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. 10 OCTOBER 1971

THE RADIO AND ELECTRONIC ENGINEER

The Journal of the Institution of Electronic and Radio Engineers

The United States' Offer

AMIDST present-day problems and argument on the economic benefits or disadvantages of international operations, it is refreshing to learn of the offer recently made by the President of the United States of America. In short, Mr. Nixon has invited discussions on the whole of the American Space Programme and its various 'spin-off' benefits. The main agency for U.S. space activity has been the National Aeronautics and Space Administration who over the past decade have demonstrated such outward-looking attitudes to international collaboration as to promise well for this latest offer.

One area has been scientific research in space in which satellites have usually been designed to carry several different experiments, some measuring extra-terrestrial and galactic phenomena, others perhaps monitoring emissions from the Earth or the properties of its diverse layers. It is not unrealistic to assume that the project management for any particular U.S. satellite could easily have obtained a full complement of experimental apparatus from university or governmental laboratories within the U.S.A.; it may also be assumed that integration of these experiments into the spacecraft would have been less fraught with complications if there were such common origins, even if only of screw-threads, not to speak of more intricate areas calling for standardization such as interfaces! Nevertheless many U.S. research satellites in the past few years have included experiments designed and engineered by British or other European universities or research establishments. It is pleasant to record that the guest experiments have generally proved successful: indeed, quite recently the British Meteorological Office's experiments to measure infra-red emission from the Earth, carried in the *Nimbus IV* meteorological satellite, were of such potential value that instead of having to apply to take part in the next project, a firm invitation has already been made by NASA for a similar experiment on a future *Nimbus*. While the provision of launch vehicles for the satellites of Britain, Canada and other countries has called for some payment in cash, it has usually been nominal in amount: thus the very effective *Ariel* series, the Canadian *Alouette* satellites and various E.S.R.O. satellites can well be regarded as having been that much better because of the advantages of U.S. experience in launching into orbit. There has also been close contact with the Americans in the design of the spacecraft themselves.

Mention must also be made of the enormous U.S. contributions made in the design, launching and exploitation of the communications satellites which have certainly revolutionized international television and have led to such a great expansion of telephone and data communications. The present operations of communications satellites are in the hands of an international commercial company but the initial impetus came from an invitation to all countries to participate in using facilities which had been set up. This is indeed the current state of affairs in two other projects of almost equally far-reaching importance—weather and navigational satellites.

The President's suggestion for co-operation and collaboration of other countries in NASA's post-*Apollo* programme has been taken up by *The Financial Times* who are planning a visit of European businessmen to NASA centres early in 1972. Those taking part will be able to learn at first-hand of NASA's technological achievements and their potential applications in other fields of activity. These include specialist subjects as diverse as: bio-medicine; measurement and control; communications; computing technology; management practices; propulsion and power generation; reliability and non-destructive testing; avionics; materials and structures; earth resources survey programme. The outcome of collaboration of this kind could well represent immense savings of research and development to any commercial firm able to seize the opportunity. It does on the other hand represent a contribution to the universal pool of knowledge by the United States which should be applauded.

Contributors to this issue

Mr. A. T. Sullivan obtained a B.Sc. in electrical engineering through a sandwich course at Woolwich Polytechnic in 1962, while with Standard Telephones and Cables, North Woolwich. He was then employed as a development engineer in the microwave Division of S.T.C., working first on waveguide components, then on i.f. amplifiers for 1800-channel radio links, and system testing. In 1966 he was appointed to a lectureship in the Department of Electrical and Electronic Engineering at Woolwich Polytechnic, now designated Thames Polytechnic.

Mr. L. E. Weaver entered the transmission laboratory of Standard Telephones and Cables in 1939, where he worked on the design of multi-channel telephone equipment, later specializing in the design of television distribution systems. Apart from leave of absence to take a B.Sc. degree at London University, he remained there until 1954, when he joined the B.B.C. Designs Department. In 1955 he was invited to form the television measurements laboratory, of which he is still the head. He is the author of two books on video measurements as well as a number of papers and monographs on television measurements and networks.

Mr. D. C. Broughton (Member 1969) qualified in electronic engineering at the Polytechnic, Regent Street, and Willesden Technical College. He joined the B.B.C. Designs Department in 1958, initially as laboratory technician, and in 1963 was promoted to engineer status. His interest in computers led him to seek further qualifications at the University of London Institute of Computer Science, where he obtained the M.Sc. degree in computer science in 1969. In 1970 he became the Designs Department Computer Applications Engineer.

Dr. A. Carrick graduated from Hatfield College, Durham University, in 1964 with a B.Sc. in chemistry and from 1964 to 1967 pursued research at Durham into the organo-metallic chemistry of germanium, obtaining his doctorate in 1967. He held a research fellowship at the National Physical Laboratory from 1967 to 1969 and he is now engaged in mass spectrometry research, including fast data acquisition, chemical information handling and related aspects of 'chemistry by computer', in the Chemical Standards Division of the Laboratory.

Mr. C. J. Kelley became senior engineer of the analogue computer Section at the Cranfield Computing Centre in 1965. During his stay at Cranfield he has designed and developed the hybrid computer complex and is currently engaged in expanding the system in both hardware and software. Although originally intending to be an industrial chemist, he became interested in electronics when he was assigned to this work during his period of national service with the R.A.F. He subsequently became head of the calibration and development section of Light Laboratories, Brighton, working on medical and industrial instrumentation. He then joined the computer section of the Government Communications Headquarters, Cheltenham, where he remained until he took up his present position.

Dr. S. H. Khan graduated with a B.Sc. degree from the University of the Panjab in 1961. He was awarded a Rhodes Scholarship at Oxford University where he read engineering science, taking the B.A. degree in 1964. He then carried out research on laser applications for a thesis for his B.Sc. degree which he obtained in 1965 at the end of his three years' scholarship period. Dr. Khan was then given a research contract from the Culham Laboratories of the U.K. Atomic Energy Authority and in 1968 received the D.Phil. degree of Oxford University for his thesis on 'The Laser Triggered Spark Gap', based on this research work, and which forms the basis of the present paper. In 1969 he returned to Pakistan and joined the Pakistan Atomic Energy Commission as a Senior Scientific Officer. At present he is head of a group working on laser design, fabrication and applications at the Pakistan Institute of Nuclear Science and Technology, Islamabad.

Biographical notes on Dr. J. G. Gardiner and Mr. G. White have been published in the May 1969 and December 1970 issues respectively.

U.D.C. 681.34.004.5

CHAMP: The Cranfield Hybrid Automatic Maintenance Program

By

C. J. KELLEY †

A preventive maintenance program for a hybrid computer complex is described which makes operational and static accuracy checks. Interfaces, logic elements and analogue elements are tested by sets of individual routines. Considerable time saving is achieved by automatic testing depending on the complexity of the test. An improvement in m.t.b.f. has been observed.

1. Introduction

A large proportion of the work published on hybrid computing has dealt with problem orientated techniques which require the participation of both analogue and digital computers. In the author's opinion, there is also great potential in the ability of digital computer, to simply control the operation of the analogue computer. Following the development of the Cranfield hybrid system, the initial impact was felt in areas such as the automatic set-up and check-out of analogue programs, and the recording of results and parameter settings. An immediate benefit of this was the ability to use the analogue computer for short periods, without the useful computation time being swamped by such time consuming operations as the manual setting of potentiometers. The resultant saving in computer time, occurring as it did at a period of heavy pressure on the machine, led to the survey of other areas which might yield benefits in time saving. One of the most obvious of these was the area of routine maintenance, which required, on average, at least one day per week to perform. The paper will describe how this particular problem was dealt with.

2. Background

The present computing centre at Cranfield Institute of Technology came into being in 1968. At this time an existing Ferranti *Pegasus* computer was replaced by an ICL 1905, which was installed in a new building together with an existing Applied Dynamics AD 256 computer, transferred from an older computing laboratory. At this time, a proposal was made to link the two computers to form a hybrid complex. As the 1905 would provide the Institute's total digital computing service, it was decided to design the interface to allow the machines to be used together as a hybrid, or separately. When the present ICL 1905 computer is fully linked into the hybrid system it will be used mainly on a time-sharing basis due to the heavy loading on the digital section.

In order to minimize disruption of the digital side during development it was decided to provide a small, relatively cheap, digital computer with which to develop the system, the machine chosen being a Digico *Micro 16*. During the initial stages it became apparent that the *M.16* would be adequate on its own for the less complex hybrid problems, setting-up procedures and problem checkout. On completion of the interfacing of the small computer and during the addition of the extra interfacing for the larger machine, the advantage of providing diagnostic routines was recognized, and the development of these commenced using the *M.16* computer. A block diagram of the complete hybrid system is shown in Fig. 1. As the programming effort proceeded, the possibility of implementing the total diagnostic program on the small computer began to emerge, a fact which would reduce the overall loading on the 1905 computer. Thus the small digital computer can now justify its inclusion in the system solely by its impact on preventive maintenance, without regard to its other uses in the set-up and check-out field.

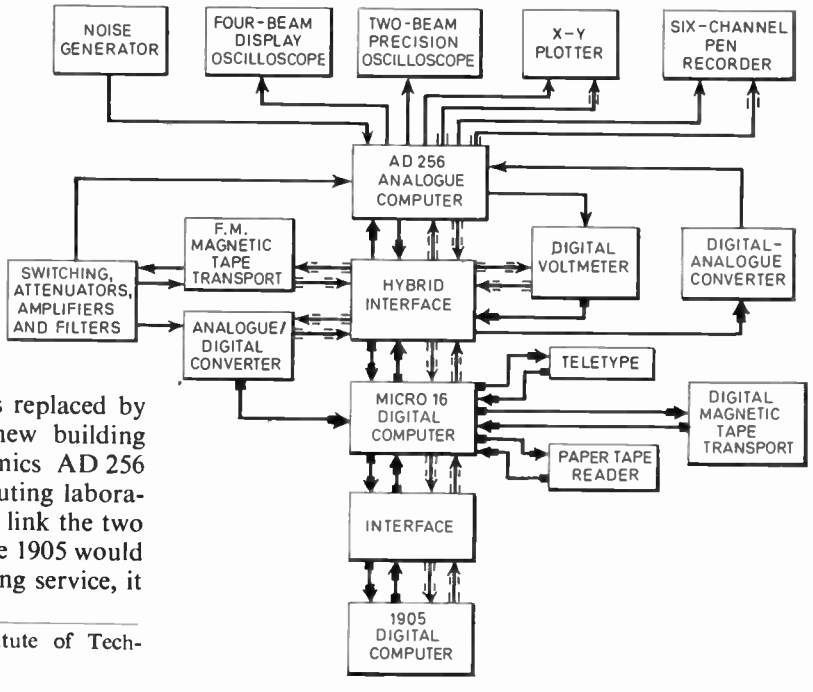


Fig. 1.
Block diagram of the Cranfield hybrid system.

† Cranfield Computing Centre, Cranfield Institute of Technology, Cranfield, Bedford.

3. Philosophy

The preventive maintenance program has been developed around a set of basic requirements and now has the following form.

The software consists of three main programs, and sets of individual test routines, for each type of component. The subroutines may be called sequentially or individually under control of the computer control panel switch-bank. The facility to call an individual test is useful to the maintenance engineer when carrying out repair work, as components may be tested rapidly before and after repairs. It has also proved useful to users who suspect that a component has developed a fault during a run. Owing to the limited size of the store in the small computer, the series of tests has had to be broken down into two groups. However, as a further development, these are being allocated a block of disk storage in the large computer and it will be possible to call either group into store almost instantaneously as far as the operator is concerned. In the initial stages the programs will be held on magnetic tape as this facility will be available on the *M.16* computer, before access is possible to the ICL 1905 disk storage.

As far as possible in these programs, all the tests are carried out on individual components. If a fault is indicated, this eliminates the need for the engineer to carry out further tests in order to separate a faulty component from a group. If more than one component is involved in a test, then a process of elimination is used. An example of this is shown in the hard limiter test. In this case the limiters are patched around amplifiers, and therefore the amplifiers are checked separately as well as being involved in the limiter test.

The previous example of combined components has some bearing on the order in which the tests are carried out. In the example given, the limiters are checked first at zero and then at ± 100 V settings in order to assure that they are set to the maximum value when the summers are tested. The prime example of this is the testing of all interface lines and logic elements prior to the testing of the analogue components. In order to do this, two logic patch boards are required, one for pure logic tests and one for the control of the analogue tests.

This brings us to the question of the number of patchboards involved. Obviously the number must be kept to a minimum for ease of operation and for economic reasons. In our case we currently use two logic boards, one trunk board and one analogue board. This set of boards is sufficient to check the interface, and logic elements, and to give operational and static accuracy tests on the analogue components.

At this time the dynamic checks, i.e. frequency response, phase shift, etc., are not included in the test program. It is intended to write further software to cover this and the extra patchboard involvement will be limited to the one interchangeable quadrant of an analogue patchboard, which is currently used for the dynamic multiplier tests.

The tests carried out with the existing software are limited to operational checks and static accuracy checks.

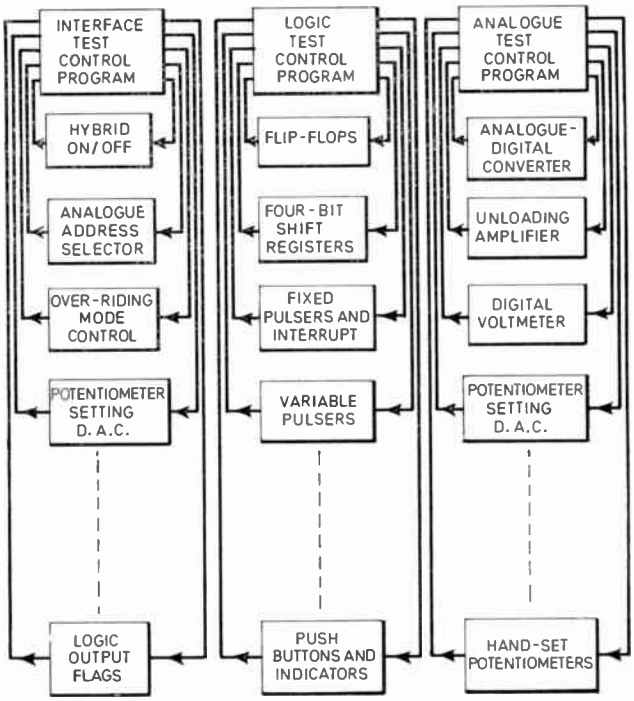


Fig. 2. Partial flowchart of the overall test program.

However, these are sufficient to cover all cases of complete component failure and many cases of component variation. For example, operational checks are carried out on trunks, potentiometers, track hold units, etc., and accuracy checks are carried out on integration capacitors, hard limiters, comparators and other components.

One of the most important aspects of any automatic system such as this, is the minimizing of operator effort. In this respect, the degree of operator intervention has been limited to the selection of the required test and whether a full print-out is required or just a print-out of faulty or out-of-tolerance components. There are also a few essential operations on components which can only be operated manually. In order to assist the operator with these operations the program prints out the necessary instructions whenever some action is required.

The overall test program is broken down into sections as shown in the following listing (see also Fig. 2).

1. Interface test
2. Logic element test
3. Analogue element test

4. Interface Test

The hybrid interface is essentially a device for matching the digital computer input/output to the control and data lines of the analogue computer. The interface test is designed to ensure that the digital computer output and input signals are correctly routed to and from the analogue computer, and that they are transmitted without error.

One method frequently used to check an interface is to provide direct feedback between its output and input lines, in order that it may monitor its own output signals.

This method is unsuitable for the hybrid interface, as the analogue computer requires certain control functions from the interface output, which have no equivalent interface input lines (i.e. analogue address lines, mode control lines, etc.). Owing to the difference between the number and type of interface input/output lines any feedback system would require changes of voltage levels and would present distribution problems

To overcome these difficulties it was decided that, rather than involve additional sources of failure, in the form of extra electronic equipment, a simple line test would be more suitable. For the test, the interface output lines were terminated with indicator lamps and the interface input lines were fed from the correct logic levels via switches. The lamps and switches are housed in a box and are arranged in rows which correspond to the various control function groups as follows:

(a) Hybrid on/off	Lamp
(b) Analogue address selector	Lamps
(c) Over-riding mode control	Lamps
(d) Potentiometer setting d.a.c.	Lamps
(e) Potentiometer setting command	Lamp
(f) Logic input lines	Lamps
(h) Interrupt	Switch
(j) Potentiometer 'set' flag	Switch
(k) Potentiometer 'not set'	Switch
(m) Logic output flags	Switches

The test consists of a main program and a series of subroutines corresponding to the above lists, which are selected by the switchbank on the digital computer control panel. Tests (a) to (f) produce patterns of output signals on the lines so that the operator may observe the indicator lamps as they are illuminated. Tests (h) to (k) request the operation of the switches and indicate the correct or incorrect result by messages on the teletypewriter. Test (m) will read the settings of the switches and display them on a bank of indicators located on the digital computer control panel.

At present two multi-way connectors, which link the interface to the analogue computer, must be reconnected to the test box. As the analogue computer is switched between hybrid and non-hybrid operation via a bank of relays, it is proposed to modify the relay bank, so that in the non-hybrid position, the interface test box is automatically connected to the interface.

5. Logic Element Test

The AD 256 analogue computer is equipped with a large complement of individual parallel logic elements, which must be tested prior to the analogue element test, in which they are used to distribute the control functions. A circuit diagram of the overall logic test is shown in Fig. 3. Each of the lower eight computer input lines is fed from a network similar to that shown on the lowest line, except for slight variations when the number of units to be tested is less than eight (i.e. variable pulsers). The upper four computer input lines together with sixteen a.d.c. lines are used to monitor the outputs from the NOR gate test. As far as possible, all types of elements are tested independently, but owing to the limited number of computer input/output lines it was necessary to use some of the NOR gates in each of the other tests. It can be seen from Fig. 3, that any of the elements shown in the network could give rise to a permanent fault condition on that line if it was in a 'permanent logic 1' condition. In order to help overcome this condition the program is provided with a slow speed alternative selected on the computer control panel switchbank, which enables the operator to observe the indicator lamps associated with each element, and thus follow the sequence of events. The test consists of a main control program and a set of subroutines, which correspond to the types of logic elements as in the following list.

(1) Flip-flops
(2) Four-bit shift registers
(3) Fixed pulsers and interrupt
(4) Variable pulsers
(5) NOR gates
(6) Preset down counters and master clock
(7) Push buttons and free indicators

5.1 *Flip-flops*

Eight flip-flops or bistable elements are available in the analogue computer, each having three inputs termed 'set', 'clear', and 'trigger', two outputs and two complemented outputs. It is considered sufficient to monitor one of the two outputs, as any fault, except a break in the other output connexions, will eventually be indicated at this output. This philosophy is applied whenever possible during the logic tests in order to minimize the number of computer input lines required.

The test commences with the clearing of all flip-flops and shift registers, which is necessary in order to avoid the presentation of spurious signals to the input lines. The flip-flops are then cycled through an operation sequence consisting of set, set, clear, trigger, trigger, clear, clear while the outputs are monitored. Any incorrect output or sequence of outputs produced by the set of inputs will cause the following message to be printed out

F.F. X.X. Y U/S

where X.X. specifies the particular flip-flop, Y shows the first operation to fail, i.e. S (set), C (clear) or T (trigger), and U/S means unserviceable.

Throughout the test, a record is kept in store of all the output states for each flip-flop. In the event of the apparent failure of all flip-flops, a comparison of the output states is made at the end of the test, whereupon it is possible to locate a common fault, caused by a failure of one of the three drivers feeding all flip-flops. On leaving the subroutine, a record of failures, in the form of input line numbers, is stored in the main control program area. At the end of the complete logic test these numbers serve to indicate any line which has a continuous fault and could, therefore, indicate a faulty NOR gate or line driver at the computer input lines.

5.2 *Shift Registers*

There are four four-bit shift registers available in the analogue computer, one in each quadrant of the logic

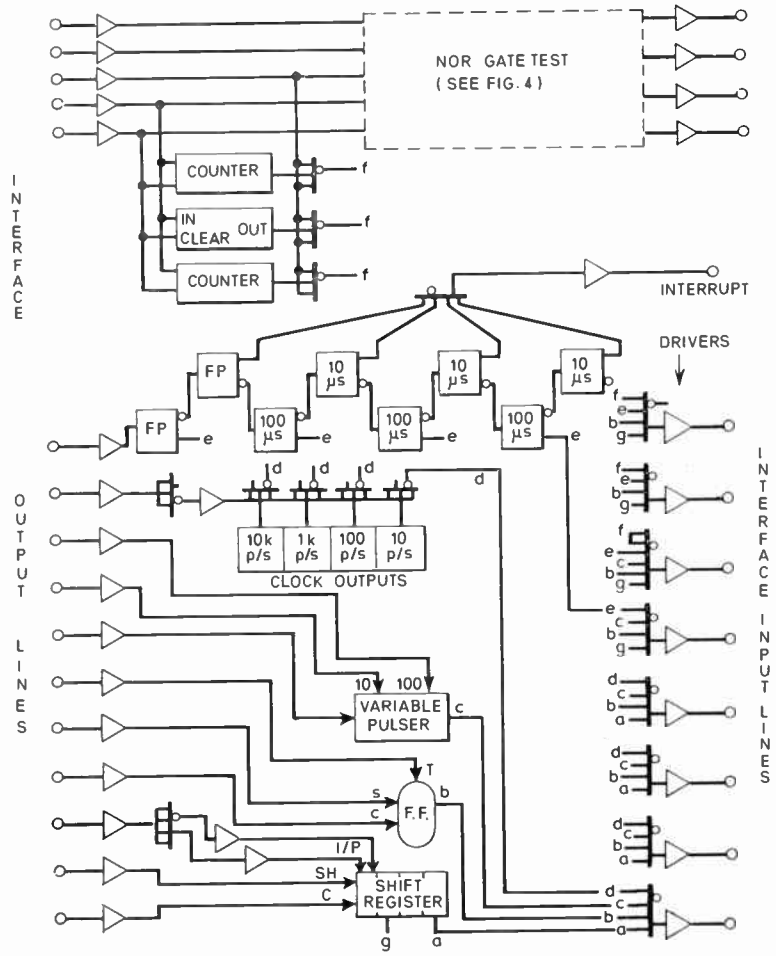


Fig. 3.
The circuit diagram of the logic element test, showing the connexion of one of each variable pulser, flip-flop and shift register.

complement. Each register requires an input and its complement and can provide the output and complemented output from each stage. The controlling inputs are shift (SH) and clear (C) where the shift requires a logic 0 to 1 transition and the clear requires a logic 1 in order to operate. Owing to the limited number of input lines to the digital computer, it is only possible to check the second and fourth stages of each register. However, this will localize any fault to two stages in any register and the isolation of the faulty stage is a fairly simple operation using the slow speed option.

The test commences, as before, with the clearing of all lines to remove any signals remaining after the previous test. Each shift register is then primed with a logic 1 at the input and given one shift instruction in order to set the first stage to a logic 1. The input is then replaced with a logic 0 and each register is given four successive shift instructions so that the initial input is cycled right through the register and finally lost at the end stage. The input is next set to a logic 1 and each register is again given four successive shifts so that all stages are set to logic 1. Finally all registers are cleared using the 'clear' input. The second and fourth stages of all registers are monitored throughout the test and an incorrect output will give rise to a printed output in the following form

SH.R. X.Y. U/S

where X is the number of the shift register and Y (L or U) indicates the lower or upper half.

The output produced by each shift instruction throughout the test is stored in the digital computer and, should all shift registers appear to be faulty, the stored results are used to locate the most likely common fault.

5.3 *Fixed Pulsers and Interrupt*

The present complement of fixed pulsers (monostable elements) on the analogue computer consists of four 10 µs units and four 100 µs second units. This is a modified version of the original system which contained eight 10 µs units, the alterations being made at the request of major users.

Owing to the relative magnitudes of the pulser time and computer cycle time (i.e. 10 µs and 6 µs) it is not possible to attempt the accurate timing of the fixed pulsers during this test. However, the operation of all pulsers, and the timing within broad limits, are checked to indicate any radical malfunction. As the 10 µs pulsers are generally used to provide the interrupts to the digital computer it was decided that they should be tested under these conditions. The implementation of the test was then determined by the fact that only one interrupt can be serviced at any time. In order to comply with this requirement the pulsers were connected in cascade, commencing with a 100 µs unit and followed by alternate 10 and 100 µs units (Fig. 3).

The fixed pulser test commences with the clearing of all lines followed by the application of a logic 0 to 1 transition at the input to the first pulser. The output of the first pulser is monitored until the first interrupt is received from a 10 μs pulser, whereupon this fact will be registered and monitoring of the second 100 μs pulser commences. If any pulser falls short of, or exceeds, its allotted time by approximately 10 μs, this fact will be recorded and indicated in the error print-out which has the following format:

P X.X Y

where XX indicates the particular pulser and Y can be S (short), L (long) or U/S (unserviceable).

A future project on the *M.16* computer is that of interfacing all test equipment such as counter/timer, waveform generator, programmable voltage standard, and digital voltmeters in order to enable their implementation in standard programs and with automatic test rigs. It is proposed to use the counter/timer in this mode in order to provide accurate timing of the short duration pulsers.

5.4 *Variable Pulsers*

The six variable pulsers have a continuously variable scale from ×1 to ×10 on three basic ranges, which are 10 ms, 100 ms and 1 s. The 10 and 100 ms ranges are selected by application of a logic 1 to holes which are labelled 10 and 100 respectively. When both lines are at logic 0 the basic pulse duration is 1 s. There are two possible methods of checking the timing of pulsers, one being the parallel method used by Hall and Peterson[5] and the other being the independent testing of each unit.

As the parallel method can only give relative timing between units and the independent method is somewhat time consuming, it was decided to combine the two methods in order to obtain the best features of each.

The method finally arrived at requires that all pulsers be activated at the same time in parallel, and the outputs are monitored either together or separately depending on whether the interval between the resetting of individual pulsers is sufficiently great to permit differentiation between them if all outputs are monitored at the same time. This means that for the maximum pulser width of 10 s, there is sufficient time to monitor all outputs together, and still note each pulser resetting individually, by using a subroutine to record the fact. We can, therefore, test all pulsers at 10 s with only one pass. On the other hand, for the minimum pulse width of 10 ms, the time interval between the resetting of any two is too short to be certain of monitoring the second one as soon as it occurs. In this case, although all pulsers are tripped at the same time, only one pulser output is monitored and timed, and thus the test requires six passes to check six pulsers. The change-over point in fact occurred in the region of 100 ms, so that we now monitor pulses up to 100 ms individually, and pulses in excess of 100 ms in parallel.

The test commences with the clearing of all lines followed by the print out of the following command

SET VAR PULSERS TO MIN & PRESS AI TO CONTINUE

On receipt of the attention interrupt† signal, the program will commence checking all pulsers at their minimum setting on each range, i.e. for pulse widths of 10 ms, 100 ms and 1 s. Any pulser which fails to operate or has a timing error of greater than ±10% produces the following print-out

V.P XX Y

where XX indicates the particular pulser and Y is S (short), L (long) or U/S.

On completion of this part of the test the following print-out is produced

SET VAR PULSERS TO MAX & PRESS AI TO CONTINUE

On receipt of this 'attention interrupt' signal, the program will commence checking all pulsers at their maximum setting on each range.

An alternative output format may be selected by the *M.16* control panel switch bank. This alternative will give a complete listing of the maximum and minimum pulse widths for all units on all ranges, as required for machine performance records.

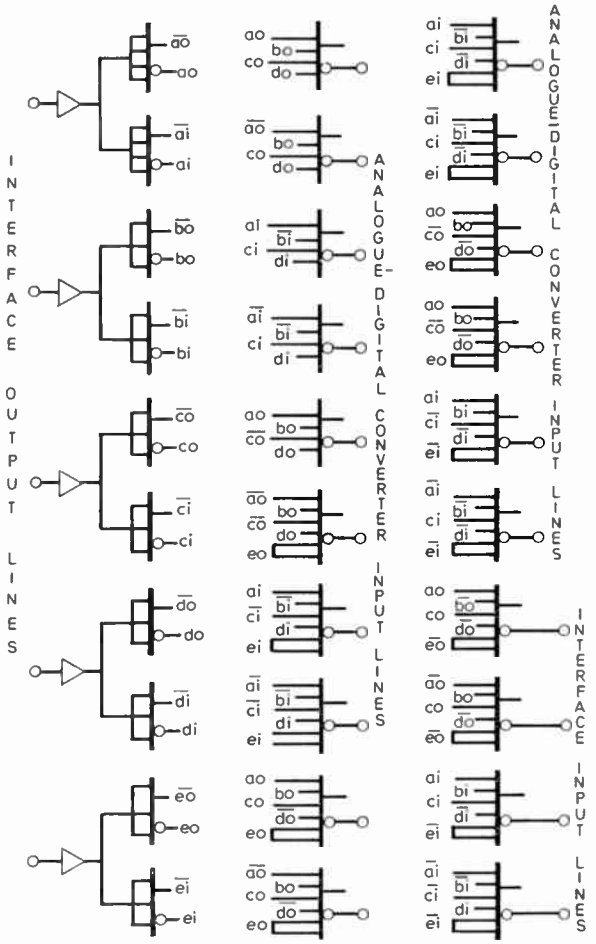


Fig. 4. Test circuit for the remaining thirty NOR gates, showing how varying numbers of inputs are dealt with.

† Operation of the attention interrupt button, which is located on the teletypewriter console, will pull down the common 'interrupt request' line and produce a computer interrupt.

5.5 *NOR Gates*

Of the total complement of forty-eight NOR gates, there are thirty which are unused in other tests. These gates are tested separately in this test. During the development of the test, an attempt was made to utilize only the remaining five computer output lines and four computer input lines. The method was based on the D-calculus as used by J. P. Roth.[9, 10] This approach was later found to be inadequate due to the complicated visual indication required during the slow speed option, and the difficulty in locating multiple faults.

By the utilization of sixteen a.d.c. lines it became possible to check twenty gates directly and the remaining ten by simple programming. The final test circuit is shown in Fig. 4. There are two input gates fed from each computer output line in order to overcome the loading difficulties which would be encountered using one gate. It can be seen that a single fault on one of the gates G_{11} to G_{30} will give rise to one easily discernible error. A fault on one of the gates G_1 to G_{10} will give rise to more than one error condition, but the pattern of the errors will indicate the particular faulty gate.

It is possible to determine the locations of several faulty gates by comparing the error patterns produced from the thirty-two combinations of input signals. However, in order to minimize the software, it was decided not to decode individual gates from multiple failures. This is a valid economy based on the fact that gate failures generally occur singly. If a multiple failure should occur the program will list all gates which could be faulty. It is then relatively simple for the operator to determine the faulty gates by observation of the indicators during the slow speed run.

5.6 *Master Clock and Preset Down Counters*

The crystal-controlled master clock and the preset down counters are part of a unit called the repetitive control unit. This unit, in addition to providing the above facilities, can be used to control the integrator initial condition, hold and operate modes. The master clock basic frequency of 100 kp/s is counted down to provide clock signals of 10 kp/s, 1 kp/s, 100 p/s and 10 p/s which are available on the logic patchboard and may also be selected as repetitive operation clock pulses by means of push buttons. For each of these basic frequencies the period may be varied over a range of 000 to 999 by thumbwheel switches. The inputs and outputs of the three counters are made available on the logic patchboard by means of three groups of three holes labelled IN, OUT and CLEAR.

As the repetitive operation periods are determined by the basic clock signals and preset counters, it is sufficient to check the operation of these two functions.

The test commences with the output of the message

SET COUNTERS TO 999, SELECT C & PRESS AI TO CONTINUE.

On receipt of the attention interrupt, the digital computer will clear all counters and input a series of pulses to them, while monitoring the outputs until all counters have indicated a completed count. During this period, the digital computer will count the input pulses and compare the final values with 999. Should any counter terminate prematurely, the following message will be printed

COUNTER X YYY

where X is the number of the counter and YYY is the value of its final count. If the input count exceeds 999 the following output will be printed

COUNTER X HIGH.

The master clock test consists of timing the output pulses from the connexions on the logic patchboard. To obtain sufficient accuracy of timing only the 10 kp/s, 1 kp/s, 100 p/s and 10 p/s outputs are timed, but any fault occurring in the primary stage will also be indicated by the test. As the clock pulses are always present, it is necessary to disable the connexions to the computer input lines except during this test. The gating of the clock pulses is accomplished by means of NOR gates which will only allow the pulses to pass through when a particular computer output line is set to logic 1 (see Fig. 3). If the timing of any stage is faulty the following message will be printed

XXXXX PPS CLOCK U/S

where XXXXX indicates the frequency of the first faulty stage.

5.7 *Push Button Switches and Free Indicators*

There are eight independent push button logic switches and twenty free standing indicator lamps installed in the analogue computer. The logic switches which can provide both the true and complemented outputs, may be connected as momentary contact or latched switches, and the pushbuttons may be illuminated from any logic 1.

In order to implement the testing of the push button switches they are connected in the latching mode, with the push button lamps supplied from the true output of the switch. A free indicator is also connected to both the true and complemented outputs of each switch. The remaining four free indicators are connected directly to logic 1 positions on the logic patchboard.

The test consists of operating all switches and observing the following events.

(a) The correct (true) free indicator should light;

(b) The push button lamp should light;

(c) The switch should latch.

The switches are then operated for the second time when the following events should occur:

(d) The correct (complemented) indicator should light;

(e) The push button lamp should be extinguished;

(f) The switch should unlatch.

As this test is obviously purely manual the computer is not involved except for the function of prompting the operator to carry out the test. This is achieved by simply printing out the following message

TEST LOGIC PUSH BUTTONS.

6. Analogue Element Test

This program controls the following subroutines:

1. Analogue-digital converter
2. Unloading amplifier (read-out buffer)
3. Digital voltmeter
4. Potentiometer setting d.a.c.
5. Servo-setting potentiometers
6. Reference and function switches
7. Hard limiters
8. Summers
9. Integrators
10. Comparators
11. Track transfer amplifiers
12. Analogue trunks
13. Diode networks
14. Multipliers
15. Sine-cosine generators
16. General purpose d.a.m.s.
17. Hand-set potentiometers

At present, the power supplies are not addressable and cannot be automatically checked, however, additional hardware is to be added to permit this.

6.1 *Analogue to Digital Converter*

Each of the sixteen channels is checked at zero and plus and minus reference. At this stage an error print-out is given for any channel requiring recalibration. The a.d.c. is then checked at frequent intervals throughout the range for linearity.

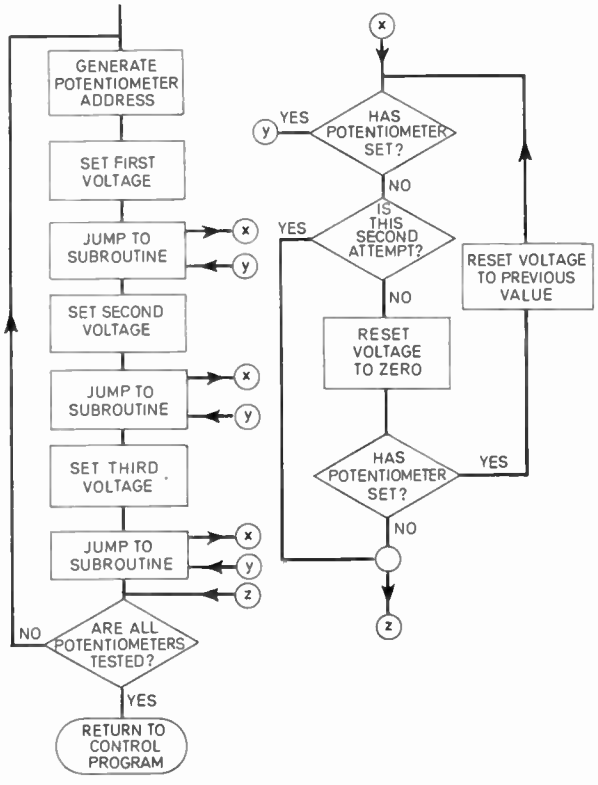


Fig. 5. The simplified flowchart of the servo-setting potentiometer test, including the SETΦT subroutine.

6.2 *Unloading Amplifier*

This amplifier is a buffer unit with an extremely high input impedance and low output impedance, followed by an inverting amplifier. When a component is addressed, the unloading amplifier is introduced in the output in order to minimize the effects of loading. The test consists of reading the input, positive output and inverted output on the a.d.c. for a range of values, comparing the three, and producing a printed output if the error specification is exceeded. Should it be desired, a hard copy of all readings may be obtained for the machine performance records.

6.3 *Digital Voltmeter*

The digital voltmeter test consists simply of comparing its output with the output of the a.d.c. for a range of positive and negative inputs. The operator may select a simple error print-out or a full print-out of the test results.

6.4 *Potentiometer Setting D.A.C. Test*

This unit which is different from the other digital to analogue converters in the system, is used as the reference against which the servo setting potentiometers are set. It is also possible to use the device as a slow speed digital to analogue converter by virtue of the fact that the output is available on the patchboard. This fact is utilized in the test, which consists of setting the unit through a range of values and comparing the reconverted output with the input for each setting. Once again the computer will indicate any setting which exceeds the specified error or will print out a complete range of settings for the performance records.

6.5 *Servo-setting Potentiometers*

In order to set a servo-setting potentiometer without the hybrid facility it is necessary to address the relevant potentiometers by push button switches, set up the appropriate reference voltage, again by push button switches, and finally to initiate the setting operation.

When testing the servo setting potentiometers this procedure must be carried out for at least three settings, i.e. upper end, lower end, and centre. Manually this is a very time-consuming process as there are 120 of these units on our machine.

The automatic servo setting potentiometer routine carries out the same sequence of events but eliminates all the manual operations, thus giving a time saving in the region of 80%.

In this test the computer print-out is always limited to giving the addresses of those potentiometers which are not setting correctly. The broad outlines of the routine are shown in the flowchart of Fig. 5. This is an example of the way in which the software is organized and serves to illustrate the general approach. Obviously, in this example there is a great deal left out, for instance, the error print-out is not indicated. However, it does show the use of a subroutine which will attempt to set a servo

setting potentiometer. This subroutine, which we call SETPϕT has been written in such a fashion that it may be used in general hybrid software whenever a potentiometer needs to be set. In order to avoid a fault being indicated in the case of a temporary failure such as dust on the track, the first attempt will be followed by an attempt to set to zero and then again to the desired value. Complete failure to set will give an error print-out, i.e. POT 3.A.04. NOT SET. It will be noted that the subroutine has two exits, one for the correctly set case and one for failure to set. The philosophy of implementing operations by general purpose subroutines has added considerably to the software library for hybrid use.

6.6 *Reference and Function Switches*

The reference switches, which can be operated logically or manually will provide either plus or minus reference voltages (i.e. ± 100 V). These switches are used to provide switching for the remaining tests in the program and must therefore be checked at an early stage. Each switch must be terminated with an amplifier, and it is the output of the amplifier which is monitored in order to check the switch operation. The automatic operation of the switches is first checked and this is then followed by a manual test under the direction of the digital computer. Instructions are given to the operator by teletype output as follows:

REFERENCE SWITCHES UP, PRESS AI TO CONTINUE

whereupon the operator must operate the appropriate switches and press the attention interrupt button. The output of each switch is checked within 10% in order to avoid a fault being indicated due to amplifier error. The correct output will cause the teletype to print out the next instruction, and an incorrect output will initiate an error print-out, i.e. RS 2A+U/S which will then be followed by the next instruction.

The function switches are manually operated single-pole double-throw units with the centre position off. These switches are checked in a similar fashion to the manual check on the reference switches with the inputs being provided from plus and minus reference voltage.

6.7 *Hard Limiters*

The limiters used in the AD 256 are active feedback types, giving exceptionally hard limiting. An equivalent circuit is shown in Fig. 6, where the adjustable voltage source is considered to have zero impedance and the diode is considered to be a perfect diode.

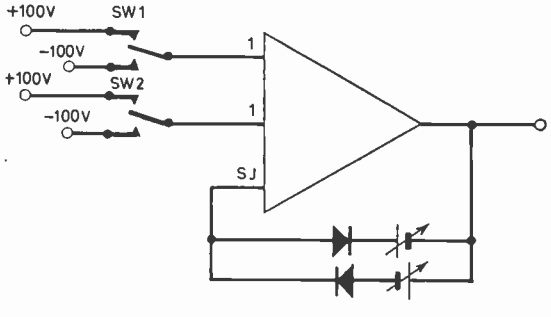


Fig. 6. Hard limiter test.

In order to use the limiters they must be patched between the output and summing junction of an amplifier, and the positive and negative limits are set by separate ten turn potentiometers.

The program used to test these elements outputs the message

SET LIMITERS TO MIN & PRESS AI TO CONTINUE

whereupon the operator must set all limiter potentiometers to zero and press the attention interrupt button. On receipt of the interrupt signal the program will apply +100 V via SW1 and SW2 to two unity gain inputs of the summing amplifier concerned and proceed to check the output voltages which should read +0·15 V. This will be followed by a reversal of SW1 and SW2 and a repeat of the read-out procedure to check the negative zero setting.

The check of the maximum limiter settings is similar to the previous check except that it is preceded by the message

SET LIMITERS TO MAX & PRESS AI TO CONTINUE

and the outputs should be $\pm 100{\cdot}15$ V.

The test will give a hard copy for performance records if desired.

This test is one of the few tests involving operator participation but this is kept absolutely minimal. At the end of the test the limiters are in the maximum condition which permits the summers to be checked.

6.8 *Summing Amplifiers*

Most of the tests in the analogue element test consist of a series of identical checks on each individual component of a component type, i.e. all potentiometers are checked for the same settings. This simplifies the programming, in that the same subroutine may be used for each component with the component address updated each time.

The summing amplifier test departs from the general approach and is broken down into a series of different tests. This is caused by the fact that the summing amplifiers are involved in the testing of several other components, they have different numbers of inputs, and some are convertible to integrators. The following list shows the amplifier configuration for which different subroutines are required and the number of summing amplifiers involved.

(i) 48 summing amplifiers convertible to integrators

(ii) 16 summing amplifiers convertible to track transfer units

(iii) 15 summing amplifiers used to terminate 30 multipliers (these summing amplifiers are convertible to 30 inverters)

(iv) 5 summing amplifiers used with limiters

(v) 12 summing amplifiers used to terminate digital–analogue switches and also to provide switched voltages for other tests.

(All the summing amplifiers are bipolar and consist of a standard summing amplifier followed by an inverter.)

The first three of the above groups are tested using the same basic subroutine with the exception of the switching which is necessary to convert them to summing amplifiers. Group (iv) does not require switching but requires an initial check to prove that the limiters are set to their maximum value. This check involves overloading the amplifiers and comparing the output with maximum limiter settings but with fairly large tolerances. Group (v) is comprised of several summing amplifiers with a variety of outputs and requires a completely separate subroutine which covers all units independently.

The general subroutine for testing the majority of the summing amplifiers consists of measuring both the polarities of amplifier output with a preset input. The summing amplifiers have different numbers of inputs ranging from two to eight, so that to test each input separately would require fairly involved switching. In order to overcome this, equal numbers of inputs are fed with plus and minus reference and the final output is determined by the minimum number of remaining inputs (i.e. one or two), see Fig. 7. This method will give an overall check on the summing amplifier performance and is supported by a secondary program which will completely check any summing amplifier which proves to be faulty or out of tolerance. For this, the summing amplifier in question has to be temporarily patched on a spare board.

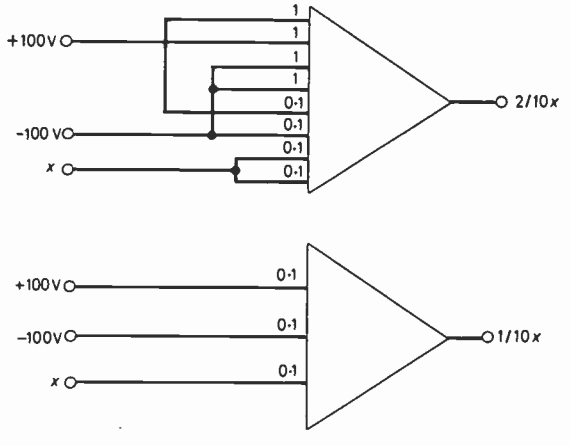


Fig. 7. Summing amplifier test.

6.9 *Integrators*

The summer/integrator units can be operated in three basic configurations, namely as summing amplifiers, as integrators or as electronic switches. When operating as integrators, the units have four switchable time scales each of which may be used in conjunction with either electronic or reed relay switching. This switching is associated with the three modes of operation of the integrators, which are initial conditions, hold and operate.

As can be seen from the previous paragraph, there are several combinations of operations which must be tested, if the summer/integrator units are to be fully tested. The operation of these units, as summing amplifiers, is checked in the overall summing amplifier test. If both the electronic and reed relay switches are operable in any state, then they are operable in all states. This fact permits the testing of some time scales on electronic switching and some on reed relay switching. The time scale tests are in fact divided equally between the two types of switching, i.e. $\times 1000$ electronic, $\times 100$ electronic, $\times 10$ reed and $\times 1$ reed. This permits the maximum amount of diagnosis with the minimum number of passes, a faulty switch giving rise to two errors and a faulty time scale giving rise to one.

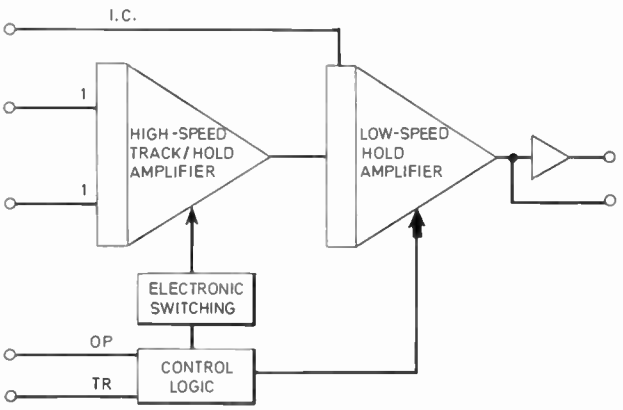
The integrator time scales are checked by the ramp method, which entails inputting a fixed level and allowing integration to take place for a fixed period of time. In order to simplify the error calculation, the integrators start with an initial condition of -100 V and an input of $+10$ V. Integration is terminated after a time period which would cause the output of a perfect integrator to become zero. The output in volts then gives the percentage error. The integration times range from 0·001 s for the $\times 1000$ time scale to 1 s for $\times 1$ time scale, the apparent factor of 10 difference between time and time scale being due to the use of a $\times 10$ input resistor which permits the use of longer time scales and increased accuracy. It may be possible to gain slightly improved accuracy by reducing the input voltage and increasing the ramp time by a factor of 10. However, this would also increase the time taken by the test. The integrator modes are also checked at the same time as the time scales by virtue of the fact that the integrator must be taken from initial conditions through operate to hold during the ramp period.

The final configuration of the summer/integrators is the switch mode. In this state the units will give an output equal to the initial conditions in one state and an output equal to the sum of the inputs in the other state.

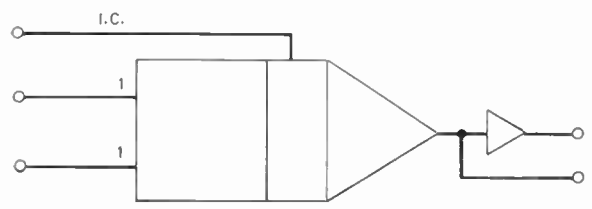
Apart from the purely operational testing on the integrators, they must also be tested for drift due to capacitor leakage. For this test the integrators are switched to initial conditions with an i.c. output of $+100$ V (reference voltage) for a period of two minutes. At the end of this time all units are switched into the hold mode and all outputs are immediately scanned by the a.d.c. through the address selector. After a further period of one minute all outputs are once again scanned by the a.d.c. in the same order and then the first reading is subtracted from the second reading for each unit. This subtraction will remove the effects of any skew introduced by the time taken to scan all outputs, sequentially, and will give a value of percentage drift. This test is repeated for each of the four time scales giving a total test time of twelve minutes. Although this may seem a long period of time it must be noted that an individual drift test on each integrator would take at least 96 hours.

6.10 *Comparators*

The comparator units are essentially analogue to digital communication devices. They each have two analogue inputs which are compared in magnitude, and give a logic output, the state of which depends upon which analogue input is the greater. In order to prevent



(a) Simplified circuit, showing component units.



(b) Circuit symbol.

Fig. 8. Track transfer unit.

instability around the switching point a hysteresis gap is introduced and with these particular units the gap is ± 10 mV. It can be seen that in order to check these units fully, both the switching point (balance) and hysteresis gap must be measured.

The method used to check the comparators makes use of the integrator ramping circuitry, one comparator input being fed from an integrator output and the other from a reference voltage. The reference voltage was again chosen to be zero, as it would give a direct reading of the errors. Each comparator must be checked independently as the results are a function of amplitude and not of time; however, to avoid the necessity of extra switching to operate each integrator individually, the output changes of the comparator are isolated by gating with binary coding from the digital computer. (This gives an example of the use of the patchable logic elements in the analogue computer.)

The test consists of providing one of the inputs, sequentially, with a positive and negative ramp, while the other input is connected to ground.

When a change of state is detected at the output of the comparator the integrator providing the input is switched to hold. At this point the output of the integrator is read by the a.d.c. It can be seen that the difference between the two readings will give a value for the gap, while half their sum will give a value of the balance offset.

Should a comparator fail to change its state, the test on that particular unit will terminate when the expected switching time has been exceeded by 50%.

6.11 *Track Transfer Units*

Sixteen of the summing amplifiers in the analogue computer are convertible to high-speed track transfer (TT) units (Fig. 8). The tracking is provided by additional operational amplifiers, and the hold facility is provided by the convertible low-speed, low-drift amplifiers. The conversion is accomplished by connecting an analogue patchboard hole labelled TT to another hole labelled RG (relay ground). The connexion is made by a relay contact at the start of the test.

The track transfer operation is divided into three modes which are classified as initial condition (IC), operate (OP) and transfer (TR). (It should be noted that the IC and OP modes are not identical with the computer modes of the same designation.) In the IC and OP modes the TT system tracks the input signal, until, on a transition from OP to TR, a sample is transferred to the output, where it is held until updated by a further transfer. Each track transfer unit has two logic control inputs labelled OP and TR respectively which are located on the logic patchboard. For automatic operation, these controls are operated from the digital computer, via the logic control lines. The automatic read-out of results is achieved via the address selector unit and a.d.c.

The track transfer tests are divided into three groups which may be classified as operational checks, transfer error checks and drift checks. An initial operational check is carried out on all three modes commencing with a check that the output follows the initial condition input in IC. The second test places the unit in the OP and checks that the present IC output is not affected by the removal of the IC input. The TR mode is checked by first applying an input to the unit, followed by operation of the TR mode, which should result in the sum of the input and initial conditions at the output.

The transfer errors are checked at zero and reference (100 V) levels. In order to check the zero transfer error, all inputs are disconnected (Fig. 8) and the TT unit is repeatedly cycled between OP and TR while the output is monitored for a drift rate of less than 0·25 mV/cycle. The reference voltage transfer error requires a signal of -100 V to be switched to one input while the other input is connected to the TT negative output (Fig. 9). The unit is once again cycled between OP and TR while the output is monitored for a drift rate of less than 1 mV/cycle. Both of the above tests run for 5 s at a rate of 20 cycles/s. An additional feature of this test is the checking of input resistor balance, which is accomplished by carrying out a transfer of -100 V for each of the two input resistors and comparing the outputs.

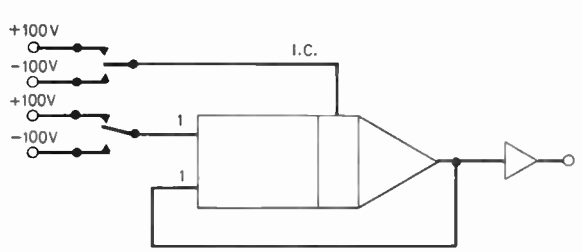


Fig. 9. Track transfer test circuit, showing the connexions for the reference voltage transfer error test.

The final test group is designed to check the TT drift at both zero and reference voltages. The zero drift test consists simply of placing the unit in OP with open circuit inputs and monitoring the output for a drift rate of less than 100 μs/s for a period of one minute. The reference drift or 'drift after transfer' test, is the same as for the zero drift test, except that −100 V is transferred to the output at the start and the drift rate for this test should be less than 0·3 mV/s.

6.12 *Analogue Trunks*

There are 96 signal channels between the analogue patchboard and trunk patchboard (distribution board) which are termed analogue trunks. To test these trunks they are linked at alternate ends to form serially connected blocks of eight, each block being connected to plus or minus reference at one end. The test addresses each trunk in turn and compares the output with the expected output. This test will obviously only locate the first open circuit trunk in each block of eight. However, this is considered acceptable, as a failure of this type is unlikely, and the possibility of two open circuit trunks in one block of eight is extremely unlikely.

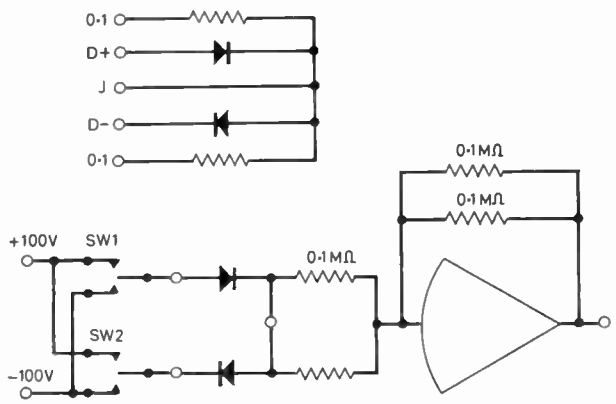


Fig. 10. Diode network test.

6.13 *Diode Networks*

Each of the 18 diode networks contains two 0·1 MΩ resistors and two diodes (Fig. 10) each of which may be used individually or in a group. The networks are tested by applying plus and minus reference alternately to each diode input, with the other input at zero. Each arm of the network is fed from an independent reference switch (i.e. SW1, SW2) but similar arms from each network are fed from the same source. The terminating amplifier is tested independently in the normal fashion but with the inputs from SW1 and SW2 being zero to avoid any diode network effects.

6.14 *Multipliers*

The Applied Dynamics multipliers installed in the AD 256 are electronic multipliers based on the quarter-squares principle, i.e.

$$\tfrac{1}{2}\left[\frac{(x+y)^2}{2}-\frac{(x-y)^2}{2}\right]=xy$$

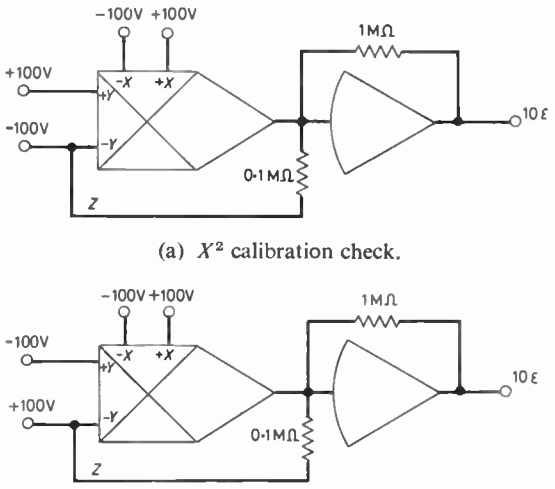


Fig. 11. Basic multiplier gain check.

The multipliers generate summing junction currents proportional to the above identity and must be terminated in the summing junctions of amplifiers. Each half of the dual squarer has two adjustments, one for gain and one for linearity. Of these two controls, only the gain generally requires adjustment. In order to check the need for the gain adjustment the circuits shown in Fig. 11 are used. The error voltage ε is monitored as 10ε at the output of the terminating amplifier and checked for a maximum of 0·2 V (i.e. ε = 20 mV). The transition from test circuit (a) to (b) is accomplished by simply reversing the polarity at both Y inputs and the comparison input (Z). A more comprehensive test which produces a set of error curves is implemented on a single 'quadrant' board, as the allocation of sufficient hardware for this test would create problems on the complete analogue test board. This test which is patched as shown in Fig. 12, generates error curves at X = +100 V, +50 V, 0 V, −50 V, −100 V for Y values ranging from −100 V to +100 V.

As shown in the diagram, the Y inputs are fed with a ramp from an integrator, the ramp operating from +100 V to −100 V. When the ramp has covered its full range the integrator output operates a comparator which signals the completion of that pass to the digital computer. At this stage the digital computer places the integrator into initial conditions and resets the two

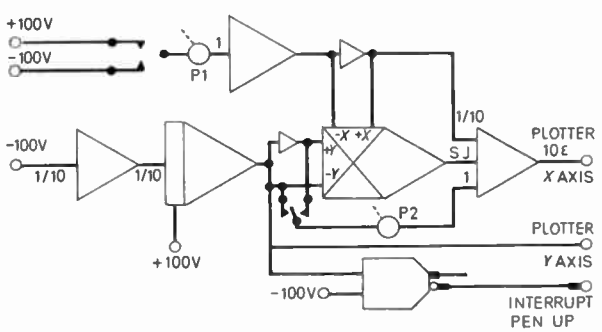


Fig. 12. Multiplier error curve circuit.

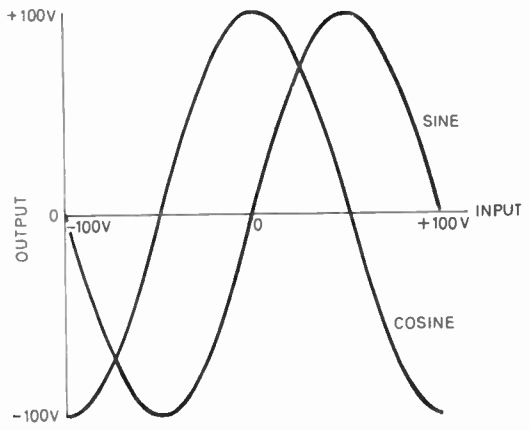


Fig. 13. Outputs of the sine-cosine fixed function generators.

potentiometers P1 and P2 to the next value for X. The potentiometer P2 is used to attenuate the initial ramp by the same factor as that provided by the multiplier, in order to provide a signal of opposite polarity, which will cancel out the ramp and leave an output equal to the error curve of the multiplier. To provide an automatic offset for each error curve plot, the X input is divided by ten and added to the error signal. This enables all the error curves for each multiplier to be plotted without operator intervention. The present X–Y plotter only takes single sheets of graph paper and thus the operator must change the sheet for each multiplier. A further improvement would be to provide an automatically operated paper roll feed which would automate the system completely.

6.15 *Sine-cosine Generators*

The three sine-cosine fixed function generators are all solid-state circuits which, when used with appropriate output amplifiers, produce output voltages proportional to sin X and cos X. The basic check, on the sin-cos generators, consists of measuring both the sine and cosine outputs for inputs of ± 100 V, ± 50 V and 0 V. The expected outputs will be ± 100 V or 0 V as shown in Fig. 13. These values are read by the a.d.c. and compared with the values held in the digital computer. Any errors occurring in the output will indicate whether adjustment of the generators is required.

A more comprehensive test of the sin-cos units is carried out on the same single quadrant board as used for the dynamic multiplier test (Fig. 14). The sine cosine inputs are fed from the digital computer via digital analogue multipliers. The input voltages are incremented from -100 V to $+100$ V in steps of approximately 0·4 V (i.e. 512 steps from 8 binary bits and a sign bit) while the outputs of the sin-cos units are monitored by analogue to digital converters. As each input level is provided to the sin-cos unit, the digital computer calculates the correct sine and cosine values. The sin-cos unit outputs are then compared with the calculated values and a difference value is produced. This difference is plotted as an error value on the XY plotter. The graphical output consists of a series of points connected by straight lines which give a good approximation of the true error curve. The final graph can be retained as part of the machine performance records.

6.16 *Digital-to-Analogue Multipliers*

The digital-to-analogue multiplier (d.a.m.) test procedure is a very simple check which is divided into three basic parts. The zero check involves setting all bits to logical one with an analogue input of zero, the output at the d.a.m. is then monitored to give a value of the zero offset. The gain and balance of the units are checked by setting the most significant bit (m.s.b) to a logical one and all other bits to zero. The analogue input is then set to $+100$ V and -100 V sequentially while the output is checked for the $+50$ V and -50 V. The final test is to check the relative accuracy of each digital bit, this test being carried out with an analogue input of $+100$ V. The method used is to commence by reading the value produced with only the most significant bit (m.s.b.) set and then dividing this value by two and comparing it with the value produced when the m.s.b. but one, is set.

The m.s.b. value is then divided by 2^2 and compared with the value produced by the m.s.b. but two. This sequence of events is continued until the value of the least significant bit (m.s.b.$-n$) is compared with the value of the m.s.b. divided by 2^n.

6.17 *Hand-set Potentiometers*

This is essentially a manually operated device and the test requires operator intervention. The test is of limited use but it does enable the operator to set all of these units at once and avoids the necessity of manually addressing each unit individually. The program prints out the message

SET HAND SET POTS TO MIN & PRESS AI TO CONTINUE

whereupon the operator zeros all the potentiometers and presses the attention interrupt button. The program will then address and monitor the output of each potentiometer in turn and checks the setting within 0·02% of full scale. On completion of the zero setting check the program will print out

SET HAND SET POTS TO MAX & PRESS AI TO CONTINUE

and on receipt of the interrupt, will check the setting of each unit within 0·02% of full scale. An additional

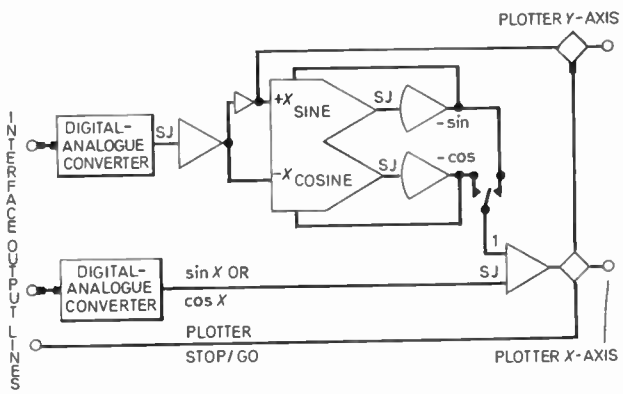


Fig. 14. Sine-cosine unit, error curve circuit.

check for noise is also provided at this setting and is accomplished by rapidly repeating the reading procedure and comparing successive values. Should an unserviceable potentiometer be located during these tests the program will output the message

POT X.X.XX U/S

with the word NOISY being substituted for U/S (unserviceable) should this be the case.

7. Conclusions

The major advantage provided by the automatic maintenance program, namely, the great improvement in the time required to perform the routine maintenance, has already been stated. There are also other benefits which may not be immediately apparent. One of these is the ability to check the 'state of the machine', even in the absence of engineering staff. This fact has proved valuable to users who suspect a machine failure when working outside normal hours. An improvement has also been observed in the 'mean time between failures' (m.t.b.f.) occurring during user time, which is directly due to the more frequent implementation of major tests.

The major problem encountered during the formulation of the program was that of organizing the analogue components. In order to accommodate all the interacting tests on one board it was necessary to specify the allocation of all components and connexions at an early stage. Consequently, any alterations required, due to problems arising from the digital program, caused considerable disruption.

In order to indicate the scale of time saving, obtained by the automation of the tests, some comparative figures for automatic and manual operation, are given below.

Test	*Automatic*	*Manual*
Servo set potentiometers	7 min	55 min
Integrator time scales	10 min	8 hours
Integrator drift	12 min	4 hours
Limiters	3 min	12 min
Comparators	2 min	18 min
Track transfer units	4 min	45 min

It can be seen that the percentage improvement is not a constant factor, and in fact, it depends largely on the complexity of the particular test. The times given for the manual test are approximate values, owing to the fact that the adjustments would normally be carried out at the same time as the check, and therefore the manual tests are designed accordingly.

The times given for the automatic tests include the time taken for a full print-out of machine performance.

The provision of the machine performance record facility has added to the understanding and anticipation of machine faults. This function is now being expanded to produce details of faults in a suitable form for processing on the ICL 1905 digital computer. It is proposed, eventually, to provide statistical component failure analysis, which will locate particularly troublesome components.

A further extension of this work is likely to be the provision of program diagnostic routines, to help locate programming errors and to carry out static checks.

8. Acknowledgments

The author wishes to extend his grateful thanks to Mr. P. G. Thomasson, Mr. J. Smeathers and Mr. H. D. Blake for helpful discussions and to Mr. K. G. Beauchamp for his constant encouragement during the preparation of this paper. In addition, the author wishes to thank the Civil Service Department for providing some of the financial support for the project.

9. References and Bibliography

1. Beauchamp, K. G., 'A hybrid linkage system', Proceedings of Fourth Australian Computer Conference, Adelaide 1969.
2. Beauchamp, K. G., Kelley, C. J., Thomasson, P. G. and Williamson, M. E., 'Development of a Hybrid System for University Research', IFIPS Congress, Yugoslavia, August 1971.
3. Holst, P. A., 'Performance testing of analog computers', *Annales de l'Association internationale pour le Calcul analogique*, January 1970.
4. Barney, G. C., Hambury, J. N. and Chow, S. P. S., 'Diagnostic routines for a hybrid computer', *Computer Bulletin*, **13**, No. 6 pp. 178–83, June 1960.
5. Hall, W. B. and Peterson, H. E., 'Hybrid preventive maintenance program', Fifth Annual Simulation Software Conference Wright Patterson Air Force Base, Ohio, U.S.A., June 1969.
6. Pelle, J., 'Hybrid computer diagnostics', AICA/IFIP Conference on Hybrid Computation, Munich, August 1970.
7. Kelley, C. J., 'Hybrid computation at Cranfield Computing Centre', *Computer Weekly*, 5th June 1969, page 6.
8. Applied Dynamics 'AD 256 Analog Computer—Users' Manuals', Ann Arbor, Michigan, U.S.A.
9. Roth, J. P., 'Diagnosis of automata failures: a calculus and a method', *I.B.M. J. Res. Devel.*, **10**, No. 4, pp. 278–91, July 1966.
10. Roth, J. P., Bauricius, W. E. and Schneider, P. R., 'Programmed algorithms to compute tests to detect and distinguish between failures in logic circuits', *I.E.E.E. Transactions on Electronic Computers*, **EC-16**, No. 5, pp. 567–80, October 1967.

10. Appendix: Configuration of the Cranfield Hybrid System

ICL 1905 Computer and Peripheral Equipment:

(a) Central Processor having 32k 24 bit words, 2 microsecond core store, floating point hardware.
(b) 2 Paper Tape Readers: 300 characters per second
(c) 2 Paper Tape Punches: 110 characters per second.
(d) Card Reader: 300 cards per minute.
(e) Card Punch: 100 cards per minute.
(f) 2 Line Printers: 300 lines per minute.
(g) Console Typewriter.
(h) Calcomp Graph Plotter (30 inch).
(i) Multiplexor.
(j) 7 Data Terminals.
(k) 7 Remote Teleprinter Consoles.
(l) 4 Magnetic Tape Transports, 556 bit/in, 20 kHz.

(m) 2 Exchangeable Disk Store Transports, each having the capacity of 4M characters.

(n) British Standard Interface, 6/7/8 bit transfer at 50 kHz.

AD 256 Computer and Peripheral Equipment:

The AD 256 is a large, general-purpose analogue computer, having solid-state switching throughout and a linear component accuracy of 0·01%.

(a) 100 Bipolar Summing Amplifiers/Integrators.

(b) 30 Quarter-squared Multipliers.

(c) 120 Servo and 30 Hand-set Potentiometers.

(d) 8 Function Generators.

(e) Non-linear Equipment including Resolvers, Limiters, Comparators and Analogue Gates.

(f) Parallel Logic including Gates, Bistables, Monostables and Shift Registers.

(g) Digital Voltmeter.

(h) Display Oscilloscope (4-beam).

(i) Precision Oscilloscope (2-beam).

(j) *X–Y* Plotter.

(k) Noise Generator.

(l) 6-channel Pen Recorder.

(m) 14-channel F.M. Analogue Magnetic Tape System.

Digico Micro 16

A small process-control computer, comprising:

(a) Central Processor, having 4k 16 bit words, 6 μs core store.

(b) Console Typewriters (2).

(c) Local and Remote Operating and Display Panel.

(d) Magnetic Tape Transport.

(e) Hardware Multiply/Divide Unit.

(f) Paper Tape Reader: 300 characters per second.

(g) D-Mac Graphics Input Device.

Manuscript received by the Institution on 25th February 1971.
(Paper No. 1410/Comp. 137.)

STANDARD FREQUENCY TRANSMISSIONS—September 1971

(Communication from the National Physical Laboratory)

Sept. 1971	Deviation from nominal frequency in parts in 10^{10} (24-hour mean centred on 0300 UT)			Relative phase readings in microseconds N.P.L.—Station (Readings at 1500 UT)		Sept. 1971	Deviation from nominal frequency in parts in 10^{10} (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 60kHz		GBR 16 kHz	MSF 60 kHz	Droitwich 200 kHz	*GBR 16 kHz	†MSF 60 kHz
1	−300·0	−0·1	+0·1	504	534·9	17	−300·1	0	+0·2	507	539·3
2	−300·0	0	+0·1	504	535·9	18	−300·0	0	+0·2	507	539·6
3	−299·9	+0·1	+0·2	503	534·8	19	−300·1	−0·1	+0·2	508	540·7
4	−299·9	+0·1	+0·2	502	534·3	20	−300·1	−0·1	+0·2	509	541·6
5	−299·9	0	+0·2	501	534·0	21	−300·0	0	+0·2	509	540·3
6	−300·1	+0·1	+0·2	502	533·5	22	−300·0	0	+0·2	509	541·4
7	−300·0	0	+0·2	502	536·3	23	−300·0	−0·1	+0·2	509	542·2
8	−300·0	0	+0·2	502	533.3	24	−300·0	0	+0·1	509	541·8
9	−300·0	0	+0·2	502	533·9	25	−300·1	−0·1	+0·1	510	542·4
10	−300·1	0	+0·1	503	534·2	26	−300·2	−0·2	+0·1	512	543·9
11	−300·0	0	+0·1	503	534·4	27	−300·1	−0·1	+0·1	513	544·6
12	−300·0	0	+0·2	503	534·6	28	−300·0	−0·1	+0·1	513	545·1
13	−300·1	0	+0·1	504	535·5	29	−300·1	−0·2	+0·2	514	547·0
14	−300·0	0	+0·2	504	541·7	30	−300·0	0	+0·2	514	547·0
15	−300·0	−0·1	+0·2	504	534·0						
16	−300·2	0	+0·2	506	539·2						

Mean monthly values: GBR: −300·03 MSF: −0·03.

All measurements in terms of H.P. Caesium Standard No. 334, which agrees with the N.P.L. Caesium Standard to 1 part in 10^{11}.

* Relative to UTC Scale; (UTC_{NPL} − Station) = + 500 at 1500 UT 31st December 1968.

† Relative to AT Scale; (AT_{NPL} − Station) = + 468·6 at 1500 UT 31st December 1968.

U.D.C. 621.384.8: 681.32

A Low Cost, High Performance Mass Spectral Data System

By

A. CARRICK,
B.Sc., Ph.D.†

Reprinted from the Proceedings of the Conference on Laboratory Automation held in London from 10th to 12th November 1970.

A modular, low cost, mass spectrometric data system based on a hybrid analogue–digital preprocessing unit and the British Standard Interface is described. The basic hardware is physically small, has a wide range of capability, and deals directly with the voltage output from electrically recording spectrometers. Mass measurements from single scans at 1 : 10 000 resolution at 10 s/mass decade show an r.m.s. error of 9 parts in 10^6, and at 1 : 400 resolution and 0·16 s/decade, an r.m.s. error of 0·06 mass units on all peaks $\geqslant 1/500$ full scale. Measured intensity values are accurate to 5% of calculated for peaks >1%. System and processing considerations, and practical arrangements are detailed.

1. Introduction

Mass spectrometry is now one of the most powerful and versatile analytical tools available to the chemist. It is capable of producing precise information on molecular composition and structure, as well as an indication of quantity of material present from sub-microgram amounts of material. In essence the technique consists of vaporizing a small amount of sample, ionizing it, usually by means of a beam of electrons, sorting out the positive ions produced in terms of their mass to charge ratio and recording the abundance of each ionic species present. The ionization energy of the process is commonly made large in comparison with chemical bond energies so that each ionized molecule may fragment producing smaller ionic units in a manner characteristic of the structure of the original molecule. Mass analysis of the ion beam can be accomplished in a variety of ways, most commonly by acceleration through a high potential, followed by deflexion in a magnetic field, in accordance with the relation

$$\frac{m}{e}=\frac{H^2R^2}{2V},$$

where H is the magnetic field strength, V the accelerating voltage, and R the radius of the path of the beam in the magnetic field. m/e values are usually determined by ratio methods using known standards, and a scanned spectrum obtained by varying H or V continuously in time. Because the exact mass of each isotope of the elements is unique, analysis of the chemical composition of an ion can be achieved by precise mass measurement and simple arithmetic.

There are many applications of mass spectral analysis, from production plant process control to analysis of pesticide residues and structure determinations of snake venoms. Unfortunately, however, the strength of the technique has until recently also been its weakness—so much information is potentially available in a mass spectrum that only a fraction can be utilized or even adequately collected by the human operator. Automatic data acquisition and processing methods have now radically altered this situation. In the past, effort has been concentrated on data acquisition from expensive high-resolution spectrometers and the degree of success achieved has been closely related to the overall cost of the system. The device now described shows that it is possible to obtain accurate data from the spectrometer during its scanning period without using a digital computer at this stage or any other very costly equipment.

2. Basic Principles

A mass spectrum is a two-dimensional information set. All the useful data are contained, in the case of an electrically-recording instrument, in an output voltage varying in time, and each voltage peak represents a discrete ionic species derived from the molecule or molecules under study. The chemist requires mass and abundance data for each species of interest, and these data may be obtained from determinations of the position of the voltage peaks as they occur in the spectral scan, and their intensity. Under the most demanding conditions, the chemist might wish to record several spectra from a component of a complex mixture while it elutes from a gas chromatographic column, or make rapid repeat scans to monitor the course of a reaction. These applications might require the measurement and processing of 10 000 discrete, chemically useful pieces of information per second.

The problems of data acquisition in these circumstances are clearly not inconsiderable.[1] They include the apparently random incidence of the narrow triangular spectral peaks during the scan and potentially high data rate required to specify these peaks when they do occur. A typical high resolution spectrum might contain up to 1000 peaks, yet even so potentially useful information is present at the output of the spectrometer during only 5% of the scan time. If peak positions are determined by reference to the whole peak profile, a fast software controlled analogue-to-digital convertor (a.d.c.) computer system must spend 95% of the spectrometer's scan time in wasteful rejection of baseline readings 'just in case' a peak should occur, and must engage in a high rate of activity during the passage of a peak in order to evaluate its position and intensity before the next peak arrives.

† Chemical Standards Division, National Physical Laboratory, Teddington, Middlesex.

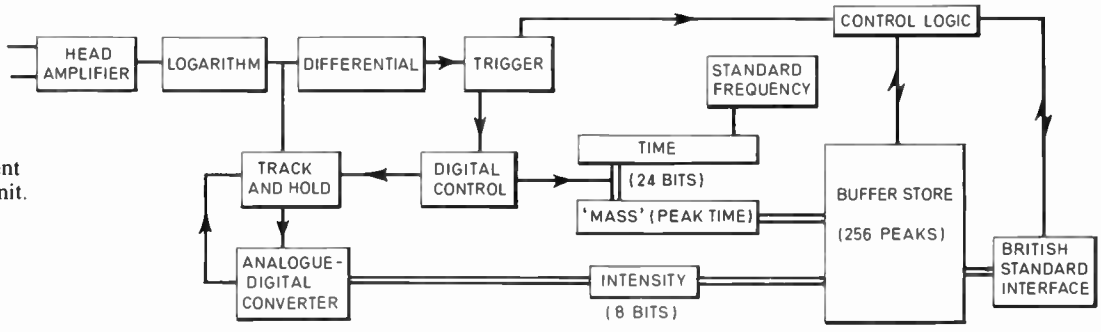


Fig. 1.
Schematic arrangement of data acquisition unit.

Such an approach[1] assumes that it is necessary to plot the profile of each peak in order to obtain the required chemically useful information, i.e. the position and intensity parameters. The system now described[2] does not. Fast analogue computation is used to take the first derivative of the signal (dI/dt), and so extract from each peak, whatever its intensity, a common reference point which immediately gives access to the position parameter by locating the peak maximum in real time. At the maximum, dI/dt is independent of peak height and is exactly zero. Thus in a single analogue stage the unit provides a signal which precisely defines each peak's position in time, and so its mass value, without any digital calculation, or indeed without reference, in a conventional sense, to the magnitude of the peak.[3] An a.d.c. can use the same maximum defining signal as a command to take a reading of the peak height at that point. For a well-shaped spectral peak with a single maximum, only the two basic parameters are recorded, and so a considerable economy of digital computational effort is effected.

Some peaks however may have multiple maxima for one of several reasons. The signal may be noisy, i.e. may contain as a significant component a signal whose period is close to the baseline width of 'genuine' peak indication. The peak itself may be an unresolved multiplet, or be of such low intensity that small numbers of ions striking the electron multiplier detector are recorded essentially separately. All these cases could produce almost identical output signals: the detector unit of the present system is fast enough to record such maxima and these 'fine structure' readings must then be dealt with by the processing software. In practice these situations have caused few difficulties, and the certain amount of inherent analogue smoothing (bandwidth limitation) helps to minimize readings from noise. The remaining cases in which readings lie close together can be treated as multiplets or singlets on a basis of the percentage of the (known) theoretical peak width (resolution) represented, in the processing stage.

3. Practical Arrangement

The basic preprocessor shown in the schematic diagram (Fig. 1) is of uncomplicated design. Micro-circuit operational amplifiers are extensively used in the analogue circuitry and transistor-transistor-logic (TTL) integrated circuits perform all the counting and control functions.

In detail, the analogue signal from the spectrometer, attenuated if necessary into the 0–10 V range, is buffered by a non-inverting twin amplifier complex with high common-mode-rejection ratio. This head amplifier feeds a logarithmic amplifier which allows full sensitivity to be obtained from the differentiating stage over at least three decades of input signal, the limiting factor here usually being the signal/noise ratio of the input signal. The differentiator at present employed is a bounded 'open loop' amplifier, the peak maximum is indicated by the detection of the voltage breakpoint of the bounding circuit using an integrated circuit comparator, which drives the digital logic directly. A track and hold amplifier monitoring the logarithmic signal is controlled by the digital indication of the peak maximum and the held voltage is buffered to an a.d.c. actuated again by the digital signal. The use of the logarithmic signal for intensity measurement means that a.d.c. quantizing error becomes uniform over the range of the signal, rather than being a maximum for the smallest signals, and so an 8 bit a.d.c. ($\frac{1}{2}\%$ quantizing error) is used.

The digital signal marking the peak maximum also provides a gating pulse for a latching register running in parallel with a 24 bit master counter, which is driven from a standard frequency source and measures elapsed time from the start of the spectrometer's scan. A clock frequency of 1 MHz is generally used. This gives quantizing errors of 1 part in 10^6 or less for all logarithmic scans up to 1 s/mass decade and is usually adequate for integer mass marking on scans as fast as 50 ms/decade.

The 32 data bits specifying each peak, 24 of time and 8 of intensity, are stored byte by byte in a random access magnetic core store. Storage of the three time bytes, which are available within nanoseconds of the arrival of a peak maximum, takes place first, with analogue–digital conversion of the peak height taking place during this period. The intensity word is available for storage, immediately after the third time byte is stored. 1024 words of store, each of 8 bits, can thus hold 256 peak readings before any output need take place, and the 4 μs core cycle time and 10 μs analogue–digital conversion time give the whole unit a dead time of 16 μs (4×4 μs) and thus an effective resolution far in excess of the capabilities of any current commercial mass spectrometer.†

† A peak width of 16 μs corresponds (approximately) to a resolution of 1 : 20 000 for a logarithmic scan of 1 s/mass decade. Fastest currently achieved is about 1 : 25 000 in 8 s/mass decade.

Data are removed from the store, again byte by byte, via a BS 4421 standard interface source channel,[4] either during the spectral scan, in an interleaved fashion, or in a block transfer mode.

4. Systems Considerations

One of the major benefits of the present approach is that the preprocessor is autonomous and produces data essentially under the control of the spectrometer. As such it may be treated as a computer peripheral standing aloof from the central processor and requesting access to it only when data need be transferred. In one commercial version of the system, for example, the preprocessor requires only 3% of c.p.u. commitment during the spectrometer's scan time, as opposed to the 100% needed for a direct spectrometer–computer link. With the limitations of the conventional close contact between computer and spectrometer lifted, interfacing the basic unit to any processor now becomes straightforward. The use of the British Standard Interface is thus entirely appropriate and adds a further dimension of versatility to the system; any suitable data acceptor fitted with a matching interface may be connected directly in a 'plug-in-and-go' fashion.

This means that a data system can be assembled in a 'building block' mode, starting with a basic kit of preprocessor (data source) and paper tape punch (data acceptor) right up to the most sophisticated of time-shared communications-network-oriented 'total-laboratory' systems. At each stage the standard interface confers particular advantages. For instance, if a small dedicated computer is available, software need only be written so that the input interface (B.S. acceptor) is serviced. No change is then necessary if data are read in from paper tape (via a reader equipped with a B.S. source) or magnetic tape, or directly from the preprocessor on-line to the mass spectrometer.

The interface's 'self-adjusting' power and independence of any particular output device confers clear economic advantages, both from the point of view of maintenance and initial cost, and at no stage does any of the equipment cease to have a useful application in a systems context.

5. Results

The performance of the author's prototype preprocessing unit has been examined under two contrasting sets of conditions: with a high resolution, medium scan-speed spectrometer† and with a low resolution, fast scanning machine.‡ Very brief examination of the unit in combination with a quadrupole mass spectrometer§ gave entirely satisfactory results, broadly similar in accuracy to those quoted for the low resolution electromagnetic instrument.

Using perchlorobutadiene as a test compound, high resolution mass measurements had an r.m.s. error of 9 parts in 10^6 for single scan determinations. No significant difference in measuring accuracy was observed over the whole sensitivity range of the unit down to the detection limit, which was usually 1/500th of full scale, signal/noise ratio limiting the range at this point. These performance figures compare favourably with those quoted for a current commercial system,† which yields an r.m.s. error of 4 parts in 10^6 for single scan mass measurements from scans at 35 s/decade over a 50 : 1 intensity range.

In many ways the r.m.s. error of 0·064 mass units for single scan measurements on singlet peaks from the fast, low resolution scans is much more significant. With a total scan time of 95 ms to cover the range *m/e* 300 to *m/e* 60, this means that data may be obtained, by averaging 9 successive scans of the same spectrum, which can yield mass measurements comparable to those hitherto obtained only from high resolution machines and this with a total scan time of less than 1 s. Table 1 lists some typical '9-average' mass measurements.

Table 1

Mass measurements, E606
'9-average' of 0·1 s scans

Species	Measured	Error (milli mass units)	*I*(TH)	*I*(Abs)
C_4CL_4	187·8724	−3	28	0·4
	189·8685	−4	37	0·6
	191·8666	−3	18	0·2
C_4CL_5	222·8693	25	62	1·0
	224·8360	−5	100	1·6
	226·8242	−14	65	1·0
	228·8337	−2	21	0·4
C_4CL_6	257·8541	41	21	0·2
	259·8827	72	41	0·2

In both these cases, high and low resolution data were processed by the familiar pseudo-linear extrapolation technique, with measurements obtained by direct linear interpolation.[5] The low resolution data have also been processed to yield integer mass marking by two methods, a direct calculation of mass from a knowledge of the exponential scan law (and particularly its time-constant variation) and empirically by a 'table-look-up' method, based on a calibration run of a known compound. Both methods yielded the required results but the latter method was the faster by at least a factor of five (essentially peripheral limited on the processor used)‡ and required less operator-supplied data.

Intensity values for both high and low resolution single scans were accurate to within 5% of calculated relative intensity for peaks >1% of base peak, and as with the mass measurements, the expected statistical improvement was obtained from '4-average' and '9-average' scans.

† A.E.I. MS9. 1 : 10000 resolution at 10s/decade.
‡ Edwards E606. 1 : 400 resolution at 0·16s/decade.
§ Finnigan Instruments model 400.

† A.E.I. Systems DS10 and DS20.
‡ *Micro-16* computer, 4k store, 16 bit words, 3 registers, 6 μs core cycle time, 20 μs add time (serial arithmetic). Digico Ltd., Letchworth, Herts.

6. Conclusions

The present approach is thought to represent a significant advance in so far that the processing computer is concerned solely with the arithmetic of the data reduction process—of maintaining and improving the accuracy of the measurements made by the preprocessor—and not at all with the data acquisition process or its control. Whilst it is not suggested that this method always provides the whole answer to all mass spectrometer data acquisition and handling problems, the low cost, modularity, small size and high performance of the system should ensure a wide range of applications not only in the research laboratory but also in routine analytical and process control situations.

7. Acknowledgment

This work forms part of the research programme of the Division of Chemical Standards of the National Physical Laboratory.

8. References

1. Halliday, J. S., 'Factors involved in the design of data acquisition systems for mass spectrometers', 'Advances in Mass Spectrometry', Vol. 4, pp. 239–54. (Institute of Petroleum, London, 1968.)
2. Carrick, A., U.K. Patent Application No. 35894/1968.
3. Carrick, A., 'A new approach to the real time digitisation of mass spectra', *Intl. J. Mass Spectrometry and Ion Phys.*, **2**, pp. 333–47, 1969.
4. B.S. 4421, 1969. 'A Digital Input/Output Interface for Data Collection Systems.' Also a large applications bibliography obtainable from NPL.
5. McMurray, W. J., Green, B. N. and Lipsky, S. R., 'Fast scan high resolution mass spectrometry', *Analyt. Chem.*, **38**, 1194–1204, August 1966.

Manuscript first received by the Institution on 21st August 1970 and in final form on 12th May 1971. (Paper No. 1411/IC 51).

U.D.C. 621.372.542.21: 621.397.23

Gaussian Filters for Pulse Shaping

By

L. E. WEAVER,
B.Sc., C.Eng., M.I.E.E.†

and

D. C. BROUGHTON,
M.Sc., C.Eng., M.I.E.R.E.†

Networks with approximately Gaussian insertion loss characteristics for the shaping and standardization of pulses used in B.B.C. television equipment have previously been designed by Dishal's method based on a Taylor series approximant. A proposal by Jones for an improved Gaussian approximant employing Laguerre polynomials suggested that more effective networks might be possible, Computer programs were written for the realization of the Laguerre approximant networks, and a series of networks was constructed and tested. The Laguerre networks were found to be superior both in their improved shaping and in possessing more convenient element values. Design data for both types are provided in tabular form.

1. Introduction

It is very desirable, particularly with television and data transmission systems, for the transitions of the pulses used to be both well-shaped and compact, that is, each transition should have the minimum overall duration consistent with the given bandwidth, together with the greatest possible degree of skew-symmetry about the half-amplitude points. At the same time they should be as free as possible from overshoots or other irregularities.

It is possible to achieve suitable shaping in the pulse generation circuitry, but a simpler and much more reliable method is to generate waveforms which have transition times shorter than is ultimately required, and then to pass them through a passive low-pass shaping network. The ratio between the initial and final transition times does not need to be at all large, say a minimum of 1 : 3, to ensure that the output waveform is substantially independent of changes in the shape of the input waveform.[1]

In order for this technique to be economically viable the networks used must have the smallest number of elements which will provide a suitable response, they must be as simple as possible to align, and at the same time they should not require capacitors and inductors with very low dissipation factors, selected or adjusted to close tolerances.

When this technique was initially considered some years ago for use in B.B.C.-designed equipment, the first thought was the application of a sine-squared pulse shaping network, often known as a Thomson filter, since this had been extensively used elsewhere in test equipment.[2] The performance of this network for the shaping of transitions leaves little to be desired, but with a possible total of eleven elements it is more expensive than one would wish, and its alignment needs a fair amount of care.

Attention was then given to Gaussian low-pass filters designed by the method of Dishal,[3] as a result of which they have long been employed in equipment for operational use. More recently, a paper by Jones[4] suggested a quite different basis for the design of such filters, which seemed to warrant further investigation. Since this paper was not taken further than the calculation of a series of proposed polynomial approximations to the ideal insertion loss characteristic, a computer program was devised for the synthesis of the corresponding networks by Darlington's method, from which a number of networks to various orders of approximation were constructed and then tested. In the course of this work it became evident that the possibility of alternative realizations of the Dishal networks had been overlooked. The various forms of these up to the 9th order were therefore computed, and some were found to offer considerable practical advantages over the realizations tabulated by Dishal. Sample networks were also constructed and comparative tests carried out between them and the previously mentioned Laguerre networks, which are similarly available in alternative realizations.

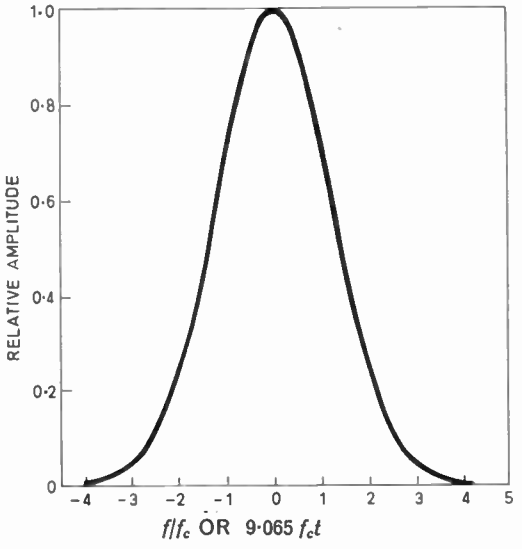


Fig. 1. Ideal Gaussian amplitude and impulse responses.

2. The Gaussian Characteristic

The Gaussian low-pass filter derives its name from the fact that the insertion loss and impulse response both, in the ideal case, have the form of a Gaussian or normal error curve (Fig. 1), which is symmetrical and free from overshoots. The step response, being the integral of the impulse response, is accordingly also symmetrical and free from overshoots (Fig. 2).

It can be predicted by simple physical considerations from these transient responses that the ideal Gaussian network is an unrealizable abstraction. In each instance, whether impulse or step response, the start of the waveform, identified as the point where the amplitude is initially zero in the strict mathematical sense, can only be found by projecting the curve backwards to negatively

† British Broadcasting Corporation, Designs Department, Broadcasting House, London W1A 1AA.

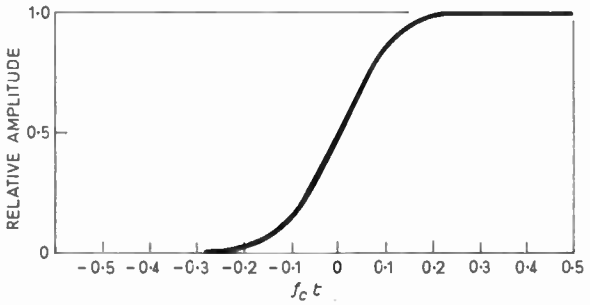


Fig. 2. Ideal Gaussian step response.

infinite time. The delay through the network is therefore infinite, corresponding to an infinite number of component reactances, and the network is consequently not realizable. The location of the centre of the transient at $t = 0$ is a convenient fiction.

Nevertheless, it is possible to approximate the ideal response as closely as may be required by means of a physical network composed of a finite number of elements, and the essential difference between the approaches to the problem made by Dishal and Jones lies in the way in which the corresponding approximating functions are derived. Paradoxically, practical filters of relatively simple form are more useful for pulse shaping than a close approximation to the ideal would be, since the restricted transmission time through the network has the effect of rendering the waveform response much more compact, in the sense in which this has been used above. Indeed, certain of them provide a close simulation of the very desirable sine-squared network response, as will be seen below. This property does not appear to have previously been pointed out explicitly.

2.1 *Taylor Series Approximation*

If the ratio between the power delivered to the output termination and the maximum possible deliverable power is $|d|^2$, the square of the modulus of the transmission coefficient, then the ideal Gaussian response would be expressed as

$$|d|^2 = [\exp -(f/f_0)^2]^2$$

where f_0 is a reference frequency. It is convenient to refer to the frequency of 3dB attenuation f_c, when the insertion loss assumes the very simple form

$$A\,(\text{dB}) = 3{\cdot}010(f/f_c)^2, \text{ where } f_c = 0{\cdot}5887\, f_0.$$

Dishal's approach to the problem was to expand the exponential into a Taylor series truncated after the desired number of terms, which makes the nth order approximant

$$P_n = 1 + 2x + \frac{2^2}{2!}x^2 + \ldots + \frac{2^n}{n!}x^n,$$

where for convenience x has been used for $(f/f_0)^2$.

The basic equation for the nth-order network having n elements is accordingly $|d|^2 = 1/P_n$.

If x is replaced by $-p^2 = (j\omega/\omega_0)^2$, the zeros of P_n may be found, remembering that since $|d|$ is a transfer function only the left-hand half-plane zeros may be retained. These have been found to be as given in Table 1.

Table 1

Zeros of the Taylor approximant

Order	Zeros
3	$-0{\cdot}8933$
	$-0{\cdot}8125 \pm j0{\cdot}5561$
4	$-0{\cdot}8351 \pm j0{\cdot}7498$
	$-0{\cdot}9584 \pm j0{\cdot}2319$
5	$-1{\cdot}0442$
	$-1{\cdot}0017 \pm j0{\cdot}4228$
	$-0{\cdot}8511 \pm j0{\cdot}9189$
6	$-1{\cdot}1032 \pm j0{\cdot}1900$
	$-1{\cdot}0333 \pm j0{\cdot}5890$
	$-0{\cdot}8632 \pm j1{\cdot}0709$
7	$-1{\cdot}1745$
	$-1{\cdot}1472 \pm j0{\cdot}3550$
	$-1{\cdot}0577 \pm j0{\cdot}7384$
	$-0{\cdot}8727 \pm j1{\cdot}2103$
8	$-1{\cdot}2286 \pm j0{\cdot}1646$
	$-1{\cdot}1816 \pm j0{\cdot}5031$
	$-1{\cdot}0773 \pm j0{\cdot}8753$
	$-0{\cdot}8805 \pm j1{\cdot}3398$
9	$-1{\cdot}2910$
	$-1{\cdot}2715 \pm j0{\cdot}3120$
	$-1{\cdot}2096 \pm j0{\cdot}6386$
	$-1{\cdot}0934 \pm j1{\cdot}0025$
	$-0{\cdot}8869 \pm j1{\cdot}4613$

2.2 *Laguerre Series Approximation*

This makes use of the approximate identity

$$\exp(2x) \simeq \sum_{m=0}^{n} a_m L_m(\alpha x) = Q_n$$

where L_m is the mth-order Laguerre polynomial and α is an arbitrary scaling factor, allotted by Jones the value of 7, which ensures rapid convergence of the series. The important difference between the two methods lies in Jones's choice of the coefficients a_m in such a manner as to minimize the remainder of the truncated series of Laguerre polynomials in an exponentially weighted mean square sense, which evidently offers the prospect of a closer approximation for a given number of terms than the plain truncation of the series.

The approximants Q_n then become, as far as the 9th order:

$$Q_3 = 0{\cdot}9744 + 2{\cdot}6656x - 1{\cdot}0976x^2 + 5{\cdot}1221x^3$$

$$Q_4 = 1{\cdot}0102 + 1{\cdot}6621x + 4{\cdot}1709x^2 - 3{\cdot}0733x^3 + 3{\cdot}5855x^4$$

$$Q_5 = 0{\cdot}9959 + 2{\cdot}1638x + 0{\cdot}6586x^2 + 5{\cdot}1221x^3 - 3{\cdot}5855x^4 + 2{\cdot}0079x^5$$

$$Q_6 = 1{\cdot}0016 + 1{\cdot}9230x + 2{\cdot}7660x^2 - 1{\cdot}4342x^3 + 5{\cdot}0197x^4 - 2{\cdot}8110x^5 + 0{\cdot}9370x^6$$

$$Q_7 = 0{\cdot}9993 + 2{\cdot}0354x + 1{\cdot}5858x^2 + 3{\cdot}1552x^3 - 3{\cdot}0118x^4 + 3{\cdot}9354x^5 - 1{\cdot}6866x^6 + 0{\cdot}3748x^7$$

$$Q_8 = 1{\cdot}0003 + 1{\cdot}9840x + 2{\cdot}2152x^2 + 0{\cdot}2180x^3 + 3{\cdot}4134x^4 - 3{\cdot}2608x^5 + 2{\cdot}5112x^6 - 0{\cdot}8246x^7 + 0{\cdot}1312x^8$$

$$Q_9 = 0{\cdot}9999 + 2{\cdot}0071x + 1{\cdot}8915x^2 + 1{\cdot}9803x^3 - 1{\cdot}2128x^4 + 3{\cdot}2158x^5 - 2{\cdot}5262x^6 + 1{\cdot}3343x^7 - 0{\cdot}3411x^8 + 0{\cdot}04081x^9$$

As before, the basic equation defining the network is $|d|^2 = 1/Q_n$. By replacing x by $-p^2 = -(j\omega/\omega_0)^2$ the zeros of Q_n may be found, as before with the proviso that since $|d|$ is a transfer function only the left-hand half-plane zeros may be utilized. These have been calculated and the values are given in Table 2.

Table 2

Zeros of the Laguerre approximant

Order	Zeros
3	−0·5353
	−0·5312 ± j0·7298
4	−0·5690 ± j0·3172
	−0·5474 ± j0·9753
5	−0·5983
	−0·5959 ± j0·5722
	−0·5591 ± j1·1883
6	−0·6272 ± j0·2618
	−0·6164 ± j0·7914
	−0·5682 ± j1·3789
7	−0·6547
	−0·6523 ± j0·4861
	−0·6327 ± j0·9868
	−0·5757 ± j1·5529
8	−0·6806 ± j0·2276
	−0·6732 ± j0·6851
	−0·6461 ± j1·1648
	−0·5820 ± j1·7139
9	−0·7060
	−0·7038 ± j0·4296
	−0·6906 ± j0·8659
	−0·6576 ± j1·3293
	−0·5876 ± j1·8645

When these zeros are compared with the corresponding values for the Taylor approximant it will be found that the locus of the former has a smaller curvature for the same order, suggesting that the phase characteristic is more linear and hence that the transient response is somewhat more symmetrical.[4]

2.3 *Network Realization*

The networks were realized by Darlington's method in the form of ladder networks with an equal resistive termination at each end. It is not proposed to describe the method in detail; for this standard texts such as Weinberg[5] should be consulted. However, it is desirable to explain very briefly how it is possible for alternative forms to arise.

If the reflexion coefficient at the input terminals of a reactance network terminated at its output is r, then the basic energy equation gives $|d|^2+|r|^2 = 1$, i.e. the sum of the transmitted and reflected powers in normalized form is unity. Since we already know $|d|^2$, it is possible to find $|r|^2$ and from that to derive $|r|$. The point of this device lies in the fact that it enables the problem to be transformed from the determination of a network having a given transmission behaviour into the synthesis of a driving-point impedance, which is much simpler.

The square of the reflexion coefficient, $|r|^2$, is obtained in the form of the quotient of two polynomials, and in order to proceed to the next stage it is necessary to discard half the zeros of each. The denominators of the transmission coefficient and the reflexion coefficient are identical, so one may only retain the zeros in the left-plane, but the numerator is not subject to the same restriction, and zeros may be taken from either or both half-planes.[6]

Fortunately, the number of arrangements is halved by the fact that selections which are mirror images give rise to the same network but seen from the other end. Nevertheless, although the number of distinct selections is only two with the 4th order, by the 9th order it has already risen to eight. All of the corresponding networks, of course, have identical performances, although different element values.

Dishal[3] restricted the networks realized from his Taylor series approximant to those derived from numerator zeros in the left half-plane, but without giving reasons for his choice. Bode[7] points out that such networks realize the highest gain-bandwidth product when used for interstage coupling, and it is clear from Table 3 that the sum of the end capacitances is greatest for this configuration. Unfortunately, this is associated in the odd orders with a ratio of the terminating capacitances which rapidly becomes embarrassingly high with increasing order, for example the 9th-order Taylor network corresponding to zeros all taken from the same half-plane has a ratio of capacitors of 40 : 1.

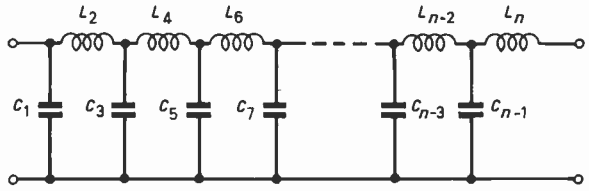


Fig. 3. Generalized network configuration.

Now if one considers the behaviour of the networks as the frequency tends to zero, it is clear that the driving point impedance corresponds to a single capacitor shunted by the terminating resistor, the effective capacitance being the sum of all the individual shunt capacitances. In other words the total shunt capacitance is a constant for a given filter, whatever the realization. One might suspect, therefore, that a large gain-bandwidth product is likely to result in one end capacitance value being smaller than would otherwise be the case, and indeed this is found to be so, to the extent that it may in some instances be comparable with the circuit stray capacitance, which is undesirable.

The element values of these networks, both Taylor and Laguerre, are listed in Table 3. With the higher orders an arbitrary choice has been made of the realization which has values of terminating capacitance approximately equal, and as high as possible. Such realizations are termed 'optimum'. As the order of the approximation decreases so does the extent of the choice, but at the same time the ratio of capacitances becomes smaller, so the difficulties of utilization are less. The realizations derived from all the numerator zeros in the same half-plane are also included in Table 3, and are termed 'maximized' types.

Table 3

Element values are given in farads and henrys for a 1 ohm filter with f_c = 1 Hz, or in microfarads and microhenrys if f_c = 1 MHz.

Filter	C_1	L_2	C_3	L_4	C_5	L_6	C_7	L_8	C_9
3TM, O	·0416	·1294	·3529						
3LM, O	·0693	·1422	·3829						
4TM	·3571	·1483	·0843	·0282					
4TO	·0650	·3978	·1115	·0437					
4LM	·3476	·1654	·1059	·0477					
4LO	·1140	·3870	·0991	·0664					
5TM	·3586	·1557	·1032	·0620	·0209				
5TO	·0349	·1207	·4024	·0970	·0453				
5LM	·3646	·1586	·1191	·0842	·0353				
5LO	·0620	·1672	·3900	·0756	·0670				
6TM	·3592	·1589	·1122	·0796	·0485	·0163			
6TO	·0355	·1102	·4064	·1167	·0780	·0279			
6LM	·3574	·1627	·1199	·0967	·0696	·0279			
6LO	·0423	·1118	·1835	·3685	·0615	·0666			
7TM	·3594	·1603	·1167	·0892	·0645	·0394	·0132		
7TO	·0312	·1115	·4068	·1183	·0943	·0591	·0217		
7LM	·3606	·1611	·1216	·0997	·0822	·0589	·0228		
7LO	·0344	·0663	·1109	·0820	·3852	·1714	·0568		
8TM	·3595	·1610	·1190	·0946	·0741	·0539	·0329	·0111	
8TO	·0236	·0705	·1234	·4083	·1173	·0870	·0563	·0197	
8LM	·3592	·1619	·1214	·1012	·0865	·0718	·0507	·0191	
8LO	·0327	·0669	·0880	·1126	·4079	·1337	·0892	·0407	
9TM	·3596	·1613	·1202	·0976	·0800	·0632	·0460	·0280	·0094
9TO	·0182	·0527	·0888	·1308	·4083	·1125	·0814	·0541	·0186
9LM	·3598	·1616	·1217	·1015	·0883	·0769	·0636	·0443	·0164
9LO	·0309	·0702	·1111	·1069	·4026	·1414	·0735	·0658	·0317

Accordingly, the various realizations are indicated by a simple alpha-numeric code in which the first figure gives the order of the approximation. This is followed by L or T for Laguerre or Taylor, respectively, and finally by O for optimum or M for maximized. Thus, 5LO signifies a fifth-order Laguerre network with the optimum selection of element values.

3. Comparison of Types

As has already been pointed out above, the primary object of the present investigation was the design of pulse shaping networks providing well-shaped transitions, for which purpose the Gaussian shape provides a convenient, although not unique, point of departure. The criterion to be adopted in judging the success of a network is therefore not the extent to which it matches the true Gaussian shape, but rather the general shape of the transition.

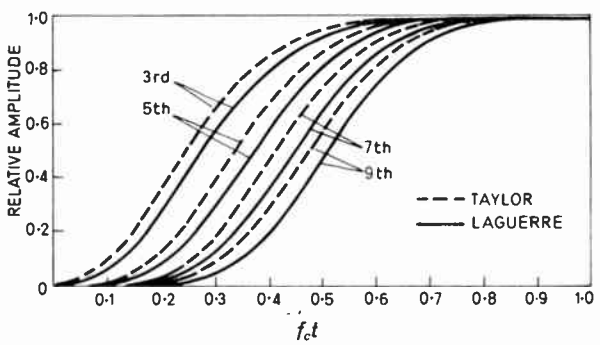


Fig. 4. Comparison of calculated Taylor and Laguerre step responses

The theoretical step responses for the odd orders up to 9 are given in Fig. 4 for both the Laguerre and the Taylor approximations; in practice the odd orders are almost invariably preferred since they terminate at each end in a capacitor, which can be allowed to absorb the circuit stray capacitances. Networks of the 3rd, 5th and 7th orders of each type were constructed according to the information given in Table 3 and tested; in each case the value of f_c was 1·5 MHz. Their impulse and step responses are shown in Fig. 5.

It is clear from these photographs that the Laguerre approximation furnishes a significantly better-shaped waveform in each instance, as one might expect, although the differences naturally become less with increasing order. A further advantage of the Laguerre type is the more favourable values of the components, particularly with respect to the final shunt capacitors in the optimum realization, which are higher than those of the Taylor type by a factor of at least 1·5. There seems every advantage, therefore, in preferring the Laguerre approximation networks to the Taylor approximation.

It is worth noting in passing that the impulse response of the 7th-order networks is almost perfectly symmetrical and a close approximation to the sine-squared shape. Since the impulse response is the derivative of the step response, this is a confirmation that even by the 7th-order, the desired skew-symmetry of the step response has already been closely approached, and for most purposes higher-order networks are not required. In some instances, even the 3rd-order Laguerre network may be found adequate.

4. Practical Considerations

The element values for the general configuration of Fig. 3 are given in Table 3 for 1 ohm terminations and f_c i.e. the frequency of 3 dB loss, equal to 1 Hz; capacitances are then in farads and inductances in henrys. Alternatively, for television purposes it is more convenient to consider the nominal cut-off frequency to be 1 MHz, when the values in Table 3 are read in microfarads and microhenrys respectively. For any other terminating resistance R ohms the capacitances are divided by R and the inductances are multiplied by R. Likewise, for any other cut-off frequency F all capacitance and inductance values will be divided by F.

The value allotted to f_c is usually found from the requirement that the output waveform shall have a given time of rise, defined as the time interval on the step response between the points at which the amplitude is 10% and 90% of the final value. The times of rise of the various networks are given in Table 4; it will be noticed that an asymptotic value of 0·341 is rapidly approached as the order of the network increases. Thus, for example, the f_c corresponding to a desired time of rise of 0·2 μs would be $0{\cdot}341 \times 10^6/0{\cdot}2$, say 1·7 MHz. For this to be true the time of rise of the input waveform must be sufficiently short, as stated in Section 2, the limiting value being dependent upon the tolerance placed upon the time of rise of the output pulse.

In certain instances one may wish to take advantage of the fairly flat group delay response, or alternatively the

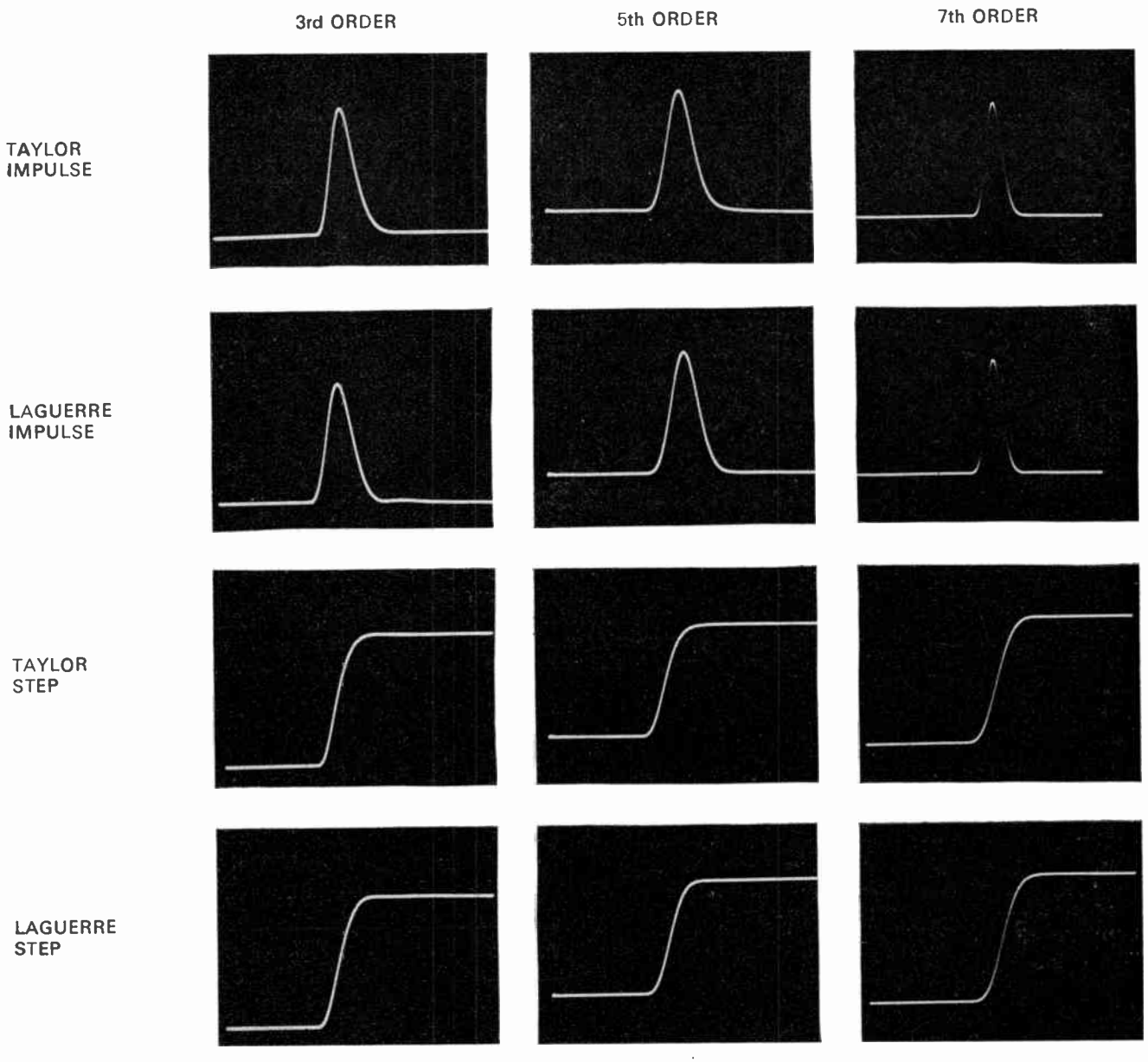


Fig. 5. Responses to impulses and step waveforms for 3rd, 5th and 7th order networks.

transmission time through the network may be needed. This is also given in Table 4 in the form of the zero-frequency group delay, which defines the location of the centroid of the impulse response. This is identical along the time axis with the position of the peak of the pulse in the case of the higher-order networks where the impulse response is symmetrical, but a slight displacement is found with lower orders. The group delay at f_c is also given in Table 4.

It might, perhaps, be mentioned that the element values for the dual of the network of Fig. 3 can also be found directly from Table 3, capacitances being read for inductances, and vice-versa. This is a consequence of the normalization of the impedance to 1 ohm.

The permissible tolerance limits on the component values naturally depend upon the extent to which the response may be allowed to depart from the theoretical, but as a practical guide $\pm 5\%$ capacitors and inductors will be found suitable for most purposes. The inductors should preferably have a Q of not less than 50. Apart from increasing the basic loss of the filter, one of the effects of dissipation is to increase the time of rise over the theoretical. A number of these networks of a range of

Table 4

Network type	Time of rise	Delay $f=0$	Delay $f=f_c$
3T	·342	·262	·227
3L	·355	·297	·269
5T	·344	·350	·328
5L	·344	·381	·362
7T	·341	·421	·404
7L	·341	·453	·438
9T	·341	·483	·468
9L	·341	·517	·504

The above times are in seconds for a network with $f_c = 1$ Hz, or in microseconds if $f_c = 1$ MHz.

orders of approximation have been constructed and tested, and in no instance has any particular difficulty been encountered.

5. Acknowledgment

The authors would like to thank the Director of Engineering of the B.B.C. for permission to publish the above work, which was carried out in the Measurement Laboratory of Designs Department.

6. References

1. Macdiarmid, I. F., 'A testing pulse for television links,' *Proc. Instn Elect. Engrs*, **99**, Part IIIA, pp. 436–44, 1952.
2. C.C.I.R. Recommendation 421, Annex IV, Geneva 1963.
3. Dishal, M., 'Gaussian-response filter design,' *Electrical Communication*, **36**, no. 1, pp. 3–26, 1959.
4. Jones, N. B., 'A Laguerre series approximation to the ideal Gaussian filter,' *The Radio and Electronic Engineer*, **40**, no. 3, pp. 151–5, September 1970.
5. Weinberg, L., 'Network Analysis and Synthesis' (McGraw-Hill, New York, 1962.)
6. *Ibid.*, pp. 594–5.
7. Bode, H. W., 'Network Analysis and Feedback Amplifier Design,' pp. 361–8. (Van Nostrand, New York, 1945.)

Manuscript first received by the Institution on 1st March 1971 and in final form on 17th May 1971. (Paper No. 1412/CC 106.)

Oral Expression of Metric Dimensions in Engineering

The subject of oral expression of metric dimensions in drawing offices and on the shop floor has been fully discussed during recent months by the Joint Committee on Metrication set up by the Royal Society and the Council of Engineering Institutions. In the Committee's opinion this matter warrants careful consideration by all sectors of industry if confusion is to be avoided from the outset. A consensus of opinion from all levels of industry—shop floor to management—is felt to be important and for this reason the Institution has been asked to give prominence to the following statement.

When in a British Engineering context, someone speaks of a 'thou' or a 'tenth of a thou', the meaning is clear and unambiguous. There is an implicit reference to the inch which does not have to be stated explicitly. Metrication will change this situation and it is timely to consider how metric dimensions and tolerances should be expressed orally, in the Drawing Office and on the Shop Floor. In particular, it is necessary to consider whether any colloquial terms could possibly be taken to relate to the inch instead of the millimetre, or vice versa, leading to errors in production.

In the Engineering industries there is a general intention to use the millimetre for most applications. For the larger dimensions the metre will be used. For very small dimensions in precision engineering the micrometre or decimal fractions of the millimetre will be used.

Synonym for Millimetre

There are three alternatives: millimetre; milli; or mil.

The full name 'millimetre' will undoubtedly be widely used and is recommended particularly in the early stages of the changeover. Of the two colloquial names, 'milli' is preferred as there is a possibility of 'mil' being confused with the American colloquial form for the one-thousandth of an inch.

Synonym for Micrometre

The only feasible alternative to the word 'micrometre' in full is the 'micron'. Although this is deprecated by some authorities it is nevertheless extensively used and is preferable to the 'micrometre' which, apart from being unwieldy, bears a resemblance to the measuring instrument, the 'micrometer'.

Name for 'O'

The alternatives here are: oh; nought; or zero.

The third alternative, 'zero', is gaining popularity and may well become the accepted international standard but this is likely to take some time. Meanwhile it is recommended that the second alternative, 'nought', be used in preference to 'oh' which is difficult for some Europeans and could cause confusion in communication with metric countries. Thus, 0·001 mm would be expressed as point nought nought one.

Decimal Fractions of the Millimetre

The main alternatives here are:

Dimension	Form 1	Form 2	Form 3
0·1 mm	point one	one tenth	one hundred microns
0·01 mm	point nought one	one hundredth	ten microns
0·001 mm	point nought nought one	one thousandth	one micron

In the Drawing Office and for formal expressions of dimensions, Form 1 is recommended.

On the shop floor it is likely that, in due course, 'one tenth' and 'one hundredth' will be used respectively for 0·1 mm and 0·01 mm but where there is a possibility of confusion with imperial measurements, Form 1 is recommended for an interim period.

On the shop floor and in the tool room the word 'micron' for 0.001 mm will undoubtedly be widely used and is recommended.

The RS/CEI Joint Committee on Metrication would welcome constructive comments from readers on the extent and intensity of this problem and on the above recommendations for colloqiual usage as soon as possible. Comments should be addressed to the RS/CEI Joint Committee on Metrication Secretariat, Institution of Production Engineers, 10 Chesterfield Street, London W1X 8DE.

U.D.C. 511.118: 621.374.4

A Versatile Cellular Array for Binary Arithmetic

By

G. WHITE, B.Sc.†

A cellular array is proposed which can be controlled to perform the following arithmetic functions; A+BD+C, A−BD−C, or A/B. This results in a saving of hardware over previous arrays, and by having more terms in each function, a reduction in the overall calculation time can be achieved.

Several iterative arrays have been described recently, which will either multiply[1] or divide[2]. Gex[3] has put forward a single array which will multiply or divide when programmed from a row of control cells. The proposed array is more versatile than those previously described, and uses a lower number of cells than the array put forward by Gex. By suitably choosing the cell logic, the boundary conditions A and C (shown in Fig. 1), can be utilized when multiplying.

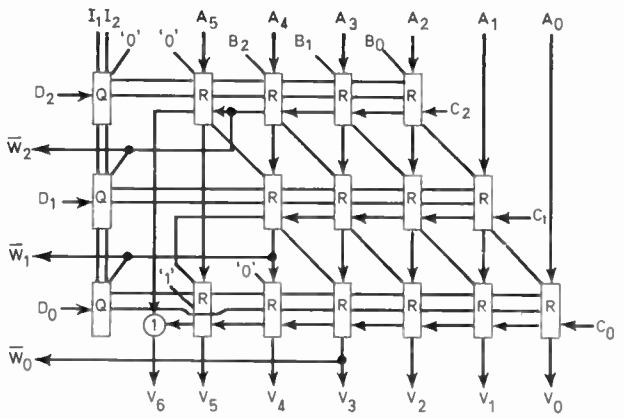


Fig. 1. Complete array.

The complete array is shown in Fig. 1 where V = A $\pm$ (B D + C) or W = A/B. The multiplier operates by forming the partial products in columns of the correct weight and summing these partial products into the final product (B × D). There is no necessity for the partial products to be added to zero, hence by using the boundary inputs A, the final product becomes A+BD. The partial products can be summed by repeated addition; or by repeated subtraction, giving a positive or negative sign to the final product (A+BD or A−BD). The cell logic, as shown in Fig. 2, is such that the multiplicand is added to, or subtracted from, the sum of the previous partial products, if the corresponding bit in the multiplier is a '1'. When the multiplier bit is a '0' the 'carries' only are 'added' to the previous partial products; hence the boundary input C can be used. Therefore in the multiplication mode the array performs the functions A+BD+C when adding, or A−BD−C when subtracting. A carry-save layout is used in the array to keep the number of cell delays to a minimum.

For division the non-restoring method is employed because this requires the cells to add or subtract, depending upon a control signal. These are the same characteristics as used when multiplying, so the same cells can be used for multiplication or division to form a versatile array. To calculate the quotient bit of 2^n, the divisor $B \times 2^n$, is subtracted from the dividend. When the answer is positive the quotient bit is '1' and the following row subtracts $B \times 2^{n-1}$ from the remainder, to calculate the next quotient bit. When the answer is negative the quotient bit is '0' and the next row adds the remainder to $B \times 2^{n-1}$. A quotient bit can only be 0 or 1, therefore $A < 2B2^n$ to obtain the correct result. The C input cannot be used when dividing because of the possible ripple of borrows across the array changing the value of A, which would not be available for a previous row.

To program the array to perform different functions a row of control cells Q are required. The inputs I_1 and I_2 of the control cells determine the functions as follows:

I_1	I_2	Function
0	0	A−BD−C
0	1	A+BD+C
1	X	A/B

The outputs e′ and d′ shown in Fig. 2, are controls to the arithmetic cells R. Output d′ is one bit of the multiplier input, when the array is used for multiplication; and is held at '1' when dividing. The control e′ determines whether the cells add or subtract. e′ depends upon the input w from the previous row when dividing, and upon the control I_2 when multiplying.

The arithmetic cell which was first described by Guild,[4] is a conditional adder/subtractor. If the input e

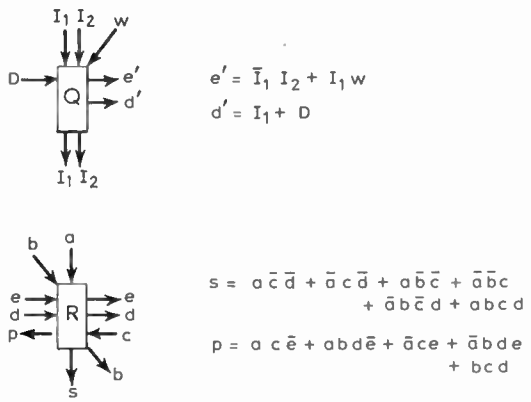


Fig. 2. Cell logic.

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

is 1, the output s = a + bd + c; when e is 0, s = a − bd − c,† The output p is the carry/borrow of s. Both cells are shown in Fig. 2 with their Boolean descriptions.

Let the number of cells in the first row be $n+1$ and the number of rows be m. (The array need not necessarily have $n = m$ as shown in the illustration.) Then the number of bits in the inputs are as follows:

When multiplying:

$$A = n+m$$
$$B = n$$
$$D = C = m$$

giving an output of $n+m+1$.

When dividing:

$$A = n+m-1$$
$$B = n$$

giving an output of m.

The array requires $nm+2m-1$ arithmetic cells and m control cells. The worst case propagation delay when multiplying is $(n+m)\tau_R$ and when dividing is

$$(nm+m-1)\tau_R+(m-1)\tau_Q,$$

where τ_R is the propagation delay of an arithmetic cell and τ_Q of a control cell.

The proposed array has a regular structure and would therefore be suitable for large scale integration. The regularity becomes more apparent than the illustration suggests when n and m are larger. The effective cost of producing such an array would be lower than previous arrays due to the higher number of functions performed. When multiplying two numbers together the result has often to be added to or subtracted from another number. With this array these operations can occur simultaneously.

† The small letters are used to define the cell's operation as distinct from the array's function, which is indicated by capital letters.

Acknowledgment

This work has been supported by International Computers Ltd. and the U.K. Department of Trade and Industry by a normal cost-sharing contract within the advanced computer technology project.

References

1. Hoffman, J. C., Lacaze, B. and Csillag, P., 'Multiplieur parallèle à circuits logiques itèratifs'. *Electronics Letters*, **4**, p. 178, May 1968.
2. Dean, K. J., 'Cellular arrays for binary division', *Proc. Instn Elect. Engrs*, **117**, pp. 917–20, May 1970.
3. Gex, A., 'Unité arithmétique et logique ultrarapide', Thesis, Ecole Polytechnique, Lausanne, Switzerland, 1970.
4. Guild, H. H., 'Some cellular logic arrays for non-restoring binary division', *The Radio and Electronic Engineer*, **39**, pp. 345–348, June 1970.

Manuscript first received by the Institution on 3rd June 1971 and in final form on 16th August 1971. (Short Contribution No. 145/Comp. 138).

U.D.C. 621.376.223

The Signal-handling Capacity of the Square-wave Switched Ring Modulator

By

J. G. GARDINER,
B.Sc., Ph.D. (Graduate)†

Expressions are derived for the maximum permissible interfering input signal power which may be applied to a square-wave switched resistively terminated ring modulator before the onset of severe cross-modulation distortion in the output spectrum. The effects of source and load mismatch are considered and good agreement demonstrated between theoretical prediction and the measured performance of a typical ring modulator using Schottky barrier diodes.

1. Introduction

The distortion properties of switching diode modulators have been considered extensively in a number of papers[1-3] and it has been demonstrated that the resistively-terminated ring modulator represents a good compromise between the usually conflicting requirements of low conversion loss and low distortion generation. It is further recognized that the rapidity with which the diodes in the ring can be switched by the local oscillator is a major factor in determining the levels of distortion products at the output (although, as demonstrated in ref. 4, some contribution exists due to the finite curvature of the diode forward characteristics) and now that switching times in the region of 150 picoseconds can be obtained with relatively straightforward wave-shaping circuits, the use of square-wave or pulse drive is possible up to switching frequencies in excess of 50 MHz. Consequently it is felt that an examination of the properties of square-wave switched ring mixers and modulators is of great interest.

In the present paper the cross-modulation performance of a typical ring mixer is considered and it is shown that restrictions may be formulated on the level of interfering signal which may be accepted by the mixer before an unacceptably rapid deterioration in performance is incurred. It is shown that the d.c. offset voltage in the diode characteristic plays an important part in determining the overload properties of the circuit even under square-wave drive conditions (the influence of this parameter in sine-wave local oscillator situations is discussed in ref. 5) and it is also demonstrated that impedance mismatching at either or both signal ports of the mixer contributes greatly to the signal handling characteristics of the circuit.

2. Overload Conditions in the Ring Modulator

The circuit of a typical modulator is shown in Fig. 1 and the effective local-oscillator circuit under small input conditions appears in Fig. 2. Overload takes place when one of the diodes biased into conduction by the local-oscillator is turned off by an equal opposing input signal current or when input signals produce an instantaneous voltage across a reverse biased diode sufficient to turn it on. It is apparent from Fig. 2 that the reverse bias appearing across any diode is determined by the forward voltage developed across the conducting

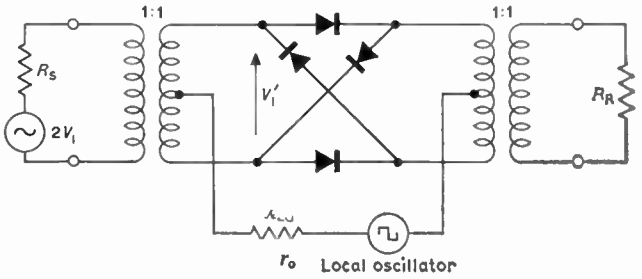


Fig. 1. Resistively terminated ring modulator.

diodes and that this in turn is determined by the magnitudes of the diode offset voltage and forward resistance and the local-oscillator source impedance.

The most commonly observed effects of overload arise from the fact that it permits the input signals to exercise an influence over the impedance condition of the diodes so that non-linear effects become apparent. For instance, if the local-oscillator frequency is ω_c and input signals are present at ω_1 and ω_2, then intermodulation products occur at frequencies $m\omega_c+p\omega_1+q\omega_2$, where p, q and m are integers and $|p|+|q| \geq 2$. Cross-modulation effects are another manifestation of the same phenomenon and arise when one of the signals (at ω_2 for instance) is a modulated interference; at the modulator output this modulation appears not only on products $m\omega_c+q\omega_2$ but also to a lesser extent on products $m\omega_c+p\omega_1$.

In a sine-wave switched modulator overload conditions of this sort arise as the local-oscillator wave approaches zero amplitude at half-period intervals and in consequence intermodulation distortion is always present to some degree in the output spectrum. Under ideal conditions of instantaneous switching the output spectrum would be completely free of distortion products until the magnitude of input signals is sufficient to change the state of a fully biased diode and thereafter any increase in input-signal level results in very rapid increase in the levels of intermodulation or cross-modulation products at the output.

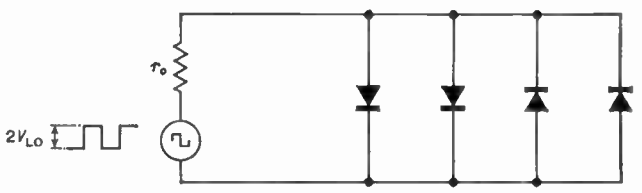


Fig. 2. Diode circuit as seen by local oscillator.

† Postgraduate School of Electrical and Electronic Engineering, University of Bradford, Bradford 7.

It has been shown previously[2] that in sine-wave switching a series of overload conditions arise during the local-oscillator polarity transition as the first overload influences the circuit conditions producing second and subsequent conditions until complete commutation is achieved; in the present treatment only the first overload point will be considered since this represents the limit of undistorted operation.

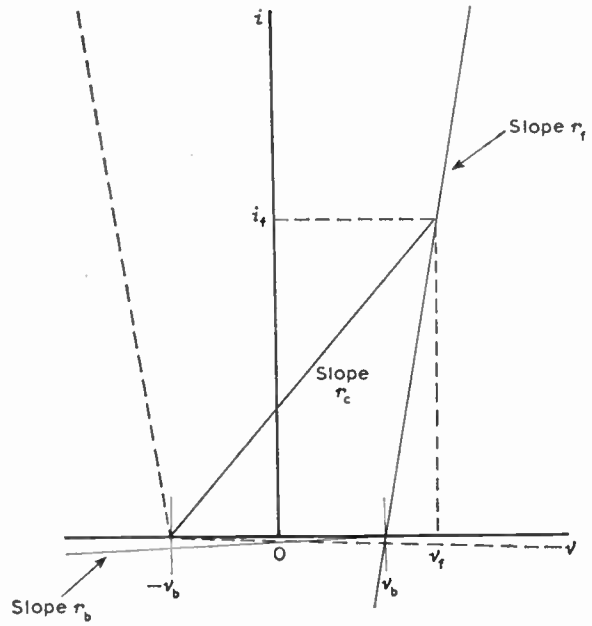


Fig. 3. Idealized diode characteristic.

The diode model used for analysis is shown in Fig. 3. It is assumed that the device possesses high resistance, r_b, up to an applied forward bias v_b and that for any applied positive voltage greater than this its resistance falls to a small value, r_f. The voltage developed across the reverse biased diodes is seen to be v_f (for a given drive current i_f in each forward biased diode) but an applied input signal must develop a further voltage v_b across the reverse biased diodes before either can be turned 'on'. Therefore we may consider the 'off' diodes to have characteristics centred on an origin $(-v_b, 0)$ so that the apparent chord resistance of the 'on' diodes, r_c, as seen by the local oscillator then becomes

$$r_c = \frac{v_f + |v_b|}{i_f} \qquad \text{.....(1)}$$

Referring to Fig. 1 it is readily demonstrated that the voltage developed by the input signal, V_1', across the conducting diodes is

$$V_{fs} = \frac{r_f(R_R + r_b)}{2r_b r_f + R_R(r_b + r_f)} \cdot V_1' \qquad \text{.....(2)}$$

and across the 'off' diodes is

$$V_{bs} = \frac{r_b(R_R + r_f)}{2r_b r_f + R_R(r_b + r_f)} \cdot V_1' \qquad \text{.....(3)}$$

It is also possible from elementary network theory to derive an expression for V_1' in terms of V_1. This is:

$$V_1' = 2V_1 \left\{1 - \frac{R_s[2R_R + r_b + r_f]}{r_b(R_R + r_f) + r_f(R_R + r_b) + R_s(2R_R + r_b + r_f)}\right\} \qquad \text{.....(4)}$$

It is apparent then that overload will occur by turning an 'off' diode 'on' if

$$i_f r_c = V_{bs} \qquad \text{.....(5)}$$

and by turning an 'on' diode 'off' if

$$i_f r_f = V_{fs} \qquad \text{.....(6)}$$

Defining

$$r_c = \alpha r_f \qquad \text{.....(7)}$$

we may now evaluate equations (5) and (6) since the maximum available input signal power is given by

$$P_i = \frac{V_1^2}{2R_s} \qquad \text{.....(8)}$$

and we may write (5) as

$$i_f \alpha r_f = \theta V_1 \qquad \text{.....(9)}$$

and (6) as

$$i_f r_f = \phi V_1 \qquad \text{.....(10)}$$

where

$$\theta = \frac{2r_f(R_R + r_b)}{2r_b r_f + R_R(r_b + r_f)} \times \left\{1 - \frac{R_s(2R_R + r_b + r_f)}{r_b(R_R + r_f) + r_f(R_R + r_b) + R_s(2R_R + r_b + r_f)}\right\} \qquad \text{(11)}$$

and

$$\phi = \frac{r_b(R_R + r_f)}{r_f(R_R + r_b)} \cdot \theta \qquad \text{.....(12)}$$

giving voltage overload of an 'off' diode at P_{iv} given by

$$P_{iv} = \frac{i_f^2 r_f^2 \alpha^2}{2\theta^2 R_s} \qquad \text{.....(13)}$$

and current overload of an 'on' diode at P_{ic} where

$$P_{ic} = \frac{i_f^2 r_f^2}{2\phi^2 R_s} \qquad \text{.....(14)}$$

Which of these two conditions (13) or (14) arises first depends on the operating point of the forward biased diodes since the power required to produce voltage overload is related to that required to give current overload by

$$\frac{P_{iv}}{P_{ic}} = \alpha^2 \cdot \frac{r_b^2(R_R + r_f)^2}{r_f^2(R_R + r_b)^2} \qquad \text{.....(15)}$$

3. Overload Conditions in the Modulator Terminated for Minimum Conversion Loss

Termination for minimum conversion loss represents a particularly simple case since, as is well known,[6] the optimum terminating resistance is given by the characteristic impedance of the lattice before overload, i.e.

$$R_s = R_R = R = \sqrt{r_b r_f} \qquad \text{.....(16)}$$

Under these conditions

$$V_1' = V_1$$

Further, by defining the diode switching characteristic in terms of a quality factor, n, given by

$$n = \sqrt{\frac{r_b}{r_f}} \qquad \text{.....(17)}$$

we may write

$$V_{bs} = \frac{n}{n+1} V_1 \qquad \text{.....(18)}$$

$$V_{fs} = \frac{1}{n+1} V_1 \qquad \text{.....(19)}$$

giving

$$P_{iv} = \frac{i_f^2 \alpha^2 r_f (n+1)^2}{2n^3} \qquad \text{.....(20)}$$

and

$$P_{ic} = \frac{i_f^2 r_f (n+1)^2}{2n} \qquad \text{.....(21)}$$

Therefore, four factors are seen to determine the maximum permissible input signal level, r_f, n, α, and i_f. The effect of varying the drive level is illustrated in Fig. 4 for a typical silicon Schottky barrier diode (HP 2900) assuming a forward resistance of 10 Ω and a value n of 50.

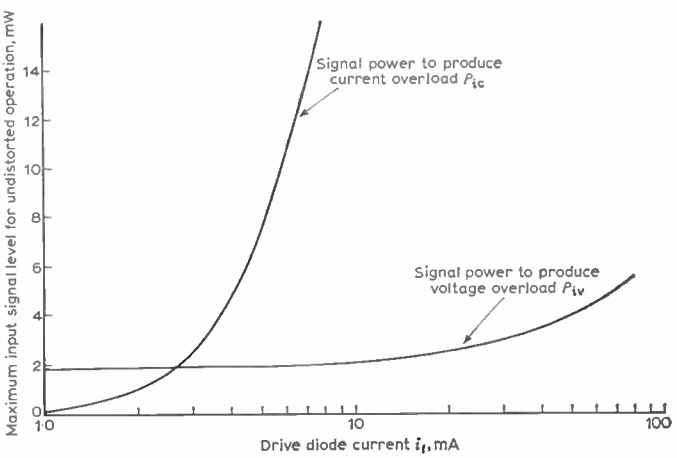


Fig. 4. Modulator terminated for minimum conversion loss; signal handling capacity.

4. Effects of Impedance Mismatch

The effects of mismatch at source and load ports of the modulator are of interest for two reasons.

(a) Variations in circuit component values and transistor parameters due to tolerances and temperature changes can result in wide variations in the impedances presented to the modulator when noise figure considerations prevent the use of attenuator pads to buffer the modulator from associated circuitry. It will be shown that this can result in severe degradation in signal handling capacity while only small changes in conversion loss are appreciable.

(b) Deliberate mismatch can result in enhanced signal handling capacity at the expense of small increases in loss and in particular much can be deduced about the nature of overload in the ring by examining the performance with varying degrees of mismatch.

4.1 *Degradation of Conversion Loss*

It may readily be shown from the work of Tucker,[6] that the modulator conversion loss in dB is given by:

$$L = 20 \log \left\{ \frac{\pi}{2} \cdot \frac{[R_R + \frac{1}{2}(r_b + r_f)][R_s + \frac{1}{2}(r_b + r_f)] - \frac{1}{4}(r_b - r_f)^2}{\sqrt{R_s R_R (r_b - r_f)}} \right\} \qquad \text{.....(22)}$$

Now defining

$$R_s = kR \qquad \text{.....(23)}$$

and

$$R_R = \lambda k R = \lambda R_s \qquad \text{.....(24)}$$

then the loss may be written

$$L = 20 \log \left\{ \frac{\pi}{2} \cdot \frac{[k\lambda n + \frac{1}{2}(n^2+1)][kn + \frac{1}{2}(n^2+1)] - \frac{1}{4}(n^2-1)^2}{\sqrt{\lambda}\, kn(n^2-1)} \right\} \qquad \text{.....(25)}$$

from which under matched conditions the minimum loss is

$$L_{min} = 20 \log \frac{\pi}{2} \cdot \frac{n+1}{n-1} \qquad \text{.....(26)}$$

The expression for loss (22) is seen to be symmetrical in R_s and R_R so that mismatch at source or load with the other termination in each case maintained at optimum results in a loss degradation ΔL_1 of

$$\Delta L_1 = 10 \log \left\{ \frac{(1+k)^2}{4k} \right\} \qquad \text{.....(27)}$$

If deliberate mismatch exists at either port then the termination at the other port may be adjusted for minimum loss by making

$$\lambda_{opt} = \frac{kn^2 + 2n + k}{k(n^2 + 2kn + 1)} \qquad \text{.....(28)}$$

and under these conditions ΔL_2 reduces to

$$\Delta L_2 = 10 \log \left\{ \frac{[(n^2+1) + 2kn][kn^2 + k + 2n]}{k(n+1)^4} \right\} \qquad (29)$$

4.2 *Signal Handling Capacity*

Equation (4) may now be written in terms of k, λ and n, giving

$$V_1' = 2V_1 \left\{ 1 - \frac{k[2k\lambda n + n^2 + 1]}{n[k\lambda n + 1] + k\lambda + n + k[2k\lambda n + n^2 + 1]} \right\} \qquad \text{.....(30)}$$

and again solving this and equations (2)–(8) for V_1, in terms of i_f, α, r_f, n, k and λ permits evaluation of P_{iv}' and P_{ic}', the powers required respectively to produce voltage overload and current overload under mismatched conditions.

Four situations are of particular interest.

4.2.1 Source mismatched, load fixed at R

Here λ in equation (24) is set to $(1/k)$.

This gives

$$V_{bs} = \frac{n}{n+1} V_1' \qquad \text{.....(31)}$$

i.e.

$$V_1 = \frac{i_f \alpha r_f (n+1)(1+k)}{2n} \qquad \text{.....(32)}$$

leading to

$$P_{iv}' = \frac{i_f^2 \alpha^2 r_f (n+1)^2 (1+k)^2}{8kn^3} \qquad \text{.....(33)}$$

and

$$\Delta P_{iv} = \frac{P_{iv}'}{P_{iv}} = 10 \log \left\{ \frac{(1+k)^2}{4k} \right\} \text{ dB} \qquad \text{.....(34)}$$

which is, as expected, identical to the increase in modulator conversion loss. A similar calculation for P_{ic} gives

$$\Delta P_{ic} = 10 \log \left\{ \frac{(1+k)^2}{4k} \right\} \text{ dB} \qquad \text{.....(35)}$$

as before.

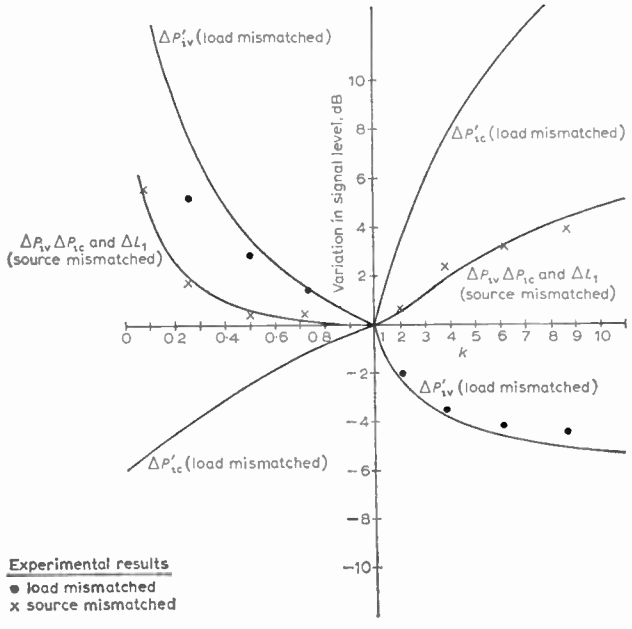


Fig. 5. Variation of signal handling capacity. One termination fixed at R_{opt}.

4.2.2 Load mismatched, source fixed at R

As indicated earlier, the variation of loss with load mismatch is identical to that observed under source mismatch conditions but to avoid confusion we will define

$$R_R = xR \qquad \text{.....(36)}$$

$$R_s = x\theta R \qquad \text{.....(37)}$$

with, for this case, $\theta = 1/x$ and R defined as before. Calculating as before

$$V_{bs} = \frac{n(xn+1)}{xn^2+2n+x} V_1' \qquad \text{.....(38)}$$

giving

$$V_{bs} = \frac{2n(xn+1)}{(x+1)(n+1)^2} V_1 \qquad \text{.....(39)}$$

i.e.

$$P_{iv}' = \frac{i_f^2 \alpha^2 r_f (x+1)^2 (n+1)^4}{8n^3 (xn+1)^2} \qquad \text{.....(40)}$$

which for $n \gg 1$ gives

$$\Delta P_{iv} = 10 \log \left\{ \frac{(1+x)^2}{4x^2} \right\} \qquad \text{.....(41)}$$

Again calculating the current overload point

$$\Delta P_{ic} = 10 \log \left\{ \frac{(1+x)^2}{4} \right\} \qquad \text{.....(42)}$$

These relationships are plotted in Fig. 5.

4.2.3 Source mismatched, load optimized for minimum conversion loss

In this situation λ takes the optimum value of equation (28) resulting in

$$V_{bs} = \frac{n[kn^3+3n^2+3kn+1]}{[kn^2+2n+k][2kn+n^2+1]} V_1 \qquad \text{.....(43)}$$

giving

$$P_{iv}' = \frac{i_f^2 \alpha^2 r_f}{2kn} \left\{ \frac{(kn^2+2n+k)(2kn+n^2+1)}{n(kn^3+3n^2+3kn+1)} \right\}^2 \qquad \text{...(44)}$$

or

$$\Delta P_{iv}' = 20 \log \left\{ \frac{kn^2(n^2+6)+2n(k^2+1)(n^2+1)+k}{\sqrt{k}(n+1)(kn^3+3n^2+3kn+1)} \right\} \text{ dB} \qquad \text{.....(45)}$$

which for $n \gg 1$, k gives

$$\Delta P_{iv}' = 20 \log \left\{ \frac{2(k^2+1)+kn}{\sqrt{k}(nk+3)} \right\} \text{ dB} \qquad \text{.....(46)}$$

Similarly

$$\Delta P_{ic}' = 20 \log \frac{1}{\sqrt{k}} \times \times \left\{ \frac{(kn^2+2n+k)(2kn+n^2+1)}{(n+1)[(kn^2+2n+k)+n(2kn+n^2+1)]} \right\} \text{ dB} \qquad \text{.....(47)}$$

which for $n \gg 1$, k gives

$$\Delta P_{ic}' = 20 \log \left\{ \frac{1}{\sqrt{k}} \cdot \frac{kn}{n+k} \right\} \text{ dB} \qquad \text{.....(48)}$$

These relationships are plotted in Fig. 6.

4.2.4 Load mismatched, source optimized for minimum loss

In this case θ takes the value

$$\theta_{opt} = \frac{xn^2+2n+x}{x(2xn+n^2+1)} \qquad \text{.....(49)}$$

$$V_{bs} = \frac{n(xn+1)}{xn^2+2n+x} \cdot V_1 \qquad \text{.....(50)}$$

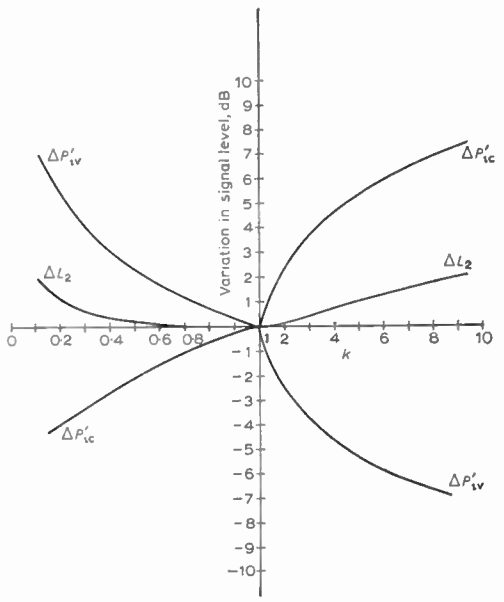


Fig. 6. Variation of signal handling capacity. Source mismatched, load optimized.

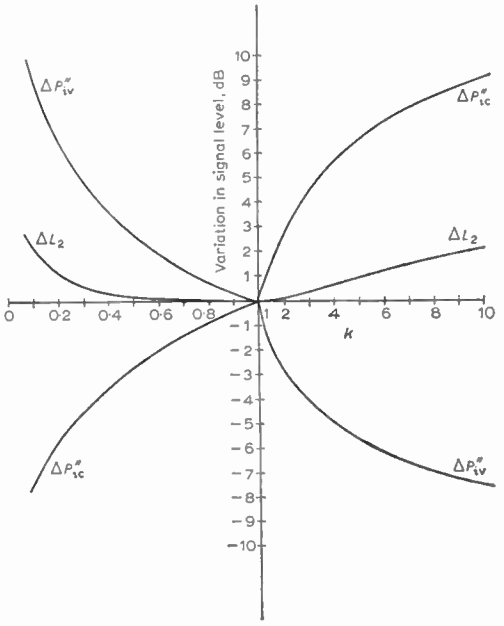


Fig. 7. Variation of signal handling capacity. Load mismatched, source optimized.

giving

$$P'_{iv} = \frac{i_f^2 \alpha^2 r_f (xn^2+2n+x)^2}{n^3(xn+1)^2 \cdot 2x\theta_{opt}} \quad \text{.....(51)}$$

which gives

$$\Delta P''_{iv}] = 10 \log \left\{ \frac{(2xn+n^2+1)(xn^2+2n+x)}{(xn+1)^2(n+1)^2} \right\} \text{ dB (52)}$$

which for $n \gg 1$ approximates to

$$\Delta P''_{iv} = 10 \log \left\{ \frac{n+2x}{x(n+2)} \right\} \text{ dB} \quad \text{.....(53)}$$

Similarly, the current overload point is given by

$$P''_{ic} = \frac{i_f^2 r_f (xn^2+2n+x)^2}{2xn\theta_{opt}(n+x)^2} \quad \text{.....(54)}$$

giving

$$\Delta P''_{ic} = 10 \log \left\{ \frac{(xn^2+2n+1)(2xn+n^2+1)}{(n+1)^2(n+x)^2} \right\} \text{ dB...(55)}$$

which for $n \gg 1$ gives

$$\Delta P''_{ic} = 10 \log \left\{ \frac{x(n+2/x)}{n+2} \right\} \text{ dB} \quad \text{.....(56)}$$

These results are plotted in Fig. 7.

5. Discussion of Results

5.1 *Experimental Ring Modulator*

The theoretical treatment of the foregoing section has assumed a single high-level interfering signal and this is a situation commonly encountered when receiving equipment is operated in the vicinity of a transmitter. In this environment the predominant effect of a modulated interfering signal is to produce cross-modulation distortion in the receiver and the level of 30% modulated interfering signal required to produce a specified degree of cross-modulation has, therefore, been taken as a means of comparing the performance of practical mixers with the predictions of theory.

It is to be expected that if fast rise-time square-wave switching is used little distortion should be apparent until the interfering signal level is of sufficient magnitude to produce an overload condition of the type mentioned above, but once this level has been exceeded a rapid increase in distortion should occur.

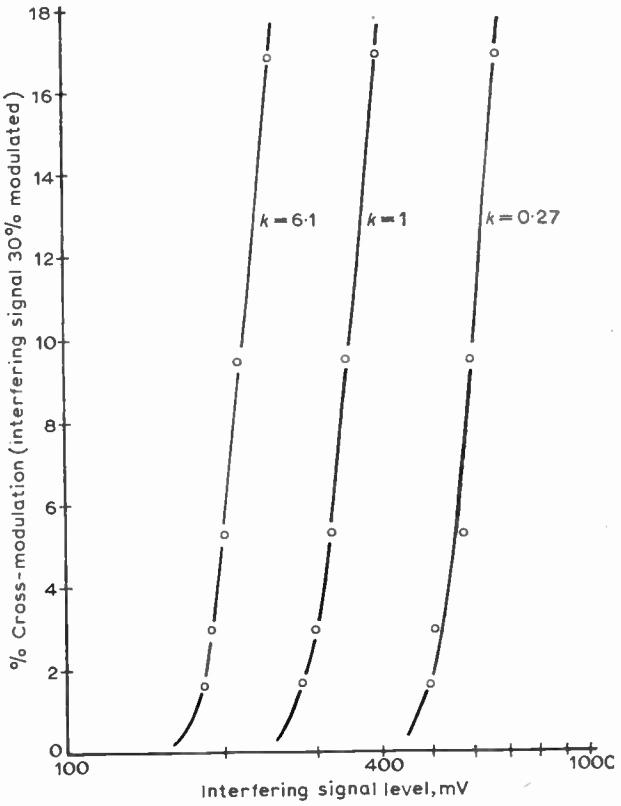


Fig. 8. Overload characteristics indicating effects of mismatch.

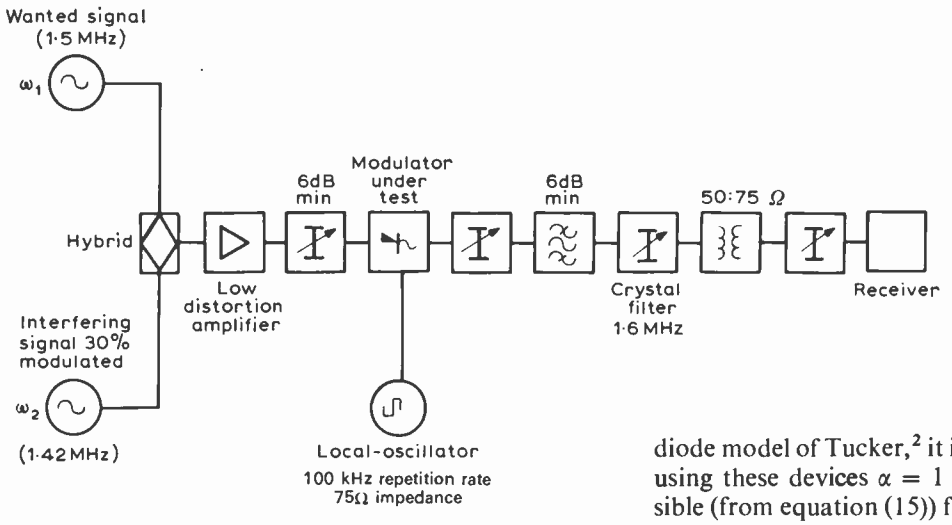


Fig. 9. Cross-modulation measuring set.

Figure 8 shows the result of an experiment in which a first-order modulation product of a wanted unmodulated signal and the local-oscillator was observed while an interfering signal 30% modulated was increased in level at the modulator input. As is apparent, at some point depending on the value of k, cross-modulation of the sidebands from the interfering signal onto the wanted product becomes significant and further increase in the interfering signal level produces very rapid increase in the extent of this effect measured as 'percentage cross-modulation'.

The experimental circuit employed the Schottky-barrier diodes type HP 2900 of Fig. 4 with transformer turns ratios of 1 : 1+1, the local-oscillator supply was provided by a pulse generator of 75 Ω output impedance capacitively coupled and operating at a repetition frequency of 100 kHz. The cross-modulation test set-up was of conventional type as discussed in ref. 3 and shown in Fig. 9 and used a 1·6 MHz crystal band-pass filter to isolate the mixer under test from the detecting receiver.

It is apparent from the experimental results plotted in Fig. 5 that good agreement is obtained between the predictions of theory and the performance of the practical circuit under mismatched conditions.

5.2 *Interpretation of Theoretical Results*

It is apparent that if a modulator is designed to work at the limit of its signal handling capability it is important to ensure that variations in the terminating conditions are kept to a minimum. The modulator is seen to be relatively insensitive to such variations as far as loss is concerned but differences in cross-modulation performance which accompany mismatch could easily exceed acceptable limits. For instance, in a modulator in which current overload predominates, a 20% mismatch at the load with matched input, corresponding to $k = 0{\cdot}8$, can result in a change in loss of approximately 0·1 dB but a deterioration in cross-modulation performance from 3% to 10% (see Fig. 8).

It is also seen that suitable mismatching can be used to equalize P_{ic} and P_{iv} and thereby increase the overload capacity of the modulator. Returning to the early diode model of Tucker,[2] it is apparent that in a modulator using these devices $\alpha = 1$ and consequently it is impossible (from equation (15)) for current overload conditions to arise before voltage overload. Tucker[6] suggests that by separating the forward and reverse transmission paths in the ring P_{iv} can be increased, by suitable biasing, to approach P_{ic} but the complicated transformers and drive circuitry required for this are difficult to realize at high frequencies.

In a practical system it is clear that if noise factor considerations are important, mismatch of the modulator is preferable to protecting the device by attenuator pads. Again referring to Fig. 5 it is seen that a 6 dB discrepancy between P_{iv} and P_{ic} can be removed by making $k = 0{\cdot}5$. This is accompanied by an increase in loss, and consequently in noise figure of only 0·5 dB while resulting in a 3 dB improvement in signal handling capacity for the same amount of output distortion.

6. Acknowledgments

The author is indebted to Mr. R. V. Stewart for obtaining the experimental results of Figs. 5 and 8, and to Mr. J. Dingley of Racal B.C.C. Ltd., whose interest in the mismatched operation of mixers and modulators prompted the present investigation.

7. References

1. Gardiner, J. G., 'The relationship between cross-modulation and intermodulation distortions in the double-balanced mixer', *Proc. Inst. Elect. Electronics Engrs*, **56**, pp. 2069–71, November 1968 (Letter).
2. Tucker, D. G., 'Intermodulation distortion in rectifier modulators', *Wireless Engineer*, **31**, pp. 145–52, June 1954.
3. Gardiner, J. G. and Yousif, A. M., 'Distortion performance of single-balanced diode modulators', *Proc. Instn Elect. Engrs*, **117**, No. 8, pp. 1609–14, August 1970.
4. Stewart, R. V. and Gardiner, J. G., 'Contribution to mixer intermodulation distortion of non-linearity in the diode forward characteristics', *Electronics Letters*, **7**, No. 10, pp. 279-81, 20th May 1971.
5. Gardiner, J. G., 'An intermodulation phenomenon in the ring modulator', *The Radio and Electronic Engineer*, **39**, pp. 193–7, April 1970.
6. Tucker, D. G., 'Modulators and Frequency Changers' (Macdonald, London, 1952).

Manuscript first received by the Institution on 18th January 1971 and in final form on 7th April 1971. (Paper No. 1413/CC107.)

U.D.C. 621.382.233

Measurement and Selection of Low Noise Avalanche Diodes

By

A. T. SULLIVAN,
B.Sc.†

This paper describes an investigation into methods of testing diodes to find those with a predictable and stable noise level, and further describes the noise characteristics of some diodes selected by these methods.

1. Introduction

It has been shown both theoretically and experimentally that p-n diodes operated in the avalanche breakdown region can be used as sources of electrical noise.[1,2,6] Special diodes have been manufactured, having guard rings and a controlled cross-sectional area to ensure uniform breakdown, which will provide a noise spectral density inversely proportional to the square root of the breakdown current and proportional to the breakdown voltage. As a source of noise these diodes have the advantages of providing high levels over a wide frequency range from a low source impedance, while requiring only a d.c. biasing circuit for operation. When facilities for manufacturing special diodes are not available, a method of selecting comparable commercial diodes would be very useful.

2. Selection of Diodes

The large majority of the diodes tested were found to contain excess noise produced by unstable breakdown areas. The voltage variation produced by these breakdowns can be large; diodes giving noise voltages in excess of 40 mV p-p were measured and this is far in excess of the intrinsic avalanche breakdown noise levels to be expected. Unstable breakdowns, or microplasmas, have been described[3,4,5] and are believed to be similar to those in a gas discharge tube, and caused by irregularities in the crystal structure of the diode reducing the breakdown voltage from the value that would be expected for a uniform junction. Such effects should not be present in a non-contaminated, linearly graded junction.

The effects of unstable breakdowns are displayed in Figs. 1 and 2, with Fig. 1 showing a portion of a diode characteristic about the breakdown voltage. As the voltage is increased from zero negligible current flows, until the breakdown voltage is exceeded. Then as the current increases the voltage decreases in an irregular manner, the slope of the *VI* relationship in this region being controlled by the impedance of the supply. As the diode current reaches approximately 70 μA the *VI* relationship becomes more stable, but a second discontinuity appears as the current reaches approximately 80 μA and the characteristic has become stable as *I* reaches 220 μA. A third discontinuity, is present at the top of the figure. Some hysteresis is apparent on all the characteristics.

Each of these areas is due to the breakdown of a fresh region of the diode as the voltage increases. The current required to flow through each microplasma to render it stable is approximately 60 μA and the effect of successive microplasmas on the characteristic become less as the total current increases, because of the shunting effects of the breakdown paths already carrying current.

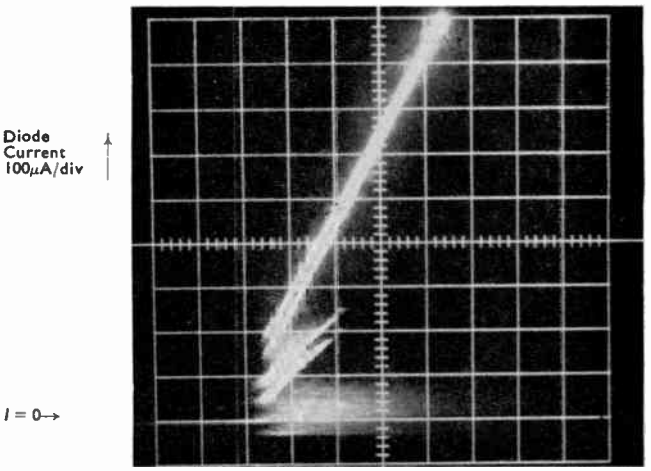


Fig. 1. Expanded diode characteristic showing discontinuities.

The initial breakdown is shown in more detail in Fig. 2. The instability at the start of current flow was found to be present in all diodes which broke down at above 9V, while in diodes breaking down below this voltage the

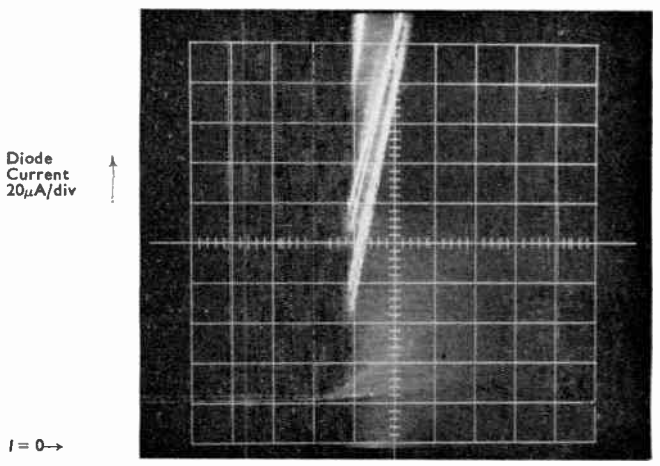


Fig. 2. Further expansion of diode characteristic.

† Department of Electrical and Electronic Engineering, Thames Polytechnic, Wellington Street, London, S.E.18.

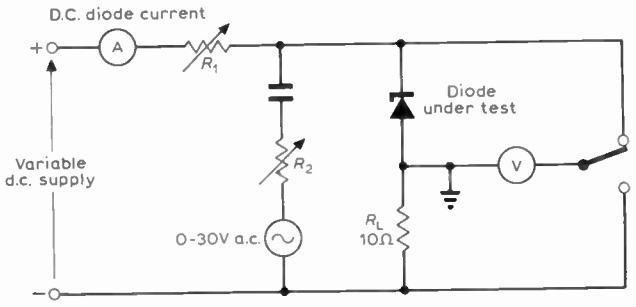


Fig. 3. Circuit for examining variation of slope resistance with diode current.

initial current was frequently found to flow more uniformly. This was thought to be due to the presence of some Zener breakdown since the knees of the characteristics were more rounded. Microplasmas were still found at higher current levels in most of the lower voltage diodes.

The diode current waveform was examined in these unstable regions, a constant voltage being applied to the diode. When this voltage is held at a value where current just begins to flow, the current waveform consists of a few random, unipolar pulses of constant amplitude, whereas when the voltage is increased the current pulses remain of the same amplitude but occur more frequently and tend to be of larger duration. At voltage levels where the microplasma is nearly stable the current remains at the level of the pulse amplitude for most of the time, switching to zero for periods which decrease in duration and regularity as the stable position of the characteristic is approached.

The diodes required to produce a predictable noise level must be free of these multiple breakdowns and three methods were used to eliminate diodes which exhibit them. The initial method used a Tektronix curve tracer on high sensitivity and with a backing-off voltage to reveal rapidly those diodes with large multiple breakdowns. Figures 1 and 2 were obtained by this method.

A more reliable, though more time-consuming, method was to examine the variation of slope resistance with diode current. Current levels at which unstabilities exist

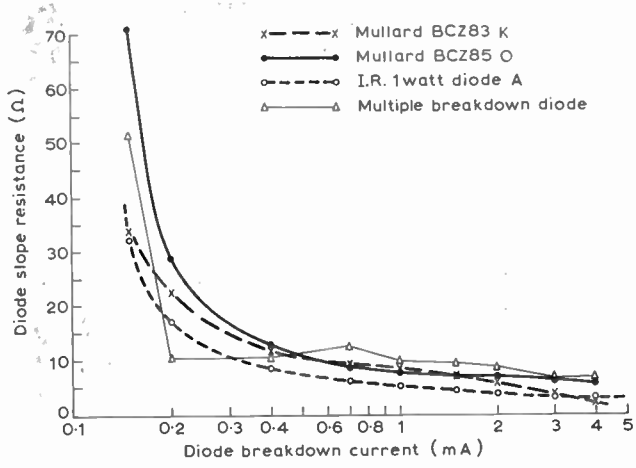


Fig. 4. Variation of diode slope resistance with breakdown current.

are revealed by discontinuities in this relationship. The circuit used is shown in Fig. 3, and Fig. 4 shows the resistance current characteristics of several diodes.

One of the diodes shown here was known to have several discontinuities, and gave the irregular characteristic marked Δ. The other diodes gave a smoothly varying resistance with diode current, except diode K which can be seen to droop slightly above 1·5 mA current.

A final check was made by inspecting the waveform of the noise current existing in the diode for a range of applied voltage. While the Gaussian distribution which is required cannot be readily identified by this method, a non-Gaussian distribution, characteristic of unstable microplasmas, frequently could be identified, leading to rejection of the diode. Figure 5 shows some typical distributions obtained.

The second of these methods was thought to be the most promising, but the percentage of commercial breakdown diodes free of instabilities is very small and the measurements are unduly lengthy.

This method is now being modified to operate on a swept basis, the graph being plotted on an oscilloscope, or X-Y plotter.

The diodes which have been found so far to be free of multiple breakdown effects are all in the low voltage region (<10 V). It is hoped that the swept measurement of slope resistance will reveal some higher voltage diodes from among the larger numbers that may be examined.

The availability of breakdown diodes free from multiple breakdown at lower voltages is possibly due to the presence of field emission effects over this range. Breakdown diodes which were manufactured to give true avalanche breakdown at these voltages were examined but none were found free from multiple breakdown effects.

3. Noise Properties of Selected Diodes

The diodes that were selected for low noise were further investigated in the following ways:

(i) variation of noise voltage with current,
(ii) variation of noise voltage with frequency,
(iii) variation of noise voltage with temperature,
(iv) amplitude distribution of the noise produced.

Since the breakdown voltages of the selected diodes were close to each other the variation of noise voltage with breakdown voltage will not give significant results until higher voltage low-noise diodes are found.

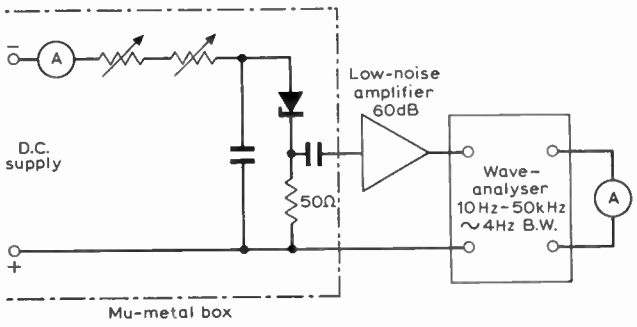
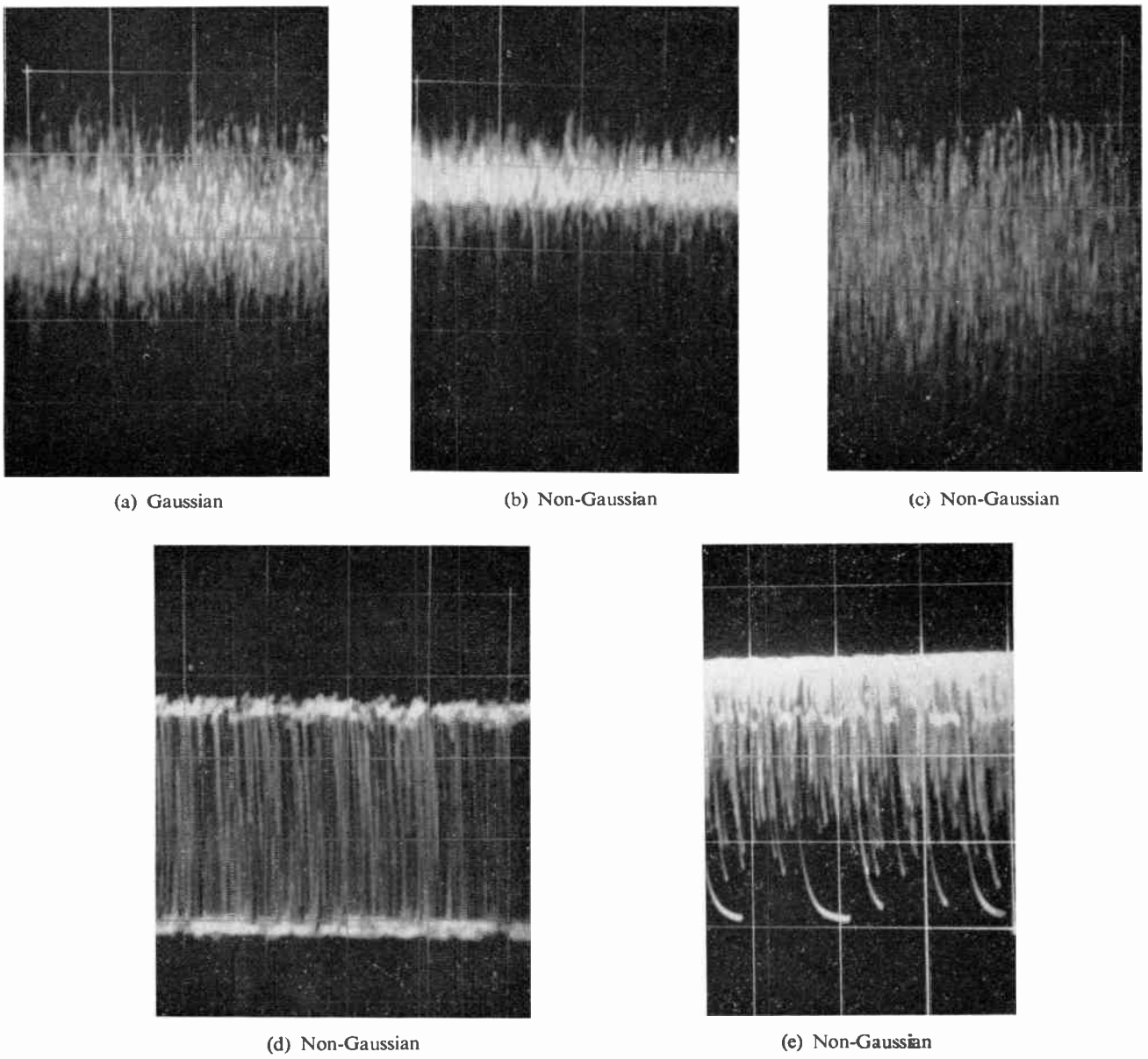


Fig. 6. Noise measurements.



(a) Gaussian (b) Non-Gaussian (c) Non-Gaussian

(d) Non-Gaussian (e) Non-Gaussian

Fig. 5. Amplitude distribution as displayed on oscilloscope.

The equipment used for these measurements is shown in Fig. 6. The biasing circuit and the diode were kept in a mu-metal box which also contained the batteries for the d.c. supply. A coaxial feed was then taken to the low-noise transistor preamplifier.

The wave analyser provided a constant bandwidth and gave a deflexion proportional to the mean rectified value of the input voltage, the scale being calibrated to show the r.m.s. value of a sine wave. A form factor of 1·05 dB was used to correct the r.m.s. reading since the noise was assumed to be Gaussian. This assumption was later justified.

The bandwidth of the wave analyser was determined using two methods, firstly by varying the source resistance connected to a low-noise valve preamplifier of known gain. The bandwidth of the measuring circuit was obtained from the rate of change of output voltage with source resistance. This result was checked later using a calibrated noise source of known bandwidth. The wave analyser indicator was small and for the measurement of narrow band noise had far too low a time-constant. A 5μA moving-coil meter, shunted to the same sensitivity, was found to have a time-constant of 14 seconds. This served to average out most of the variations in output indication and enabled a fairly accurate reading to be taken without too great a delay in time.

The variation of open-circuit spectral noise density with breakdown current is shown for three diodes in Fig. 7. The diodes are the same three which exhibited smooth resistance–breakdown current relationships.

All three diodes show a fairly smooth decrease of noise with current between 250 mA and 1·5 mA, while diode K,

which was seen to have a change of slope above 1·5 mA in the earlier graph, shows an increase in noise level starting at this point. The peak noise obtained from diode K at 3 mA current is lower than the average for microplasma noise. Diode O gives a relationship between the spectral noise density, and the breakdown current I_B of $4{\cdot}2\times10^{-10}\ V_B\ I_B^{(-0.4)}$

The relationship, theoretically obtained, [6] is $a\ V_B\ I^{(0.5)}$ where a is between 2·6 and $3{\cdot}4\times10^{-10}$

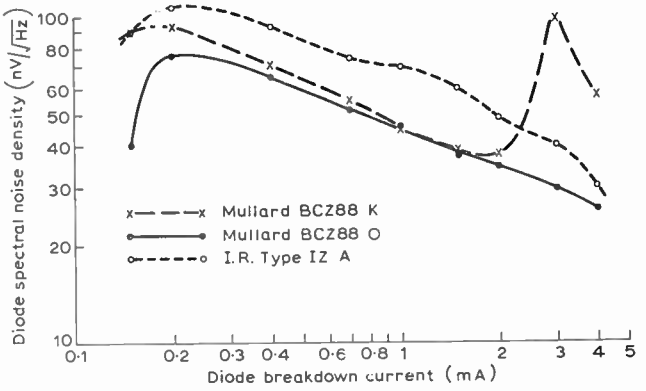


Fig. 7. Variation of spectral noise density with breakdown current.

The variation of this noise voltage with frequency was investigated between 10 Hz and 50 kHz and was found to be flat, for 1 mA current, within the limits of measurement.

The variation of spectral noise density with temperature was found to be less than measurable over the temperature range $-10°$ to $+90°$C.

The results of measurement of the amplitude distribution of the noise waveform are shown in Fig. 8 where it can be seen to be almost indistinguishable from a Gaussian distribution. These measurements were made with a probability density analyser.

4. Conclusions

A small number of diodes had been obtained in which the noise is close to that predicted theoretically. The fraction of commercially available diodes where this is the case is extremely small, at present none have been found with a breakdown above 8·5 V.

The test of slope resistance for varying breakdown currents seems the most effective way of selecting the diodes, and when done on a swept measuring basis should enable sufficient numbers of these diodes to be selected. This measurement could in addition be used to

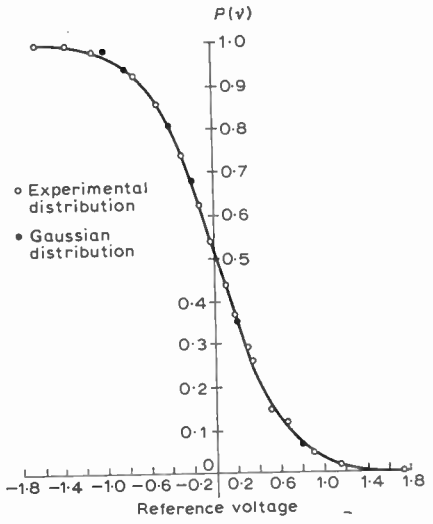


Fig. 8. Amplitude distribution function.

select diodes with a lower slope resistance over a specific current range for normal circuit use. Since the presence of multiple breakdowns may be regarded as a sign of unreliability the method of selection recommended here may prove more useful in providing a source of lower noise diodes of improved reliability than in enabling diodes to be selected as generators of predictable noise levels.

5. References

1. Haitz, R. H., 'Noise of a self-sustaining avalanche discharge in silicon: low-frequency noise studies', *J. Appl. Phys.*, **38**, No. 7, pp. 2935–46, June 1967.
2. Haitz, R. H. and Voltmer, F. W., 'Noise of a self-sustaining avalanche discharge in silicon—studies at microwave frequencies', *J. Appl. Phys.*, **39**, No. 7, pp. 3379–84, June 1968.
3. Rose, D. J., 'Microplasmas in silicon', *Physical Review*, **105**, No. 2, pp. 413–8, January 1957.
4. Shockley, W., 'Problems related to p-n junctions in silicon', *Solid State Electronics*, **2**, No. 1, pp. 35–67, 1961.
5. Haitz, R. H., 'Model for the electrical behaviour of a microplasma', *J. Appl. Phys.*, **35**, No. 5, pp. 1370–6, May 1964.
6. McIntyre, R. J., 'Multiplication noise in uniform avalanche diodes', *I.E.E.E. Trans. on Electron Devices*, **ED-13**, No. 1, pp. 164–8, January 1966.
7. Yim, Y. D. and Misra, R. P., 'Noise spectral density as a diagnostic tool for reliability of p-n junctions', *I.E.E.E. Trans. on Reliability*, **R-18**, No. 4, pp. 197–200, November 1969.

Manuscript received by the Institution on 18th January 1971. (Paper No. 1414/CC108.)

U.D.C. 537.52: 535.374

The Laser Triggered Spark Gap

By

S. H. KHAN,
D.Phil.†

Investigates the feasibility of the laser beam as a trigger for high-voltage gaps. The formative times are studied as a function of the working voltage for different values of pressure, gap spacing and laser power. A low jitter was obtained, of the order of ± 1 ns for a delay of up to 20 ns, making the laser suitable for triggering multiple arc gaps.

1. Introduction

There is considerable interest in switching high voltages in the range of 10–1000 kV, with the minimum of delay and jitter. Delays of a few nanoseconds with assured, reasonable freedom from jitter are of considerable importance in work requiring synchronization, such as the pulsed line acceleration of electron rings.[1]

The general requirements of a spark-gap switch are two-fold: (a) it must be able to take the full voltage applied to it; and (b) it must be able to conduct on demand. For this it is necessary that enough primary electrons be present to initiate the discharge and the gap voltage must be sufficiently high. The primary electrons are generally provided by ultra-violet or other irradiation of the gap, while a third or fourth trigger electrode is used to supply a voltage pulse that would disturb an existing voltage distribution and provide sufficient gap voltage. Such methods work quite well up to 70–80 kV, beyond which the problems posed by proper insulation and voltage division are quite enormous. The laser offers an alternative triggering device which is basically safer and simpler to use as it is isolated electrically from the spark gap. Moreover, a two-electrode configuration has obvious advantages from the engineering design point of view.

2. General Theory

For several years now, the laser has been a source of copious ionization from metallic targets as well as from gases when a sufficiently high-powered laser is focused in them.[2–10] The emission is intense and has a definite temporal relation with the laser spike, delay times of the order of a few nanoseconds being obtained when the laser is operating in the Q-switched mode. It was the possibility of such ultra-fast synchronization between the laser and the emission pulses that first made the laser look promising as a trigger for spark gaps.[11–14] Basically the laser trigger operates by injecting a plasma blob into the gap such that the concentration of space charge serves to distort the applied electric field, E, in the gap. When focused in the gap, a multiphoton process could generate as many electrons as there are atoms in the gas, with electron temperatures as high as 60 eV and velocities of the ionization front approaching 10^7 cm/s.[9] Similarly when the laser is focused on an electrode, a current density of several amperes per square centimetre can be emitted with electron velocities in excess of 10^6 cm/s.[1–4] The latter emission is believed to be thermionic in origin. A local hot spot is produced at the laser focus whose temperature has been estimated at several thousand degrees Celsius.[6]

There is a definite laser power threshold when the laser is focused in a gas below which gas breakdown does not occur. Generally less power is required to initiate breakdown when the laser is focused on an electrode.[11] Economy of power is desirable for multiple-gap switching, hence focusing on an electrode surface was investigated.

3. Triggering Mechanism

When focused on an electrode, the laser is believed to supply the two essential requisites for fast gap breakdown: (1) the primary electrons and (2) an over-volting mechanism. Copious plasma emission takes place, whose violent nature can be judged by comparing with conventional ultra-violet irradiation of spark gaps, where the strongest possible sources (auxiliary spark gaps) are capable of generating photo-electric currents of the order of 10^{-12} A/cm^2 only. Lasers generate several A/cm^2. Furthermore, a current of the order of 10^{-12} A/cm^2 is regarded as adequate in removing statistical time delay and so the delay in laser triggered spark gaps may be regarded as being entirely formative in nature.

It is suggested that the intense laser-produced plasma at a target electrode would serve to distort the applied electric field resulting in either an effective over-volting of the gap or the formation of a streamer in the gap. The following may happen:

(a) The high-speed high-conductivity plasma generated by the laser would reduce the effective gap spacing due to the extension of the virtual cathode in the gap.[12]

(b) There could be a reduction in the effective static gap breakdown voltage, V_s, due to the intensity of the laser light. This would be similar in origin but far greater in magnitude than the reduction in V_s mentioned by, say, Brinkman and Meek.[15,16]

(c) The critical carrier number, N_c is reached at the point in the avalanche started by the laser at which a streamer is generated with velocities of 10^8–10^9 cm/s, compared with avalanche speeds of about 10^7 cm/s. The formative time may then be regarded as being made up entirely of the time taken to reach this transition point and would be given by the equation[17,18]

$$t_f = \frac{\log N_c/N_o}{pvA \exp(BE/p)} \qquad (1)$$

where t_f = formative time
N_o = initial electron number
N_c = critical electron number (10^9)

† Formerly Department of Engineering Science, University of Oxford; now with Pakistan Institute of Nuclear Science and Technology, Islamabad.

v = avalanche drift velocity
E = applied electric field
p = gas pressure
A, B are constants.

The relationship between t_f and E/p would be expected to be inverse exponential while t_f would vary inversely as the gas pressure p, for a constant value of E/p. Because of the effect of the laser-induced plasma bridging the gap, shorter gap spacings may be expected to give smaller formative time delays. The effect of laser power can be deduced from a simplified evaluation of the temperature of the spot where the laser is focused. The one-dimensional theory of the growth of temperature at the laser focus can be shown[13,21,22] to predict the surface temperature as

$$T = \frac{2SF}{(\pi\rho ck)^{\frac{1}{2}}}[t^{\frac{1}{2}} - (t-\tau)^{\frac{1}{2}}] \qquad \ldots.(2)$$

where S = absorption coefficient of the solid
F = laser power density in W/cm^2
ρ = density of the electrode material
c = specific heat
k = thermal conductivity
t = time after the start of the laser pulse
τ = duration of the laser pulse

The term temperature evidently needs re-defining, since according to equation (2) even a refractory material like tungsten would have vaporized within 5 ns, when a laser pulse of power 7 MW and duration 20 ns is focused to a 1 mm diameter spot. According to the energy balance drawn up by Ehler,[3] the laser energy after vaporization is utilized in heating and maintaining the plasma temperature. He estimated that only 3% of the laser energy is used to heat the target while 86% is absorbed in the plasma produced, the remainder being lost by reflexion and scattering.

The emission current would then be given by the well-known Richardson–Dushman equation:

$$J = aT^2 \exp(-b/T) \text{ A/cm}^2 \qquad (3)$$

where a, b, are constants and T is the surface temperature.

J is more strongly dependent on the exponential term than on the term aT^2. Thus a 1% change in T would produce a 2% change in aT^2 but a 20% change in the value of $\exp(-b/T)$. Since the temperature is directly dependent on the laser power density F (equation (2)) and hence the laser power P for a given focusing condition, the formative time t_f may be expected to vary in an inverse exponential manner with the laser power. These predictions will now be compared with the experimental results.

4. Experimental Arrangement

Figure 1 shows the various assemblies and components of the experimental method. The Q-switched laser was capable of powers of up to 15 MW in a single pulse of duration of about 15 ns. Q-switching was achieved by using an approximate 10^{-6} M (molar) solution of cryptocyanine in methanol. The same solution could be used for several hundred shots before any deterioration in the laser pulse was observed. The laser power could be measured either with a calibrated carbon-cone calorimeter or with an ADP crystal device[19] which allowed a direct read-out of the laser power, its calibration being 15 mV/MW. The laser power was measured at the beginning and at the end of every set of observations; a fast photodiode/oscilloscope system provided a direct comparison of the power from shot to shot.

The spark gap had a working range of 5–55 kV and 1–5 atmospheres. By means of a 10 cm lens, the laser radiation could be focused at grazing incidence when 5 cm diameter hemispheres were used as electrodes; but when plane parallel electrodes were used the laser radiation was fired axially through a 2·5 mm hole in the earthed electrode. The gap could be varied between 1–15 mm and its spacing measured accurately to 0·1 mm. In the preliminary experiment, a 1000 pF, 100 kV capacitor was charged through a 200 MΩ resistor. Later a 100 kV coaxial cable was used instead, and a much sharper pulse front (with a rise-time of about 5 ns) was obtained.

Breakdown was monitored by means of a resistive voltage divider which, together with the Tektronix 519

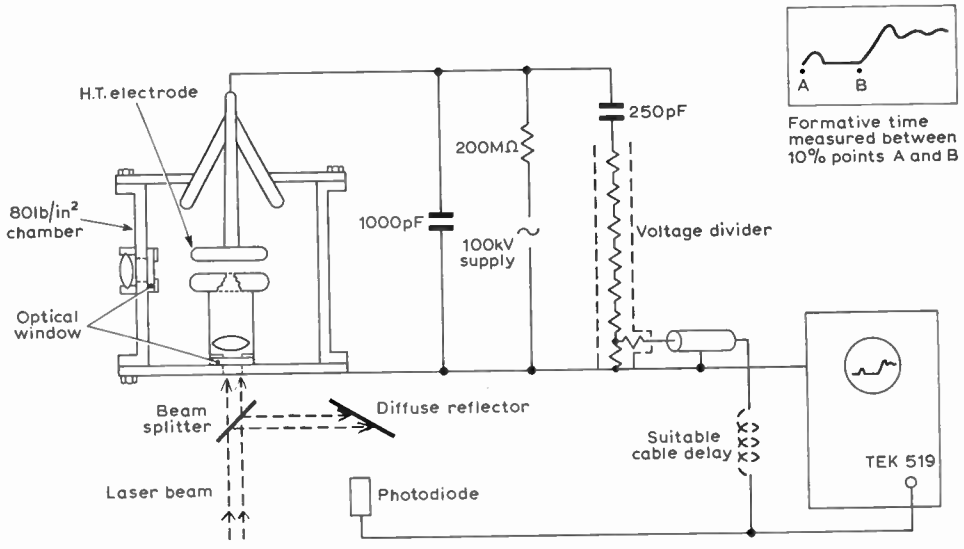
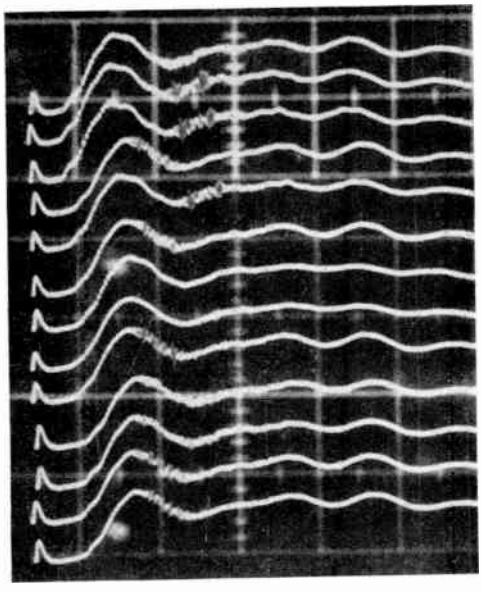


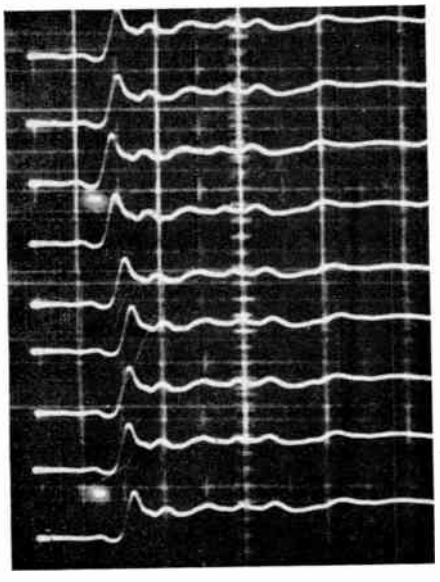
Fig. 1.
Experimental arrangement.



(a) Formative time vs. working voltage. Air-gap: 9 mm; V_s = 29·0 kV; 1 atm.

No.	kV	No.	kV
1–3	27·5	7–9	25·0
4–6	26·0	10–13	24·0

5 MW; 50 ns/cm; cable delay 23·6 ns.



(b) Typical jitter study. Air-gap: 4·8 mm; 3 atm; V_s = 39·0 kV.

No.	kV
14–17	29·0
18–22	27·5

7·8 MW; 20 ns/cm; cable delay 13·5 ns.

Fig. 2. Typical oscillograms.

oscilloscope, gave a response time of better than 3 ns. The attenuation factor was about 300 : 1. Both the photo-diode signal monitoring the laser radiation and the breakdown signal from the divider could be shown on the same sweep, through T-connectors, or the breakdown signal alone could be shown, the sweep being triggered externally by the laser signal (Fig. 2). The delay was measured between the 10% points as at A and B; the cable plus optical delays were subtracted to give the correct formative time delay t_f. The gases were filtered and dried before being introduced into the gap.

5. Experimental Results with Air in the Gap

The major series of investigations was carried out with clean dry air in the gap. Two possible focusing angles—grazing or axial—were studied, the results being shown in Fig. 3. The laser power was about 5 MW. Delays of less than 1 ns were obtained when the laser was focused at axial incidence on the cathode at a working voltage of 90% of V_s; while at 80% of V_s the delay was about 13 ns for the axial mode and about 40 ns for grazing incidence. Quite fast switching was also obtained with grazing incidence on the earthed anode but axial focusing on the cathode seemed to give the fastest switching. It may be explained on grounds of the greater laser power density which leads to faster surface heating of the electrode. Axial focusing on the h.t. cathode was thus used throughout the remainder of the experiment, and the effect of gap spacing d, the gas pressure p, and the laser power P, on the formative time t_f was studied as shown in Figs. 4–7.[22]

5.1 *Effect of the Gap Spacing d*

A set of results was obtained for air, where the pressure was constant (1 atmosphere) while the formative times were plotted against V/V_s for three different values

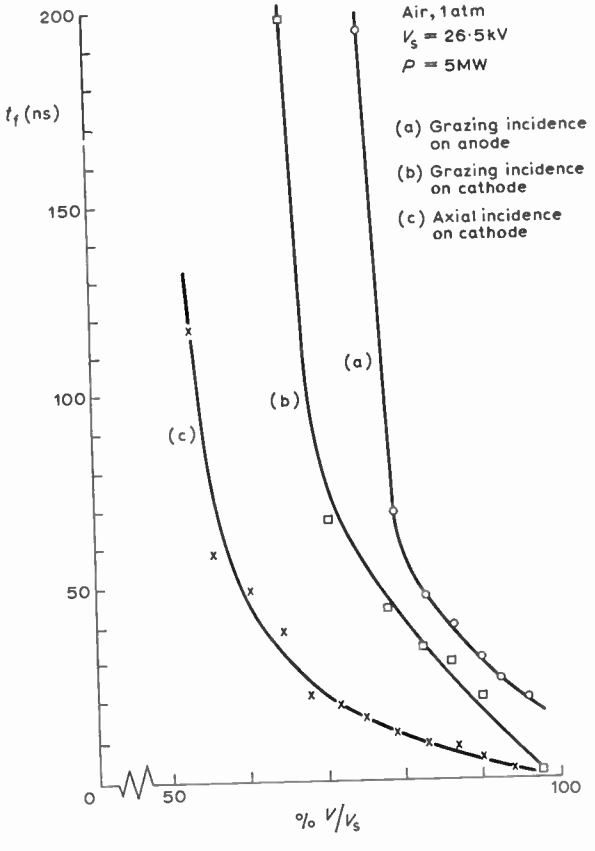


Fig. 3. Effect of laser incidence angle on the formative time, t_f.

of the gap spacing (Fig. 4). Each data point is the mean of three observations, typical oscillograms for the 9 mm gap being shown in Fig. 2. The method of introducing both the laser signal (from the photo-diode) and the breakdown signal from the voltage divider on to the same sweep was advantageous as it allowed a study of possible variations in the laser pulse itself. The laser was found to show a worst possible variation of $\pm 10\%$ in amplitude or duration.

The curves in Fig. 4 show that V/V_s varies with d. This may be due to the increased distance of travel for the avalanche, and points to the possible applicability of the 'virtual cathode' theory.

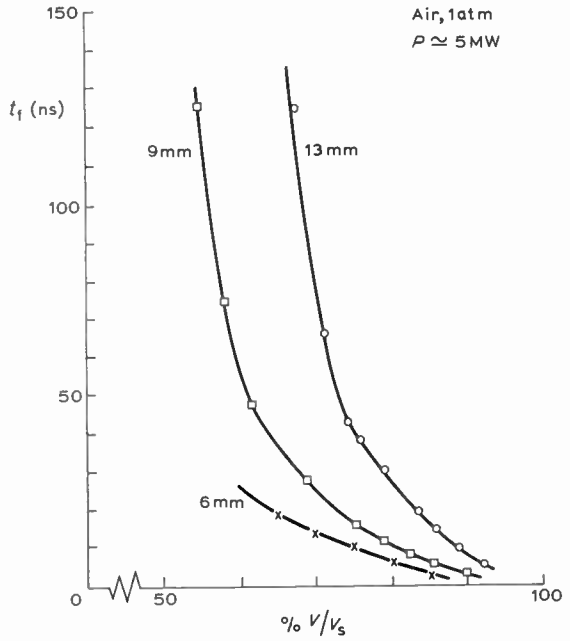


Fig. 4. Effect of gap spacing, d, on the formative time, t_f.

5.2 *Effect of Gas Pressure p*

The effect of gas pressure on the formative time can be seen from Fig. 5. The laser power was constant at 5·5 MW, while a 6·4 mm gap was used. The dependence on pressure is a linear one as shown by the inset in Fig. 5, where the working voltage at each pressure point was such that E/p was constant at 31 V/cm-torr.

5.3 *Effect of Small Gap Spacings*

Figure 6 shows the results obtained for gap spacings of 1·8 mm and 3·6 mm. Each data point is again the average of three observations and the electric field in the gap was uniform since plane parallel electrodes were used. At a pressure of 4 atmospheres the curves for gaps of 1·8 mm and 3·6 mm coincide. This seems to contradict the results of Fig. 4 which shows an increase in t_f with d, at constant pressure. However, this anomalous result may be explained from photographs of the plasma 'blob' generated at an electrode surface which is seen to extend about 3 mm from the electrode surface. It would seem therefore that because of the plasma cloud the gap spacing has an effect on the formative times only for values of d greater than 4 mm or so.

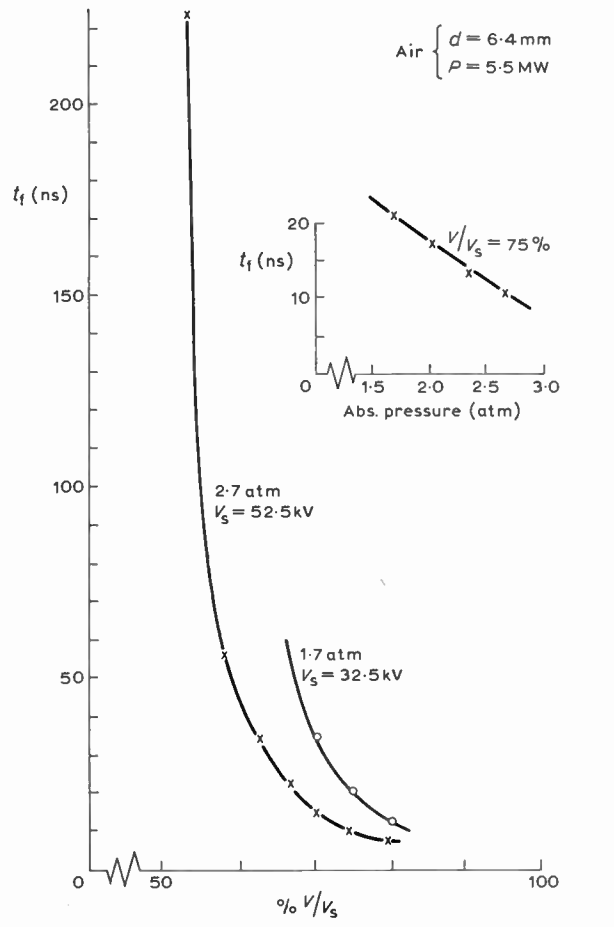


Fig. 5. Effect of gas pressure, p, on the formative time, t_f.

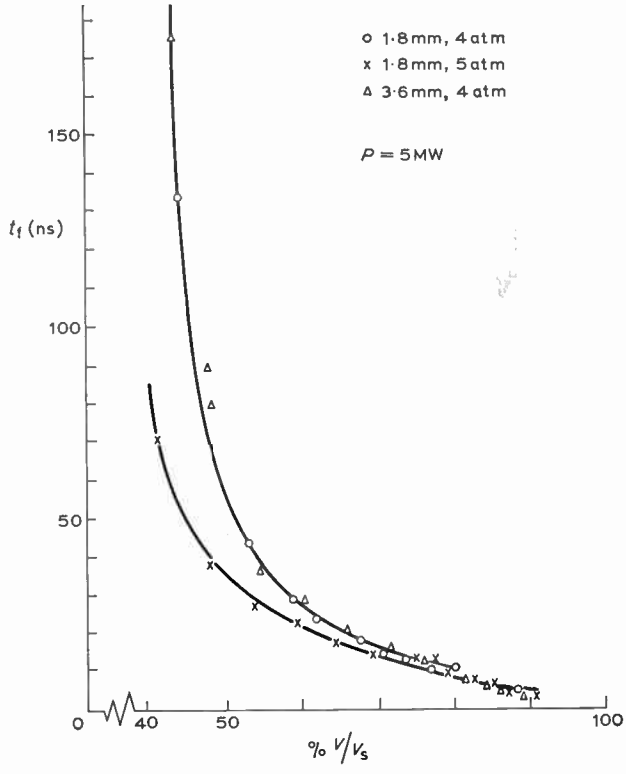


Fig. 6. Formative times for short gap spacings.

5.4 *Effect of Laser Power*

The effect of changing the laser power is shown in Fig. 7. The laser radiation was incident axially on the cathode which was made of tungsten instead of the brass used in the previous experiments. The inset shows the vatiation of t_f with laser power for a fixed value of the working voltage ($V/V_s = 75\%$), the laser power being controlled by interposing various dielectric filters in the beam path.

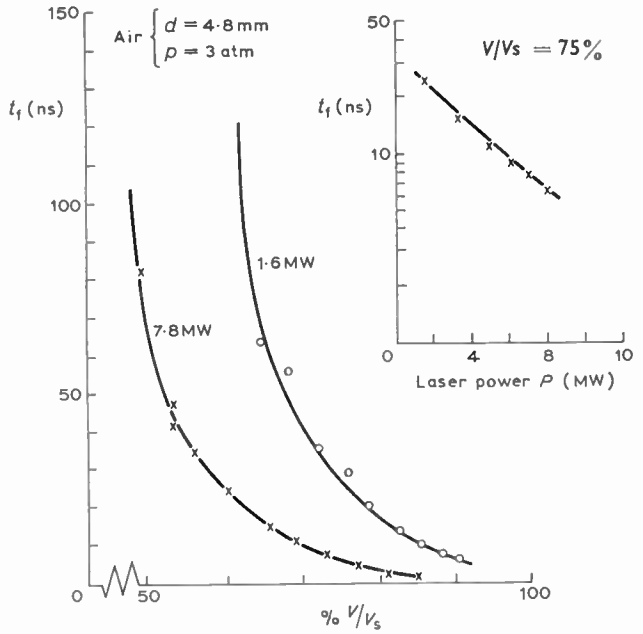


Fig. 7. Effect of laser power, P, on the formative time, t_f.

5.5 *Study of Jitter*

For reliable operation, the laser-triggered spark gap must have small jitter, i.e. the random variations in the formative time delay should be small. The 'jitter' has been computed as the standard deviation about the mean formative time from oscillograms such as those shown in Fig. 2. The results were quite encouraging. Figure 8(a) indicates that the jitter is only about ± 1 ns at 60% of V_s, increasing rapidly below that value. Another way of evaluating the jitter is as a function of the formative time as shown in Fig. 8(b), in which it appeared to be no more than ± 1 ns for formative times up to 20 ns.

The biggest cause of jitter was believed to be in the laser itself. It is necessary here to distinguish between two kinds of jitter There is first of all the random appearance of the laser pulse with respect to the flash-tube that excites the ruby crystal. This depends on the method employed for Q-switching. It is about 50 μs when a rotating mirror is used, about 10 μs for the passive cryptocyanine cell employed in the experiments above, a few tens of nanoseconds for the Kerr cell and about 7 ns for the Pockell's cell (Korad Model k-QS2). This type of jitter was not considered in these experiments as the formative delays were measured with respect to the arrival of the laser pulse at the electrode surface. The other cause of jitter, and the one studied here, was due to variations in the laser power from shot to shot. The worst possible variation in the laser amplitude was $\pm 10\%$. This was believed to be the biggest cause of jitter, since variations in the working voltage (once it is set by the d.c. supply) or the gap pressure (kept constant by the pressurized gas cylinder and the inlet valve) may be neglected. Adequate cooling and a constant rate of laser fire would reduce this.

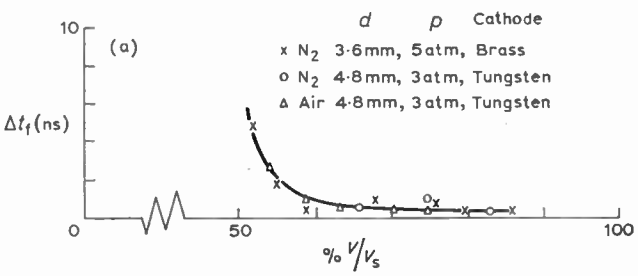
There seems to be a 'fast' region of laser triggered breakdown lying between 60–90% of the static breakdown voltage V_s and showing small jitter.

6. Conclusions

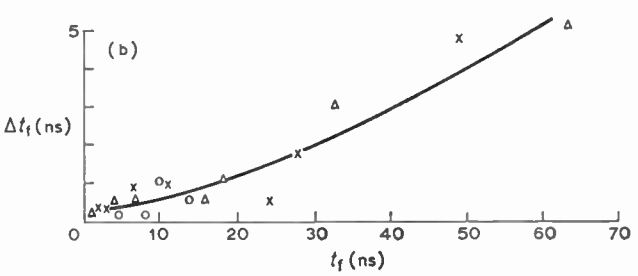
The laser appears to be a very useful trigger for spark gaps. The laser power required is not very high, 5–7 MW being ample when the beam is focused axially on the cathode. For a given value of pd (pressure × gap spacing) faster switching is obtained for smaller gaps and higher pressures, though no advantage seemed to accrue by reducing the gap below 3 mm. The jitter is small, being $\pm 0{\cdot}1$ ns for a formative delay of 1 ns and ± 1 ns for a delay of up to 20 ns. The results can be explained qualitatively by considering the distortion in the applied electric field by the laser-induced plasma in the gap. The low jitter makes the laser extremely suitable for triggering multiple-arc gaps without any transit time isolation between them. Since the working voltage in the gap is generally much lower than the static breakdown voltage, the laser triggered gap also appears to be free from premature breakdown which had posed problems in the fast, low jitter field distortion gaps.[20]

7. References

1. Hartwig, E. C., 'Pulsed Line Acceleration of Electron Rings', Symposium on Electron Ring Accelerators, Lawrence Radiation Laboratory, Berkeley, California, 1968.



(a) Jitter (Δt_f) versus $\%V/V_s$.



(b) Δt_f versus t_f.

Fig. 8. Jitter study.

2. Bernal, G. E., Ready, J. F. and Levine, L. P., 'Ion emission from laser irradiated tungsten', *I.E.E.E.J. Quantum Electronics*, **QE-2,** pp. 480–2, September 1966.
3. Ehler, A. W., 'Plasma formed by a laser pulse on a tungsten target', *J. Appl. Phys.*, **37,** p. 4962, 1966.
4. Khan, S. H., Richards, F. A. and Walsh, D., 'Time resolution of laser induced electron and ion emission', *I.E.E.E.J. Quantum Electronics*, **QE-1,** pp. 359–60, November 1965.
5. Isenor, N. R., 'Metal ion emission velocity dependence on laser giant pulse height', *Appl. Phys. Letters*, **4,** p. 152, 1964.
6. Richards, F. A., 'Study of the Interaction of a Laser Beam with Clean Metal Surfaces', D.Phil. Thesis, Oxford University, 1967, Dept. of Engineering Science.
7. Wright, J. K., 'Theory of the electrical breakdown of gases by intense pulses of light', *Proc. Phys. Soc.*, **84,** p. 41, 1964.
8. Meyerand, R. G., Jr. and Haught, A. F., 'Optical-energy absorption and high density plasma production', *Phys. Rev. Letters*, **13,** p. 7, 1964.
9. Mandelstam, S. L., Pashinin, P. P., Prokhorov, A. M., Raiser, Yu. P. and Sukhodrev, N. K., 'Investigation of the spark discharge produced in air by focussing laser radiation', *Soviet Phys.*, (*J.E.T.P.*), **22,** p. 91, 1966.
10. Gold, A. and Bebb, H. B., 'Multiphoton ionisation of rare gases and hydrogen atoms', *Phys. Rev. Letters*, **14,** p. 60, 1965.
11. Pendleton, W. K. and Guenther, A. H., 'Investigation of a laser triggered spark gap', *Rev. Sci. Instrum.*, **36,** p. 1546, 1965.
12. Deutsch, F., 'Development of a laser triggered spark gap', CERN Internal Report MPS/Int. MA, 67–2.
13. Khan, S. H.; 'The Laser Triggered Spark Gap', 1st Conf. on Laser Applications, Paris, July 1967. Also: 'The Laser Triggered Spark Gap', D.Phil. Thesis, Oxford University, 1968.
14. Baumhacher, H., Maischberger, K., Hofmister, F. and Schmitter, K. H., 'Investigations on Triggering of Two Electrode Spark Gaps by the Focused Beam of a High Power Ruby Laser', 5th Symposium on Fusion Technology, Oxford, July 1968.
15. Brinkman, C., *Z. Phys.*, **111,** p. 737, 1939.
16. Meek, J. M., 'A theory of spark discharge', *Phys. Rev.*, **57,** p. 722, 1940.
17. Fletcher, R. C., 'Impulse breakdown in air', *Phys. Rev.*, **76,** p. 1501, 1949.
18. Guenther, A. H. and Bettis, J. R., 'Laser-triggered Megavolt Switching', I.E.E.E. Conference on Laser Engineering and Application, June 1967.
19. Ward, J. F., 'Absolute measurement of an optical rectification constant in ADP', *Phys. Rev.*, **143,** p. 569, 1966.
20. Barnes, P. M., Gruber, J. E. and James, T. E., 'High-current Switching using Low-Inductance Field Distortion Spark Gaps Closely Connected in Parallel', U.K.A.E.A. Culham Report CLM–R7.
21. Carslaw, H. S. and Jaeger, J. C., 'Conduction of Heat in Solids', p. 56, 2nd ed. (Oxford University Press, 1947).
22. Khan, S. H. and Walsh, D., 'Effect of various gases on formative times of the laser-triggered spark gap', *Electronics Letters*, **6,** pp. 551–3, 20th August 1970.

Manuscript first received by the Institution on 17th November 1969, in revised form on 8th March 1971 and in final form on 22nd April 1971. (Paper No. 1415/CC 109.)