R_u} \frac{C_c}{(C_u + C_c)}$$

Leakage current due to  $R_c$  (see Fig. 8(c))

$$t \cdot i_{L2} = -i_{C_u} \cdot t = -V \cdot C_u$$

$$\frac{V}{t} = -\frac{i_{L2}}{C_u} \quad \dots\dots(3)$$

and  $t \cdot i_{R_c} = \frac{V}{R_c} \cdot t = V(C_u + C_c)$

$$\frac{V}{t} = \frac{V}{R_c(C_u + C_c)} \quad \dots\dots(4)$$

Equating (3) and (4)

$$i_{L2} = -\frac{V}{R_c} \frac{C_u}{(C_u + C_c)}$$

Resulting leakage current  $i_L$

$$i_L = i_{L1} + i_{L2}$$

Therefore

$$i_L = \frac{C_c}{R_u(C_u + C_c)} - \frac{C_u}{R_c(C_u + C_c)}$$

$$= \frac{V}{(C_u + C_c)} \left( \frac{C_c}{R_u} - \frac{C_u}{R_c} \right)$$

The current  $i_L$  is opposed by the reference bias current  $i_b$  (Fig. 3) which is itself proportional to  $V$ , such that

$$i_b = \frac{k \cdot V}{R_b}$$

and at balance

$$i_L = i_b$$

Therefore

$$\frac{k \cdot V}{R_b} = \frac{V}{C_u + C_c} \left( \frac{C_c}{R_u} - \frac{C_u}{R_c} \right)$$

and

$$R_b = \frac{k(C_u + C_c)}{\frac{C_c}{R_u} - \frac{C_u}{R_c}}$$

which may be re-written as

$$R_b = k \cdot R_u \frac{1 + \frac{C_u}{C_c}}{1 - \frac{R_u C_u}{R_c C_c}} \quad \dots\dots(5)$$

or

$$R_b \cdot C_c = k \cdot R_u \cdot C_u \frac{1 + \frac{C_c}{C_u}}{1 - \frac{R_u C_u}{R_c C_c}} \quad \dots\dots(6)$$

The factors  $R_u C_u$  and  $R_c C_c$  may be replaced by their equivalent time-constants  $T_u$  for the unknown and  $T_c$  for

the coupling capacitor and the equations (5) and (6) can therefore be simplified with respect to low or high values of unknown capacitors.

For 'low values', when  $C_u \ll C_c$ , equation (5) becomes

$$R_b = k \cdot R \frac{1}{1 - \frac{T_u}{T_c}} \dots\dots(7)$$

For 'high values', when  $C_u \gg C_c$ , equation (6) becomes

$$R_b \cdot C_c = k \cdot R_u \cdot C_u \frac{1}{1 - \frac{T_u}{T_c}} \dots\dots(8)$$

In the first case, when  $C_u$  has a low value, the measurement is independent of the actual value of the capacitor and of the parameters of the source and detector. Equation (7) shows that the relation is purely resistive and that the measurement is directly proportional to the value of the bias resistor  $R_b$  which may therefore be calibrated in terms of insulation resistance of the unknown in MEGOHMS.

In the second case, when  $C_u$  has a high value, the measurement is still independent of the parameters of the source and detector but equation (8) shows that the leakage is now expressed in the form of a ratio of time-constants; the time-constant of the combined elements  $C_c$  and  $R_b$  can therefore be calibrated in terms of the time-constant of the unknown, in SECONDS or in MEGOHMS  $\times$  MICROFARADS.

The alarm signal, given when the leakage current exceeds the bias current, is therefore obtained when

$$R_u < \frac{R_b}{k}$$

and

$$R_u \cdot C_u < \frac{R_b \cdot C_c}{k}$$

respectively. However, both equations (7) and (8) have a factor

$$\frac{1}{1 - \frac{T_u}{T_c}}$$

which must be close to unity to validate the balance considerations of the above equations. Since the time-constant of the unknown is beyond the user's control, one must ensure that the time-constant of the coupling capacitor  $T_c$  is always greater than  $T_u$ ; this calls for a very high quality capacitor with extremely high insulation resistance and especially so in the case of  $C_u \gg C_c$  when  $C_u > 10C_c$  or often more.

The insulation resistance of the terminals, or the circuit such as one might find when the instrument is connected to a sorting machine, is cancelled by an

additional 'set-zero' bias current  $i_z$  which is provided in a similar way as the measuring bias current and adjusted to suit particular circumstances.

Since the principle of operation of the circuit calls for the unknown capacitor  $C_u$  to be either 'much greater or much smaller' than the coupling capacitor  $C_c$ , but never equal, the instrument is provided with two operating ranges when  $C_c$  is made respectively either smaller or greater than 30 nF; at the same time appropriate correction is made to the bias control factor  $k$  so that either insulation resistance or time-constant designation may be applied to the bias controls.

Besides the validity of the measurement of resistance and time-constant which has been proved, the other desirable requirement is that the applied voltage must not be allowed to drop appreciably during the measurement. To analyse this factor one must consider the equation relating the various definitions of electrical charge, such as:

$$Q = C \cdot V = I \cdot t$$

and

$$dQ = C \cdot dV = I \cdot dt = \frac{V}{R} dt$$

The latter term is valid if one assumes that the current  $I$  will remain substantially constant during the discharge, which is true in this case. The relative percentage voltage drop will therefore be given by:

$$\frac{dV}{V} = \frac{dt}{R \cdot C} 100\%$$

in which  $dt$  is fixed at 30 ms, i.e. 10 ms settling time in addition to the 20 ms measuring time, and in the case of  $C_u < 30$  nF,  $C_c$  is set at 0.11  $\mu$ F and the minimum value of  $R_u$  is limited to 10 M $\Omega$ ; therefore, in the worst case:

$$\frac{dV}{V} = \frac{30 \times 10^{-3} \times 100}{10^7 \times 1.1 \times 10^{-7}} = \frac{3}{1.1} \% = 2.82\%$$

In the case of  $C_u > 30$  nF, the minimum time-constant is 10 seconds and the maximum percentage ripple will therefore be:

$$\frac{dV}{V} = \frac{30 \times 10^{-3} \times 100}{10} = 0.3\%$$

**5. References**

1. Lemouzy, J., 'High resistance and high insulation measurements', in 'Instruments et Laboratoires' (Dunod, Paris, 1957).
2. Tait, D. A. G., (Culton Instruments Limited) Patent application No. 20185/70.

*Manuscript first received by the Institution on 4th February 1971 and in revised form on 3rd June 1971. (Paper No. 1423/IC55).*

© The Institution of Electronic and Radio Engineers, 1971

# A Cellular Array which Represents the Inertial Motions of Objects

By

R. BAKER, B.Sc., M.Sc.†

Arrays of identical, uniformly connected circuits are described, which can represent the inertial motions of a group of objects. The objects are represented by 'object words', which are propagated within a planar array. Applications for traffic control systems, image-processing and artificial intelligence are proposed.

## 1. Introduction

A number of applications exist for data-processing systems which can represent the motions of several objects in space. For instance:

Traffic control systems (vehicles on fixed paths as in rail systems, or in open media as in air traffic systems).

Robot systems used in artificial intelligence studies (mobile robots or simulated hand-eye systems). Such systems are described by Burstall,<sup>1</sup> Raphael,<sup>2</sup> Feldman and Sproull<sup>3</sup> and others.

Image-processing systems, such as those used in tracking radar signals.

Such systems must exist within the central nervous systems of many living organisms including humans, but their mechanisms are not yet understood: the construction of machines to perform similar processing tasks could contribute to an understanding of their biological counterparts.

## 2. Mapping a System of Moving Objects

This paper discusses the possibility of representing a region of space containing moving objects by forming a map of the region on a cellular array of electronic circuits. To introduce this notion some elementary features of conventional map-making will first be mentioned. Consider the problem of representing a region of space containing several distinct but immobile objects (we restrict the discussion here to two-dimensional regions). An effective method of mapping then consists of representing the region of space by a plane piece of storage medium (e.g. blank paper), and representing the objects by symbols stored at corresponding points on the medium (e.g. by ink marks on the paper). If the objects are mobile, then physically movable symbols can be used (e.g. coloured pins). To avoid the need for physical movement in the map itself, an erasable storage medium can be used, with electronic means for entering symbols at appropriate points (the phosphor surface of a long-persistence c.r.t. has this property). Alternatively the map may be made up from a finite number of storage elements, usually provided with display devices. Such maps have also been formed using a portion of the store of a general-purpose computer. In the present paper it is proposed that maps should be formed from uniformly interconnected, regular arrays of identical electronic circuits (cellular arrays).

When the represented objects are in motion there is a need regularly to update the map; this is achieved through

† Twickenham College of Technology, Department of Engineering Technology, Egerton Road, Twickenham, Middlesex.

a sensor-system. It may happen that many or even all of the represented objects are in motion, requiring a high rate of information flow into the map; the sensor-system may itself be in motion, as in mobile robot systems, causing apparent motion of all the represented objects. Unfortunately there may be factors which limit the information rate from the sensor-system to less than that required adequately to update the map. These factors include a limited scanning rate, temporary obstructions in the line of sight and temporary loss of recognition. Automatic recognition devices for visual scene analysis can take seconds or even minutes to recognize each object. A solution of this dilemma is to provide a mechanism whereby the map representation of objects can be updated, even when the corresponding sensor information is temporarily unavailable; to achieve this the dynamics of the object-system must be known (or assumed). Human operators can perform this task, but are notoriously inaccurate and unreliable in so doing. Alternatively a fast general-purpose computer can update the position of each object-representation in turn, according to assumed dynamics. The task can also be achieved by mapping onto a cellular array which incorporates the assumed dynamics: this technique is described in the next Section.

## 3. Cellular Array as a Map

To obtain the speed required for real-time processing of maps with many moving objects the following approaches may be used:

- (1) Very high-speed time-sharing processor.
- (2) Parallel-processing, for each moving object.
- (3) Parallel-processing, for each occupyable region of the map.

The last approach is considered in this paper; this would lead to a system containing a large amount of hardware, but this will be offset by the possibility of constructing it from a regularly connected array of identical and relatively simple cells. This possibility of cellular structure leads to a conceptually simple, but very versatile system. The cellular structure is further suggestive of biological models for perception.

It is therefore proposed that a map should be constructed from an array of identical cells. Planar arrays will be described which map plane sections of an environment, or process 2-dimensional images. There is a possibility of 3-dimensional maps, but these are ruled out here as being probably too costly; the planar map will be quite useful for 3-dimensional environments, as some thin slabs of the environment can be considered as

approximately planar (e.g. most human activity is confined to a slab about 2 metres thick and resting on the walking surface). A 2-dimensional mapping system working in a 3-dimensional environment will be aided by an ability to change the map scale and orientation so as to select the most significant section for processing. Each cell of the map-array corresponds with the smallest occupiable region of the space which is mapped. This suggests that a very close mesh of cells is required, with consequent high cost; this is not necessarily so. Division of the space into cells which are either occupied or vacant implies a restriction on the number of objects which can be mapped at one time, and a loss of positional accuracy in the representation of those objects. In practice, formal systems of traffic control usually rest on the temporary allocation of a distinct region for each vehicle, and the aim is to prevent two vehicles entering or approaching the same region; the number of such regions would rarely exceed  $10^4$ , which implies a map-array of only  $100 \times 100$  cells. In the case of robot-guidance, the environment is likely to be considerably less formalized, especially if guidance in real-life environments is considered; comparison with human perception suggests that attention is in fact directed to a small number of objects in the environment when these are moving, and even then judgement of the absolute position and speed of these is not especially good in most cases. The important features appear to be the ability to select 'significant' objects for attention, and rapidly to assess the spatial relationships of these. If this is true, then an array which can process several hundred objects in real time (or faster), with a uniform, if rather low, accuracy, may well be of value in robot guidance. For image-processing applications, the complexity problem could be severe; for scientific or commercial applications it might be possible to 'thin' the data arriving at the map, cutting out automatically some less significant objects; in the extreme case recognizable objects would be replaced by symbols or codes.

#### 4. Assumed Dynamics

The system proposed would certainly be impractical if the processors were external to the map since this would require multi-way cables linking every cell with its processor. It is therefore proposed that the processors should be placed within the map cells. In this way the map becomes dynamic, and thus becomes a *model* of the environment, as well as a map of it. We now come to the problem of what to put in the cell, and what to leave out. To avoid undue complexity of the system, only geometric and simple dynamic properties should be incorporated in the cells; sophisticated dynamics, and non-geometric properties can be processed externally to the array. In the present paper, the array represents only the properties of continuity of motion and constancy of velocity in the absence of collisions. Specifically, it is assumed that the object-system represented is inertial and force-free in the absence of collisions or outside intervention. That is, objects which are temporarily unobserved maintain their last observed velocities. The array which performs this function is described here as an 'inertial array'. Although the dynamics assumed are so simple,

they are expected to give a first-order approximation useful in real-world situations. It is significant here that humans are proficient in manoeuvring in, and manipulating, real-world systems, without any formal knowledge of dynamics.

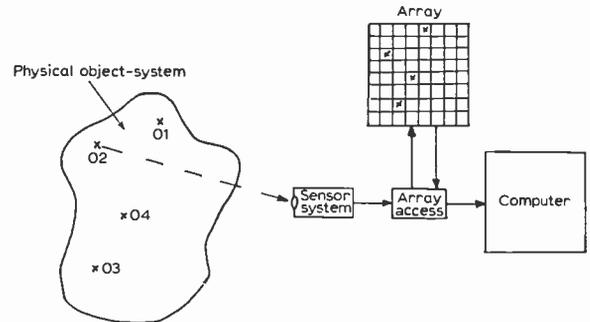


Fig. 1. Perceptual system using inertial array.

#### 5. Characteristics of the Cell

Each cell contains a number of binary storage locations (bistable elements). One of these stores the 'occupancy bit' ( $C$ ) which equals '0' when the corresponding region of space is empty, and '1' when it is occupied. This bit can be set to '0' or '1' by 'external signals' originating from the sensor system. A possible system is illustrated in Fig. 1; the computer at the right is to process non-geometric information. The occupancy bit can also be set to '1' by a 'propagate signal' ( $P$ ) arriving from an adjacent cell. (Zeros are not propagated, since only objects, not spaces, are inertially propagated.) Additional 'identity bits' ( $I$ ) may also be retained: these can identify the kind of object, or individual objects; in image processing applications these bits could carry other information, e.g. colour. Further information can be retained to specify the velocity of the represented object. All this information makes up an 'object-word' ( $W$ ) characterizing the object. When an object is observed to move from one region to another, the sensor system deletes the corresponding word from one cell and writes it into another. If contact is temporarily lost with a moving object, the inertial property of the array requires that the corresponding word shall be automatically propagated through the array with a rate and direction the same as that which was last observed. It would be possible to require the sensor system to measure the velocities of objects and supply a rate-signal to the array: in the arrays described in this paper however, the rate and direction are automatically calculated by the mutual action of the array cells. 'Velocity-processing circuits' are provided within each cell for this purpose. The possible directions of intercell propagation are defined by the intercell connexions: several arrangements are described in later sections; all intercell connexions are between adjacent cells.

#### 6. Simple Linear Array

A simple array is first described, consisting of a single chain of cells. Motion in one direction, at a fixed speed is represented. This array is too simple for practical use,

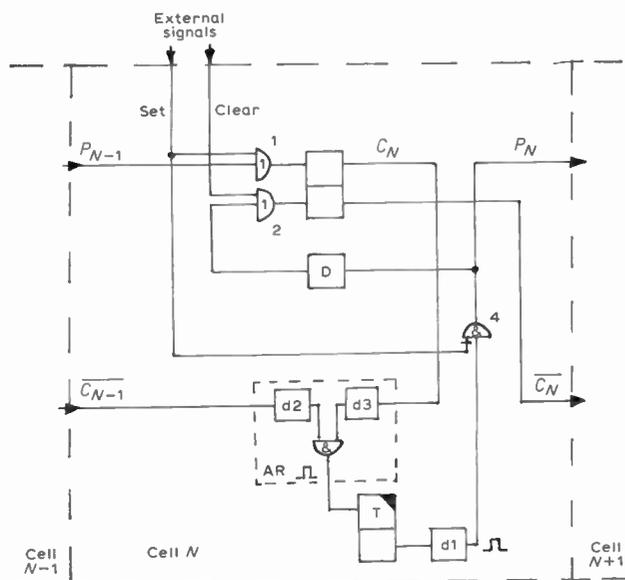


Fig. 2. Cell of simple linear array.

but is described to illustrate some of the main features of more complex arrays. Figure 2 shows the logic of the  $N$ th cell of the array. The occupancy bistable element retains the occupancy bit  $C_N$ . Only one kind of object is represented, so there are no identity bits. The value of  $C_N$  is set to '1' or '0' by external signals, via gates 1 or 2 respectively. Or '1' can be set in by a 'propagate-pulse'  $P_{N-1}$  from cell  $N-1$ . Arrival of occupancy in cell  $N$  from cell  $N-1$  is detected by the 'arrival-circuit' (AR). If as a result of external signals occupancy disappears from cell  $N-1$  and appears in cell  $N$  at nearly the same moment, then it is assumed that the represented object has moved between corresponding regions, and hence is in motion. Therefore, in the absence of further external signals, occupancy must be transferred to cell  $N+1$  after a preset time. The 'arrival-circuit' will respond also to arrivals caused by intercell propagation, so once the occupancy bit starts moving, it continues through the array until further external signals arrive. The arrival-circuit consists of two differentiating circuits (which emit a positive pulse in response to a rising edge). The arrival causes pulses to appear simultaneously at the inputs of gate 3, which then triggers monostable element  $T$ . (Some lack of synchronism in the edges presented to differentiators  $d2, d3$  is accommodated by the finite duration of their output pulses.) The fact of arrival is retained in the monostable element for a period  $T$ , at the end of which a rising edge is presented to  $d1$ . This generates a 'propagate-pulse'  $P_N$  which sets the following cell, and after a brief delay ( $D$ ) clears  $C_N$ . It may happen that occupancy arrives at cell  $N$  and remains there; in this case gate 4 inhibits propagation and clearing. This mechanism also serves to stop existing propagation when an external signal indicates that this is required.

In all the cells illustrated (Figs. 2 onward), the connections to other cells are shown in their correct physical positions. For this reason, no inter-cell wiring is shown: by packing the cells together in a rectangular array the correct inter-cell connexions are guaranteed. For this reason also, the incoming lines are not always labelled:

corresponding lines always enter and leave at corresponding points on opposite sides of the cell.

### 7. Propagation of Identity Bits

Before proceeding to more sophisticated arrays we will consider the means by which identity bits are propagated in the array. The mechanism described can be added with suitable modifications to any of the arrays. Figure 3 shows circuits for processing two bits, and this circuit is simply added to the cell of Fig. 2. These bits ( $I_{N1}, I_{N2}$ ) can be set to '0' or '1' by external signals; external signals can thus overwrite the contents of these bistables. The bistables can also be set by inter-cell signals; since propagation is normally into empty cells, only the 'set' is propagated. Read-out is initiated by the outgoing propagate pulse, and after read-out the bistable elements are cleared by the delayed pulse from  $D$ .

### 8. Representation of Different Speeds

A simple means of representing objects with different speeds will now be described. The circuits described are applicable to the linear register of Section 6, their application to planar arrays will be described in later sections. The minimum requirement is an ability to detect the speed of an object word at one part of the array (defined by external signals) and to ensure that this speed is maintained subsequently. (At this stage *direction* of motion is not discussed, since the array is linear and unidirectional.) Figure 4 shows a cell with suitable circuits. The simplest way to measure speed is in fact to measure the transit-time of a word through a cell, and generate a measure of this, the 'transit-time signal' ( $T$ ). This signal becomes part of the object-word, and is transmitted with it to an adjoining cell. In Fig. 4 the transit-timer ( $T$ ) is simply a timer which is cleared and started by the rising edge of  $C_N$ : it then runs until  $C_N$  returns to '0' (this is precisely the moment when  $T_N$  will be required by the next cell). Since only one cell is illustrated, we now consider what use is made of a transit-time signal  $T_{N-1}$  when it is supplied to cell  $N$ . The arrival circuit of cell  $N$  detects an arrival, as previously described; the circuit emits a brief pulse which simultaneously triggers the monostable elements and admits  $T_{N-1}$  from the previous cell (switches are shown as SW). Signal  $T_{N-1}$  then modulates the period of the monostable element, so that it precisely reproduces the

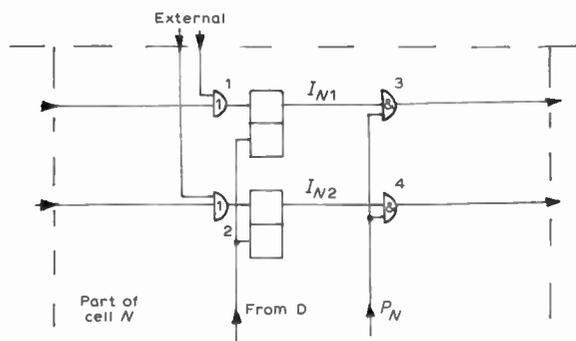


Fig. 3. Propagation of identity bits in a linear array.

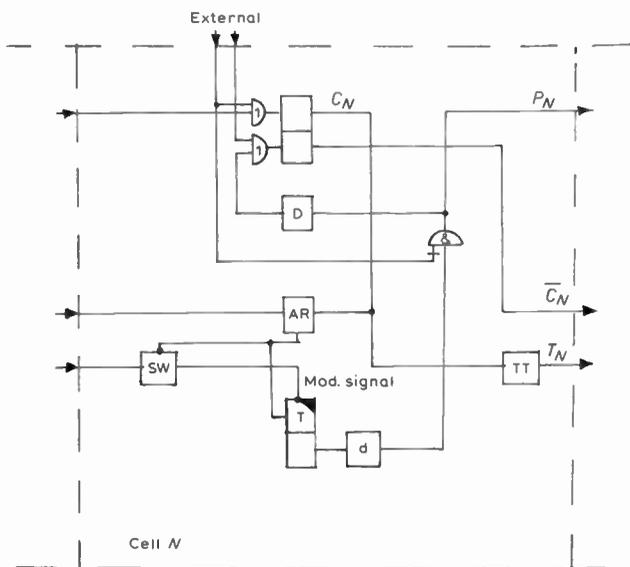


Fig. 4. Cell of linear array (variable speed).

transit-time of the previous cell. The rest of the cell mechanism is as Fig. 2, and as before, propagation will continue in the absence of external signals. It will be seen that we have in effect an asynchronous shift register, the individual parts of which are 'clocked' by the local contents.

Many of the elements described for the cell of Fig. 4, and also for the cells to be described later, are clearly digital, but some others could be realized by either digital or analogue means. In Fig. 4 the switch, timer and even the monostable element are of this kind. In analogue form the timer becomes an integrator, with an analogue switch, and the monostable element is modified from a standard C-R type. In purely digital realization the timer and monostable element become counters (which can operate from a common clock for all cells); the switch is then realized with AND gates.

Planar motion in *any* direction cannot be realized with this cell, but it is applicable to motion along fixed paths. Motion in and out of the cell in two or more directions can be represented by the addition of extra logic, switches and arrival-detectors; the other circuits are unchanged in number. A crude planar array is even possible, with up to six directions of motion; this would be achieved by an array of hexagonal cells. Traffic networks of any complexity can also be represented, provided that they are made only from interconnected one- or two-way tracks.

We now come to the problem of representing motion in any direction in a plane. The basic problem is that a single layer of regularly interconnected cells is necessarily anisotropic, whereas the space to be represented is isotropic. Some possible solutions are:

- (1) A single layer of randomly-connected cells; the randomness of connexions would ensure that over a large group of cells there was no preferred direction.
- (2) A multi-layer network, with many levels, each representing a different direction; this can be made as nearly isotropic as required.

- (3) A regular array which nevertheless appears isotropic to propagated signals.

Method (1) is not considered practical for hardware implementation; (2) and (3) are described below.

### 9. Multi-layer Array

This is an array to represent motion in any direction within a plane. It is provided with many different directional layers, so as to be as nearly isotropic as is needed for a given application. Its design is derived from that of the linear arrays previously described. Since a layer is required for each direction of motion it is inefficient in its use of hardware; it is described here mainly to suggest biological models. The upper layer consists of an array of rectangular compartments containing all the non-directional elements of the cell, together with the external connexions (layer 0 in Fig. 5.). This layer is characterized by the absence of inter-cell connexions. The operation is based on that of the linear cell described in Section 8. Below this layer lies a stack of layers (from layer 1 onwards). Each of these layers processes a particular direction of motion, and so contains only the elements processing directional motion. The function again resembles that of the corresponding parts of the linear cell. Each level is in fact a rectangular array of compartments all of which are aligned in the same direction. If, for example, it is required to define directions of motion within  $10^\circ$  then 36 uniformly rotated directional layers are needed. Uniform rotation of the levels is not strictly necessary: some directions may require more or less angular resolution. Due to the rotation of levels, not all parts of a cell lie directly below the corresponding compartment of layer 0; this results in some loss of positional accuracy in representing objects, and this must be made up by using a closer mesh of cells.

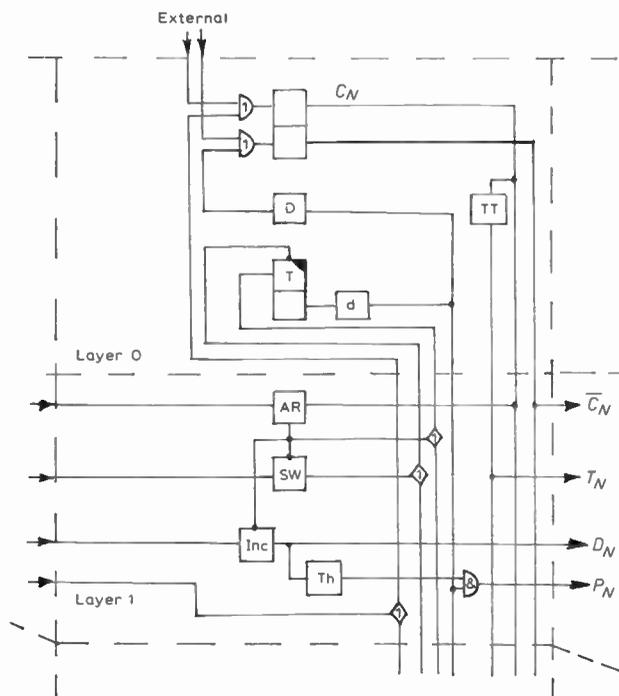


Fig. 5. Cell of multilayer array.

The mode of operation is that a propagate-pulse arrives at one of the lower layers, and sets the occupancy bistable element via a wired OR gate. The rising edge of  $C_N$  is then sent down the vertical 'cable' and picked up by the appropriate arrival-detector. The output of this triggers the monostable element and admits the appropriate transit-time signal. In due course the propagate pulse is sent down the cable, where it is routed to the correct level by an AND gate. The correct level is decided by a 'direction-defining circuit': this consists of the incrementer (INC) and threshold element (TH). The aim of these circuits is to establish in which layer a chain of arrivals of specified length has been detected. Each arrival pulse increments by a fixed small amount the value of the 'direction signal' (D). Through a chain of such cells the value of D eventually rises from zero to exceed a threshold value which is detected by all subsequent threshold elements. It will be seen that objects at rest are represented in the upper layer, while propagation takes place in the appropriate lower level.

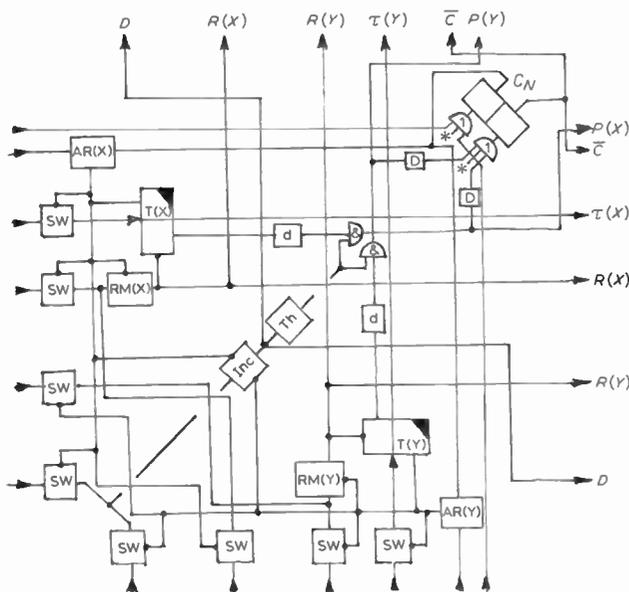


Fig. 6. Cell of multi-directional planar array.  
\* Denotes external input.

value in measuring the components of velocity of an object-word. Instead a method based on measuring the average rate of  $X$  and  $Y$  shifts is used. As a result of external signals the representation of an object appears successively in a chain of adjacent cells. The sensor system is to be adjusted so that the chain of cells concerned is linked at all points by vertical or horizontal displacements. The occupancy bistable element (Fig. 6) is set or cleared initially by external signals. Arrivals in the  $+X$  or  $+Y$  direction are detected by the corresponding arrival-detector, and fed into ratemeters (RM). The outputs ( $R$ ) of the ratemeters modulate the monostable elements with the characteristic that the period  $T$  is inversely proportional to the modulating signal. The arrival-circuit also triggers the monostable element which in due course initiates a propagate-pulse  $P(X)$  or  $P(Y)$ .

The ratemeter is unconventional: the meter within each cell receives either one or no pulses. This is clearly insufficient: the meter therefore picks up the last setting of the corresponding meter in the previously occupied cell, via the corresponding switch. (A simple R-C circuit is sufficient to perform this function for  $R$  in analogue form.) After a certain number of shifts the ratemeters will settle to a reliable value, and only then should inertial propagation take place in the array; a 'distance circuit' therefore inhibits propagate-pulses until a suitably long chain of occupancy is established.

The timing action of the monostable elements should ensure shifting at the measured rates in the  $X$  and  $Y$  directions: however there is a problem. Whichever monostable element restores first causes propagation in the corresponding direction. By the time the other monostable element restores it cannot contribute to the shifting, since the scene of activity has already transferred to another cell. It is therefore necessary that when one monostable element restores, the unexpired time of the other monostable element is transferred (as part of the object-word) to the new cell: there it should be set into the appropriate monostable element and allowed to run out. The path of the 'unexpired-time signal' ( $\tau$ ) is shown in Fig. 6. (Using conventional voltage-controlled monostable circuits this is simply achieved by transmitting the timing-voltages between the respective monostable circuits.) The input and output lines of the cell are again shown in their correct physical positions, so that it is merely necessary to assemble the cells into a rectangular array. It is interesting to note that in transferring the scene of activity along a chain of identical circuits, the function of some circuits is effectively changed. For example, we may note that the monostable elements in occupied cells are never quiescent: their action is equivalent to that of an *astable* multivibrator (with a high mark/space ratio) which is transferred along the line of occupancy. The ratemeters are similarly effectively transferred along the line of occupancy.

So far it has been implied that access to the array is via lines from every cell. Alternatively, an addressed system is possible. The array will then act like a matrix store, reading or writing being accomplished by the appropriate command accompanied by a cell address. Alternatively,

### 10. Multi-directional Planar Array

We will now consider an array which can represent motion in any direction in a plane, using only a single layer of cells. The basis of operation is that the velocities are resolved by the array into  $X$  and  $Y$  components. In accordance with the principle of resolution of vectors these components are processed entirely separately; in consequence the array *appears* isotropic to moving signals. Although this results in a rather complex cell it will be seen that very full use is made of the hardware. In consequence the array could perform fast, precise processing of large numbers of objects. Figure 6 illustrates the cell needed for processing velocities in the positive quadrant. Motion in all directions will require twice the number of switches (SW) and arrival-detectors (AR) also logic to process sign bits for the  $X$  and  $Y$  components of velocity. Measurement of a single transit-time is of no

interrogation by identity code could yield the corresponding addresses as a reply. The array with access unit becomes a complex package (Fig. 1) but one with very few input and output lines.

It will be seen that many of the elements in this cell are simply switches. Nevertheless an array of such cells constitutes a highly parallel system, with correspondingly high hardware costs. The uniform cellular structure should lend itself to microcircuit techniques with consequent reduction of costs. Ultimately the construction of an array with access unit on a single chip might be possible.

### 11. Applications

The arrays may be used in a supervisory or in a predictive capacity. Reference to the former has been made in Section 1, where the arrays were proposed as a means of maintaining an up-to-date map of a system when the input rate from the sensors is limited. Warning of imminent collision can be given by cell logic which detects occupancy of adjacent cells. For true collision warning the *predictive* capacity of the arrays can be used. In this application the main array, or a secondary array, can be run ahead in fast time to warn of the time and place of possible collisions. In this case actual collision can be allowed to occur (in the fast-time array): very simple cell logic can detect such a collision, which is represented by propagation into an already occupied cell.

The application of arrays in artificial intelligence was proposed in Section 1, where mapping in real time was described; the predictive property is equally relevant. A further application is now suggested. It is a character-

istic of human intelligence that not only can the observed present and expected future states of the environment be internally represented, but also 'supposed' situations and events can be represented, past, present or future. These supposed situations differ from real ones by the introduction of objects or interactions which are not found in the real environment. Supposed situations can readily be represented by the arrays described, by partially or wholly isolating the array from the sensor system, and injecting 'supposed' objects into the array from a suitable generator. (In this connexion, arrays can be designed which represent other properties of physical objects, such as the ability to adhere and form rigid assemblies.) The value of the 'suppose' feature lies in its ability to represent object-systems which are temporarily or permanently inaccessible, because of time or distance and also systems which never can or never have existed physically. These latter, for example, 'typical' or 'idealized' systems are a cornerstone in our understanding of the physical world.

### 12. References

1. Burstall, R. M., 'A survey of some recent work and objectives in the Department of Machine Intelligence and Perception', MIP-R-76, 1970, University of Edinburgh.
2. Raphael, B., 'Programming a robot', in 'Information Processing 68', pp. 1575-81, (North Holland, Amsterdam, 1969).
3. Feldman, J. A. and Sproull, R. F., 'System support for the Stanford hand-eye system', pp. 183-190 in Proc. Second International Joint Conference on Artificial Intelligence. (British Computer Soc., 1971.)

*Manuscript first received by the Institution on 1st June 1971, in revised form on 17th August 1971, and in final form on 27th September 1971. (Paper No. 1424/Comp. 139.)*

© The Institution of Electronic and Radio Engineers, 1971

## The Author



Mr. R. Baker studied at Queen Mary College, University of London, and obtained his B.Sc. degree in physics in 1959. For the next eight years he worked on research and development in industry, his appointments including that of senior engineer with Elliott-Automation and scientific staff member at the Hirst Research Centre of the General Electric Company. He was appointed a lecturer in electronics at Twickenham College

of Technology in 1967 and subsequently was appointed to a research fellowship at the College. Mr. Baker obtained the M.Sc. degree of the University of London for research on adaptive logic, and he is currently concerned with research into artificial intelligence in the Department of Engineering Technology at Twickenham.

# Maximum Gain, Mutual Coupling and Pattern Control in Array Antennas

By

JOHN F. McILVENNA,

B.S., M.S.†

and

CHARLES J. DRANE, Jr.,

B.S., M.S.†

This paper adds the concept of interelement mutual coupling to a previously presented design technique for aerial arrays. With mutual coupling accounted for, one can realistically design electronically scanned arrays of wire elements that have both maximum gain and a number of independent, steerable nulls in the radiation pattern.

## 1. Introduction

In an earlier paper, the authors presented a technique for maximizing the gain of an aerial array, while at the same time placing a number of independent nulls in the far-field radiation pattern.<sup>1</sup> Such a method effectively reduces unwanted interference or jamming without unduly decreasing the overall aerial performance. Although it was pointed out in that paper that the mathematical technique is generally applicable even when mutual coupling between the aerial elements is accounted for, the sample computations were restricted to isotropic or idealized point source elements and no inter-element coupling. Since that time, a generalized method for the design of arrays of thin wires (the element type most often encountered in arrays) has been combined with the previously presented maximum gain-constraint technique, and, most importantly, mutual coupling effects are accounted for.

Mutual coupling exists in all arrays. At radio frequencies the current-carrying conductors of an array will interact through their external electromagnetic fields, i.e. the current distribution on each will be modified because of the presence of the others. (If all the elements are fed from a single source, then the feeding network itself introduces additional inter-element effects.) Although coupling effects can be minimal for some types of elements and array geometries, e.g., short dipoles in a linear array, one may not neglect the electromagnetic mutual coupling in an array of long elements, which is electrically scanned over a broad range of angles. Since electrically scanned arrays of wire elements represent a widely used class of arrays, they form the basis of the analysis in this paper.

## 2. Mutual Coupling in Arrays of Wire Antennas

The design method briefly outlined below is due to Strait and Harrington; it has been reported in detail.<sup>2-5</sup> Each wire element in the array is subdivided into a number of segments (ten per wavelength has proved to be sufficient for far-field design problems) whose end-points are treated as the terminals of a multi-port electrical network. If one defines column matrices  $I$  and  $V$ , whose components are the  $M$  segment currents and voltages, one can write

$$V = \bar{Z}I, \quad \dots\dots(1)$$

where  $\bar{Z}$ , a symmetric, complex square matrix, is the

† Air Force Cambridge Research Laboratories (AFSC), L. G. Hanscom Field, Bedford, Massachusetts 01730, U.S.A.

mutual impedance matrix relating the voltage and current on any segment to the voltage and current on every other segment. The elements in  $\bar{Z}$  depend only on the segment length (a common assumption is that all segments are equal) and the intersegment distances. Relations between the elements in the  $\bar{Z}$  matrix and methods to simplify their computations are available in the literature.<sup>6</sup> Thus, once an array geometry has been specified, the mutual impedance matrix  $\bar{Z}$  can be calculated; and upon choosing a set of driving voltages, the currents on every segment, with the effects of coupling accounted for, can be determined by

$$I = \bar{Z}^{-1}V = \bar{Y}V,$$

where  $\bar{Y}$  is the mutual admittance matrix. In terms of these currents, the far-field amplitude radiation pattern is given by

$$E(\theta, \phi) = K_0 \sum_{n=1}^M I_n \exp [jk(x_n \sin \theta \cos \phi + y_n \sin \theta \sin \phi + z_n \cos \theta)] \quad \dots\dots(2)$$

where  $x_n$ ,  $y_n$  and  $z_n$  are the Cartesian coordinates of the centre of the  $n$ th segment,  $M$  is the total number of segments in an  $N$ -element array,  $\theta$  and  $\phi$  are the angular coordinates of the far-field point,  $K_0$  is a constant, and the  $I_n$  are the segment currents (the components of  $I$ ). Defining an angle-dependent row vector  $F$  with  $M$  elements, i.e.,

$$F_n(\theta, \phi) = \exp [jk(x_n \sin \theta \cos \phi + y_n \sin \theta \sin \phi + z_n \cos \theta)], \quad 1 \leq n \leq M$$

and replacing  $I$  by  $\bar{Y}V$  leads to the compact matrix form for eqn. (2),

$$E(\theta, \phi) = K_0[F(\theta, \phi)\bar{Y}V]. \quad \dots\dots(3)$$

Recall that the power pattern of an aerial is

$$P(\theta, \phi) = \frac{|E(\theta, \phi)|^2}{\eta_0},$$

and that aerial gain is

$$G = 4\pi \frac{\text{power radiated in a particular direction } (\theta_0, \phi_0)}{\text{total power input to array}}$$

where  $\eta_0$  is the impedance of free space.

For an array of centre-fed wires

$$\begin{aligned} G &= K_1 \sin^2 \theta \frac{V_D^*(F\bar{Y})_D^*(F\bar{Y})_D V_D}{V_D^*(\text{Real Part } \bar{Y})_D V_D} \\ &= K_1 \sin^2 \theta \frac{V_D^* \bar{A} V_D}{V_D^* \bar{B} V_D}, \quad \dots\dots(4) \end{aligned}$$

where the subscript D denotes the  $N$ -element remnants of  $V$  and  $F\bar{Y}$  and the  $(N \times N)$ -element remnant of  $\bar{Y}$  after all terms not associated with the driving points have been discarded<sup>7</sup>, superscript \* denotes complex conjugate transpose and  $K_1$  is a constant. Equation (4) is the desired starting point for the constraint technique. It is an expression for the gain of an array of centre-fed wire elements, which includes all the effects of mutual coupling. It also verifies a claim made in our earlier paper that the gain relation with mutual effects included is still a ratio of quadratic forms,  $\bar{A}$  is a one-term dyad,  $\bar{B}$  is positive definite, and both matrices are Hermitian. Thus, as described in our earlier paper, the procedures for maximizing gain while placing nulls in the pattern can now be applied to equation (4).†

### 3. An Example

Consider the problem of designing a linear array of straight wire elements which provides maximum gain and whose beam can be scanned all the way from broadside to endfire in the principal H-plane. In addition, at every scan angle in that same plane, the radiation level must be held at or below  $-30$  dB over a  $4^\circ$  angular sector adjacent to the main beam. A practical application of this design is to a search or tracking array being jammed by a source whose angular location is not accurately known. Note also that because of the large scan angles involved, it is important to include mutual effects in the solution.

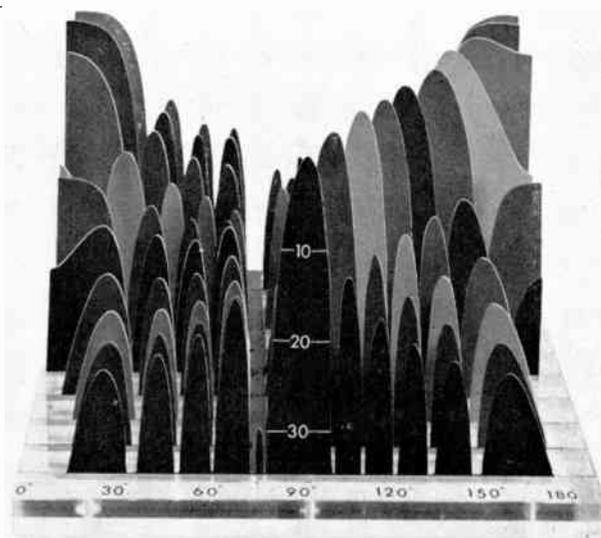
One of the most common linear array configurations utilizes half-wavelength elements uniformly spaced at half-wavelength intervals. This geometry was selected for use in the above problem; there were twelve elements and five segments per element. The minimal radiation level sector was created by using two constraints that placed nulls in certain directions in the H-plane ( $\theta = 90^\circ$ ); these nulls were separated by  $4^\circ$ , and their locations were fixed as the beam was scanned in  $10^\circ$  increments from broadside ( $\theta = 90^\circ, \phi = 90^\circ$ ) to endfire ( $\theta = 90^\circ, \phi = 180^\circ$ ). The results of the computations are shown in Fig. 1, where the principal H-plane pattern at each of the scan angles is shown. The horizontal axis is the scan angle, while the vertical axis represents decibels below the main beam maximum value. Each pattern has been normalized by the peak value for comparison purposes.

Note the presence of a well-defined trough that is at least  $4^\circ$  wide at every scan angle and in which the highest radiation level is  $-30$  dB for the broadside pattern and at least  $-35$  dB for the other scan angles.

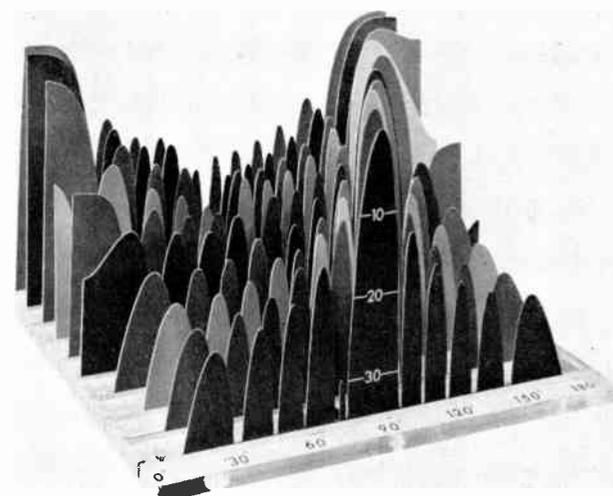
Figure 1 also shows the gradual (and expected) broadening of the main beam and the build-up of the endfire lobes as scan angle increases, until, at  $\phi = 180^\circ$ , the pattern consists of two main beams, oppositely directed. This splitting of the main beam into two equal beams is a characteristic of the half-wavelength spacing.

† The use of centre-fed wires in this discussion does not imply any restrictions on the mutual coupling method. Indeed, techniques that handle arrays of arbitrarily bent wires with resistive or reactive loadings at any point along their lengths have been developed and are available in the form of computer programs. See references 6 to 9.

Table 1 provides the data necessary for a comparison of the gains and driving-point voltage distributions for the unconstrained and constrained patterns. Note in particular in Table 1(a) that little gain must be sacrificed to achieve the desired degree of pattern control in the constrained patterns. The largest loss occurs in the broadside pattern, where it is necessary to reduce the first (and largest) secondary peak from  $-13$  to  $-30$  dB; yet this results in only a 4% loss in gain. For all the other scan angles, the gain decrease is even less. Generally speaking, the decrease in gain value caused by pattern constraints is proportional to the value of the unconstrained pattern in the immediate vicinity of the



(a)



(b)

Fig. 1. Principal plane radiation patterns for an array of dipole elements. The main beam is scanned from broadside ( $90^\circ$ ) to endfire ( $180^\circ$ ) in  $10^\circ$  increments. All patterns have maximum gain and include all mutual coupling effects. The trough represents a  $4^\circ$  sector in which the power level was constrained to be below 30 dB. (see a). Changes in pattern structure as scan angle increases include main beam broadening and build-up of the endfire lobes (see b).

**Table 1.** Comparison of unconstrained (U) and constrained (C) array properties.

(a) Array Gain.

Scan Angle 90° (broadside)	100°		110°		120°		130°		140°		150°		160°		170°		180° (endfire)			
	U	C	U	C	U	C	U	C	U	C	U	C	U	C	U	C	U	C		
<b>GAIN</b>	25.92	24.97	25.29	24.93	23.59	23.39	21.26	21.09	18.80	18.54	16.59	16.26	14.85	14.76	13.71	13.42	13.34	13.22	13.32	13.20

(b) Voltage Distribution (Amplitude).

Element Number 1	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
2	0.97	0.83	0.84	0.74	0.76	0.73	0.77	0.79	0.82	0.97	0.92	1.15	1.01	1.06	1.05	1.17	1.09	1.16	1.10	1.14
3	0.97	0.83	0.94	0.97	0.86	0.94	0.77	0.84	0.73	0.75	0.82	0.85	1.00	1.03	1.07	1.19	1.12	1.21	1.15	1.22
4	0.97	0.90	0.87	1.01	0.86	0.89	0.83	0.78	0.76	0.83	0.73	0.89	0.99	1.01	1.07	1.12	1.14	1.16	1.18	1.18
5	0.97	1.09	0.90	0.98	0.83	0.77	0.82	0.90	0.82	0.87	0.69	0.74	0.97	1.02	1.06	1.19	1.15	1.23	1.19	1.24
6	0.97	1.17	0.90	0.85	0.84	0.90	0.80	0.80	0.82	0.89	0.71	0.82	0.95	0.96	1.04	1.13	1.14	1.22	1.20	1.26
7	0.97	1.21	0.89	0.92	0.84	0.92	0.81	0.83	0.78	0.82	0.77	0.88	0.93	0.99	1.02	1.10	1.14	1.17	1.20	1.20
8	0.97	1.08	0.89	0.97	0.86	0.80	0.80	0.86	0.75	0.83	0.84	0.92	0.90	0.90	0.98	1.07	1.12	1.20	1.19	1.26
9	0.97	0.96	0.92	1.06	0.83	0.85	0.78	0.74	0.79	0.83	0.88	1.05	0.87	0.93	0.94	1.05	1.10	1.17	1.18	1.22
10	0.97	0.80	0.85	0.88	0.82	0.92	0.83	0.92	0.86	0.94	0.87	0.91	0.82	0.80	0.88	0.89	1.06	1.07	1.15	1.15
11	0.97	0.84	0.96	0.89	0.92	0.89	0.88	0.87	0.85	0.90	0.79	0.96	0.74	0.80	0.80	0.94	1.00	1.10	1.10	1.18
12	1.00	0.97	0.85	0.84	0.75	0.73	0.68	0.68	0.64	0.65	0.58	0.54	0.58	0.55	0.65	0.63	0.89	0.91	1.00	1.00

(c) Voltage Distribution (Phase).

Element Number 1	0°	0°	0°	0°	0°	0°	0°	0°	0°	0°	0°	0°	0°	0°	0°	0°	0°	0°	0°	0°
2	-18°	-16°	-15°	-9°	-7°	2°	2°	12°	6°	17°	9°	9°	9°	5°	8°	1°	5°	2°	4°	1°
3	-10°	3°	-11°	1°	-12°	-7°	-6°	-5°	3°	2°	15°	19°	14°	14°	12°	13°	8°	3°	5°	5°
4	-15°	6°	-10°	-3°	-8°	-10°	-6°	-1°	-3°	7°	14°	12°	16°	23°	16°	13°	10°	9°	6°	6°
5	-12°	7°	-11°	-12°	-9°	-3°	-3°	3°	-2°	-1°	8°	10°	18°	19°	19°	17°	11°	9°	7°	4°
6	-13°	0°	-10°	-10°	-11°	-3°	-4°	-3°	1°	8°	3°	3°	19°	16°	22°	20°	12°	13°	7°	7°
7	-13	-7°	-11°	1°	-11°	1°	-4°	5°	2°	4°	0°	0°	21°	21°	44°	42°	13°	12°	7°	6°
8	-12°	-13°	-10°	1°	-11°	-8°	-3°	-2°	-2°	6°	2°	3°	22°	20°	26°	23°	14°	12°	7°	5°
9	-15°	-17°	-9°	-4°	-10°	-1°	-4°	0°	-5°	-4°	6°	5°	23°	22°	28°	28°	23°	16°	6°	7°
10	-10°	-6°	-9°	-11°	-15°	-13°	-7°	0°	-3°	8°	12°	16°	24°	23°	30°	27°	15°	13°	6°	4°
11	-18°	-5°	-12°	-8°	-12°	-11°	-1°	0°	4°	3°	19°	13°	24°	22°	31°	27°	14°	12°	4°	3°
12	0°	21°	5°	18°	0°	10°	8°	20°	9°	26°	19°	29°	26°	20°	30°	35°	12°	15°	0°	2°

(d) Linear Progressive Phase Script Required to Scan Beam.

	0°	31°	62°	90°	116°	138°	156°	169°	177°	180°
--	----	-----	-----	-----	------	------	------	------	------	------

enforced nulls. For example, as seen above, at broadside the constraint region includes a significant secondary pattern peak and the gain loss is largest; when the unconstrained pattern value in the constraint region becomes smaller the gain loss is less.

An examination of the voltage amplitude distributions in Table 1(b), (each distribution is normalized to the voltage for its first element) reveals that these maximum gain voltages, constrained and unconstrained, are all non-uniform and most are non-symmetric about the centre of the array. The only two symmetric distributions (the broadside and endfire unconstrained cases) correspond to the only two radiation patterns that are symmetric about the broadside angle  $\phi = 90^\circ$ . Note that in some cases the voltage required for element 12 differs significantly from the voltages for all other elements. It also appears that for certain scan angles, e.g.  $150^\circ$ , the difference between the unconstrained and constrained amplitude distributions is small. As was true for the earlier comparison of the corresponding gain values, the explanation here is that at some scan angles the unconstrained pattern values are either very small in the constraint region, or the direction for one of the unconstrained nulls is nearly the same as that of one of the constraints. The constraint condition therefore requires only a slight perturbation of the unconstrained voltage distribution in these cases. Finally, note that in no case is the dynamic range of the driving-point voltages excessive. Compared to the ten-to-one variation usually taken as the practical limit of realizability, the range for both constrained and unconstrained optimum distributions is less than about two-to-one, and the distributions should offer no construction problems. The phase distributions for these maximum gain arrays can be considered to consist of the superposition of several components. The first is the uniform progressive phase taper that steers the beam to the desired scan angle and is common to all electronically-scanned arrays, constrained or unconstrained. The second taper is that required to maximize the gain. And in the constrained cases, an additional taper is required to place the pattern nulls as prescribed. Table 1(c) shows the unconstrained and constrained normalized phase distributions, after the beam steering phase distribution (shown for reference in Table 1(d)) has been removed. As in the case of the voltage distributions, note here also that the phase distributions are non-uniform and that most are non-symmetric. Again, small differences between unconstrained and constrained phase distributions are related to constraints being applied in directions where the unconstrained pattern levels are already small. Finally, keep in mind that in the examples discussed in this paper, and indeed in most expected applications, even though the designer directly controls the sidelobe level in only some local region of the radiation pattern, additional control, as a direct result of the application

of these constraints, will not become necessary over the remainder of the sidelobe region.<sup>10</sup> This is due to the general phenomenon that reasonably low sidelobes are an intrinsic concomitant to maximization of gain.

#### 4. Conclusions

We conclude that even with mutual coupling accounted for, the maximum gain-constraint technique provides realistic and practical solutions to array design problems requiring localized pattern control. The technique is applicable to practically all arrays of wire elements, regardless of their geometric arrangement and their electrical loading characteristics. The required relations and equations can be cast in matrix form and are ideally suited to computer manipulation, making them useful in adaptive or real-time situations requiring rapid pattern reconfiguration.

#### 5. References

1. Drane, C. J. and McIlvenna, J. F., 'Gain maximization and controlled null placement simultaneously achieved in aerial array patterns', *The Radio and Electronic Engineer*, 39, No. 1, pp. 49-57, January 1970.
2. Strait, B. J. and Adams, A. T., 'Analysis and design of wire antennas with applications to EMC', *I.E.E.E. Trans. on Electromagnetic Compatibility*, EMC-12, pp. 45-54, May 1970.
3. Chao, H. H. and Strait, B. J., 'Radiation and scattering by configurations of bent wires with junctions', *I.E.E.E. Trans. on Antennas and Propagation*. To be published.
4. Strait, B. J. and Hirasawa, K., 'On long wire antennas with multiple excitations and loadings', *I.E.E.E. Trans. on Antennas and Propagation*, AP-18, pp. 699-700, September 1970.
5. Hirasawa, K. and Strait, B. J., 'Analysis and Design of Arrays of Loaded Thin Wires by Matrix Methods', Scientific Report No. 12 on Contract F19628-68-C-0180, Report No. AFCRL-71-0296, May 1971.
6. Strait, B. J. and Hirasawa, K., 'Computer Programs for Radiation, Reception and Scattering by Loaded Straight Wires', Scientific Report No. 3 on Contract F19628-68-C-0180, Report No. AFCRL-69-0440, October 1969.
7. Strait, B. J. and Hirasawa, K., 'Computer Programs for Analysis and Design of Linear Arrays of Loaded Wire Antennas', Scientific Report No. 5 on Contract F19628-68-C-0180, Report No. AFCRL-70-0108, February 1970.
8. Chao, H. H. and Strait, B. J., 'Computer Programs for Radiation and Scattering by Arbitrary Configurations of Bent Wires', Scientific Report No. 7 on Contract F19628-68-C-0180, Report No. AFCRL-70-0374, September 1970.
9. Mautz, J. R. and Harrington, R. F., 'Computer Programs for Characteristic Modes of Wire Objects', Scientific Report No. 11 on Contract F19628-68-C-0180, Report No. AFCRL-71-0174, March 1971.
10. Sanzgiri, S. M. and Butler, J. K., 'Constrained optimization of the performance indices of arbitrary array antennas', *I.E.E.E. Trans. on Antennas and Propagation*, AP-19, pp. 493-8, July 1971.

*Manuscript first received by the Institution on 26th July 1971 and in final form on 12th October 1971. (Paper No. 1425/AMMS 44.)*

© The Institution of Electronic and Radio Engineers, 1971



