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 47 No. 11 November 1977

The Radio and Electronic Engineer

The Journal of the Institution of Electronic and Radio Engineers

Engineering under the Microscope

ALREADY this year two 'high power' reports have appeared dealing with the problems of the British engineering profession, looking at the broad issues of management in industry and the education of engineers. These were respectively 'Industry, education and management', a discussion paper published in July by the Department of Industry*, and 'Education, Engineers and Management', a study prepared under the aegis of the British Association at the University of Aston in Birmingham.† Apart from these investigations, a full-scale Government Committee of Inquiry, chaired by Sir Montague Finniston, is now inviting evidence under fairly wide terms of reference‡ and added point will certainly be given to an important personal view of the engineer himself when the 1977 CEI Survey of Salaries is published towards the end of this year.

Between them these various analyses have, or no doubt will, cast a critical eye on what have been vexed problems for many years—why is the performance of many parts of industry so disappointing; should we blame management; would management be better if there were more engineers in the higher levels; are engineers educated and trained properly; why is the status (and the pay) of engineers so low; and so on. All these factors must, in our view, be inter-related and as far as identifying the facts goes, there is fair agreement between the two reports already published. (Curiously, the D of I put forward 'questions for discussion' while the B.A. report recommends 'actions'.)

Although after a cursory inspection, cynics may regard the contents of first two reports as being 'the mixture as before', this is hardly fair: there is, in both, evidence of careful examination of statistics and expert evidence, and the analyses highlight a number of proposals which will have the approval of many IERE members. The Institutions are urged to concentrate their efforts on monitoring education and training schemes rather than the individual, who nevertheless must have opportunities for retraining in the interests of making fullest use of his talents, particularly in managerial skills. The avenues for the technician engineer to reach chartered engineer status must too be kept open.

Collaboration between the schools and both industry and the universities has long been paid lip service, and this is an area where a genuine belief in its importance is vital. The BA report's call for more emphasis on mathematics will upset no one: but the suggestion that the second A-level for university entrance to an engineering degree course need not necessarily be physics is controversial.

While engineering generally is experiencing problems, production engineering seems to be especially neglected, and the passage in the D of I report pin-pointing this area suggests that the quality of engineers entering production is low. But there is no single, simple way to remedy any of these difficulties; action by Government, by industry itself and the co-operation of educationalists are needed. The concepts of Total Technology and of the Teaching Company, as well as financial incentives, will all be steps in the right direction.

Finally, we are most of us aware that there are certain social factors behind the present situation: unlike in Europe and elsewhere, the two cultures in Britain, based respectively on arts and science, have historically lead to lower status for the third, the art of making things. If the Finniston Committee can draw upon the facts and deductions of these reports and other sources, and come up with proposals of how the engineering profession can overcome this built-in obstacle, it will not only be the engineers who should feel grateful.

F.W.S.

* Single copies obtainable from Department of Industry, Room 601b, 1 Victoria Street, London SW1H 0E5.
† See page 536 of this issue.
‡ See page 345, July 1977 issue.

Contributors to this issue

Gordon Burrows (Member 1958, Graduate 1951) served a radio engineering apprenticeship with Pye Limited, Cambridge, from 1940 to 1944 and after service in the Royal Navy, returned to Cambridge in 1947 to work initially as a Test Engineer with Pye and later as a Research Assistant in the University Engineering Laboratory; during this period he was for a year a part-time lecturer at Cambridge Technical College. In 1950 he obtained the HNC in radio engineering and from 1950 to 1952 he worked at a hospital in Surrey on medical electronics. He then went back to teaching, first at the Northern Polytechnic, London, and later at Hatfield Technical College where he was in 1957 appointed a Senior Lecturer; while at Hatfield he started a research project on tropospheric reflections and in 1963 obtained the Diploma of Imperial College. In 1967 Gordon Burrows was appointed to a lectureship in the Department of Electrical Engineering at Imperial College and he subsequently became Senior Tutor. In 1968 he obtained his Ph.D. by means of a thesis on tropospheric propagation. He has presented some nine papers at AGARD meetings on aspects of his research and in 1968 he published a book on 'VHF Radio Wave Propagation in the Troposphere'. In January of this year he went to Bristol to take up the Headship of the Department of Marine Electronics at Brunel Technical College. Dr Burrows has served on the Institution's Papers Committee since 1969.

Norman Kenyon graduated in physics and electrical sciences from Cambridge University in 1963 and he stayed on to do research with Professor A. H. W. Beck on millimetre-wave tubes, gaining his Ph.D. in 1967. For the next four years he was at Bell Laboratories working on millimetre-wave solid-state oscillators, and he returned to England in 1972 to continue this work at the Post Office Research Centre. Dr Kenyon became closely involved with the millimetric waveguide project, and now heads a section responsible for terminal and repeater equipment.

Keith Baker studied both Electrical Engineering and Mathematical Physics at undergraduate level and he obtained the D.Phil. degree for research in nuclear physics at the University of Sussex. In 1970 he entered the computer industry as a Systems Analyst with the Software Systems Division of the Plessey Company. Subsequently he returned to the University of Sussex to help establish a computer science degree. Later on, whilst on leave from the University, he was appointed Project Manager and Supervisor of Software Engineering with the Burroughs Corporation based at Seneffe, Belgium, where he was responsible for various aspects of a microprocessor-based business data processing system. Since returning to Sussex he has been concerned with the application of multimicroprocessor systems to data processing problems.

Brian Penney was formerly a physicist, taking a Ph.D. in High Energy Nuclear Physics at Imperial College, and doing post-doctoral research at Imperial College and the Rutherford Laboratory. He is now a computer scientist with research interests in computer communications, multiprocessor computer architecture, distributed computer control and information retrieval. His current appointment is that of Lecturer in Computer Science at Imperial College London with teaching responsibilities mainly in operating systems and microprocessors. Dr Penney is currently chairman of the Elliott Computer Users Association (this includes the GEC range of computers), and he has acted as consultant to various organizations in the UK.

Professor Dennis Towill (Fellow 1970) was appointed to the Chair of Engineering Production in the Department of Mechanical Engineering and Engineering Production, in the University of Wales Institute of Science and Technology, at the beginning of 1970. He is a graduate of the Universities of Bristol and Birmingham and before joining UWIST in 1966 he was a senior lecturer in automatic control at the Royal Military College of Science. He was the founder of the Dynamic Analysis Group at UWIST. He has held a number of industrial consultancies and was for some years with the then Bristol Aircraft Company as a dynamic analyst. Professor Towill is the author of numerous papers and books, and five of his papers have gained Institution Premiums. He is a member and past chairman of the IERE Automation and Control Systems Group Committee.

UDC 621.372.823.029.65: 621.395.632.64

Terminal and repeater equipment for the first operational waveguide system

N. D. KENYON, M.A., Ph.D., C.Eng., M.I.E.E.*

SUMMARY

The paper presents a fairly concise descriptive account of the millimetric-band waveguide equipment envisaged for operational use in trunk transmission by the British Post Office. It explains how the repeater spacing is chosen and how the circuit capacity could be built up gradually over many years; it also shows how the service requirements of the system will be met. The functions of the various equipment modules are described, together with an outline of the way it is expected that they will be realized.

* *Post Office Telecommunications Headquarters, Research Department, Martlesham Heath, Ipswich IP5 7RE.*

1 Introduction

In order to confirm that circular waveguide can be developed into a practical and economic means of high-capacity trunk transmission, and to investigate some of the options in the design of such a system, the Research Department of the British Post Office has installed a 14 km field trial system near its laboratory in Suffolk. This experiment has been quite successful, and Post Office Telecommunications has decided to install its first long-distance operational waveguide system; this will be between Reading and Bristol and is to be ready for service in 1983. Consideration is now being given to the technical requirements of such a system: details of installation, commissioning and maintenance are being worked out, and specifications for all the component parts are being drafted.

The operational system will differ in a number of ways from that of the field trial, both to remedy the known shortcomings of the latter and to make the system more reliable and easier to maintain. The purpose of this paper is to give a brief but connected description of the terminal and repeater equipment as it is currently envisaged.

2 System Capacity

The waveguide system will be so designed at the outset that in years to come it may gradually be expanded in capacity as the need arises, without introducing any extra repeater stations or at any time interrupting the traffic already carried. For this purpose the bandwidth of the 'ultimate', fully-loaded system is taken as 30–110 GHz; by this is meant that, although the guide will certainly support waves at 25 GHz or 115 GHz, we do not expect it to become practical even in the more distant future to use those frequencies, because the repeater stations we are installing at the beginning will be too far apart.

Even on the most optimistic predictions for traffic growth, the ultimate bandwidth will not be needed for many decades: why then should we insist on making as much as 80 GHz available? If we contented ourselves with, say, 40–80 GHz the repeaters could be 30% further apart. The answer here is that the cost saving would be relatively small, since repeater stations are only a small component of the initial system cost, and the preferred solution is to keep open the option of very large ultimate capacity.

Just how much is this ultimate capacity? Using quadrature phase-shift keying (q.p.s.k.) at an information/bandwidth ratio of about 0·9 bit s^{-1} Hz^{-1}, the figure is 36 gigabits/second in each direction. (Should it ever become necessary, more sophisticated techniques could probably increase capacity still further.) Taking a digital block as 139 264 kbit/second (enough for 1920 voice channels, or 2 colour-television channels, or a few dozen Viewphones), this becomes 256 both-way digital blocks. The 80 GHz total bandwidth will be divided into eight bands of 10 GHz each (See Fig. 1); it is possible to transmit any one band either in one direction or the other, so any

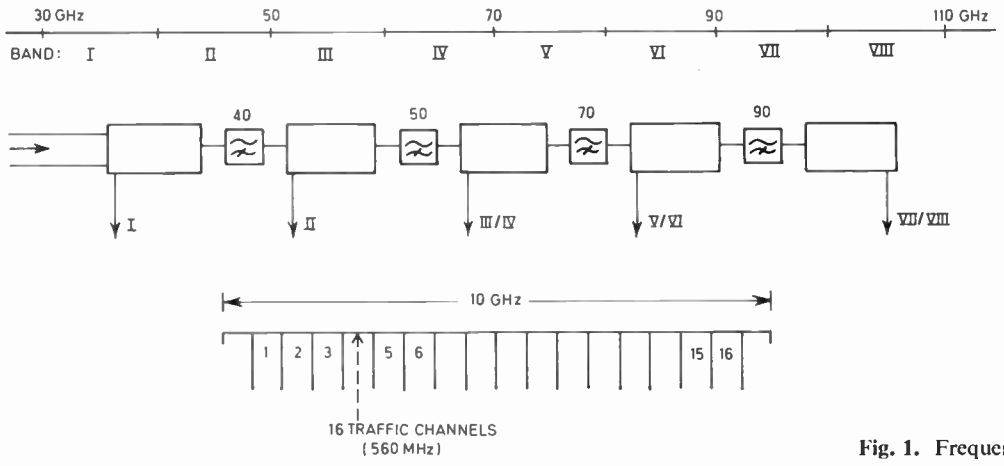


Fig. 1. Frequency plan.

pair of bands may be chosen for the two-way transmission of anything up to 64 digital blocks.

The growth of capacity on the system is envisaged to involve the following stages:

Stage 1. Initial provision of a few digital blocks only, in Bands I (30–40 GHz) and II (40–50 GHz), using binary phase-shift keying (b.p.s.k.).

Stage 2. Gradual filling up of Bands I ('go') and II ('return'), still using 2-phase modulation, giving a maximum of sixteen carriers each way, carrying two digital blocks each. (Fig. 1.)

Stage 3. Still in Bands I and II, upgrading the carriers in turn to four digital blocks each. Maximum capacity is now 64 bothway blocks.

Stage 4. Gradual filling of Bands III (50–60 GHz, 'return') and IV (60–70 GHz, 'go').

Stage 5. Use of Bands V to VIII.

3 Repeater Spacing

As already noted, the repeaters must not be so far apart that signals at 30 GHz or 110 GHz disappear irretrievably into the noise of the system. (This rather loose statement will be amplified later, in Section 7.) According to present calculations the maximum spacing is 24 km; but the maximum planned repeater spacing will be 23 km, leaving one kilometre in hand to cope with unforeseen circumstances, such as an enforced detour around some obstacle during installation of the system, or a subsidence of some part of the route.

Other determinants of repeater spacing are the overall route length (e.g. a system 125 km long will have an average repeater spacing of 21 km) and the need to find suitable sites for building the repeater stations.

Taking all factors together, the expected repeater spacings are in the range 18–23 km, and we would almost never have to stretch repeater 'gain' to its limits. This point is developed further in Section 7.

4 Utilization of the Waveguide

In considering the potential waveguide capacity, it was tacitly assumed that the traffic would be digital. It would not be impossible to transmit analogue signals over a waveguide, but there are good reasons why no further consideration has been given to this aspect. Firstly, the largest assembly of analogue telephone signals in common use is the 900-channel hypergroup, occupying less than 4 MHz of bandwidth; even 12 of these together as for the 60 MHz cable system take only a small part at a time out of the bandwidth of the medium, and the congestion of equipment at repeaters would prevent the use of more than 10–15% of the ultimate capacity.

More importantly, the distortions occurring in the system, arising from the use of millimetre-wave components, cannot economically be reduced to the point where the performance of a long analogue system with many repeaters could reach an acceptable standard. Indeed the raising of transmission performance is itself a strong motivator in the trend towards the use of digital encoding for transmission systems in general, and with the future prospect of telephone exchanges also 'going digital' the demand for a new analogue transmission system becomes negligible.

Given a potential one-way capacity of 36 Gbit/s, and a hierarchy of digital (t.d.m.) traffic whose highest generally accepted level is 139·264 Mbit/s, how may the two best be matched?

If on the one hand we were to assign a separate carrier for every digital block, the number of equipments would reach 256 at full system capacity, and this capacity would be limited to 128 both-way blocks.

On the other hand, very high bit rates become unattractive because the necessary techniques, though realizable in the laboratory, are a long way from a commercial production capability; they would also be more expensive, more power-consuming and more prone to impairments; finally a scheme in which the minimum increment of capacity were, say, sixteen blocks would probably be uneconomical at the present time.

The signalling rate for at least the earlier stages of waveguide utilization will therefore be around 290 Mbaud, enabling two digital blocks to be transmitted on one carrier by 2-phase modulation (b.p.s.k.) or four blocks by 4-phase

modulation (q.p.s.k.). This is the assumption used in Section 2 for stages 1 to 3 in the growth of traffic capacity. However, recognizing the possibility that, in the many years to pass before reaching stage 4, developments at higher signalling rates may make them more attractive, no steps will be taken at the outset which would preclude the use of 600 or even 1200 Mbaud signalling.

It should be noted that since q.p.s.k. at 290 Mbauds carries four digital blocks, it should be possible (via a code-converter) to link up with any other 560 Mbit/s system.

5 Service Requirements

To the outside world the waveguide system need only appear as a body of equipment by means of which (suitably supplied with electric power, nitrogen,† etc.) digital blocks may be cheaply conveyed over distances of tens to hundreds of kilometres. There are three ways in which system imperfections might become apparent to the disinterested outsider to whom the digital blocks are ultimately delivered: he may find his data corrupted with errors, he may find his data arriving at a less constant speed than it was sent, or he may find that for too much of the time it does not arrive at all. The waveguide system is planned and organized in such a way as to provide acceptable service on all three counts, and we here outline how this is to be done.

5.1 Errors

The internationally agreed target for data transmission is that not more than 1 error in 10^8 bits should be added for each 100 km of transmission distance, with an additional allowance of 1 in 10^9 for each multiplexing or demultiplexing stage.

The waveguide system will monitor continuously the error rate on each digital block, and will give a visual indication whenever the performance fails to meet this target. Steps would then be taken to remedy the situation at a convenient opportunity, though it would not be treated as an emergency. At some higher value of error-rate, for example one in 10^3, the signals would automatically be re-routed through a spare channel of the system: this so-called 'automatic channel-protection switching' action would be completed within 70 milliseconds of the high error-rate being detected, so that telephone connections etc. would not be cleared down.

5.2 Jitter

We may visualize a reference time-scale consisting of uniformly spaced instants 7·1806 ns apart. Then ideally the input bit stream will have all its transitions coincident with such instants, the tolerance on the long-term average bit length being ± 15 parts in 10^6, and the output stream likewise. In practice both input and output signals will depart from their ideal, having transitions occurring before (and after) the reference instants by some quantity Δt. Without worrying at this juncture about the details of the behaviour of Δt, we remark that the waveguide system equipment should not cause any deterioration in this respect, and that for a proper input the output will meet stringent tests. In contrast to the error-rate, however, there will not be any attempt to monitor this property on an operational system.

5.3 Non-availability

A 139 264 kbit/s digital path through a waveguide system contains a large number of components, and from time to time one of them may fail. It is impossible to guarantee that any one digital path shall always be in service, but it is expected that a 100 km path will be available for more than 99·99% of the time.

Though individual components are set a very high standard of reliability, there are so many of them that it seems unlikely that we shall meet the target through every waveguide channel. But by providing the automatically switched spare a very high proportion of individual channel failures would not be noticed outside the system; only when a second channel fails before the first has been repaired will traffic be actually lost.

There might occasionally be a failure of a more serious kind affecting all or several channels simultaneously: failure of power supplies, for example, or a malicious assault by a mechanical excavator. The probabilities of all such occurrences are counted into the 0·01% non-availability, and although one has to recognize the conjectural nature of such a numerical exercise, it does provide an operating framework and a scale against which the system can eventually be assessed.

6 Equipment Organization

During the research and development phases of waveguide work, attention has quite naturally been focused on the separate components involved, with particular emphasis on those only recently invented or emerging from progress in solid-state technology. However, in tackling the problems of an operational system, where specification and procurement, commissioning, maintenance and repair come to the fore, it is appropriate to conceive the system in a functional building-block form. We are, for example, no longer interested in how a transmitter source is optimized, or in how the modulation is achieved; rather, we need only to be sure that its output is strong and pure enough to serve our purpose, and that when it fails it can rapidly be replaced.

This 'modular' approach is also a convenient starting point for a description of the terminal and repeater equipment. Figure 2 shows the basic modules of a both-way waveguide channel with their interconnections, and

† Details of waveguide construction and installation may be found in a paper by W. K. Ritchie and C. E. Rowlands entitled 'The millimetric waveguide systems; the design production and installation of the waveguide', *Post Office Electrical Engineers, Journal*, 69, part 2, pp. 79–86, July 1976.

the function of each module is described briefly in the following paragraphs. Binary p.s.k. modulation is assumed, as appropriate to stages 1 and 2 of the implementation programme.

Two digital streams at 139 264 ± 2 kbit/s enter a synchronizer where they are brought up to a common rate of 141·04 Mbit/s. The two streams pass to a channel-protection and interleaving module (c.p.i.); here an extra parity bit is inserted at the end of every 25 traffic digits in order to provide for error monitoring in the system. This raises the bit rate to 146·68 Mbit/s, and the two streams are interleaved to give a single 293·36 Mbit/s stream.

At the transmit head a millimetre-wave carrier source is phase-reversal modulated by the 293 Mbit/s information, after which this and many other modulated carriers are frequency-division-multiplexed in a channelling unit (c.u.) covering the whole of the relevant 10 GHz waveguide band. Finally the various bands launch their signals into the 50 mm circular waveguide (in the 'go' direction) or accept signals from the waveguide (in the 'return' direction) through a broadband filter structure known as the band-branching unit (b.b.u.).

After traversing the waveguide the modulated carrier finds its way through the b.b.u. and c.u. to the correct receive head where it is down-converted to a 1·4 GHz intermediate frequency and preamplified. It is passed on to the equalizer, which has a circuit to compensate for the delay-distortion of the circular waveguide and an attenuator for setting the output signal power to a standard level. Further processing is carried out in a module containing amplifier, demodulator and regenerator (a.d.r.), the output of which is the original 293 Mbit/s stream; this proceeds either to a transmit head for onward transmission or, if required to leave the waveguide system at this station, to a c.p.i./receive module (complementary to c.p.i./transmit), where it is de-interleaved into two 146·68 Mbit/s streams. The parity bits are found, checked, and removed; finally the two streams—now back at 141·04 Mbit/s—are presented to the desynchronizer (complementary to the synchronizer) whence they emerge as fair copies of the original 139 264 kbit/s streams.

If at the c.p.i./r. the error rate is found to exceed a predetermined threshold a message is sent via the telemetry channel to the c.p.i./t., which then sets up a parallel path through the spare channel.

6.1 Synchronizer/Desynchronizer

The 139 264 kbit/s signals enter the synchronizer in coded-mark-inversion form (c.m.i.), so the first operation is to convert them to binary and put them in buffer stores. From there they are read out again as needed and inserted into 'frames' of 828 bits at the output rate of 141 040 kbit/s. Before reaching the output, however, the signal is reconverted from binary into c.m.i. form.

It is a desirable feature of the digit streams passing through the waveguide system that they should have approximately equal numbers of binary 'ones' and 'zeros', taking an average over a few milliseconds; this eases some of the circuit requirements in the regenerators. Since there is no guarantee that the incoming 139 Mbit/s

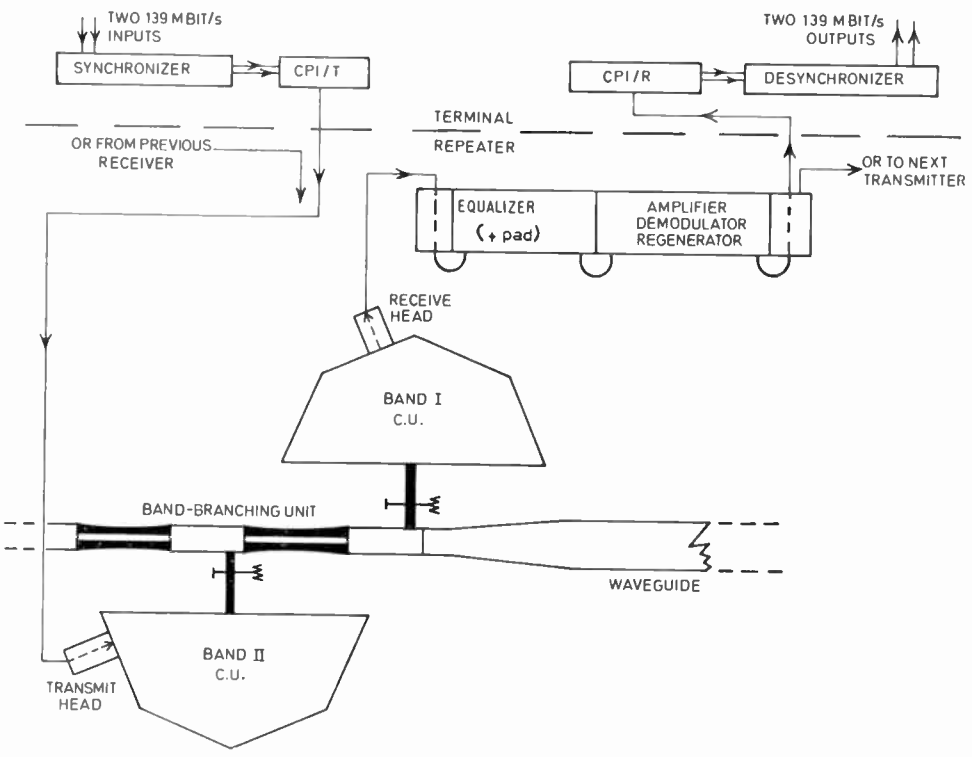


Fig. 2. Modular organization of equipment.

streams have this 'balanced' property, and indeed for some types of traffic they could be expected not to be balanced, the streams are 'scrambled' as they are read out of the buffer store and inserted into the frame. This is achieved by combining the stream with a 2^9-1 pseudo-random sequence, using a bit-by-bit exclusive-OR operation, the sequence having a specific relationship to the frame.

Each frame starts with an easy-to-recognize eight bit 'frame alignment' word, of which four bits are put in each tributary. Next comes a 'service' bit, used for the transmission of alarms, etc. The remainder of the frame consists of five 'justification control' bits and either 818 traffic bits or 817 traffic and one dummy bit, making 828 in all. The 139 264 and 141 040 kbit/s rates both have tolerances of ± 15 parts in 10^6: the difference is between them is therefore 1934 ± 4 kbit/s. When the latter is on the low side there will be fewer frames carrying a dummy bit, and conversely when it is near its upper limit most of the frames will have a dummy. In this way input streams of, say, 139 263·2 and 139 265·1 kbit/s are both brought to a common rate of (say) 141 197·5 kbit/s, though of course they contain different quantities of dummy information.

If bit number 691 in the frame is a dummy bit, then all five control bits will be set to 'one'. These control bits are spaced out through the frame, so that even when the 'traffic' is all zeros some frames at least will have a 'one' every 137 bits—this helps to keep the clocks in the regenerators synchronized.

The desynchronizer performs operations the reverse of those described above. The input $141\,040 \pm 2$ kbit/s streams are decoded from c.m.i. into binary; the frame alignment signals are recognized; the traffic bits are descrambled by exclusive-OR with the same 2^9-1 pseudo-random sequence, and entered into a buffer store; the justification control bits are checked, and if three or more of them are 'ones' then the bit number 691 is not treated as traffic but thrown away. This use of a majority decision on five control bits makes it possible to operate the system at moderate error rates without the danger of inserting extra bits by mistake into the output bit stream (even at a bit error rate of 1 in 10^4, only one frame a week is likely to have three erroneous control bits). The output is read smoothly from the buffer store, and since it contains only the original traffic bits it must also be at the same average rate as the input; it is reconverted into c.m.i. form before appearing at the output socket.

6.2 Channel Protection and Interleaving (Transmit)

At the input of this module the two synchronized 141 040 kbit/s streams† encounter a switch matrix: this is the point at which they will be diverted to the 'spare' channel if their allocated 'worker' channel fails. They are then converted from c.m.i. to binary form and divided into blocks of 25 bits (this is done without any particular relationship to the incoming frame structure); the parity of each block is assessed, and a 26th bit is added so that the overall 26-bit word contains an odd number of 'ones'. This operation increases the bit-rate of the streams by 4% to 146 681 kbit/s.

The two streams are interleaved by taking a half-bit sample from each in turn at a high-speed toggle circuit, and the 293 363 kbit/s result is once again converted to c.m.i. The 26-bit words in the two tributaries are staggered by about a half-word.

6.3 Channel Protection and Interleaving (Receive)

The 293 363 kbit/s c.m.i. input is converted to binary de-interleaved into two 146 681 kbit/s streams. These are passed to circuits which locate the 26-bit word structures of the two streams, more-or-less on a trial and error basis: a sequence of 52 bits (from the high-speed input stream) should have the parity bit from one tributary in slot 24 and that from the other in slot 51; if the start of the sequence was incorrectly chosen then parity violations will occur in those two slots for about 50% of the checks of successive 52-bit sequences; the sequence start is then slipped by one slot and another attempt is made, continuing until a start is found where parity violations fall to a low level at slots 24 and 51. If no such start can be found the channel is deemed to have failed: a signal is despatched back to the transmitting end to initiate the switching over to the spare channel, and an alarm indication signal, consisting of all binary 'ones', is sent on instead of the traffic.

Having latched onto the 26-bit word structure the number of parity violations is a good indication of the binary error-rate of the system. To make use of this information the following scheme can be envisaged. The violations from both tributaries are fed to the same counter which is reset every 100 ms; if the count reaches 300 then without waiting for the 100 ms to finish an alarm signal is given to activate automatic channel protection. In this way a sudden error-rate of 1 in 10^3 is recognized in 1 ms, 1 in 10^5 is recognized in 100 ms, and fewer errors than this do not cause protection switching to take place. Moreover the system does not over-react to sudden isolated bursts of errors less than 300 in number.

A second counter in parallel to that above maintains a count over 10 seconds at a time; if a count of 30 is reached a lamp is lit permanently on the front of the module, indicating that the channel, though not yet 'failed', is in need of attention as soon as can conveniently be arranged.

The parity bits are removed from the tributaries, bringing their rate back down to 141 040 kbit/s. After conversion to c.m.i., they pass to the output through another switch matrix which, if the traffic had been diverted onto the spare channel, would bring it back onto its allocated path.

† For a q.p.s.k. system *four* synchronized 141 040 kbit/s streams would be processed together.

6.4 Telemetry

Means must be provided for the transmission, in both directions along the entire route, of housekeeping information, including alarms, error-rates, and so on. A dedicated r.f. channel might be assigned for this purpose, or alternatively it may be found more economical to carry this data by amplitude modulation of some of the traffic carriers close to 40 GHz.

6.5 Spare Channel

In most respects the equipment for the spare channel is identical with that for any worker channel. It is normally either used for low-priority traffic or left 'free-running', but on receipt, from the remote-end of the system, of the signal to initiate protection switching, the switch matrix is activated to take the 141 Mbit/s streams to both the spare and the worker channels in parallel. The control circuits associated with these activities therefore differ slightly between spare and worker.

At the receive end of the spare, the 26-bit word structure would be sought in the normal way, and only when word alignment is found would the switch matrix be changed over to feed the output from the spare in place of the failed or failing worker channel.

6.6 Band-branching Unit

The waveguide will enter the terminal or repeater building below ground level, and from a tensioning and pressurization point of view will terminate in a special-purpose chamber. From there to the equipment room a feeder waveguide will be extended, being perhaps a simple straight 50 mm section passing through a common wall, or a longer and more tortuous connection depending on the accommodation available. The b.b.u. itself will be mounted horizontally, either along a wall, or near to floor or ceiling level with access along both sides. It consists of an in-line cascade of filters giving five rectangular waveguide outputs along its length, in ascending order of frequency as indicated in Fig. 1. Each of the four boundary frequencies is defined by the exact size of the reduced-diameter semicircular waveguides which link the five output couplers. It will be noted that Bands III and IV are brought out at a common port, as are Bands V and VI, and VII and VIII: this arrangement helps to keep down the total filter loss at the highest frequencies. At each rectangular port there is a 20 dB cross-guide coupler which can be used for some test purposes.

6.7 Channelling Unit

The channelling filters are built into a single structure having conveniently, an approximately fan-shaped outline: at the centre a single rectangular port is connected to the corresponding b.b.u. port, either directly or with as short an insert as will enable access requirements to be met (Fig. 2). Around the perimeter there are sixteen ports in the same rectangular waveguide size, each carrying a particular r.f. channel with a bandwidth of about 560 MHz. The channelling filter is basically a four-stage commutating network, which contains no resonant or reflecting elements and is consequently of low loss and rather well matched at all frequencies. At the first stage the signal is split into two equal parts which travel different distances before recombination in a four-port coupler; alternate channels emerge at the two ports of this coupler and are passed to the following stages for further separation. This type of network is readily made on a computer-controlled milling machine.

6.8 Transmit Head

Two methods of forming a transmit head were investigated in the waveguide field trial, and still others are possible. However a strong preference has emerged for direct phase-modulation of a millimetre-wave carrier by means of a reflection switch, since the method makes efficient use of the available power and does not tend to produce spurious sidebands.

A b.p.s.k. transmitter is shown schematically in Fig. 3. A pure carrier of tens to hundreds of milliwatts is generated by an impatt oscillator; the frequency of the carrier is kept very stable ($\sim$ 1 MHz) by controlling the temperature of the oscillator body very carefully, or alternatively by coupling a high-Q stabilizing cavity to the circuit. The power passes out through an isolator: this is there to mop up the (modulated) leakage signal coming back from the modulator, which would otherwise tend to perturb the frequency of the oscillator.

The modulator itself is a circuit having only one millimetre-wave port, and is switched into one of two states; in both the incident carrier is about 90% reflected, but the phases of the reflections in the two cases differ by 180°. The critical element in the modulator is a tiny p–i–n diode which is biased into forward conduction for one state and into total depletion for the other; the driver amplifier needed to accomplish these bias changes at 293 Mbit/s has to be very fast and powerful. The logic of this amplifier is such that a binary 'one' at its input is made to change the modulator state, while a 'zero' leaves it unchanged (this is known as differential phase-shift keying, d.p.s.k.).

The signal reflected from the modulator is separated from the incident carrier by means of a millimetre-wave circulator; it then passes to the output of the head through a coupler which takes a small sample of the output to a separate port for test purposes.

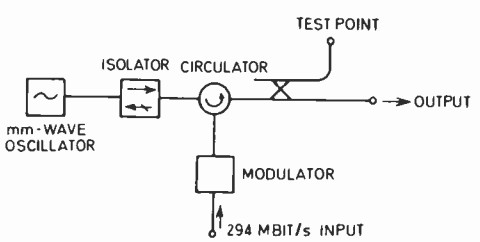


Fig. 3. Binary phase-shift keying transmitter.

The whole transmit head is bolted directly onto the channelling unit.

6.9 Receive Head

Only a few milliwatts of local oscillator power are needed, but the purity and frequency stability of the output must match that of the transmitter oscillator; in practice it may be most convenient to use an identical oscillator under conservative operation, rather than employ a different device, such as a transferred electron oscillator.

The local oscillator feeds into a balanced millimetre-wave mixer through one waveguide port; the modulation signal from the channelling filter enters through the other, at a level of around −50 dBm. Both are passed through a tiny microwave integrated circuit to two matched gallium-arsenide Schottky-barrier diodes, which perform the mixing.

The resultant intermediate-frequency signal, now a b.p.s.k.-modulated 1·4 GHz carrier, is immediately amplified by 45 dB or so in a broadband transistor amplifier.

The noise figure of this receiver is basically determined by the conversion loss of the mixer diodes (5–10 dB) and the noise figure of the transistors (2–4 dB) since the diodes themselves contribute negligible extra noise.

Normally the local oscillator frequency is 1·4 GHz lower than the incoming signal. However if this were so for channels 1–4 of Band II, the 'image' frequencies would lie in the top of Band I; for example, any leakage from channel 16 of Band I onto the mixer of Band II channel 3 would be converted to the same i.f. and since it is at the transmitter end of I/16 the leakage might be comparable in strength to the wanted II/3 signal. The problem is solved by placing the local oscillator 1·4 GHz higher than the incoming signal; however the i.f. spectrum is now inverted, and some slightly different equalizer will have to be used.

It is very important that the i.f. be accurate, namely 1400 ± 5 MHz: it is set to ± ½ MHz when the waveguide link is commissioned, by a fine adjustment on the local oscillator. The stability of the transmitter and local oscillators should be such that the correct i.f. is maintained for months or years.

Like the transmit head, the receive head is bolted directly onto the channelling unit. Connection to subsequent modules is made via a long length of semi-rigid microwave cable, permanently installed in the station and terminating on a front-facing panel in the equipment racks.

6.10 Equalizer

As a module the 'equalizer' is intended to include any i.f. signal processing which is particular to the channel in question. The most obvious factor here is that the group-delay slope across a channel depends both on the length of the waveguide section and on the carrier frequency: for 24 km at 30·8 GHz the required delay slope in the equalizer is +160 ns/GHz, and for 18 km at 75 GHz it is only +8 ns/GHz, while for the 'inverted spectrum' channels II/1–4 one would need about −65 ns/GHz.

The i.f. power level, too, will vary from one situation to another, not only because of the waveguide length and carrier frequency, but also because of tolerances on transmitter power, channelling losses, receiver gain, and the loss of the delay equalizer. Provision will be made in this module for bringing the power level to some standard value, probably somewhere near ½ μW.

The delay equalizer itself is a passive component, a composite slow-wave structure made of copper tape totally surrounded by polypropylene. It is proposed that a range of fourteen delay slopes be provided, such that any delay slope can be achieved to an accuracy of ±2 ns/GHz by cascading only two units: this accuracy is considered to be adequate for b.p.s.k. modulation, though perhaps a little coarse when it comes to q.p.s.k.

Connection to the equalizer from the station semi-rigid cable, and from equalizer to a.d.r. module (q.v.) will be by means of U-links on the front panel, to provide access to the signal for commissioning and maintenance purposes.

6.11 Amplifier-Demodulator-Regenerator Module

In contrast to the equalizer, the a.d.r. is identical for all b.p.s.k. 293 Mbit/s channels. The i.f. signal is amplified to about 20 mW and band-limited to minimize the thermal noise and interference components reaching the demodulator. In the demodulator the signal is split into two parts, one of which is delayed by one time-slot; the two are then compared in phase and a baseband signal is formed whose polarity indicates whether the relative phase is greater (positive) or less (negative) than 90°. Choosing an optimal instant to sample this baseband waveform, the regenerator makes a clear decision as to its polarity: if positive, the assumption is that a 180° phase-change had been made at the transmitter between the two time-slots now being compared and that therefore the input to the modulator had been a 'one'; if negative, very probably no such change was made, corresponding to a binary 'zero' input.

Of course, in its journey from transmitter modulator to receiver demodulator, the signal has been *impaired*, by distortion in the various non-ideal components it has passed through, by interference from some of the other channels or local oscillators, and by thermal noise. The first of these will not by themselves spoil the signal phases by more than a few degrees, but occasionally a random thermal noise excursion will cause the relative phase between two adjacent time-slots to be caught on the wrong side of the 90° 'threshold' just at the moment of regenerator decision, and an error in the output bit stream results. In practice, and taking into account the distortion and interference, a signal-to-noise (S/N) ratio at i.f. of 15 or 16 dB will ensure that the error rate is less than 1 in 10^9.

The importance of frequency stability to this differential p.s.k. system arises from the one-slot delay involved in the phase comparison: this delay is 1800° at 1400 MHz, so every 1 MHz drift in the i.f. will reduce the 90° threshold by 1·3 degrees.

7 Margin

To maintain a bit error rate of less than 1 in 10^9, the signal-to-noise ratio reaching the demodulator must be at least 15 dB for b.p.s.k., and at least 22 dB for q.p.s.k. using differential phase detection. This takes into account the distortion of the signal and a small amount of inter-channel interference, but assumes that everything is optimally adjusted. In order that the system should at least be tolerant to reasonable drifts in such parameters as i.f., threshold setting, timing, output power, and so on, the actual S/N is increased by an amount known as operating margin. The value of this margin depends on whether or not the channel is a critical one in relation to repeater spacing.

The millimetre-wave power required at the input to the receive head is $kTBF.(S/N)$ where k = Boltzmann's constant, T = absolute temperature, B = equivalent noise bandwidth, F = receiver noise factor. This is obtained from the transmitted power P, attenuated by a factor A_c for the total channelling filter losses at both ends and by a factor A_{WG} for the circular waveguide itself.

Case (1) Uncritical: For example, at 39 GHz we can take the filter loss as 7½ dB, the loss of 24 km of waveguide as 55 dB, and $kTBF$ as −77 dBm. Then a mere 6½ dBm of transmitter power will give a handsome operating margin of 6 dB for b.p.s.k., surely enough for all reasonable contingencies. In fact there is a case for relaxing the noise factor requirement for the receiver under these circumstances. Even 4-phase d.c.p.s.k. can be operated over 24 km at this frequency, with an operating margin of 7 dB, without requiring an unreasonable transmitter power.

Case (2) Critical: At 30·8 GHz, the lowest traffic-carrier frequency, the attenuation of 24 km of waveguide is expected to be 76½ dB, and filter losses will be 6½ dB; with some development of the receiver, and selection of mixer diodes, $kTBF$ may be held to −80 dBm. A very powerful transmitter would be needed for four-phase operation. Even with an output of 200 mW S/N would be down to 20 dB; differential detection would be difficult at this level, and one would have to resort to coherent detection (or perhaps to use only b.p.s.k. in this eventuality).

Figure 4 shows how the various parameter values would be allocated for repeater spacings up to 24 km. It is seen that from the most critical channels the transmitter power drops rapidly while S/N rises to a maximum plateau and the noise factor requirement is relaxed. This is consistent with the expectations of the semiconductor device fabrication, namely that processing of a batch of devices yields a few of the very best (highest power impatts, lowest loss

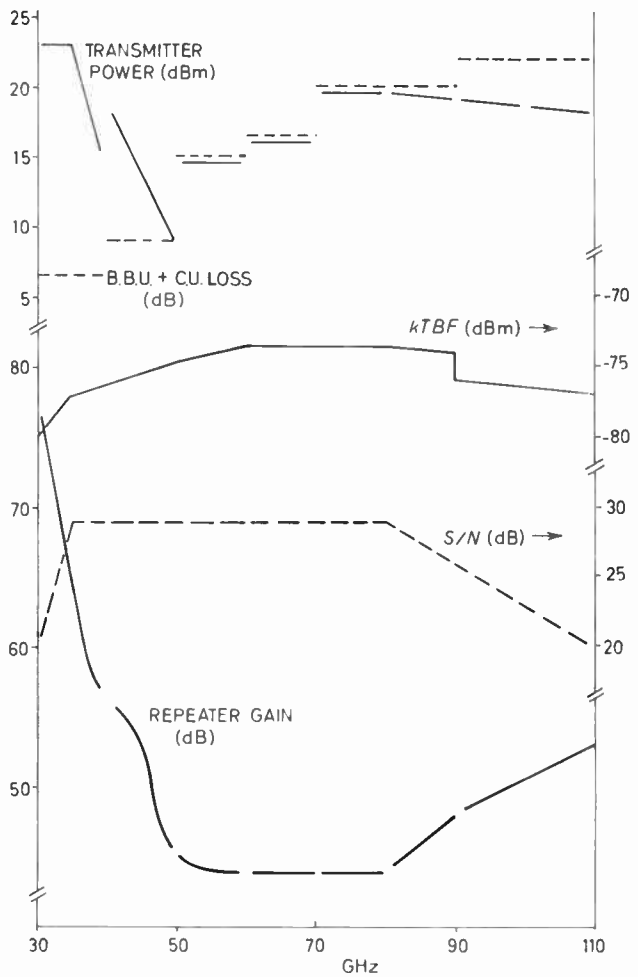


Fig. 4. Repeater parameters for 24 km spacing.

mixers and p–i–n modulators) and a larger quantity of devices of more modest performance, all usable over a fairly wide range of frequencies.

8 Conclusion

Once the waveguide is in the ground and repeater stations provided at up to 24 km spacing, the system provides for traffic growth in increments of two or four digital blocks, at low marginal cost and at short notice. The ultimate capacity is far greater than would be needed for telephony alone, but would clearly be well able to cope with an up-turn in the demand for transmission capacity for other purposes, should the need arise. The transmission performance will meet Post Office and international requirements, and the system should need little maintenance effort.

9 Acknowledgment

Acknowledgment is made to the Director of Research of the Post Office for permission to publish this paper.

Manuscript received by the Institution on 17th May 1977 and in final form on 17th August 1977. (Paper No. 1790/Comm. 156)

UDC 681.32

Functional decomposition on multi-micro-processor systems

KEITH D. BAKER,
B.Sc., D.Phil., M.I.E.E.E., M.A.C.M., M.B.C.S.*

Based on a paper presented at the Conference on Computer Systems and Technology held at the University of Sussex from 29th to 31st March 1977

SUMMARY

Increased sophistication ofpresent microprocessors and other l.s.i. devices allows creation of unconventional computer architectures. The work reported here is part of a project concerned with the application of multi-processor architectures to business data processing. An application is considered as a number of interacting functional units. The operation of a system consists of the allocation of functional units to processing units at execution time. A processing unit contains a microprocessor, r.o.m., r.a.m., firmware and interface devices for communication with adjacent processing units and peripherals. Basic modes for interconnection of processing units are discussed. More generalized cellular structured architectures are considered and software problems examined.

* *School of Engineering and Applied Sciences, University of Sussex, Falmer, Brighton BN1 9QT.*

1 Introduction

The work reported in this paper represents the first part of the continuing investigation and development of new architectures for computing systems applied to business data processing. The final objective of the investigation is the implementation of a dynamically reconfigurable multi-processor system. The realization of this objective is planned in several phases. Each phase will correspond to the completion of a particular multi-processor architecture to satisfy the requirement of a particular class of business data processing applications. An assessment of the characteristics of these systems shows that the optimal configuration is application dependent. On the other hand an architecture suitable for all applications and which at the same time is extendable to allow for expansion of the users computing needs requires a modular system construction. Each module may be regarded as a basic computation resource consisting of one or more processors, memory, I/O ports, etc. and some firmware. A particular processing task is accomplished by organizing the resources so as to realize as closely as possible the optional configuration. Extension of the computing capability by connecting together several primitive computation resources can be regarded as the formation of a cellular structure.

2 Hardware Trends

It is of interest to examine briefly the developments that have taken place during the past three decades. In the early days of computers the basic building blocks available to the system designer consisted of the basic circuit components, resistors, capacitors and thermionic valves. This was followed by the development of transistors which increased the reliability of the circuits and reduced the power requirements. Following this was a period of substantial development in the field of solid-state circuitry resulting firstly in small-scale integration and later, medium-scale integration. The result was an increase in the sophistication of the basic building blocks available to the system designer. This saw the development of large powerful general-purpose central processors generally following the now classical von Neumann architecture. The high cost of these processors required methods to increase the central processing efficiency relative to the processing carried out in peripheral parts of the system. This lead to the development of large multi-tasking operating systems, the overlap of I/O operations with central processor activity, virtual memory systems etc. This allowed the central computing power to be shared concurrently by many users both in batch and interactive modes of operation. The price of this flexibility is the overhead of a sometimes very large complex operating system to schedule the resources amongst the various classes of users.

Restricting our attention now to the past decade it can be seen that there have been further dramatic advances in

solid-state technology. The fabrication of high-density gate circuitry with semiconductor chips of up to the equivalent of 20,000 transistors is now economically viable. The emergence of large-scale integration techniques has had, and will continue to have, profound effects on the developments in the computing industry and the electronics industry generally. Initially the new fabrication techniques were applied to the production of memories, the latter being a regular structure with high gate-to-pin ratios. Since about 1968 it has been observed[1] that the bit density of memory chips has more or less doubled each year. It is now possible to have 16 kbits memory available on a single chip. Using the conservative estimate of bit density doubling every 1·5 years, chips with capacities exceeding 64 kbits can be expected towards the end of this decade. An upper limit based on calculations for present fabrication techniques[2] can be expected at about 128 kbits.

Following the success of l.s.i. memories the circuit designers turned their attention to producing the central processor part of the computer with l.s.i. techniques. The first processor-on-a-chip, the Intel 4004, appeared on the market late in 1971. Since then several manufacturers have produced microprocessors of 8, 12 and 16-bit word lengths and with both fixed and microprogrammable instruction sets. At the present time the most successful of these, in terms of market share, has been the Intel 8080. The 8080 processor has 78 instructions and the instruction cycle time is 2 μs. Recently Intel have announced a new processor chip to succeed the 8080 which will be faster and will contain memory on the same chip.

Further developments are currently taking place with charge-coupled devices which provide a solid-state alternative to mechanical rotating memories such as the disk. These devices can achieve very fast data transfer rates in the region of 16 Mbit/s. At the present time however, the conventional secondary storage devices, disks and mini-disks are being used in conjunction with the microprocessor. Activity can also be seen in the interface circuitry between the processor and the peripheral devices. Again the functions carried out by the interface circuits have been integrated in some cases on a single chip.

The system designer now has available some very sophisticated building blocks. The fact that a processor can now be realized on a single 40-pin, dual in-line package, means that the general-purpose computer, once occupying a complete room, can now be considered as another circuit component on a printed circuit board. The fabrication techniques of l.s.i. have meant a dramatic decrease in the cost of processing power and memory. Consequently the application areas where the use of microprocessors is now viable is increasing and areas previously realized with random logic may now be found using programmable logic. As far as the computer user is concerned there has been a move away from the central computing facility and more intelligence has appeared in the peripheral devices. There has been a significant increase in the so-called intelligent terminal. Here an interactive terminal containing a microprocessor is programmed to carry out a given function such as data entry and point of sale. General-purpose key-to-disk terminals, with and without hard copy facilities, together with special purpose terminals for banking and industrial data collection are gaining an increasing proportion of the computer hardware market. It is estimated[3] that in the U.S.A. alone terminals will account for 24% of the total computer hardware cost by 1980.

Because of the decrease in the cost of processing power it has now become a viable proposition to consider the possibility of radically different hardware structures based on a number of communicating processors. Apart from a relatively small class of applications requiring very high processing speeds, the large processing power of a conventional computer is consumed in processing many jobs concurrently. It would therefore seem acceptable to have several processors concurrently processing mnay jobs so long as the same throughput is maintained. Several examples of multi-processor organization are already in existence. One such example is the array processor where a number of processors are arranged logically in the form of a two-dimensional array. In general the whole structure operates synchronously with the same instruction being broadcast to and executed by each processor in the array. This type of processor is particularly suited to problems where the same sequence of operations must be carried out on multiple data items. In a further type of multiprocessor arrangement the architecture of the system has a network topology. Except for the case of a communication network it is generally true for systems of this structure that a number of smaller computers communicate with a large host machine. Most multi-processor systems currently operating are based on the use of minicomputers.[4] It must therefore be concluded that the full potential of interconnecting a number of low-cost processors and memories has yet to be fully examined.

3 Software Trends

From what has been discussed in the previous Section it can be seen that the typical computer installation at the present time consists of a large central processing unit servicing the needs of multiple users via a complex multitasking operating system. Experience has shown, time and time again, that the specification, design and implementation of any medium or large-scale software product is a complex process. A superficial inspection indicates several causes of the problems encountered. Incompleteness and ambiguity in the product specification, the inherent difficulties in the logical design of correct software and the indeterminate nature of the implementation phase are but a few of the more obvious. In general changes to the specification of a software product will occur throughout the design and implementation phases and also throughout the lifetime of the eventual product. It has become a fact of life that any software product

remaining in use is subject to constant change. This follows as a direct consequence of the basic economics of the situation. Any medium or large-scale software product represents a substantial investment for the manufacturer. As such there is significant pressure to change and modify the existing product to suit new applications. Changes or extensions to the original specification introduce further errors which at best are associated only with that software relevant to the change. In general, however, the process of eliminating software errors uncovers further errors. If the product specification remains constant then this is generally a convergent process. If the change in the specification requires a partial restructuring of the software then significant degradation of the product must be expected.

Application of conventional product management procedures are not adequate in the case of software production. In an effort to gain an understanding of the laws governing the processes of software engineering Lehman *et al.*[5–7] have carried out detailed studies of the histories of several products over the period of several years. These workers have successfully identified a number of global variables enabling the programming process and the superimposed project management process to be modelled as the interaction between two complex dynamic systems. Some remarkable observations have been reported by these workers.

Three systems have been studied, a large multi-purpose operating system, a transaction-processing system and a small system executive. All three systems fall into the class of systems to which the work reported here is directed. Although the manufacturer and environment for each of the systems was different and the size of the systems spanned some two orders of magnitude the similarity of the observations is indicative of an underlying common set of processes.

A convenient measure of size of a software system is the number of modules. An equally convenient measure of time is the release sequence number since it represents a point in time when a new functional specification emerges. The following observations have been reported.

(a) *Product Growth.* The size of the product showed a steady growth with release sequence number. Any increase in size above some natural increment between releases produced a decrease in the subsequent release.

(b) *Structural Complexity.* The complexity of the product is a measure of the degree of interaction between its components, in this case the interaction between modules. The number of modules handled for each change was observed to increase with time. When complexity is defined as the fraction of modules handled then the increase in complexity was found to be quadratic.

(c) *Limits to Growth.* An apparent limit to the growth of the product was observed despite the continued need for implementation of further changes. The increase in complexity of the product was so great that further extension could only be obtained by a restructuring of the design.

What conclusions may be drawn from these observations? Firstly, the software design process is not well understood. The advent of structured programming has made some impact in this area but has yet to receive wide application. Certainly, medium and large-scale software products do undergo continuous changes and in some instances extensions to the specification do lead to unforeseen consequences and severe degradation of the product quality. This can be viewed as the onset of some kind of instability in that the complexity of the system has increased to the point where it becomes unmanageable. Increasing the capability of a system leads to a corresponding increase in the complexity.

4 Multi-Processor Architecture

The system software currently in operation on large general-purpose computers must be considered as a complex dynamic system that evolves with time. If the capability of such systems is increased and hence its complexity increased, then it is likely that a point will be reached where it becomes unmanageable. In this event it will become unstable with the occurrence of unforeseen random events that degrade the overall system performance. Since the demand for functional capability of a computing system is likely to increase then the system designer must employ architectures that seek to minimize the increase in system complexity. On the other hand the decrease in cost of hardware and the availability of the processor on a chip suggests the need for an investigation into the use of multiple processor architecture in the solution of certain data processing problems.

We first make the observation that current microprocessors do not match the processing power of the large third generation central processors. Some microprocessors are, however, approaching the speed of the conventional minicomputer. The processing power of the large machines is diluted because of its multiprocessing operation. Also, for the majority of tasks found in business data processing programs absolute speed is not the most essential attribute. In making this statement the comparison is being made with the large scientific programs where arithmetic calculations dominate.

If we consider the starting point of the design of any system to be the functional specification then it is generally the case that the design process proceeds by sub-dividing the specification into a number of functionally related areas. These areas form the sub-components of the system and the total system is realized by the dynamic interaction between the sub-components and the outside world. The fact that sub-systems may be identified is indicative of the fact that the interaction between the sub-system and the

rest of the system is less than the interactions occurring within the sub-system. In this case the sub-system is almost independent of the rest of the system and only requires occasional inputs and outputs. The meaning of occasional here depends on the time-scale being considered.

This functional sub-division of the specification may be continued at a lower level and is typical of the top down design process. Each functional sub-system is divided into a number of almost independent sub-sub-systems, and so on. The complexity of the overall system is handled by the construction of a hierarchy which assists in the identification of the interactions between the various system functions. Now, at some stage in the decomposition process we may make an association between sub-system function and a processor on which that function will be executed. No consideration is given at the present time to determine the level in the decomposition at which the association is made. Suffice to say that at each stage in the operation of the system the code associated with each identified sub-system function must be bound to a processor for execution. Unless a particular sub-system function is almost independent of other sub-system functions then it will be subject to a number of constraints. These include the time at which the sub-system code is executed and the necessary configuration of the processor, or processors, on which it executes.

4.1 Implementing a Functional Partition

An analytical approach to the problem of partitioning the total system function into a finite number of interacting functional sub-systems has yet to evolve. We have chosen to follow the advice offered by Dijkstra[8] and make use of the observation reported by Simon[9] concerned with nearly decomposable systems. We may define the functional sub-system at any level of decomposition of the total functional specification as a functional unit. There are at least two functional units at each level of the decomposition. The depth of the decomposition process may be identified by assigning the level number to the functional units occurring at that level. The deeper the decomposition process proceeds the smaller and simpler are the functional units at the lowest level. Carrying the decomposition to the limit, then, at the lowest level, a functional unit becomes an instruction to be excuted by a processor. The optimal level of the decomposition for a given application and a given configuration of basic computational resources is the subject of further work associated with this project.

In the first phase of the investigation reported here it was decided that small and medium-scale business-oriented systems would be analysed in the first instance. The functional decomposition was taken to the point where each functional unit was sufficiently simple to be executed on a single processor and at any time a processor is concerned only with the execution of one functional unit. In a typical on-line business computing environment the total function of the system may be concerned with the input of data from one or more terminals, the update of a central database, the processing of information contained in the database, dealing with enquiries and generating reports. Each of these areas may be considered as a functional unit at the first level of decomposition. In turn each of these units may be decomposed into a number of simpler functional units. Some functional units will therefore require an interaction with the outside world. That is, the input of data from an operator or the output of data to an operator or other user. At the time of execution therefore, such functional units must be bound to processing units having the necessary input output resources. This need not be absolutely the case in the final versions of the architectures under investigation since processing units may be connected in a cell like structure. Access to the processing unit with the required input output resources will therefore be possible. Other functional units may be identified that require no interaction with an operator. The execution of these functional units must be synchronized either implicitly or explicitly with the execution of other functional units.

4.2 Basic Processing Unit Hardware

Certain aspects of the basic processing unit structure are considered in this Section. Whenever possible standard manufactured i.c. chips have been used. All programmable devices are used in accordance with the manufacturers' published specifications. In essence the functional operation of a system in a particular application is contained in the execution of the functional units. The functional units are purely a software concept but must be brought into contact with a processor at execution time. The basic processing unit provides the basis for the execution of a functional unit.

The basic processing unit consists of a microprocessor together with an extendable amount of r.a.m. and r.o.m. Also contained in this unit is the circuitry necessary for the communication with other processing units and with various peripheral devices. Typically these would include such things as keyboards, visual display screens, printers etc. The I/O ports on each processing unit include facilities for both serial and parallel communication. One of the more important access channels on the processing unit is the direct memory access (DMA) port which allows fast transfer of information between processing units and between a processing unit and a peripheral device. Several DMA ports may be contained in each processing unit. Figure 1(a) shows the diagramatic representation of a basic processing unit with four DMA channels. Figure 1(b) represents a basic processing unit with two DMA channels and three slower communication channels. The objective of the total investigation is the interconnection of processing units to form a cellular structure as shown in Fig. 1(c).

At the present time each processing unit contains a single Intel 8080 microprocessor. Each unit operates

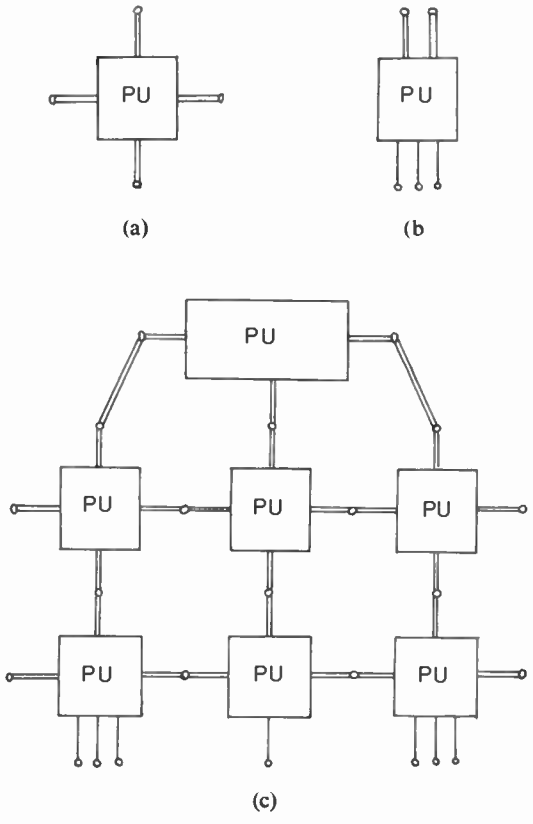


Fig. 1. Processing unit configurations.

asynchronously except during the period of an inter-unit DMA transfer. At this time the activity is synchronized and controlled via a handshake mechanism between the units. Special logic is necessary to implement the DMA process. A single bidirectional bus connects two processing units. Each unit may gain access to the bus in order to request a transfer of data. Bus contention is resolved by the DMA logic and error checks are carried out on the data transferred. On completion of a DMA transfer, status information is contained in the DMA status registers associated with each processing unit. During the transfer of data the processors may either be active or passive. In the first case the transfer is carried out on a cycle-stealing basis. Alternatively, each processor is put into a HOLD state for the duration of the transfer. In the second case it is necessary to implement a mechanism to free the processors at the end of the transfer. Furthermore it is also necessary to free each processor in the event of an abnormal termination of the data transfer. In the case of slower communication channels such as those with keyboards and printers, standard interface chips have been used. Ideally each processing unit is identical in construction to the next. From an economic standpoint each processing unit will contain the interface logic specific to the application in which it is used.

It is not envisaged to implement circuitry to switch peripheral devices between processing units automatically. In this context it is seen that dynamic configuration of a system will include only that concerned with the allocation of a set of processing units to satisfy the requirements of a given set of functional units. Allowing dynamic switching of hardware increases the complexity of the system which is contrary to one of the basic objectives of the investigation. Instead it is expected that a particular structure of processing units can be used in several different applications. The functional operation of a system is contained in the application software which is decomposed on the hardware structure. At a later stage it is planned to allow for manual transfer of peripheral devices such as keyboards and printers between processing units. Each processing unit will have knowledge of its own configuration.

4.3 Basic Processing Unit Firmware

A basic distinction is made between the software concerned with the functional operation of the system and that concerned with the operation of the hardware. The former is that to implement the particular application. The latter concerns the details of the hardware which is independent of the application and may be termed system software. The system software in the basic processing unit is sealed in r.o.m. forming the firmware. The firmware is completely modular in structure and consists of a number of port handlers together with a basic monitor. The structure is illustrated in Fig. 2.

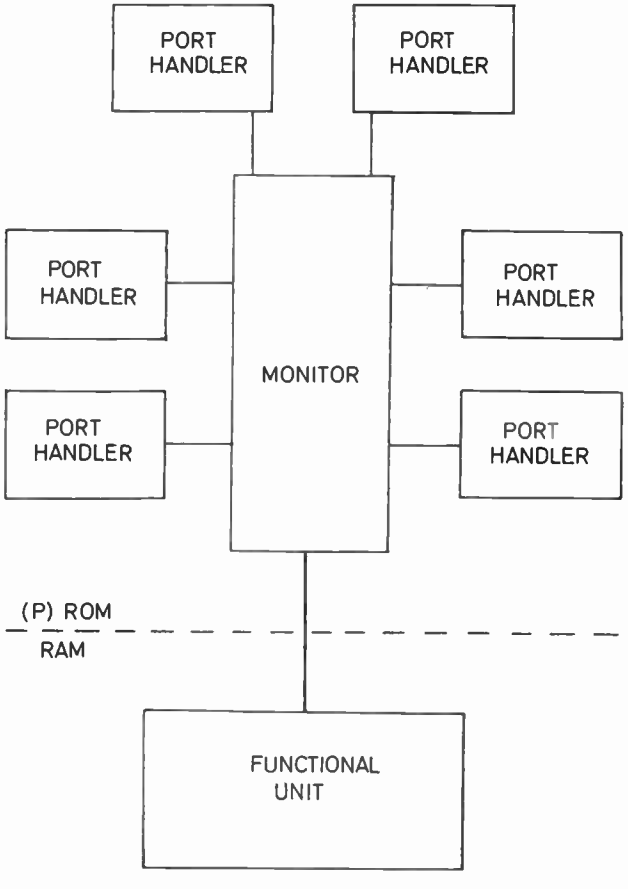


Fig. 2. Software/firmware organization.

Each port handler is specific to the particular device with which it interfaces. Its function is to carry out the particular information transfers to and from the device it controls. All the details necessary to carry out the processing on a particular device are contained in the associated port handler. In other words the handler provides the interface between the operational software and the device hardware. There is a standard format for the interface between the monitor and each port-handler. In general this software interface will contain a command and associated data for a device or a status request from the device controller.

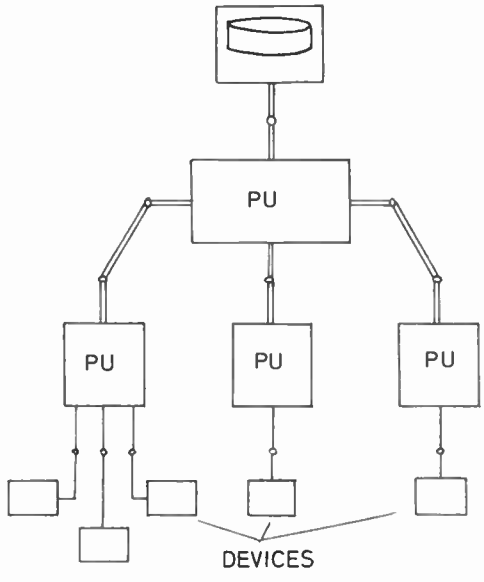
The operational function of the monitor is to establish communication between the functional unit and the various devices connected to its processing unit. This includes communication with another processing unit. Communication between the functional unit and the monitor is via a standard interface consisting of a set of user macros and subroutine calls. The functional unit may issue requests for data from a keyboard or a record from a file. The monitor establishes contact with the necessary port handler to initiate the request. Once a command has been executed the monitor must check for error conditions and take any necessary action before allowing the execution of the functional unit to proceed.

The system function of the monitor includes the establishment of the configuration of the processing unit. The transfer of functional units in and out of the processing unit is also under the control of the monitor. In the initial systems considered the functional unit can be considered almost independent of the execution of other functional units in real time. In the second phase of the developments of these systems this is no longer the case. It is necessary for additional information to be associated with a functional unit describing the functional environment necessary for its execution to proceed. This will include the communication with other functional units. It will be the responsibility of the monitor to ensure that the required functional units are located in adjacent processing units. Since the partition of the application into functional units will have been accomplished before the application is loaded then the above exercise ensures the correct communication between them.

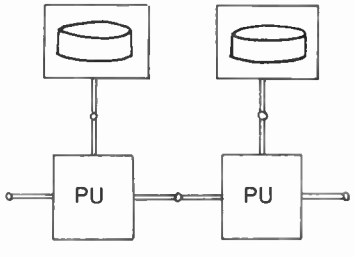
5 A Basic Application

Initially a number of simple business applications have been considered where the functional decomposition of the system is fairly well defined. The configuration of the processing units is correspondingly simple and reflects the structure of the application. To illustrate the points made here we consider a typical system allowing several kinds of data entry to a central database. The database is composed of a number of files stored on direct access devices. On-line enquiries may be made to the database by a user from a v.d.u. terminal. The data contained in the files must be processed and reports generated on a printer. Each data entry requirement may be considered as an almost independent functional unit. This is generally true since and even if the data entered at one terminal is correlated with that entered at another, as for example in an on-line seat reservation system, resolution of any possible conflict may be handled by a further internal functional unit. On-line enquiries may be considered to form a further functional unit and so also the report generating procedures. No further decomposition of the functional units is made in this case. A configuration of processing units capable of supporting these functional requirements is shown in Fig. 3(a). This is a simple star configuration that has been specifically constructed for the particular application. In a cell structure topology of processing units the functional units would be associated with appropriate processing units at execution time.

In the representation given the central database has been shown connected to a single processing unit. In practice however, for reasons of reliability, this processing unit may contain two units. At the lowest level of simplicity of operation the functional units may be associated with a given processing unit. This allows no flexibility of operation in the event of component failure. Each processing unit is reduced to the function of a



(a) Configuration for basic data processing system.



(b) Distributed files.

Fig. 3

dedicated intelligent terminal. Moving to the next level of sophistication the functional units are stored in object form on secondary storage. A request from a terminal connected to a processing unit causes the appropriate functional unit to be moved from secondary storage, first into the database processing unit and then to the requesting terminal processing unit. Since any functional unit interacting with an operator may be selected from any terminal processing unit the overall reliability of such a system is increased compared to single processor equivalent. In a star-connected configuration the main node can be a compound processing unit.

It is not necessary to have the data files located in a central disk store. Secondary storage devices may be connected to any processing unit. Furthermore a central database may be partitioned and access to it is controlled byseveral communicating processing units (Fig. 3b). The system software for handling the disk storage devices is an order of magnitude more complicated than that handling access to simpler peripheral devices. Not only must it maintain the structure of the data file access methods but also resolve conflicts of simultaneous access to data by several users. Internal functional units may be executed in any processing unit. In the basic systems however they are allocated to a specific unit which can be either the database processing unit or a further unit connected to it.

6 Further Developments

In order to provide a specific system, in terms of its operational function, on a non-specific hardware configuration, greater definition of the functional units is necessary together with ability of the system software to organize functional units. Each functional unit must contain a description of the processing environment necessary for its execution. Since each processing unit has a knowledge of its configuration then the compatiblity between functional and processing units is possible. If the decomposition of the function is taken to one or more lower levels then a given operation function may be considered to be represented by several elementary functional units. For instance the data entry operation may be decomposed to input of raw data, data validation, pre-processing and update of a file. Of these only the elementary functional unit concerned with the entry of raw data need be executed on a terminal processing unit. The data entry functional unit is therefore composed of four elementary processing units. Work is in progress to investigate the possible defintion of a set of primitive functional units. A given functional unit will therefore be represented as a combination of primitive functional units.

A request for a specific system operation will require a 'working set' of functional units allocated to appropriate processors. The breakdown of the operation into functional units is made prior to the system being loaded. The system software, using the functional unit descriptors, is responsible for correct allocation to processing units in the correct sequence to satisfy the overall processing task. The ability to perform this kind of processing depends very much on the decomposition of the task into functional units. Although a number of guidelines may be drawn up for the purpose, a unique decomposition is not possible. A great deal more work is necessary to formalize the procedures.

7 Concluding Remarks

Current trends in hardware devices resulting from l.s.i. technology have made possible the investigation of new approaches to computational architectures. In particular the system user is concerned with reliability and extendability of his computing facility. Reliability is concerned with both the hardware operation and the software operation. Hardware reliability can be achieved by redundancy. Software reliability can be greatly improved by reducing the complexity of the system functions. The low cost of current processors allows the move away from the complex and large general-purpose system software.

The investigation forming the subject of this paper is concerned with the decomposition of a system functional specification to enable it to be executed on a number of interconnected microprocessors. The initial phase of the work has examined the decomposition of small and medium business data processing requirements since these areas are best suited to the processing power of currently available microprocessors. New microprocessors have however been announced by the manufacturers with instruction cycle time approaching that of the conventional minicomputer. The availability of these processors will allow the same techniques to be applied in applications where greater processing speed is required.

A particular application of the concepts was given using a small business application on a specific hardware configuration. A generalization of the concepts to include the definition of a set of primitive functional units was discussed. Each functional unit contains a descriptor which allows its processing environment to be realized at execution time. The main complexity in the system software is that concerned with the allocation of functional units to processing units. In a fully dynamic system the allocation is made before the functional operation is excuted. The allocation remains fixed until the operation is complete. During the period of the operation one or more processing units are dedicated to the execution of the operation. No multi-tasking is encountered. Furthermore the overall reliability is increased since the system is still able to function with other parts idle. Extension of the functional capability is effected by the addition of further processing units to accommodate the extra functional units.

8 Acknowledgments

The author wishes to acknowledge the many interesting conversations with Mr. D. L. Wood and Mr. A. Wilkinson both of Hyperon Consulting Engineers Ltd., Hove,

Sussex, who are involved with the development of hardware for systems related to the concepts expressed here.

9 References

1 Kornstein, H., 'Future trends in memories and microcomputers', Symposium on Microprocessors at Work, University of Sussex, England, September 1976.

2 Lee, I., 'LSI microprocessors and microprograms for user oriented machines', 7th Annual Workshop on Microprogramming, Palo Alto, California, 1974.

3 Salzman, R. M., 'The computer terminal industry: a forecast', *Datamation*, **11**, pp. 46–50, November, 1975.

4 Casaglia, G. F., 'Distributed computing systems: a biased review', *Euromicro Newsletter* **4**, No. 2, p. 5, 1976.

5 Lehman, M. M. and Parr, F. N., 'Program evolution and its impact on software engineering', 2nd International Conference on Software Engineering, San Franscico, 1976.

6 Lehman, M. M. and Parr, F. N., 'Program evolution dynamics and its role in software engineering and project management', European Computing Conference on Software System Engineering, 1976.

7 Belady, L. A. and Lehman, M. M., 'The evolution dynamics of large programs', IBM Research Report, Yorktown Heights, New York, 1975.

8 Dijkstra, E. W., 'Notes on Structured Programming', EWD 249, Technical University of Eindhoven, Netherlands, 1969.

9 Simon, H. A., 'The architecture of complexity', *Proc. Am. Phil. Soc.*, **106**, No. 6, pp. 467–82, 1962.

Manuscript first received by the Institution on 1st December 1976 and in final form on 19th July 1977.
(Paper No. 1791/Comp. 182)

Standard Frequency Transmissions—September 1977

(Communication from the National Physical Laboratory).

September 1977	Relative Phase Readings in Microseconds NPL—Station (Readings at 1500 UT)		
	MSF 60 kHz	GBR 16 kHz	Droitwich 200 kHz
1	3·2	4·5	22·3
2	3·1	4·6	21·8
3	3·2	4·5	21·2
4	3·2	4·4	20·3
5	3·2	4·4	19·6
6	3·3	4·3	18·9
7	3·3	4·3	18·2
8	3·3	4·2	17·7
9	3·2	4·2	17·0
10	3·2	4·4	16·6
11	3·2	4·6	16·1
12	3·3	4·4	15·6
13	3·4	4·5	15·2
14	3·3	4·4	15·1
15	3·3	4·4	14·8
16	3·4	4·4	14·8
17	3·3	4·4	14·4
18	3·3	5·1	13·8
19	3·3	5·5	13·0
20	3·3	5·0	11·9
21	2·9	4·7	11·2
22	2·7	5·0	10·2
23	2·8	3·8	9·3
24	2·9	4·3	6·7
25	2·6	4·2	6·1
26	2·6	4·1	5·7
27	2·6	4·5	5·4
28	2·4	4·2	5·1
29	2·4	4·3	4·7
30	2·5	5·6	4·7

Notes: (a) Relative to UTC scale $(UTC_{NPL}-Station) = +10$ at 1500 UT, 1st January 1977.

(b) The convention followed is that a decrease in phase reading represents an increase in frequency.

(c) Phase differences may be converted to frequency differences by using the fact that 1 μs represents a frequency change of 1 part in 10^{11} per day.

UDC 62–52

Dynamic testing of control systems

Professor D. R. TOWILL,
D.Sc., C.Eng., F.I.E.R.E., M.I.Prod.E.*

Based partly on a lecture given to the Institution's Automation and Control Systems Group in London on 4th December 1975.

SUMMARY

The paper reviews the manual, computer-controlled, and computer-based dynamic testing of control systems, and the reasons for the popularity of such tests. It is shown that correlation techniques are necessary in either time or frequency domains in order to reduce uncertainty due to measurement noise. Engineering guidelines for choosing those test measurements, such as spot frequencies and cross-correlation function time delays, which contain the most useful information for system checkout are included, so that the test time necessary to achieve user confidence in the particular product may be considerably reduced. Hardware examples are used throughout to illustrate the approach.

* *Dynamic Analysis Group, Department of Mechanical Engineering and Engineering Production, University of Wales Institute of Science and Technology, King Edward VII Avenue, Cardiff CF1 3NU.*

Principal Symbols

s	Laplace operator
$H(s)$	transfer function of SUT
a_i, b_i	coefficients of s^i in system transfer function numerator and denominator respectively
n, q	order of polynomials in s
$X(t)$	SUT input signal
$Y(t)$	SUT output signal
T_i, T_j	SUT time-constants
ζ_i, ζ_j	SUT damping ratios
ω_{ni}, ω_{nj}	SUT undamped natural frequencies
z_i	SUT zeros
p_j	SUT poles
A_j, ϕ_j	residue terms in SUT transient response
λ	ratio of undamped natural frequencies
T	transpose (in test feature analysis)
T	correlation time
F_i	ith feature used to check-out SUT
$\pm g_i$	check-out gate width set on ith test feature
t_{sp}	time to reach peak of step response of nominal system
t_{Ip}	time to reach peak of impulse response of nominal system
α_j	jth parameter affecting performance of SUT
ω, f	excitation frequency
ϕ	SUT phase lag
$\lvert H(j\omega)\rvert$	SUT amplitude ratio
Φ	correlation function
$n(t)$	measurement noise
τ	time delay used in correlation function
a	peak amplitude of p.n.s. pulse and of sinusoidal stimulus
n	harmonic number
N	$\omega T/2\pi$ (in sinusoidal testing)
N	(2^R-1) (in p.n.s. testing)
S_p	peak step response of SUT
I_p	peak impulse response of SUT
R_p	peak ramp response of SUT
R	number of stages in p.n.s. shift register
Δt	p.n.s. clock period
σ	standard deviation
ρ	spectral density of white noise
r	number of sequences of p.n.s. over which correlation takes place
r	individual spectral line due to p.n.s. excitation
$S(\omega)$	spectral density of signal
A_i	$R_i+\mathrm{j}Q_i$
θ_i	system input
θ_0	system output
ϵ	system error
$\hat{(\cdot)}$	estimate

1 Introduction

Very many systems and sub-systems such as amplifiers, servoactuators, radar trackers, autopilots, compensation networks, and prediction devices are required to respond to some time varying input ('the message'), and to reject other time varying inputs ('disturbances') in order to perform the operational role for which they are intended. For example in an aircraft automatic terrain-following control system, the terrain profile plus set clearance represents the 'message'. 'Disturbances' will arise from such sources as sensor noise, internally generated noise due to discrete signals in the command computer, and wind gusts tending to deflect the aircraft off course. The total system is designed to discriminate adequately between the 'message' and 'disturbances' so as to achieve the desired level of mission effectiveness which in this case is the best balance between the probabilities of crashing into the ground and 'mission abort'. In order to achieve an adequate design for such complex systems, the system designer must represent both message and disturbances by realistic time-varying signals. It is intuitively obvious (and substantiated by many company-confidential system simulation studies) that dynamic tests are needed to rapidly establish the operational status of the system. The purpose of this paper is to review currently used techniques suitable for dynamic testing and to show how these techniques may be used in a practical environment, particularly for production and maintenance purposes. In particular, there is often a need for test-time reduction in order to increase test station productivity. This paper therefore includes proven guidelines for the selection of a reduced number of measurements (or test features, to use pattern recognition terminology) which have proved most useful for shop floor testing. Although the detailed examples chosen to illustrate the paper clearly reflect the author's practical involvement with servomechanisms and similar systems, the results are much more widely applicable as referenced throughout.

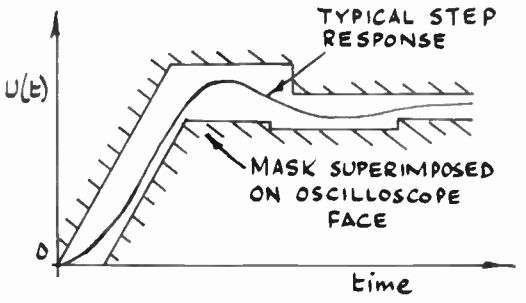


Fig. 1. Manual check-out of SUT via step response.

Although there is a swing towards computer-controlled automatic test equipment (ATE) implementation of dynamic testing because of the advantage of speed at which test data becomes available, there are plenty of situations in which it is acceptable to have dynamic testing undertaken either in the manual or the built-in-test mode of operation. Indeed, for a given system, the philosophy used may vary according to the purpose of the test. This is illustrated in Table 1, which lists five different test modes used in the maintenance of a large electro-mechanical system.[1] Note that the system control computer is involved in two of these tests. There are two other instances in which computers appear as an integral part of the test scene.[2] Firstly, a computer may be dedicated to the automation of test instruments, and secondly the computer, supplied with appropriate interface, may replace instruments in addition to automating the test. These themes will be developed later in the paper.

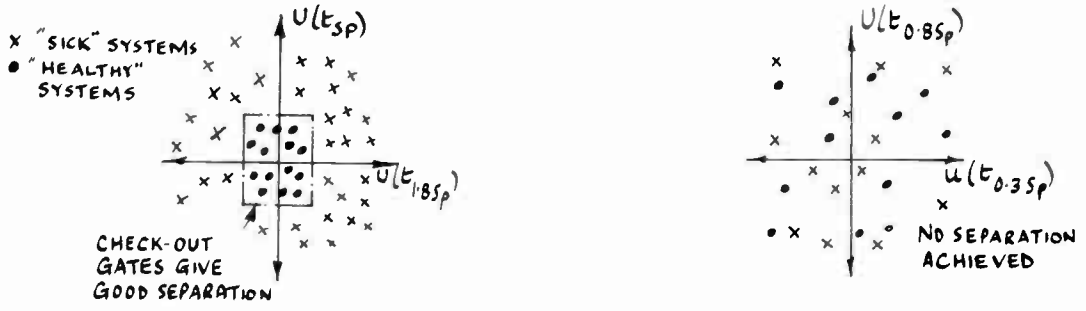
2 What is Dynamic Testing?

The dynamic response of the system under test is defined as the behaviour of the system when stimulated by a time-varying input such as the unit step or one of the many

Table 1

Current maintenance test philosophies for a large electro-mechanical system[1]

Test philosophy	Test Stimulus	Display	Function of Test	Frequency of Test
Automatic test equipment (ATE)	Direct voltages etc.	Lamp and numerical displays for go/no-go; with 3 tolerance levels	Routine maintenance	Daily
Manually-controlled built-in test equipment (BITE)	Step, ramp, harmonic injection to rate and position loops	Pen records of system error	Routine maintenance	Bi-monthly
System computer control	Fixed level sine wave	Lamp display for go/no-go	Confidence immediately prior to use	As required
System computer control	Simulated real inputs	Pen records of system error	'Fingerprints' performance definition and system comparison	On commissioning before/after refit
Manually controlled (mechanical)	Steady torques	Displacement dial gauge	Quality assurance in manufacture and special maintenance	During build, before/after refit



(a) Good features.

(b) Bad features.

Fig. 2. Selection of test features for check-out confidence.

alternative signals which will be considered in a later section. Consequently, a dynamic test is any test which yields information on the dynamic response of the system-under-test (SUT), even if the data yielded do not completely describe the dynamic behaviour of the system. Thus if only the final value of the response of the SUT is measured, then the test would be classified as a static test only, whereas if the behaviour of the SUT is continuously observed or sampled at various times during the transient response, then the test would be classified as a dynamic test.

Figure 1 shows the step response of a servomechanism as recorded on an oscilloscope. Superimposed is a check-out 'mask' which the test technician uses to categorize the system into 'healthy' or 'sick' status according to whether or not the response crosses the boundary, 'healthy' being the pattern recognition term for a system which is operational. In automatic tests, the response of the servomechanism is sampled at a few discrete points in time,[3] and a judgment made by comparison with a set of check-out gates, or by post-test data processing and subsequent reference to a decision surface, using, for example, the nearest neighbour rule. The choice of test features (in this example the features are the delay times t_1, t_2, etc. at which the step response is sampled) is crucial in arriving at a high level of correct classification. In the language of pattern recognition theory, we are seeking those test features which readily discriminate between 'sick' and 'healthy' systems, as shown in the two-dimensional case of Fig 2(a), which give unambiguous boundaries, and wish to discard test features such as those shown in Fig. 2(b) which only serve to confuse the status of the SUT because there is no clear segregation of those systems which are truly operational from those which are not. In Fig. 2, terms such as $u(t_{sp})$ means the value of the step response of the SUT sampled at the time delay where the nominal system step response is at the first maximum.

Fortunately, in engineering, the SUT is created by a known designer, as distinct from the human SUT of medicine, so that mathematical and functional models exist prior to test schedule design. Hence provided the necessary analysis is undertaken at the system design stage by the design authority it is relatively straightforward to select adequate test features by simulating the response of samples of 'sick' and 'healthy' systems Critical regions of the response can then be identified, as shown in Fig. 3, and competitive feature sets compared on a statistical basis. It is also clear from Fig. 3 that the static test only tells us that the system is *operating*, not whether or not the system is *operational*, i.e. will function in the real life operational role.

Ideally, the four gate widths and test features shown in Fig. 3 need confirmation from field trials or simulation studies before final commital within the test schedule. One technique for initial gate width selection is to assign realistic tolerances to all parameters, and then compute the expected boundaries of performance variation of 'healthy' systems using sensitivity functions and assuming the parameter independence rule to hold.[4] Hence if α_j is the jth parameter, then the gate width $\pm g_i$ set on the ith measurement is given by

$$\pm g_i = \left\{ \sum_{j=1}^{j=J} \left(\frac{\partial F_i}{\partial \alpha_j / \alpha_j} \right)^2 \left(\frac{\Delta \alpha_j}{\alpha_j} \right)^2 \right\}^{1/2} \qquad (1)$$

where

$$\left(\frac{\partial F_i}{\partial \alpha_j / \alpha_j} \right)$$

are sensitivity functions relating the rate of change of a measurement to the parameter causing the change, and $(\Delta \alpha_j / \alpha_j)$ is the percentage tolerance set on the jth parameter, there being J parameters in all which are considered by the system designer to affect system performance. Sensitivity functions may be determined

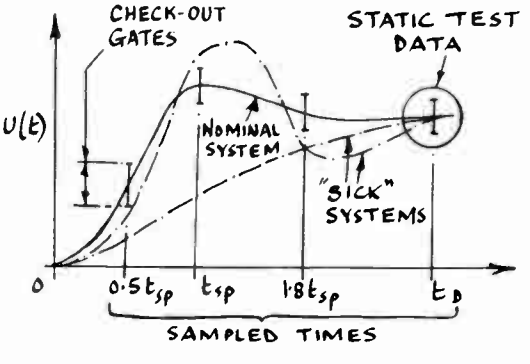


Fig. 3. Sampling SUT step response for automatic check-out.

analytically, via simulation, or via breadboard testing, and in addition to their use in setting 'gates' on acceptable performance, can indicate areas in which parameter tolerances need adjustment. An alternative and popular approach to the above method of gate setting is to use Monte Carlo simulation to predict performance boundaries and to base gate widths on the resulting envelope. It is also common for major system design authorities to ask for such a Monte Carlo simulation from sub-system vendors as part of the design certification requirements.

3 Why Use Dynamic Testing?

The main reasons for using dynamic testing are as follows:

(i) *User Confidence.* Very many systems are designed to respond to real time varying input signals, so that a high degree of correlation exists between operational success and the setting of suitable dynamic performance specifications. For example, in analysing the integrity of aircraft all-weather landing systems required to achieve better safety standards than one accident in 10^7 landings, it is possible to allocate a fraction of the permissible touch-down error to the dynamic response of the system.[5] After extensive simulator studies, it is then possible to determine a frequency response for the system which satisfies the 'false abort' case due to guidance signal noise, and to check that this design produces suitable dynamic recovery of the aircraft when disturbed by a wind gust. A dynamic performance specification for the all-weather landing system means that dynamic tests are essential for all the major sub-systems such as servo-actuators, flight control computers etc.

(ii) *Spares Inventory Reduction.* If only static performance tests are used to establish the status of a system which has to meet a dynamic operational capability in order to confirm the integrity of the system, it is found that static tolerances have to be made excessively tight. Consequently many 'healthy' systems are wrongly categorized as 'sick', resulting in the setting up of an excessively large spares inventory in order to provide a reasonable level of system availability.[6]

(iii) *Repair Costs Reduction.* It follows from the previous paragraph that the repair load on the maintenance depot will be reduced because a smaller proportion of 'healthy' SUTs will be wrongly sent back for stripdown and repair. In addition, for SUTs correctly classified as 'sick', although beyond the scope of this paper, it is possible to infer from dynamic test results the likely causes of failure,[7,8] thus reducing fault location time and hence repair costs.

(iv) *Increased System Reliability via Component Reduction.* Dynamic tests can be designed to reduce the need for intermediate access points used to inject or monitor signals needed in static test schemes. Since it is argued that provision of access points in itself degrades system reliability, their omission will be beneficial in this respect.[9]

(v) *Increased System Reliability via Failure Prediction.* In the integrity analysis of the *Concorde* autopilot expected failures may reasonably be categorized with equal probabilities of 'catastrophic' changes and 'drift' changes in performance.[10] Although as yet there is no direct evidence that dynamic testing may help in predicting impending 'catastrophic' changes in performance, it has been suggested that prediction of gradual degradation via regular testing and time series analysis is feasible[11] so that if suitable documentation is kept, potentially 'sick' SUTs may be removed prior to failure and restored to a satisfactory condition.

(vi) *Performance Matching for Selective Assembly.* Improved matching during selective assembly is possible if dynamic performance data are made available for individual sub-systems. As a hardware example, the use of frequency response plots to match the correct value of tuning resistor to an individual amplifier in order to obtain the desired performance from an R-C coupled oscillator has been detailed elsewhere.[12]

(vii) *Control System Tuning.* Reasonable component tolerances must be permitted if systems are to be produced at realistic costs. It is therefore often necessary to 'fine tune' a complex control system on final assembly especially those with mechanical or hydraulic resonances so as to obtain the best possible performance for that specific system. Dynamic testing greatly assists this fine tuning procedure.

(viii) *Design Proving.* At the system development stage, extensive dynamic testing is required to verify the various transfer function models used during the design phase. This will necessitate the use of appropriate identification and parameter estimation packages with the dynamic test data as input. At this stage, it is also necessary to establish the likely variation in these mathematical models which may be expected during normal operation, and any stability limitations inherent in the design.

Of the eight preceding reasons for using dynamic testing, it is the last, the design proving phase, that has received most attention in the literature. In general, this is the application which most needs and justifies extensive testing in order to provide a mass of data, and will frequently, but not necessarily be reserved for the testing of a few pre-production prototypes. In contrast, the other reasons for dynamic testing generally apply equally to every relevant item manufactured, i.e. the high volume end of the market with an added emphasis on fast-turn round to achieve high productivity from production and maintenance test stations. The test engineer is therefore expected to select the test method most suited to the economic procurement of the absolute minimum data necessary to achieve user confidence. By this time, the system and possible failure modes are better understood, so that this experience can be utilized in test time

reduction, so that only a few measurements will suffice. This problem, although of universal importance, has received far less attention in the literature, and the aim of this paper is therefore to provide guidelines to assist the test engineer in measurement selection.

4 Transfer Function Techniques

For linear systems, a transfer function model of the form,

$$H(s)=\frac{\sum_{i=0}^{i=q} b_i s^i}{\sum_{i=0}^{i=n} a_i s^i} \qquad (2)$$

where s is the Laplace operator, defines the response of the SUT for all possible inputs so may be used as a unifying factor in studying dynamic test techniques. For a test stimulus $X(t)$ with Laplace transform $X(s)$, the response of the SUT may be written

$$Y(t)=\mathscr{L}^{-1}X(s)H(s). \qquad (3)$$

Of particular interest in pre-flight dynamic testing are the following standard responses:

(a) the unit impulse response (or weighting function) defined by

$$h(t)=\mathscr{L}^{-1}\,H(s)$$

(b) the previously met unit step response, defined by

$$u(t)=\mathscr{L}^{-1}\frac{H(s)}{s}$$

and (c) the unit ramp response, defined by,

$$r(t)=\mathscr{L}^{-1}\frac{H(s)}{s^2}.$$

Of these responses, the step response is the one most commonly met mainly because of the ease of signal generation and the engineer's intuitive understanding of displays. Responses are readily available in standard form for a wide variety of transfer functions.[13] It should be noted that if the system is mildly non-linear, which is usually the case with present-day design and manufacturing skill levels, the transfer function model may still be a satisfactory representation of the SUT under the stipulated test conditions, but care is then needed in interpretation of the results to other test domains.[14]

Direct impulse testing of hardware systems is rarely favoured due to the problems of signal generation and excessive disturbance of the SUT, but as we shall see later, the impulse response may be estimated indirectly via pseudo-noise sequence injection and output-input cross-correlation which overcomes these difficulties. Ramp testing is frequently used in tracking systems since the ramp function may be regarded as similar to operational inputs often experienced during at least part of the mission. The steady state ramp error occurring after the 'lock-on' phase is then of particular interest.

Although in theory the impulse, step, and ramp responses contain the same information on system performance, in practice the extraction of the information can be made difficult by an unsatisfactory choice of test stimulus. This is particularly true if the steady state ramp error is inferred from the SUT step response, since any integration inaccuracies will affect the final estimate of a relatively small quantity which is often regarded as a figure of merit of fundamental importance. When estimated from direct observation of the ramp response, the steady state error is, of course, obtained from a single measurement. A less obvious, but equally useful observation on test stimulus selection concerns the amplification of SUT secondary resonances via impulse testing (there will be applications where secondary resonances due to drive mechanisms, structural deflections etc. will need to be assessed as part of the dynamic test, and other applications where this need not be done). To see why this is so, we write equation (2) in SUT pole-zero form:

$$H(s)=\frac{K\prod_{i=1}(1+T_i s)\prod^{i=q}\left(1+\frac{2\zeta_i s}{\omega_{ni}}+\frac{s^2}{\omega_{ni}^2}\right)}{\prod_{=1}(1+T_j s)\prod^{j=n}\left(1+\frac{2\zeta_j s}{\omega_{nj}}+\frac{s^2}{\omega_{nj}^2}\right)} \qquad (4)$$

where $\prod$ is the product sign. We need not commit ourselves at this stage on the specific breakdown between real and complex factors in equation (4), so the intermediate limits are left open.

Using Laplace Transform or the classical technique, the solution for $Y(t)$ may now be written in terms of the residues A_i (of which ϕ_j is a constituent part providing the second arbitrary constant for a complex mode). The SUT time domain behaviour is therefore

$$Y(t)=\sum_{j=1} A_j\exp(-s/T_j)+\sum^{j=n} A_j\exp(-\zeta_j\,\omega_{nj}t)\times \times\sin(\omega_{nj}\sqrt{[1-\zeta_j^2]}t+\phi_j)+ +\text{steady state terms which depend on the test input, } X(t). \qquad (5)$$

The residues obviously depend on the SUT poles and zeros, but also on $X(t)$, which is under the control of the test designer, and it is their dependence on $X(t)$ which can be turned to advantage in system testing.

As an example, consider the case of a fourth-order SUT, in which there are two complex modes of damping ratio ζ_1 and ζ_2 respectively, ζ_2 typically being lightly damped. In the s plane, the frequency separation is $(\omega_{n2}/\omega_{n1})=\lambda$, where $\lambda\gg 1$ by definition of a secondary resonance. If A_1 and A_2 are the residues at the system poles, it has been shown that attenuation effects dependent on the choice of $X(t)$ can be readily estimated from the scalar parts of equation (5) alone, so that if A_1 and A_2 are the scalar factors in the residues, the attenuation of the secondary modes is given by[15]

$$\frac{|A_1/A_2|_{u(t)}}{|A_1/A_2|_{h(t)}}=\frac{1}{\lambda}. \qquad (6)$$

The ramp residues are similarly attenuated with respect to the step response by a factor λ. Consequently, the secondary mode, even when lightly damped, is heavily attenuated each time the test stimulus is integrated, as shown in Fig. 4 for which $\lambda = 5$. Therefore one advantage of the impulse-like test is the exposure to view of the higher resonances so that the system designer is forced to consider their implication (if any) on operational performance. Although the theory behind equation (6) is based on a fourth-order transfer function, these attenuation effects are present irrespective of the order of the system, as confirmed in the testing of a high-order hardware electro-hydraulic servomechanism typically used in copying machines, numerically-controlled machine tools, and for the precise movement of control surfaces.

5 Selection of Check-out Features for Time Domain Testing

It is clear from Section 4 that high-frequency modes are generally severely attenuated in the system step response, so that the decision to implement such a dynamic test implies that secondary modes are (a) not present, (b) are unimportant, or (c) are checked via a separate test. As a consequence, the selection of suitable test features based on sampled values of the step response is simplified, so that even if a simulation study is implemented to find a 'best' set of features as suggested in Section 2, it should be possible to start with a near optimum solution.

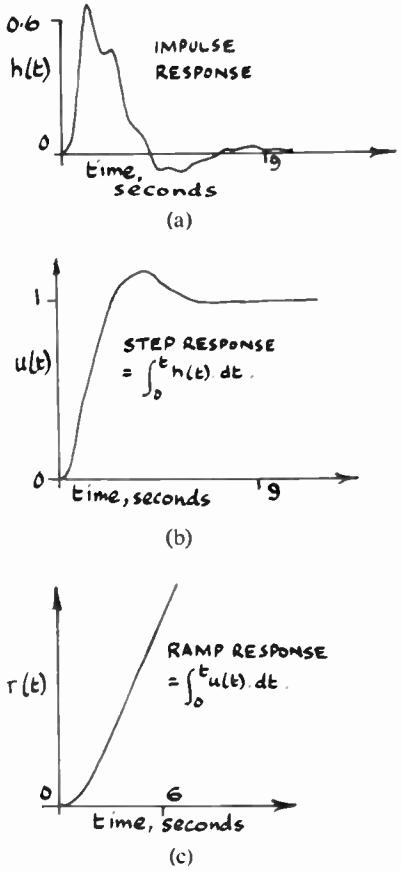


Fig. 4. Attenuation of secondary mode of S.U.T. by integration.

Based on a selection of studies on various SUTs the present author suggests that four sample times should prove satisfactory, since these time delays are generally sensitive (as a set) to important parameter changes. These are shown in Fig. 3.

$$F_i = [u(0{\cdot}5t_{sp});\ u(t_{sp});\ u(1{\cdot}8t_{sp});\ u(t_D)]^T \qquad (7)$$

where t_{sp} is the time to first peak overshoot of the nominal SUT, so that $u(t_{sp})$ is the SUT step response at time t_{sp} and so on, and t_D is the practical SUT decay time; F_i is shorthand notation for the feature vector (set of measurements) used in check-out decision making. Even if the test designer is not operating under constraints imposed by system test time or computer capacity, the discrete feature vector of equation (7) should not be made needlessly long, since there is a danger of factors important to the operational efficiency of the SUT being submerged in a wealth of unimportant and redundant detail.

If high-frequency secondary oscillatory modes are not observable on the impulse response, feature selection for check-out is as straightforward as for the step response, and in like manner a feature vector with four elements will often prove adequate. It has been found that a reasonable set of measurements for initial investigation is

$$F_i = [h(t_{IP});\ h(2t_{IP});\ h(3t_{IP});\ h(t_D)]^T \qquad (8)$$

where t_{IP} is the time to impulse response peak of the nominal SUT. When high-frequency lightly-damped secondary modes are observable such as in Fig. 4(a) the problem becomes much more complicated because acceptable changes in the dominant mode can render gates set to constrain the secondary mode ineffective. A possible solution is to supplement the feature vector of equation (8) with a further three features chosen close together and within the first observable period of secondary oscillation.

It should be noted that unless there are many dynamic performance requirements written into the SUT operational specification, it is extremely unlikely that the number of test features needed for satisfactory check-out will approach the $(n+q+1)$ minimum needed for transfer function identification because the information on whether or not the system is operational is contained in just a few measurements. The needs of testing for operational check-out and design proving are therefore significantly different and this fact is already exploited by test designers since it is known that many existing test schedules call for as few as 4 data points as suggested above.

6 System Frequency Response

If a stable linear system with transfer function $H(s)$ is excited by a sinusoidal signal $a \sin \omega t$, it is well known that after an initial transient phase, the system will settle down to a steady sinusoidal response of the same frequency as the input. In general there will be a phase

shift ϕ and the output waveform will be of a different amplitude to the forcing function, so that in the absence of measurement noise the output signal may be written as $ka\sin(\omega t+\phi)$. It is also readily shown that the solution of the differential equation describing the steady state behaviour is obtained by substituting $j\omega$ for s in equation (2). This gives a rotating vector

$$H(j\omega)=|H(j\omega)|\exp(j\phi) \qquad (9)$$

which after writing in the form $[(A+jB)/(C+jD)]$ can be put in standard polar notation as

$$\left.\begin{aligned}|H(j\omega)| &= \sqrt{\left(\frac{A^2+B^2}{C^2+D^2}\right)}\\ \phi &= \tan^{-1}(B/A)-\tan^{-1}(D/C)\end{aligned}\right\} \qquad (10)$$

Equation (10) gives the information available from the steady state response to a single sinusoidal input, ϕ being the aforementioned phase shift and $k=|H(j\omega)|$ being the system amplitude ratio. Because both phase and amplitude ratio are available we have two test features per test frequency. It is important to make use of both features if test time needs to be reduced. In calculating the theoretical system frequency response, discrete values of ω are substituted into $H(j\omega)$ and the results plotted as a function of ω.

If an individual sine wave are injected into the SUT, one amplitude ratio and one phase estimate is made available, so that if the test is repeated, a series of discrete points may be plotted in exactly the same manner as for the calculated values. This is the serial method of test, in which we wait until the system has settled at each test frequency before a measurement is made. Total test time is therefore some non-linear function of settling time multiplied by the number of test frequencies needed. At the design and development stage of a system, a wide ranging frequency response plot such as shown in Fig. 5 is essential for proving the design, and for finger-printing purposes. Test time for such wideband information can be considerably reduced still using an essentially serial mode technique in which the sinusoidal input has a slowly time varying frequency of excitation,[16] and this has proved an extremely useful method in the past even when using such crude displays of the return signal as chart recorders.

The slow sweep method is based on the fact that by careful choice of frequency sweep characteristics, the envelope of the system output, when plotted as a function of time, approximates to the system amplitude ratio, and the phase shift may be recovered as well. With such a slow sweep technique a satisfactory test time for a second-order system with natural frequency ω_n radians per second would be of the order of $100/\omega_n$ seconds. The slow sweep frequency may be obtained using special purpose instruments, or from a computer-controlled Fourier Response Analyser (FRA) of the type to be described later.

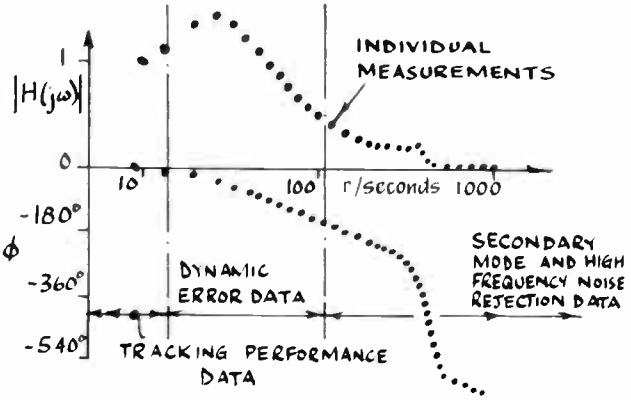


Fig. 5. Frequency response of SUT showing 'partitioning' effect of frequency domain.

7 Time/Frequency Domain Interchange

An interesting feature of the frequency domain is the ready separation of system response into regions which correlate well with operational performance. For example, in the response shown in Fig. 5, steady state tracking errors correlate with low frequency behaviour, dynamic errors correlate with the mid-band behaviour, and secondary resonance quenching plus high frequency noise rejection correlate with behaviour at frequencies well beyond bandwidth. This correlation also leads to simple and effective design techniques based on shaping the $|\theta/\epsilon|$ plot.[17,18]

Time domain and frequency domain behaviour of a system are of course formally related via the Fourier Transform. However, except in simple cases, the mathematical complexity of the Fourier Transform does not permit a great deal of light to be shed on the important engineering guidelines to be adopted when transforming performance specifications from one domain to another, which is more readily done by simulation of transfer function models if sufficient data are available.[19] Nevertheless certain simple conclusions emerge as suggested previously, and extensive empirical rules have been developed to aid the system designer,[20–22] so that frequency domain design methods may be used to synthesize systems to meet time domain specifications. Fortunately in system check-out of a known system, we can be more precise since the family of 'healthy' systems is sufficiently constrained to allow us to use small perturbation theory to relate any time domain performance criteria to any particular frequency domain criteria. Figure 6 shows how changes in the ramp peak error (R_p) are related to the three frequency domain criteria M_p, ω_p and ω_B for a family of electro-hydraulic servomechanisms. The numerical equation applicable to this particular family is

$$10^3\Delta\hat{R}_p=-15{\cdot}1\Delta M_p-0{\cdot}526\Delta\omega_B+0{\cdot}566\Delta\omega_p. \qquad (11)$$

Equation (11) can also be rewritten in terms of performance criterion variances, thus permitting performance envelopes to be established in any desired test

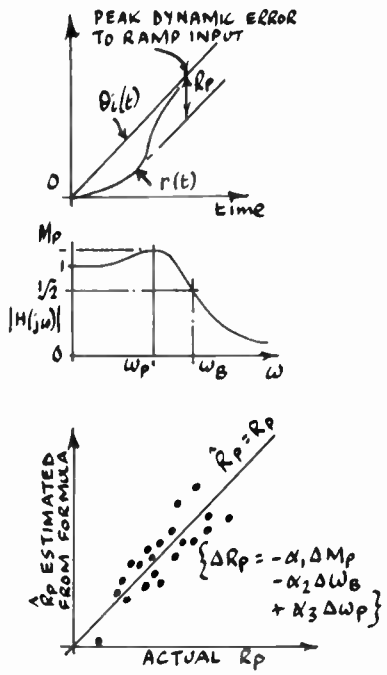


Fig. 6. An example of time domain prediction from frequency domain criteria.

domain. Such predicted envelopes have been found to agree well with envelopes based on measurements for a batch of hardware systems,[23] and are thus perfectly satisfactory for setting check-out gates to be used in conjunction with new test equipment differing in concept from obsolete instruments being replaced.

8 Selection of Check-out Features for Frequency Response Test

When a serial mode frequency domain test technique is used in situations where there is limited SUT test time, it becomes extremely important to select a few test frequencies which will yield the necessary confidence in the operational status of the system. Fortunately, as we have already seen from Fig. 5, the frequency domain partitions the operational function of the SUT into three distinct regions. We also make full use of wide-band testing undertaken during SUT development to place confidence in the mathematical structure, or reliable transfer function describing the family of SUTs,[19] to isolate critical areas for inclusion in the test schedule. The present author's experience is as follows.

(a) *Low-frequency region* where the SUT is generally expected to track the input closely. In order to obtain the necessary resolution it is useful in a feedback system to monitor the error signal, ϵ, directly rather than estimate the error from input and output measurements. This may require the provision of access points specifically for this purpose.[24] Essentially, a test in this region is checking the low frequency loop gain, which can usually be inferred from just one measurement of error for a low frequency sinusoidal stimulus at, say, a frequency $\simeq 0{\cdot}10\,\omega_B$. In some systems, the true error cannot be observed unless the SUT is made to respond to a real or simulated operational stimulus so that the loop is properly closed, and all components activated. A partial solution here is to use an available system control computer to generate typical loop signals based on previous operational trials, such as could happen in the system for which Table 1 is applicable, or by the provision of suitable aerodynamic and kinematic simulators based on the dynamic response of the nominal system.[25]

(b) *Mid-frequency region* encompassing peak amplitude ratio (M_p) and bandwidth (ω_b). As we have already seen, this region primarily determines such important time domain performance characteristics as I_p, S_p, and R_p (impulse peak, step peak, and ramp peak error respectively), and the times at which these peaks occur, proving that on a practical basis time and frequency domain test methods are interchangeable. Experimental and theoretical work shows that these time domain criteria can be adequately constrained by measurements made at three test frequencies f_{45}, f_{90}, and f_{135}, i.e. the frequencies at which the *nominal* SUT will have phase lags of 45°, 90°, and 135°. It is not necessary to design the test to select the frequencies at which an *individual* SUT has these phase lags,[26] thus saving equipment development costs and SUT test time. The recommended measurement vector, based on extensive studies, is

$$F_i = \left[\begin{array}{l} |H(j\omega)|_{f_{45}};\ \phi_{f_{45}};\ |H(j\omega)|_{f_{90}};\ \phi_{f_{90}};\ \\ \qquad\qquad |H(j\omega)|_{f_{135}};\ \phi_{f_{135}} \end{array} \right]^T. \quad (12)$$

(c) *High-frequency region* well beyond bandwidth encompassing noise rejection and secondary resonance requirements. In the time domain, the high-frequency noise rejection is compressed into the region around $t=0$, and the secondary modes are superimposed in the manner already seen in Fig. 4, but in the frequency domain the secondary mode of Fig. 5 is well separated from the dominant mode, and can be detected for check-out purposes by three further test frequencies, one at the nominal value of modal frequency, plus one in relatively close proximity either side. High-frequency noise rejection beyond bandwidth can be ascertained from one or two test frequencies depending on the rate of roll-off sought, so that an extensive test schedule for a complex system may appear as shown in Table 2. If the secondary resonance is very lightly damped, or there is more than one important secondary resonance, use of the swept frequency method is indicated.

9 Test Implementation

Dynamic test requirements are shown in schematic form in Fig. 7 and are applicable to both manual and automatic testing.[2] In manual testing, the human operator is essentially responsible for the functions shown in blocks (1) and (5), and the confidence in the test is dependent on the skill of the human operator plus the efficiency of aids such as flow charts which are provided to assist him in his task. Many instruments, such as

Table 2

Complete frequency domain test schedule for check-out of electro-hydraulic servomechanism

Purpose of Test	Constrain Tracking Errors	Constrain Dynamic Errors			Constrain Secondary Resonance			Constrain High Frequency Noise Transmission
Test frequency	0·8 Hz	4 Hz	8 Hz	16 Hz	42·5 Hz	50 Hz	57·5 Hz	100 Hz
Return signal measured	ϵ	θ_0	θ_0	θ_0	θ_0	θ_0	θ_0	θ_0
Amplitude ratio	<0·015	1·20–1·40	0·80–1·20	0·32–0·52	<0·35	<0·35	<0·35	<0·05
Phase lag	Not measured	32·5–57·5°	65–115°	110–160°	205–260°	260–330°	345–395°	>500°

Fourier Response Analysers (FRAs) and Time Domain Analysers (TDAs) are currently available for executing blocks (2) and (3), and standard packages such as hard-wired Fast Fourier Transform (FFT) devices can be procured for performing block (4). Additionally, blocks (1), (2), (3), (4), and (5) can be completely automated via a digital computer provided with suitable interface to the SUT.

It is thus clear that the logistics of test implementation are highly dependent not only on the particular SUT, but also on the maintenance concept adopted. This point is illustrated in Fig. 8, which shows how the choice is influenced by the availability of manpower and prime cost. The cost figures quoted are naturally suspect due to rapid advances in instrumentation technology, but are felt to be indicative of present trends. In particular, there may be cost effective instances where any necessary computation can be undertaken cheaply via a desk calculator.[22]

In Fig. 8, the sophistication in return signal analysis greatly increases when correlation techniques are used. These techniques referred to as 'modern' measurement techniques,[28] are essential in both time and frequency domain test techniques when the return signal is contaminated even with modest measurement noise. The remainder of the paper deals with the exploitation of these correlation techniques.

10 The Fourier Response Analyser

As far back as 1961, serial mode frequency response method of system testing was a recommended test for USAF equipment,[29] and has mushroomed ever since. The method may be implemented in a wide variety of ways, as partially indicated in Fig. 8, including the use of oscillators, phase-variable filters and oscilloscope displays, in the manual test mode, and sampling plus counting techniques in digital-computer-aided test stations. However, the most universally used technique involves correlation, and the frequency domain test instrument designed to exploit the correlation principle has become known as the Fourier Response Analyser (FRA) because gain and phase estimators are identical to the Fourier series coefficients used to describe a repetitive waveform. To emphasize the importance of correlation techniques in both time and frequency domains, the approach adopted in Ref. 28 will be used here.

By definition, the cross-correlation function $\Phi_{XY}(\tau)$ of two signals $X(t)$ and $Y(t)$ is

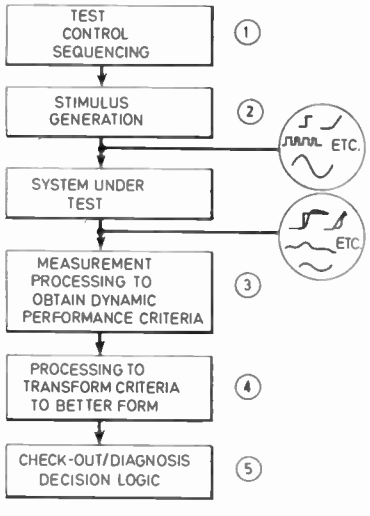


Fig. 7. Computational requirements for dynamic testing.[2]

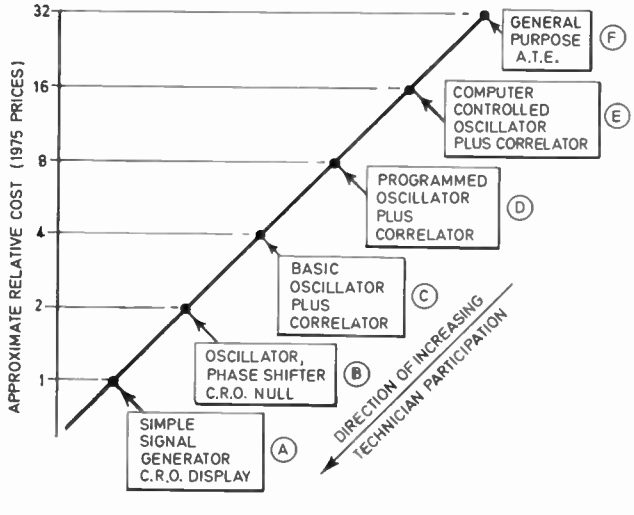


Fig. 8. Spectrum of methods for sequential sinusoidal testing.

$$\Phi_{XY}(\tau)=\lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{+T}X(t-\tau)\,Y(t)\,\mathrm{d}t. \tag{13}$$

For sinusoidal excitation of the SUT we have

$$X_s(t)=a\sin\omega t \tag{14}$$

where X_s means a sinusoidal signal, and subsequently X_c will mean a cosinusoidal signal.

There will be a noisy return signal from the SUT of

$$Y(t)=|H(\mathrm{j}\omega)|\,a\sin(\omega t+\phi)+n(t). \tag{15}$$

We also require the cosine signal,

$$X_c(t)=a\cos\omega t \tag{16}$$

to be generated, which is readily mechanized using a slave oscillator or a time delay unit. It is implicit that the integration time, T, is chosen to be an integer multiple of the input sine wave, so that $T=2\pi N/\omega$. We now correlate the return signal [eqn. (15)] with the sine and cosine signals of equations (14) and (16) respectively and set $\tau=0$. The result is

$$\left.\begin{aligned}\Phi_{X_sY}(0)&=\left\{\frac{|H(\mathrm{j}\omega)|a^2}{2}\right\}\cos\phi+\frac{a}{T}\int_0^T n(t)\sin\omega t\,\mathrm{d}t\\ \Phi_{X_cY}(0)&=\left\{\frac{|H(\mathrm{j}\omega)|a^2}{2}\right\}\sin\phi+\frac{a}{T}\int_0^T n(t)\cos\omega t\,\mathrm{d}t\end{aligned}\right\} \tag{17}$$

for the in-phase and quadrature components respectively.

On the assumption that noise may be neglected, the gain and phase estimators become

$$\left.\begin{aligned}|\hat{H}(\mathrm{j}\omega)|&=\frac{2}{a^2}\{[\Phi_{X_sY}(0)]^2+[\Phi_{X_cY}(0)]^2\}^{1/2}\\ \hat{\phi}&=\tan^{-1}\left[\frac{\Phi_{X_cY}(0)}{\Phi_{X_sY}(0)}\right]\end{aligned}\right\} \tag{18}$$

A block diagram of the FRA principle is shown in Fig. 9. Having been available commercially in analogue form for many years, it is now available from several manufacturers in digital form with computer control capability for use in an automatic test set.[29,30]

The correlation process may be regarded as a filter through which we can observe a cleaned-up version of a noisy return signal, thereby providing reasonable estimates of $|H(\mathrm{j}\omega)|$ and ϕ (Ref. 2). Some specific results for noise rejection obtained in Ref. 28 are as follows:

(a) For sinusoidal noise of frequency ω_N, and amplitude a_n

$$\left.\begin{aligned}\sigma_R&=\frac{a_n\sqrt{2}}{\pi N}\left[\frac{1}{(\omega_N/\omega)^2-1}\right]\sin\pi N\left(\frac{\omega_N}{\omega}\right)\\ \sigma_Q&=\left[\frac{\omega_N}{\omega}\right]\sigma_R\end{aligned}\right\} \tag{19}$$

for in-phase and quadrature standard deviations respectively. In the complex plane, the uncertainty due to sinusoidal noise is therefore an ellipse centred at the true value of amplitude ratio and phase, as shown in Fig. 10(a).

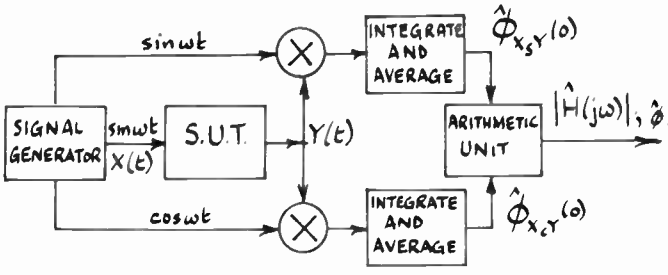


Fig. 9. Mechanization of Fourier Response Analyser.

(b) For white noise of spectral density ρ,

$$\sigma_R=\sigma_Q=\frac{a}{2}\sqrt{\left(\frac{\rho\omega}{\pi N}\right)} \tag{20}$$

so that the uncertainty is now a circle centred at the true value of amplitude ratio and phase, as also shown in Fig. 10(a).

Unfortunately, noise characteristics are rarely known at the time the SUT test schedule is designed, and often vary widely between apparently similar SUTs, so that some flexibility in choice of N is needed. Some practical results obtained using an FRA in a maintenance depot are shown in Fig. 10, for an SUT where the contamination is due to at least two sinusoidal sources plus some wideband noise. The reduction of measurement variance with integration time is self-evident. It can also be seen that the central limit theorem (σ proportional to $1/\sqrt{N}$) is here a conservative estimate of the rate of variance reduction, and furthermore, although correlation greatly reduces uncertainty, in the practical situation, perfect filtering is unlikely to be achieved.

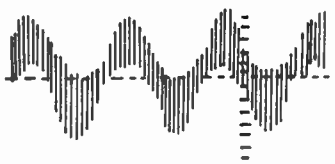
11 Indirect Impulse Testing via P.N.S. Excitation and Cross-Correlation

11.1 Development of the Convolution Integral in Terms of Input Autocorrelation Function

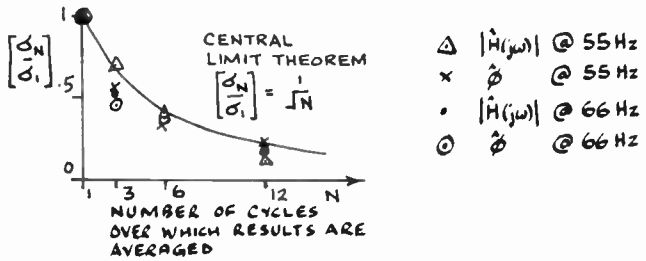
In recent years pulse testing time domain and serial mode frequency domain techniques have been rivalled by the appearance of pseudo-noise test signals, which for *linear* systems, can, via the cross-correlation principle, yield under specified conditions a realistic approximation to the system impulse response without the physical injection of an impulse stimulus.[31,32] The theoretical basis for this work is the convolution integral, so that if $X(t)$ and $Y(t)$ are the SUT input and output signals at time t, and $h(\tau)$ is the SUT impulse response as before, then

$$Y(t)=\int_{-\infty}^{\infty}h(\tau)X(t-\tau)\,\mathrm{d}\tau. \tag{21}$$

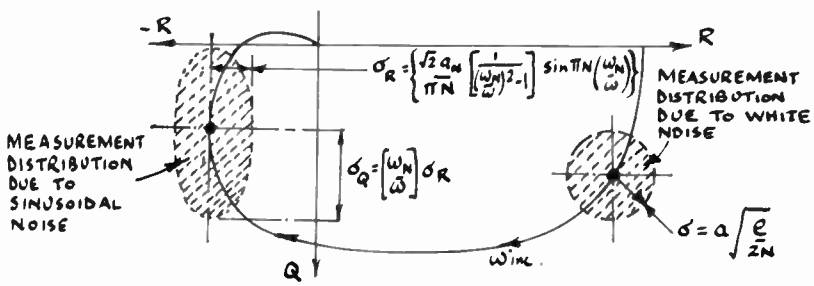
As shown in Ref. 33, equation (21) can be manipulated into an alternative form, which can finally be written as,



(a) Typical return signal to FRA



(b) Reducing measurement spread by averaging over N cycles.



(c) Theoretical variation in FRA measurements due to noise.

Fig. 10. Application of Fourier Response Analyser to noisy SUT

$$\Phi_{XY}(\tau) = \int_0^T h(\tau_1)\Phi_{XX}(\tau-\tau_1)\,d\tau_1 \qquad (22)$$

where $\Phi_{XX}(\tau)$ is the autocorrelation function of the input signal averaged over measurement time T. Equation (22) is called the Wiener–Hopf equation. *If $\Phi_{XX}(\tau)$ is the unit impulse, then $\Phi_{XY}(\tau)=h(t)$ and equation* (22) *is an impulse response estimator for the SUT.* Our search for a test stimulus to obtain the system impulse response is therefore aimed at signals with an autocorrelation function suitably approximating to the unit impulse.

11.2 P.N.S. Characteristics

White noise is one test stimulus which has the unit impulse autocorrelation function required by equation (22) but unfortunately infinite measurement time is implied for satisfactory estimates. As can be seen from Fig. 11 the uncertainty due to the test signal can be removed by using a pseudo-noise sequence (p.n.s.) with precisely defined statistical properties. Two-level sequences (p.r.b.s.), are particularly attractive since they are easily generated by shift registers incorporating the necessary feedback and operating on modulo 2 arithmetic. The resultant test signal and autocorrelation function are shown in Fig. 12 and the schematic mechanization is shown in Fig. 13, the method of test being either serial or parallel mode depending either on the number of delay lines provided, or on the provision of intermediate storage prior to computation. $\Phi_{XX}(\tau)$ is not a perfect impulse, being a function of the clock period ΔT, and the amplitude of the pulse $\pm a$, as follows:[33]

$$\Phi_{XX}(\tau) = \frac{(N+1)}{N}\,a^2\Delta t\,.\,\delta(\tau) - \frac{a^2}{N} \qquad (23)$$

where $\delta(\tau)$ is the Dirac delta function. The 'spike' in equation (23) repeats with a periodicity $N\Delta t$, where $N=(2^R-1)$, R being the length of the generator shift register. $R=10$ is a common length, giving a test sequence periodicity of 1023 Δt but success with a specific

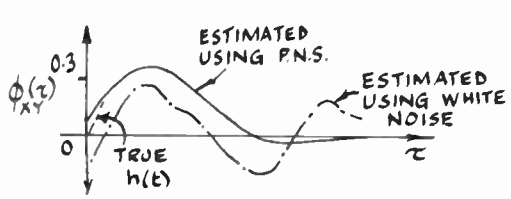
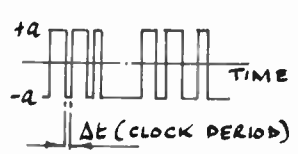
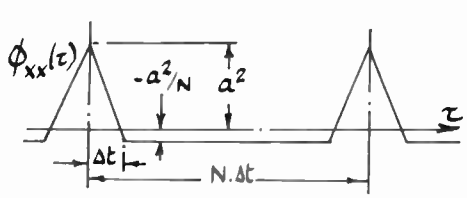


Fig. 11. Removing uncertainty in $\hat{\phi}_{xy}(\tau)$ by using p.n.s. stimulus.



(a) Time sample of p.n.s.



(b) p.n.s. $\phi_{xx}(t)$.

Fig. 12. P.n.s. characteristics.

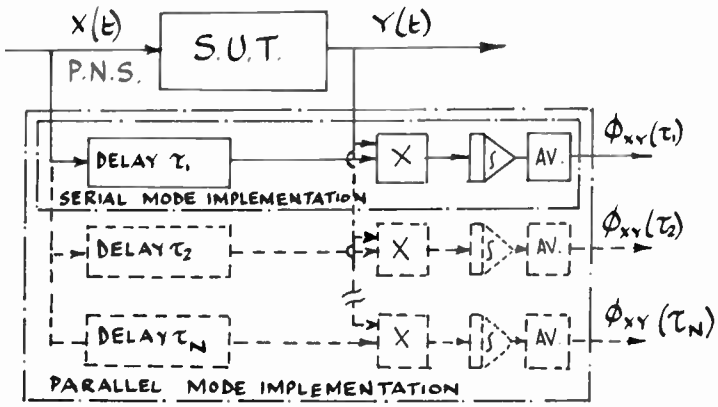


Fig. 13. Implementation of impulse testing via p.n.s. injection and cross-correlation.

range of SUTs has been achieved with R as low as 6, giving $N=63$ bits. As R increases, so is the error in estimation of $h(t)$ due to d.c. offset reduced, and the need for post-test correction avoided. However, this is a minor reason in the choice of N, since post test correction is not difficult, and the problem may be avoided altogether by the use of inverse repeat sequences.

11.3 Matching the P.N.S. to the SUT

If the system delay time is less than $N\Delta t$, the cross-correlation function for the SUT excited by two-level p.n.s. is,[33]

$$\Phi_{XY}(\tau)=\frac{a^2(N+1)\Delta t\, h(\tau)}{N}-\frac{a^2}{N}\int_0^{N\Delta t} h(t)\,\mathrm{d}t \qquad (24)$$

and it is clearly useful to reduce or eliminate the second term on the right-hand side of equation (24) by making N large, or by using inverse repeat sequences. However, in selecting p.n.s. characteristics for system testing, it is useful to place the graphical interpretation of Fig. 14 on equation (22) so that the various sources of error other than d.c. offset may be understood. These errors can be dealt with in three parts.

(a) Equation (22) implies integration over all time: in the practical situation this means that the product of the two functions must be zero outside the time span of integration, which in turn requires the second spike to occur after the response is over. *The p.n.s. sequence length must therefore be somewhat longer than the decay time of the system.*

(b) *The initial value will, in general, be in error*, because the triangular autocorrelation function is *centred* at the origin, so that the product will differ from the true impulse response. Even for a narrow autocorrelation function, there will be an error between the actual and estimated impulse response near the origin as can be seen in Fig. 11. This particular source of error disappears for $\tau>\Delta t$.

(c) *Errors due to the finite width of the autocorrelation function clearly depend on the behaviour of $h(t)$ in any time interval $2\Delta t$.* From Fig. 14 it is clear that distortions can take place in regions of high rates of change of $h(t)$. In particular, oscillations present on the impulse response which have a period comparable with the pulse width are removed by the multiplying and integrating action of convolution as will be seen in Section 11.5. As would also be expected from an information theory approach, Δt must be chosen by considering the highest frequency likely to be present in the SUT impulse response, and the following analysis has proved helpful in making the choice.

In a theoretical study[34] undertaken to detect oscillatory modes, it has been shown that in order to detect the peak to within 1%, the ratio of (modal period/clock period) must be about 20 : 1. As this ratio decreases, the accuracy of estimation falls off rapidly, as shown in Table 3, a ratio of 10 : 1 appearing to be a reasonable compromise choice for good resolution. The existence of secondary modes can, of course, be detected with a much lower ratio of (modal period/clock period) than 10 : 1, as a number of case studies have shown, but the secondary mode is then greatly attenuated compared to the true impulse behaviour.

11.4 Noise Rejection Characteristics

If the corrupting measurement noise is white, then $\Phi_{nn}(t)=\rho\delta(t)$, where ρ is the spectral density, and $\delta(t)$ is the Dirac delta function. It may then be shown[31] that the resulting uncertainty becomes,

$$\sigma=a\sqrt{\left(\frac{\rho}{T}\right)} \qquad (25)$$

Table 3

Accuracy of detection of sine wave amplitude using p.n.s. and cross correlation[34]

resonance period / p.n.s. clock period	estimated peak amplitude / true peak amplitude
10·00	97·5%
3·33	74%
2·00	41%
1·43	14%
1·10	1·5%

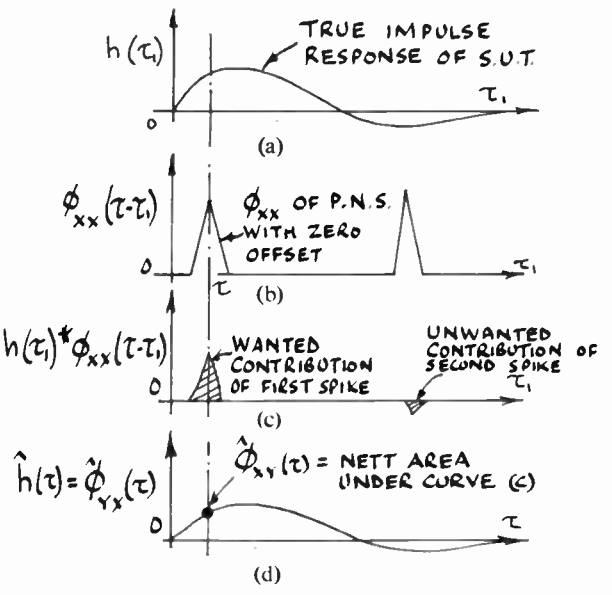
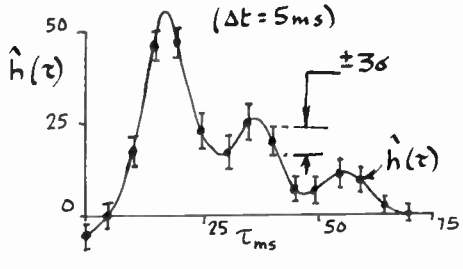
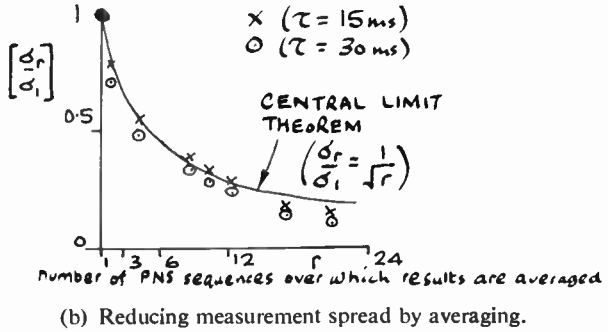


Fig. 14. Graphical interpretation of Weiner–Hopf equation as an aid to matching p.n.s. and SUT

so that the measurement standard deviation is proportional to the signal level $\pm a$, and is inversely proportional to the square root of measurement time divided by spectral density, i.e. the central limit theorem applies. A further interesting conclusion from equation (25) is that the error variance is independent of delay time. Figure 15 shows the cross-correlation function for the SUT previously studied in Fig. 5 and measured in a typical maintenance environment of *a priori* indeterminate noise level and it can be seen that the variance error is reasonably constant and is reduced by approximately $(1/\sqrt{r})$ if correlation takes place over r sequences of p.n.s. Some success in reducing measurement uncertainty in the presence of sinusoidal noise has been reported,[28] the rules of p.n.s. selection basically following those for secondary mode filtering, as will be illustrated in the next section.



(a) Estimated $\hat{h}(\tau)$ and noise spread for SUT excited by p.n.s.



(b) Reducing measurement spread by averaging.

Fig. 15. Indirect estimate of electro-hydraulic servosystem impulse response via p.n.s. injection and correlation.

11.5 Example on P.N.S. Selection

Suppose we wish to identify the SUT of Fig. 4(a) using p.n.s. and cross-correlation to estimate the impulse response. The procedure is as follows:

(a) From the observed decay time of the complete impulse response of 12 seconds, $N\Delta t > 12$.

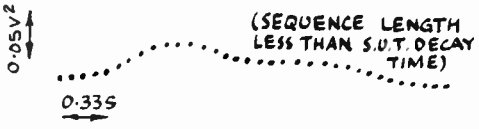
(b) From accuracy considerations in identifying the secondary mode, $\Delta t < 2\pi/(5 \times 10)$, if we choose 10 clock intervals per period of the secondary mode.

If Δt is made equal to 0·05 seconds then N should be > 240, so that $R = 8$ giving $N = 255$ would be satisfactory for accurate identification if the instrumentation is sufficiently flexible to permit such a choice. In fact for the tests shown in Fig. 16, the best practical compromise was found to be $R = 7$ and $\Delta t = 0{\cdot}1$ s. Figure 16 also shows the effect of varying Δt and N, and confirms the chosen p.n.s. parameters.

(a) $R = 7$; $f_c = 10$ Hz; $a = \pm 1$V.

(b) $R = 7$; $f_c = 3$ Hz; $a = \pm 1$V;

(c) $R = 7$; $f_c = 1$ Hz; $a = \pm 1$V.



(d) $R = 5$; $f_c = 10$ Hz; $a = \pm 1$V.

Fig. 16. Practical example of matching p.n.s. to SUT

11.6 Test Time using P.N.S.

For correlation over r sequence lengths, total correlation time in equation (25) is $T=(rN\Delta t)$. It is customary to allow one complete sequence of p.n.s. to 'initialize' the SUT prior to correlation commencing, so the total test time is $\{(r+1)N\Delta t\}$. In the early days of p.n.s. testing via special-to-type instruments, only one delay line was available, so that a test time of $(r+1)N\Delta t$ was required per each point on $\Phi_{XY}(\tau)$. More recent instruments have typically provided 100 delay lines, so that 100 points on $\Phi_{XY}(\tau)$ can be estimated from a test time $(r+1)N\Delta t$ in the so-called 'parallel' mode of Fig. 13. The p.n.s. technique can obviously be implemented directly by digital computer, but care must be taken to design the test schedule so as not to exceed the computer capacity.

12 Spectral Analysis Methods

The use of computers as signal generators, return signal processors, and data transformers (blocks 2, 3, and 4 of Fig. 8) is particularly suited to implementation via spectral analysis techniques which effectively involve testing in the parallel mode, i.e. estimates for gain and phase at many frequencies are obtained simultaneously. The auto spectral density $S_{xx}(\omega)$ is defined by

$$S_{xx}(\omega)=\frac{1}{2\pi}\int_{-\infty}^{\infty}\phi_{xx}(\tau)\exp(-j\omega\tau)\,d\tau \tag{26}$$

and the cross spectral density similarly defined by

$$S_{xy}(\omega)=\frac{1}{2\pi}\int_{-\infty}^{\infty}\phi_{xy}(\tau)\exp(-j\omega\tau)\,d\tau \tag{27}$$

and until recently system frequency response has been determined via these relationships,[35] since

$$\hat{H}(j\omega)=\frac{S_{xy}(\omega)}{S_{xx}(\omega)} \tag{28}$$

is a frequency response estimator. However, it is generally preferable to compute $H(j\omega)$ directly from the Fourier transform of the input and output signals using the relationships,

$$\left.\begin{aligned}X(\omega)&=\int_{-\infty}^{\infty}X(t)\exp(-j\omega t)\,dt\\ Y(\omega)&=\int_{-\infty}^{\infty}Y(t)\exp(-j\omega t)\,dt\end{aligned}\right\} \tag{29}$$

The SUT transfer function is then given by

$$\hat{H}(j\omega)=\frac{G_{yx}(\omega)}{G_{xx}(\omega)}=\frac{Y(\omega)X^*(\omega)}{X(\omega)X^*(\omega)} \tag{30}$$

where * means complex conjugate. In practice, the calculations will be performed on discrete data, and the resultant implications, together with a detailed proof of the equivalence of the direct method with the correlation method are to be found in reference 36. Computer implementation as suggested for testing electronic SUTs is shown in Fig. 17, the Fast Fourier Transform (FFT) being used in view of the enormous reduction in computing effort thereby achieved. Reference 12 presents a number of broad spectrum frequency response results obtained using spectral analysis methods as typified by Fig. 18, although no recommended input stimulus is given in that paper. P.n.s.,[37] white noise[38] and a fast frequency sweep[39] have been used as test signals in spectral analysis.

In order to check the influence of measurement and extraneous noise on spectral analysis estimates, the coherency function,

$$\gamma^2=\frac{|G_{zx}|^2}{G_{xx}G_{zz}}=\frac{G_{yy}}{G_{yy}+G_{nn}} \tag{31}$$

is used, values of $\gamma^2=1$ resulting from tests on a completely noise-free linear system in which case $G_{yx}(\omega)=G_{zx}(\omega)$, which is the quantity actually observed. An example of the transfer function testing of a communication system is shown in Fig. 18. Here the coherency function is seen to decrease at frequencies above 0·7 MHz, but it is not clear whether this is due to an abnormally high noise level in that part of the band or a normal noise level coupled with a low system gain. Determination of $|H(j\omega)|$ then resolves the situation, as also shown in Fig. 18, since this plot clearly separates the possible causes. For such measurements, the signal-to-noise ratio for the SUT is then

$$[\gamma^2/(1-\gamma^2)]$$

which may be used at the SUT development stage as a guide to choosing a suitable test signal spectrum. Details of the accuracy of transfer function estimates using spectral methods are given in reference 40, errors occurring due to the effects of noise, frequency

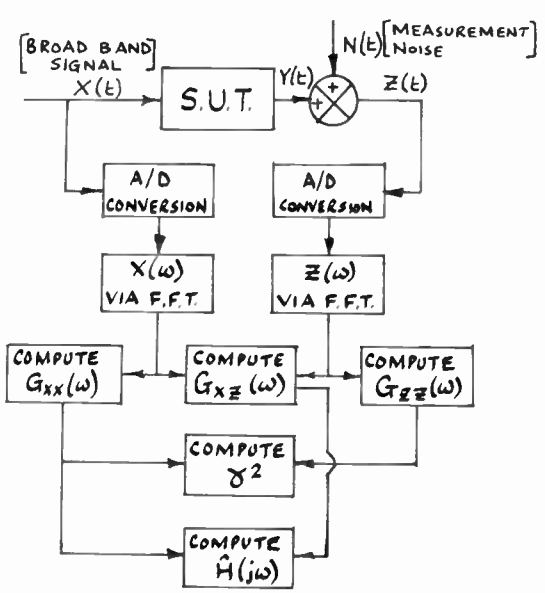
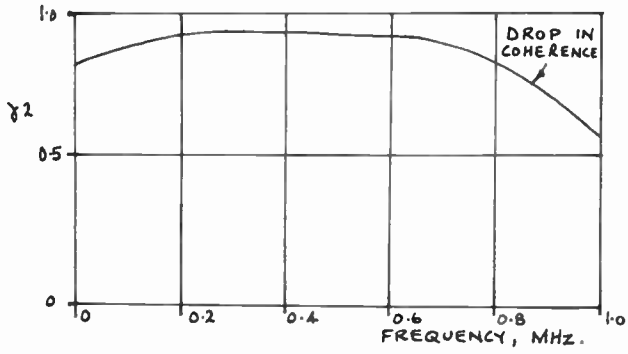
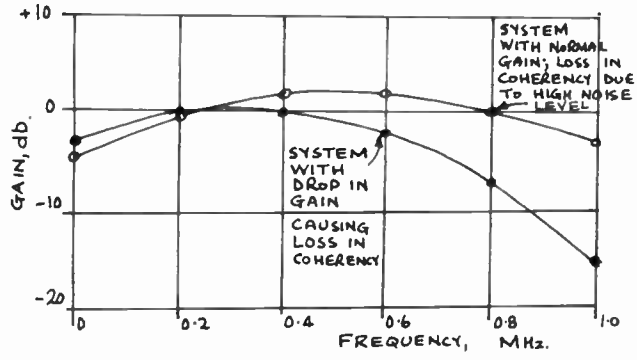


Fig. 17. Block schematic of spectral analysis method.



(a) Coherency.



(b) Two explanations of coherency loss.

Fig. 18. Transfer function testing of communication link.[12]

resolution, and aliasing. When digitizing data prior to transfer function estimation, aliasing errors can only be avoided if the highest frequency present in the test data is less than the Nyquist, or folding frequency, given by the formula (1/twice the sampling interval) in Hz.

13 Rapid Frequency Response Determination Directly from P.N.S. Injection

As can be seen from Fig. 19 p.n.s. stimuli have a precisely defined frequency domain $[(\sin x)/x]^2$ line spectrum with spectral lines occurring at $(1/N\Delta t)$ Hz; $(2/N\Delta t)$ Hz, etc. Nodes occur at integer multiples of the clock frequency $(1/\Delta t)$ Hz, etc. Because the input spectrum is so precisely defined, p.n.s. may be regarded as a parallel-mode frequency stimulus. Although the frequency response could be obtained from Fourier transforming $\Phi_{XY}(\tau)$ in the normal way, if only a few frequency data points are required for check-out as suggested in Section 8, significant reductions in computing time result from taking advantage of the known spectral characteristics, A_{ri} and Φ_{ri} which have been tabulated over the frequency range of interest.[34] The approach is as follows:

At the rth spectral line (at frequency ω_r), the in-phase and quadrature components of the p.n.s. input are

$$\left.\begin{aligned} R_i(\omega_r) &= A_{ri}\cos\phi_{ri} \\ Q_i(\omega_r) &= A_{ri}\sin\phi_{ri} \end{aligned}\right\} \tag{32}$$

where A_{ri} and ϕ_{ri} are known *a priori*. If the return signal

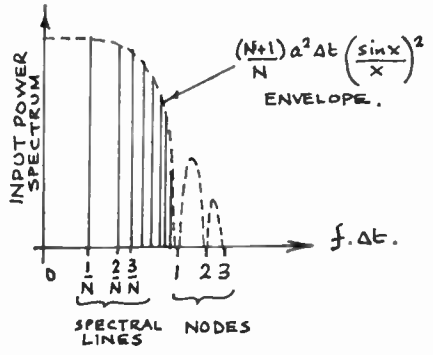


Fig. 19. Defined power spectrum of p.n.s. stimulus.

is correlated with $\sin\omega_{ri}$ and $\cos\omega_{ri}$ in turn, as outlined in Section 10, then the in-phase and quadrature components of the SUT output are estimated as

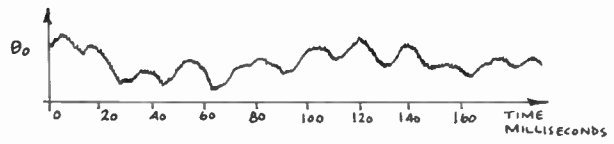
$$\left.\begin{aligned} R_0(\omega_r) &= A_{r0}\cos\phi_{r0} \\ Q_0(\omega_r) &= A_{r0}\sin\phi_{r0} \end{aligned}\right\} \tag{33}$$

the required measurement data $|\hat{H}(j\omega_r)|$ and $\hat{\phi}_{\omega_r}$ can now be evaluated from:

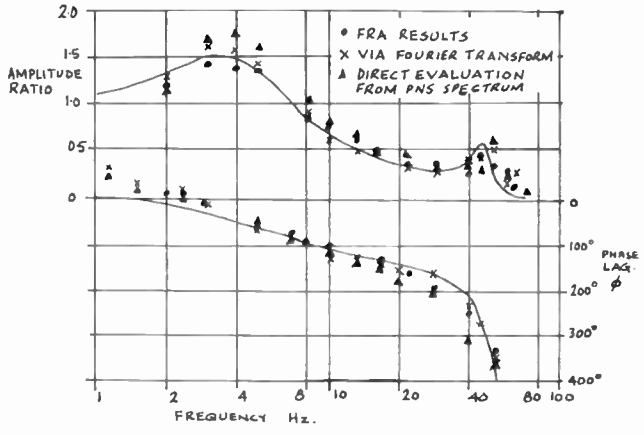
$$|\hat{H}(j\omega_r)| = \frac{A_{r0}}{A_{ri}} = \sqrt{\left(\frac{R_0^2+Q_0^2}{R_i^2+Q_i^2}\right)} \tag{34}$$

$$\hat{\phi}_{\omega_r} = [\phi_{r0}-\phi_{ri}] = [\tan^{-1}(Q_0/R_0) - \tan^{-1}(Q_i/R_i)]. \tag{35}$$

As an example of the benefit obtained using this approach, if the FFT algorithm is incorporated, a reduction in computation time of 30 : 1 is estimated if only three frequencies are required compared to the method of obtaining the complete cross-correlation function $\Phi_{XY}(\tau)$ first, in a simulation on a noise free, linear, system. As a result of the tremendous reduction in computational effort the method is suited to the needs of automatic test stations. Figure 20 compares results on a noisy hardware SUT for the three different methods of determining frequency response discussed in this apper. It can be seen that the direct p.n.s. method is not as accurate under these circumstances as the Fourier Transform method, but the computational simplicity can outweigh this disadvantage. The return signal from the SUT is also shown in Fig. 20, and it is clear that there is no impulse-like peak in the return signal, thus confirming that p.n.s. testing is a relatively gentle way of obtaining dynamic test data. The reduction in system peak excitation levels using p.n.s. varies enormously with the specific application. For the SUT of Fig. 20, the peak motion of the output due to p.n.s. excitation alone is about 10% of the peak output motion expected for a unit impulse input. Fourier Transform methods of estimating transfer functions need not require a conventional digital computer. An instrument is commercially available, and has already been used experimentally on the SUT of Fig. 20.[41]



(a) Response of SUT to p.n.s. with 400 Hz clock speed.



(b) SUT frequency response.

Fig. 20. Transfer function testing of electro-hydraulic servo system.[34]

14 Conclusions

Dynamic testing is now a universal method of assessing the operational status of a wide variety of systems ranging from amplifiers at one end of the spectrum to complete aircraft autoland systems at the other. The advent of the FFT algorithm, coupled with the ready availability of digital computers, has had a considerable effect on the implementation of dynamic test techniques. It cannot be emphasized too strongly that it is the dynamic test data which are fundamental and the method of obtaining these is secondary to the objective of selecting those test features which adequately discriminate between 'sick' and 'healthy' systems. It is far better to undertake a manually-controlled dynamic test of very simple form rather than to have no dynamic test at all. It is hoped that this paper has adequately reviewed the basic guide lines to be adopted in dynamic test stimulus and measurement feature selection.

15 Acknowledgments

Much of the work described in this paper has been supported by the Science Research Council, the Ministry of Defence (Procurement Executive) and UK Industry. This support is gratefully acknowledged as is the opportunity to participate in such a highly stimulating activity.

16 References

1 Harley, H. V. and Towill, D. R., 'Automation of large electromechanical system testing', IERE Conf. Proc. No. 30 'Advances in Automatic Testing Technology', 1975, pp. 169–84.

2 Towill, D. R., 'Dynamic test techniques utilised in ATE', *REME Journal*, No. 22, pp. 35–43, 1972.

3 Levadi, V. S., 'Pattern recognition applied to fault detection', Proc. IEEE Convention on Military Electronics, Washington D.C., pp. 197–202, Sept. 1965.

4 Towill, D. R. and Mehdi, Z., 'Prediction of the transient response of high order linear systems using low order models', *Trans. Inst. Meas. Control*, 3, No. 1, pp. T1–T9, 1970.

5 Gill, F. R., 'The Integrity of a Civil Blind Landing System with Particular Reference to the Azimuth Channel', RAE Tech. Report No. 65022, Feb. 1965.

6 Parker, A. H. and Stewart, A. H., 'Automatic test systems for production and maintenance', Symposium on ATE, Proc. Convention on Electronics in the 1970s, pp. 283–311 (IERE Conference Proceedings No. 12, July 1968).

7 Sriyananda, H. and Towill, D. R., 'Fault diagnosis using time domain measurements', *The Radio and Electronic Engineer*, 43, No. 9, pp. 523–33, September 1973.

8 Sriyananda, H., Towill, D. R. and Williams, J. H., 'Voting techniques for fault diagnosis from frequency domain test data', *IEEE Trans. on Reliability*, R-24, No. 4, pp. 260–7, 1975.

9 Garcia, R. F., 'Fault Isolation Computer Methods', NASA Contractor Report CR-1758, 1971.

10 Fielding, W. F. (Transl.), 'Flight Control System for the Concorde Supersonic Civil Transport Aircraft', RAE Lib. Trans. 1615, 1972.

11 Allen, R. J., 'Failure prediction employing continuous monitoring techniques', *IEEE Trans. on Aerospace*, AS-1, No. 2, pp. 924–30, August 1963.

12 McCarthy, J. P. and LaClair, P. W., 'Transfer function testing', IEEE ASSC Record, pp. 1–5, 1972.

13 Towill, D. R., 'Transfer Function Techniques for Control Engineers' (Iliffe, London, 1970).

14 Towill, D. R. and Lamb, J. D., 'Pseudo-random signals test non-linear controls', *Control Engineering*, 17, No. 11, pp. 59–63, 1970.

15 Towill, D. R., Cooper, J. S. and Lamb, J. D., 'Dynamic analysis of fourth order feedback control systems', *Int. J. Control*, 10, No. 2, pp. 121–46, 1969.

16 Reed, W. H., Hall, A. W. and Barker, L. F., 'Analogue Techniques for Measuring the Frequency Response of Linear Physical Systems Excited by Frequency Sweep Inputs', NASA Report No. IND-508, 1960.

17 Winston, G. C., 'Design of a precise control system', *Control Engineering*, 17, pp. 84–9, March 1970.

18 Towill, D. R., 'Coefficient plane models for tracking system design', *The Radio and Electronic Engineer*, 45, No. 9, pp. pp. 465–71, September 1975.

19 Payne, P. A., Towill, D. R. and Baker, K. J., 'Predicting servomechanism dynamic performance variation from limited production test data', *The Radio and Electronic Engineer*, 40, No. 6, pp. 275–88, December 1970.

20 Jaworsky, Z. E., 'Empirical transient formulae', *Electronic Engineering*, 26, pp. 396–400, September 1954.

21 Biernson, G. A., 'A general technique for approximating transient response from frequency response asymptotes', *Trans. Am. Inst. Elect. Engrs (Applications and Industry)*, 75, Pt. II, pp. 253–73, 1956.

22 Axelby, G. S., 'Synthesis of feedback control systems with a minimum lead for a specified performance', *IRE Trans. on Automatic Control*, PGAC-1, pp. 56–73, May 1956.

23 Brown, J. M., Towill, D. R. and Payne, P. A., 'Predicting servomechanism dynamic errors from frequency response measurements', *The Radio and Electronic Engineer*, 42, No. 1, pp. 7–20, January 1972.

24 Hayes, K. A., 'The Necessity for Test Points in Servomechanisms', Ministry of Technology Report REMC/28/FR, May 1969, p. 346.

25 Lamb, J. D., Pankhurst, A. R. and Towill, D. R., 'Correlation techniques applied to the dynamic testing of aircraft autoland systems', Proc. 7th Int. Aerospace Symposium, Cranfield, pp. 23.1–23.7, 1972.

26 Towill, D. R. and Payne, P. A., 'Frequency domain approach to automatic testing of control systems', *The Radio and Electronic Engineer*, **41**, pp. 51–60, January 1971.

27 Youngberg, R., 'Team calculators with instruments', *Electronic Design*, **24**, pp. 176–9, 22nd November 1974.

28 Lamb, J. D. and Payne, P. A., 'Noise rejection properties of modern measurement techniques in control engineering', Proc. 5th Asilomar Conf. on Circuits and Systems, pp. 548–54, 1971.

29 Gibson, J. E., Rekasius, Z. V., McVey, E. S., Sridhar, S. and Leedham, C. D., 'A set of standard specifications for linear automatic control systems', *Trans. Am. Inst. Elect. Engrs (Applications and Industry)*, **80**, pp. 65–77, 1961.

30 Elsden, C. S. and Ley, A. J., 'A digital transfer function analyser based on pulse rate techniques', *Automatica*, **5**, No. 1, pp. 51–60, 1969.

31 Hughes, M. T. G. and Noton, A. R. M., 'Measurement of control system characteristics by means of a cross-correlator', *Proc. Instn Elect. Engrs*, **109**, Pt B, No. 43, pp. 77–83, 1962.

32 Graupe, D., 'Identification of Systems', chap. 4 (Van Nostrand Reinhold, New York, 1972).

33 Davies, W. D. T., 'Random signal testing for evaluating system dynamics', *Wireless World*, **17**, No. 8, pp. 407–12, August 1966.

34 Lamb, J. D., 'Use of Pseudo-Random Binary Signals for the Production Testing of Dynamic Systems', Ph.D. Thesis, UWIST, 1970.

35 Driefke, G. E., Hougen, J. P. and Mesmer, G., 'Effects of Truncation on time to frequency domain conversion', *ISA Transactions*, pp. 353–68, October 1962.

36 Newland, D. E., 'An Introduction to Spectral Analysis', chap. 10 (Longmans, London, 1975).

37 Polk, D. R., 'On-line checkout of linear dynamic systems by identification', Proc. 15th Joint Automatic Control Conference, pp. 389–397, 1974.

38 McCalister, D. A. and Bickford, L. L., 'Frequency response determination using white noise excitation', *J. Spacecraft Rockets*, **5**, No. 7, pp. 873–4, July 1968.

39 Skingle, C. W., 'A Method for Analysing the Response of a Resonant System to a Rapid Sweep Input', RAE Technical Report 66379, December 1966.

40 Hughes, M. T. G. and Wellstead, P. E., 'Accuracy of Transfer Function Estimates Derived from Numerical Fourier Transforms', IEE Conf. Publication No. 106 on Digital Instrumentation, November 1973, pp. 166–73.

41 Payne, P. A., 'Dynamic Testing of a Rate Servomechanism Using Various Test Instruments', UWIST Technical Note DAG 51, September 1972.

Manuscript first received by the Institution on 16th August 1976 and in revised form on 20th June 1977.
(Paper No. 1792/ACS 19)

UDC 681.325.5

The implications of microprocessor architecture on speed, programming and memory size

BRIAN K. PENNEY,

B.Sc., Ph.D., D.I.C., M.B.C.S.*

Based on a paper presented at the Conference on Computer Systems and Technology held at the University of Sussex from 29th to 31st March 1977.

SUMMARY

This paper gives the results of benchmark tests of several different microprocessors. The results are given in terms of speed of execution, size of source program (number of statements) and of memory requirements for the compiled programs. An analysis of these results is given in terms of the microprocessor architectures.

* *Department of Computing and Control, Imperial College of Science and Technology, 180 Queen's Gate, London SW7 2BZ.*

1 Introduction

The advent of microprocessors with their very low cost offers the potential of low-cost computer and computer control systems. So low is their cost (commonly between £10 and £100 in unit quantities) that the comparative proportions of the various cost factors such as programming, memory and engineering costs, have become significantly different than those of the previous generation of the smallest computers, mini-computers. Also, there is a wider diversity of architecture and of performance amongst the 50 or so microprocessors currently marketed than there was for minicomputers.

These factors make it difficult to assess the merits of the various different microprocessors available. One approach, previously used for minicomputers†, is to give a figure of merit for each of several architectural and technological factors (such as typical instruction execution times) and to combine these figures to derive an overall merit factor. As microprocessor architecture is both different to, and more diverse than, minicomputer architecture, this approach is of questionable merit. A different approach has been adopted in which programs were written to perform several standard tasks for a wide range of microprocessor architectures. The programs were characterized by three factors; these were (a) the number of source language statements, (b) the size in bytes of the compiled program, and (c) the speed of execution. The first of these largely determines the cost of programming. The second factor, amount of memory required, is of significance, as in any practical system the cost of the memory is likely to be many times the cost of the microprocessor itself and will be the dominant element in the total hardware cost. The third factor, program execution time, is a much better measure of the power of a processor than the cycle time or the execution time of any particular instruction. For a given semiconductor process technology (p-m.o.s., n-m.o.s. etc.) the cycle times encountered are, in practice, very nearly equal, while it will be shown that programs for different microprocessors have a very wide range of execution times, much wider than can be accounted for by differences in cycle times or individual instruction times.

It will be shown that this approach identifies very clearly the microprocessors which are the most effective (according to the criteria just described) as well as some architectural factors which contribute to the effectiveness.

2 The Tests

Several tasks were devised as being representative of the type of job for which a microprocessor might be used. Corresponding programs were then written in the appropriate assembler language for each of several processors and the size of code and speed of execution were calculated using the manufacturers' published data. The range of tasks selected includes 8-bit addition, 16-bit

† 'Minicomputers and Europe', Report of the Electrical Research Association, Leatherhead, Surrey, 1972.

addition, 8-bit multiplication, and two output routines, one for programmed output and the other for interrupt controlled output.

2.1 Eight-bit Addition

This is a common operation and occurs in microprocessors in several applications such as I/O handling and character manipulation. It is assumed that the operands and result are in arbitrary memory locations rather than in registers. As all processors examined have an addition instruction, this test is dominated by the addressing operations required to access the operands and results. In writing the code it was assumed that the operands represented either signed or unsigned integers depending on which was easier to handle on the microprocessor being considered.

2.2 Sixteen-bit Addition

This operation also occurs often, especially in arithmetic work. Processors with an 8-bit (or less) arithmetic logic unit may have to do this as a sequence of smaller operations. Similar assumptions were made as for 8-bit addition.

2.3 Eight-bit Multiplication

The majority of microprocessors, especially those with an 8-bit (or less) a.l.u., do not have multiply instructions, and multiplication must be performed using an algorithm based on repeated addition and shifting. This is an excellent overall test of the data manipulation operations and is not much influenced by the addressability of the main store as most of the intermediate operands and results are kept in registers.

2.4 Program-controlled Output

A major problem with all computers is moving information between the computer and its I/O devices. With microprocessors the two major methods used are program-controlled I/O and interrupt-controlled I/O. In contrast, larger computers use interrupt-controlled I/O and d.m.a., but rarely use program-controlled I/O. To test the basic ability of microprocessors to perform I/O two simple tasks were devised.

For programmed I/O it is assumed that:

The processor will be dedicated to the task of transferring data between the main store and an I/O device.

The I/O is sufficiently fast that the limiting speed is determined by the processor rather than by the I/O device.

A fixed number of characters is to be output.

The number of characters to be output is sufficiently large that initialization overheads may be ignored.

Figure 1 shows a flow chart of the program. Note that only the central loop has been coded and that initialization is ignored.

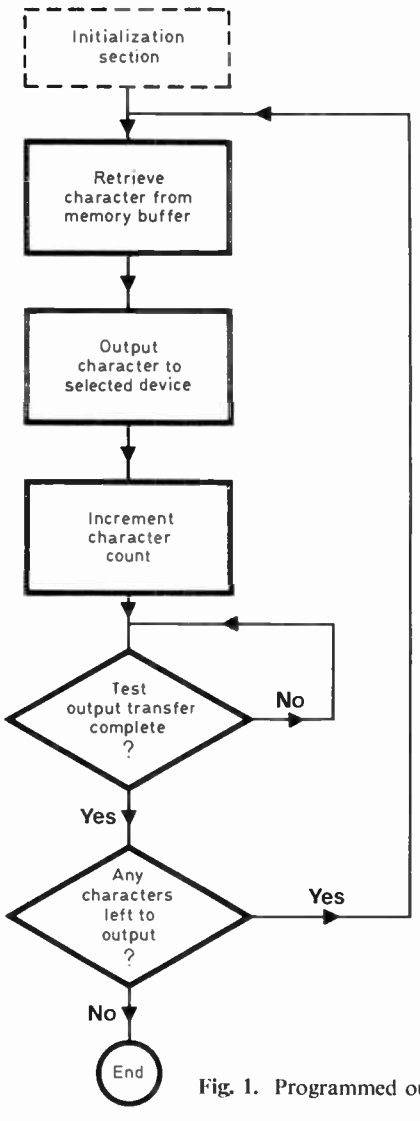


Fig. 1. Programmed output.

2.5 Interrupt-controlled Input/Output

Similar assumptions to those for program-controlled output were made together with the following:

It is assumed that another, independent, program would be running at the time that interrupts occur, and that consequently on interrupt certain registers would need to be saved.

Other I/O devices are also in use. Where vectored interrupts are not used it is necessary to poll the I/O devices to identify the source of the interrupt.

It is assumed that there are four I/O devices and that on average the third device polled will be the one requesting service.

Figure 2 shows the flow chart for the two cases of vectored and non-vectored interrupts. In a real application it would probably be necessary to switch buffers at the end of output but for this test it was assumed that there is only one buffer and that initialization overheads are negligible. The code included all the device identification code and register saving necessary for a system

Table 1. Basic parameters of microprocessors assessed

	Process technology	A.l.u. size (bits)	Memory width (bits)	Instruction length (bits)	Number of instructions	Stack	Registers	Instruction cycle time (μs)	Add time (μs)	Interrupts
Intel 4040	p-m.o.s.	4	8	4, 8, 12	60	7×12	24×4+Acc	10·8	10·8	1
Fairchild F8	n-m.o.s.	8	8	8, 16	48	none	64×8+Acc	2	2–5	many
Intel 8080A	n-m.o.s.	8	8	8, 16, 24	78	external	6×8+Acc	2	2	1†
Motorola MC6800	n-m.o.s.	8	8	8, 16, 24	72	external	Index (1×16)+Acc (2×8)	1	4	1
N.S.C. SC/MP	p-m.o.s.	8	8	8, 16	46	external	3×16 Base+Acc+Extension	2	38	1
Intersil IM6100	c-m.o.s.	12	12	12	40+	none	Acc (1×12)+MQ (1×12)	0·5	2·5	1†
G.I. CP1600	n-m.o.s.	16‡	10 or 16	10, 20, 30	87	external	6×16	0·4	4	1†
N.S.C. PACE	p-m.o.s.	16	16	16	45	10×16	4×16	2	8	6
T.I. 9900	n-m.o.s.	16	16	16, 32	69	external	16×16 (multiple copies)	0·33	7	16

† Extendable to multi-level with external logic. ‡ Organized internally as byte serial.

Table 2 Results of microprocessor benchmark tests

	8-Bit Addition			16-Bit Addition			8-Bit Multiplication			Output (Programmed)			Output (Interrupt)		
	N.S.	Bytes	Time (μs)	N.S.	Bytes	Time (μs)	N.S.	Bytes	Time (μs)	N.S.	Bytes	Time (μs)	N.S.	Bytes	Time (μs)
Intel 4040	14	19	324	14	19	561	—	—	—	14	19	124[a]	26	32	294[a]
Fairchild F8	6	12	51	20	27	94	40	49	514	8	10	40	19	23	87
Intel 8080A	4	10	28	5	11	38	20	35	374	7	14	32	17	25	102
Motorola MC6800	3	9	13	6	18	26	12	24	206	10	21	41	33	53	98
N.S.C. SC/MP	4	7	120	7	13	230	27	40	2306	6	12	188	56	64	1060
Intersil IM6100	4	4[b]	10	9	9[c]	23	44	44[b]	500	7	7[b]	22[b]	10	10	35[b]
G.I. CP1600	8	28	28	3	12	12	18	52	182	5	16	16[d]	23	74	95[d]
N.S.C. PACE	3	6	24	3	6	24	15	30	546	7	14	62[e]	14	28	136[e]
T.I.9900	3	12	22	3	12	22	3	12	39	5	4	39[f]	4	10	29[f]
Average 8-bit	4·25	9·5	53	9·5	17·25	97	25	37	850	7·75	14·25	75·25	29	41·25	334
Average 16-bit	4·66	15·3	24·66	3	10	19·33	12	31·33	255	5·66	11·33	39	17	37·33	87

N.S. = number of program statements.

(a) 4-bit characters.
(b) 12-bit characters.
(c) 24-bit addition.
(d) 16-bit words.
(e) either 8- or 16-bit units, wired option.
(f) either 8- or 16-bit units, program option.

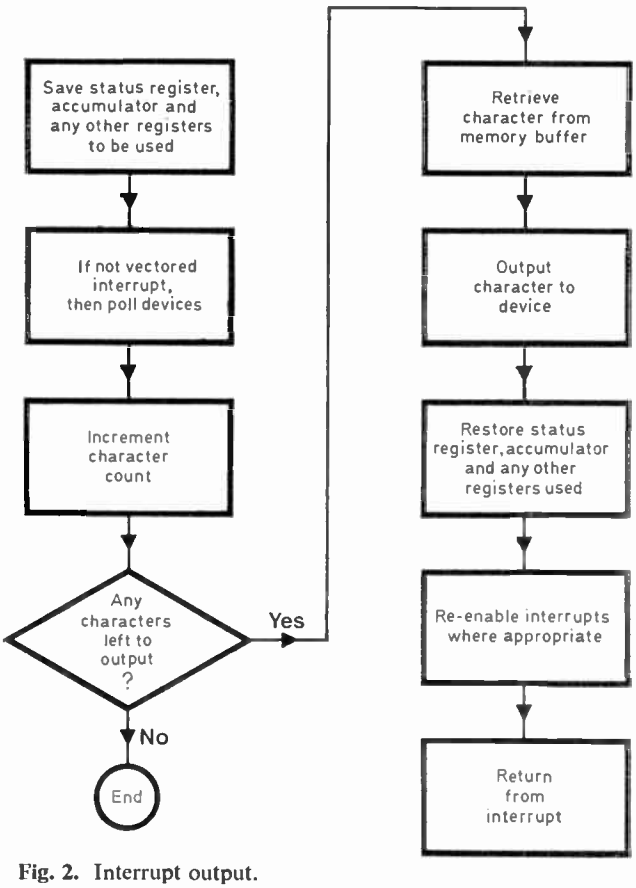


Fig. 2. Interrupt output.

with four devices, but only one device service routine was included.

3 Benchmark Programs

In any programming exercise there are a number of design compromises to be considered, the major ones being speed vs. size vs. simplicity. It was decided to optimize simplicity and to achieve some balance (necessarily subjective) between speed and size. Clever tricks were avoided. For example, on one processor the stack pointer could be used to speed arithmetic programming at the expense of disabling the interrupt system but such a technique was considered unnecessary and undesirable.

There is a wide variation in processor and system architecture and where slight changes to the bench mark standards could significantly improve its performance the appropriate changes were made.

Several processors had architectures which limited the range of addresses accessible without reloading pointer, page or index registers to a small value. In these cases it was assumed that all operands would lie inside that range. In real situations this ideal would not be reached and overheads would be incurred in resetting the appropriate registers.

Nine different processors spanning the word length/cost/performance range were assessed. Table 1 shows the processors assessed and some of their basic parameters. In all, 44 programs were coded in the appropriate assembler languages for the 9 different processors. The 45th program (8-bit multiplication for the Intel 4040) was abandoned as being too difficult to write.

4 Variation from the Benchmark Standards and Assumptions Used

4.1 Intel 4040

The 8-bit multiplication was found to be too difficult to code in the time available.

In the output programs it was assumed that:

(i) Output would be as a string of 4-bit 'nibbles' instead of 8-bit bytes.

(ii) The output buffer would be contained in a 20-word data r.a.m. memory, and that initialization overheads would be negligible.

(iii) In the interrupt service routines no registers would need to be saved. In the 4040 the registers are arranged as three banks with each bank containing eight 4-bit words. Two of these banks share the same address space and a select instruction is used to access the required bank. With this arrangement it is natural to dedicate one bank to the interrupt service routine and one to the program level routine. This would only be permissible in the very smallest of real systems. A full register save and restore would take 2400 μs.

Both output routines assumed use of a 16 character buffer, corresponding to one register file in a data r.a.m. chip. In both cases the initialization overheads would in practice be quite heavy.

In any practical system, therefore, the service times would be considerably longer than those in Table 2.

4.2 Intel 8080A

It was assumed in the interrupt output program that there was external logic to supply the service routine address and hence form a vectored interrupt system.

4.3 Fairchild F8

This processor has a large register file on the c.p.u. chip; there are 64 registers of which 48 are used as a general-purpose set and the rest have special properties. As with the Intel 4040 it was assumed that a set of registers could be dedicated to the interrupt service routine. For a truly general-purpose system it would be necessary to save the entire register set on interrupt which would take approximately 500 μs.

4.4 Intersil IM6100

As this is a 12-bit processor (in fact a PDP8/E equivalent), the tests were redefined to apply to 12- and 24-bit operands instead of 8- and 16-bits. The output programs were coded assuming use of a PIE interface chip.

5 Results

All of the programs were coded. None of these was run; instead, the performance was calculated using the manufacturers' data for instruction execution times.

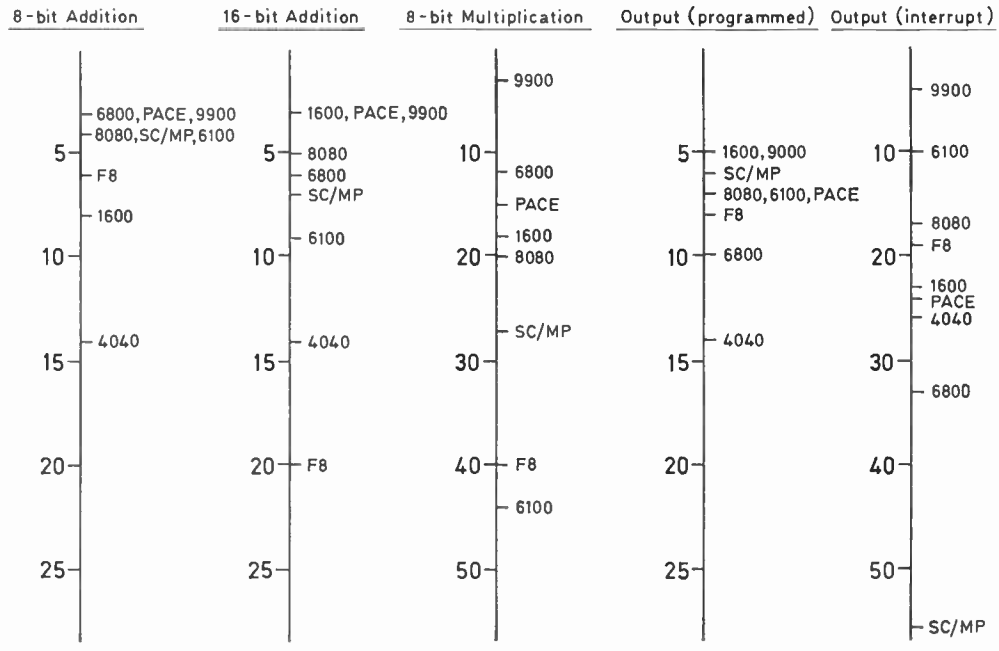


Fig. 3. Number of assembler statements in benchmark programs.

Where faster versions of the microprocessors exist (such as the Intel 8080A-1), use of the standard version was assumed. Table 2 tabulates the results and Figs. 3, 4 and 5 show the results graphically. Table 3 presents the results as averages for 4-bit, 8-bit and 16-bit processors. The 4-bit average is in fact an average of one result. The Table also shows the ratios of average 8-bit to average 16-bit results.

5.1 Number of Instructions

Here there is a strong relationship with word length. The Intel 4040, the only 4-bit processor examined, had on average about three times as many statements as did the 16-bit processors, and the 8-bit processors averaged about 80% more statements than the 16-bit processors.

The cost of writing a program is directly related to the number of program statements and, in practice, costs

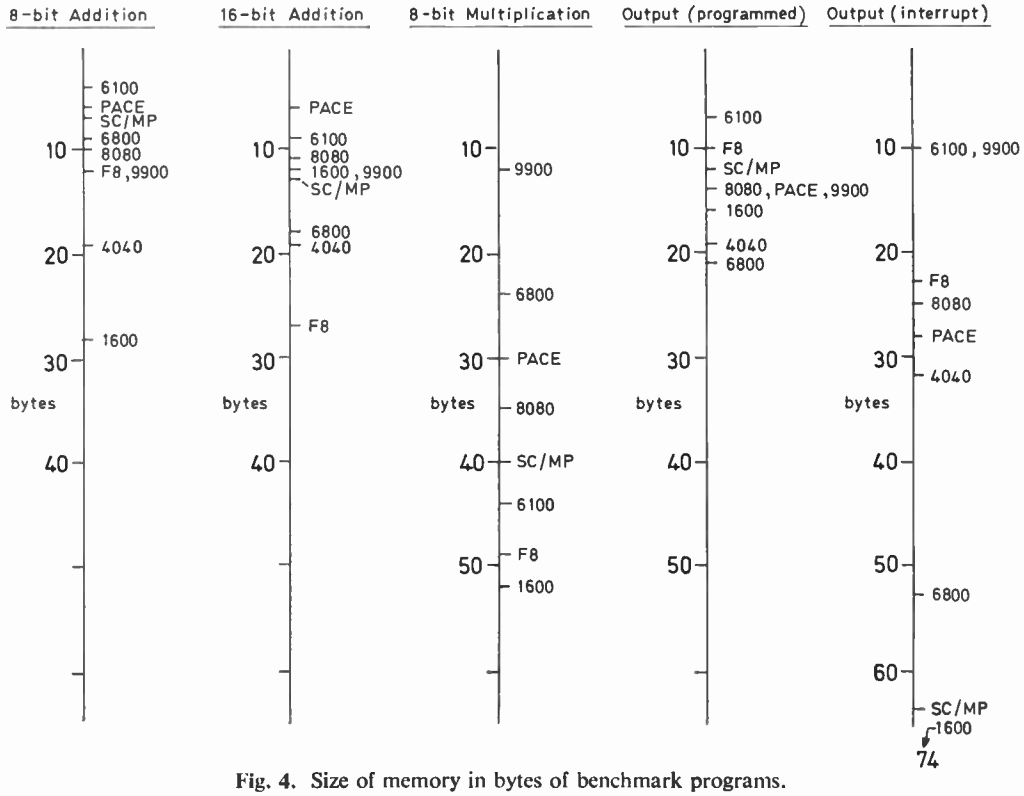


Fig. 4. Size of memory in bytes of benchmark programs.

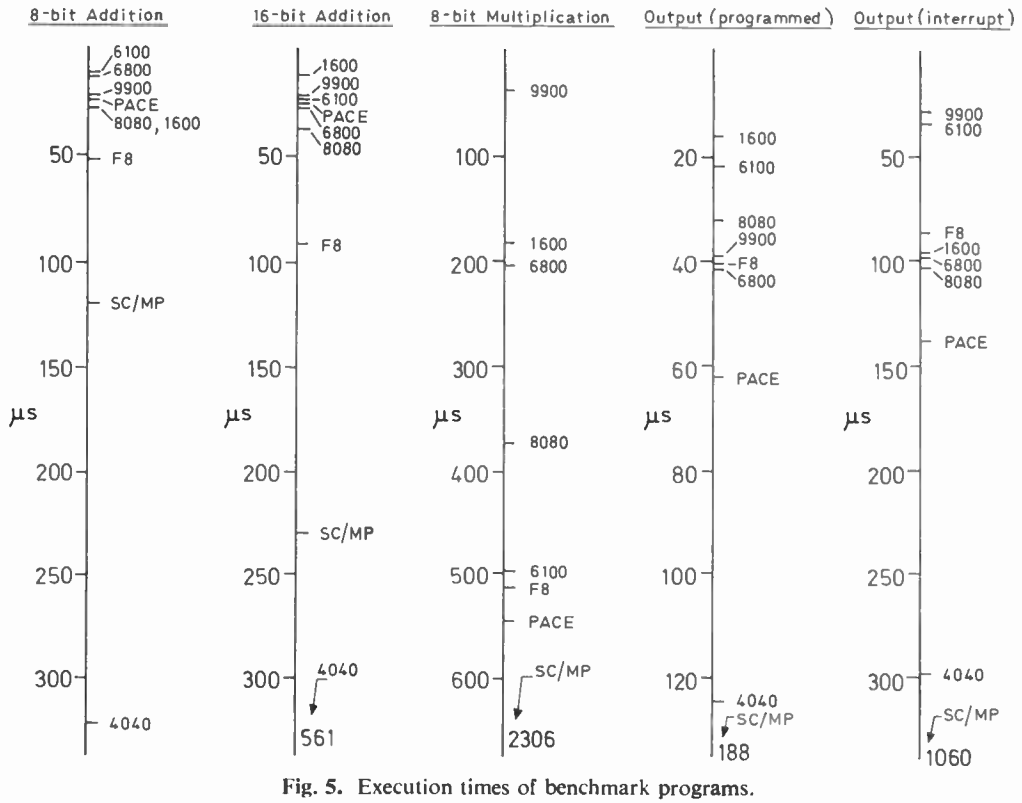


Fig. 5. Execution times of benchmark programs.

Table 3

Average results of benchmark tests

Test	4-bit	8-bit	16-bit	8-bit/16-bit
	Execution time in µs			
1	324	53	25	2·11
2	561	97	20	4·35
3	—	850	255	3·33
4	124	75	39	1·93
5	294	334	87	3·84
			average	3·11
	Number of source statements			
1	14	4·3	4·7	0·91
2	14	9·5	3	3·17
3	—	25	12	2·08
4	14	7·8	5·7	1·37
5	26	29	17	1·70
			average	1·84
	Memory size in bytes			
1	19	9·5	15·3	0·67
2	19	17·3	10	1·73
3	—	37	31	1·19
4	19	14·3	11·3	1·26
5	32	41·3	37·3	1·11
			average	1·19

increase faster than linearly with size due, amongst other things, to the extra structural requirements of bigger programs. Of course, program statements in one assembler language may take longer to write (and hence cost more) than those in another language. It might be supposed that the larger number of statements required for the 4- and 8-bit processors compared to those for the 16-bit processors would be counterbalanced by being correspondingly simpler. In fact, the converse was found to be true and much more time was spent per statement in coding programs for processors with shorter word lengths.

5.2 Memory Requirements

Here there is a smaller spread of values than for numbers of program statements, but there is still a definite trend for the longer word to require less memory.

It might be supposed that for a given task the program memory requirements would be smaller for a processor with a smaller word length than for one with a bigger word length; for example, the Intel 4040 4-bit processor has a majority of instructions occupying one byte of memory, whilst the Texas Instruments 9900 16-bit processor has a majority of instructions occupying 4 bytes of memory and this might indicate that programs for the 9900 would need more memory than those for the 4040. However, inspection of the results shows that, on average, programs for the 4040 require more memory than do the 8- and 16-bit processors, and that programs for 8-bit processors occupy more memory than do the 16-bit processors.

5.3 Execution Time

The execution times showed the strongest variations, with the average 16-bit processor being three times as fast as the average 8-bit processor and between four and twenty-five times faster than the Intel 4040.

6 Discussion of Results

The tests have shown that in respect of the three factors examined, size of source program, memory requirement of compiled program and execution time there is a definite improvement with increasing word length, and that execution times have the most striking differences.

The assumptions about addressing restrictions used in the coding of the various programs would tend to minimize these differences, and it can be expected that larger, 'real' applications programs would show these differences more strongly. These assumptions were that where the addressing scheme used a limited number (usually 8) of bits for the address field in the instruction word, all operands in the program would lie within range of the page, index or pointer register used to extend the address range. In a real application this assumption would not always be valid and extra instructions would be required to reset such registers from time to time. Processors with a 16-bit address field have a strong advantage in this respect.

Another assumption was that where these addressing schemes exist a certain amount of dedicated allocation of registers to interrupt handlers would be allowed. In a real system this might not be allowable and, again, this would open up the gaps between the 4-, 8- and 16-bit processors.

One factor which was identified as contributing to the differences was register structure. It was much preferrred to have a simple, uniform register structure such as the register files on the 9900, than complex schemes with different registers having different properties. For example, the algorithms used for programmed multiplications need two operands, a double length result and a loop counter. In the SC/MP processor it was very difficult to decide the register allocation; the existence of three pointer registers (one with special properties), an accumulator and an extension register, all with different properties, made allocation difficult and time consuming.

If a processor had only a limited number of memory reference instructions then extra instructions were needed to load the operands into registers before operating on them. Related to this is the ability to leave the result of an operation directly in a memory location, absence of which would require additional store instructions. An example of these limitations is the Intel 4040 which does not allow a memory location to be incremented; this causes a marked increase in the complexity of the interrupt handling program, for example.

One distinct lack in several of the 8-bit processors (and the 4-bit processor) is the ability to operate on 16-bit operands. Even though in 8-bit processors the majority of the data items might be 8-bit words, there is often a requirement to load, store and increment 16-bit registers (page, base, index etc.). If the appropriate 16-bit instructions are not provided then these operations may require a sequence of 8-bit instructions. This is particularly true in the case of the SC/MP processor.

A related point is the ability of 16-bit processors to address and operate on 8-bit operands. This requirement applies, for example, in character handling and manipulation. Absence of this facility (as in the G.I. CP1600) means that the required operands must be accessed via load, mask and shift operations which greatly increase complexity.

One reason for using smaller word-length c.p.u.s is that they cost less than the bigger word-length c.p.u.s. However, other costs that must also be taken into account when considering the total system are those of memory, programming and engineering. This paper has considered the first two costs and showed that they are less for c.p.u.s with bigger word lengths. If the engineering costs are independent of word length then as program size increases there will be a point at which the total system cost is less for a processor with a longer word length.

It is shown elsewhere† that the semiconductor industry is advancing in such a way that by 1980 the cost of all microprocessors will tend towards the same low value of less than £10. As this is less than the other system costs for all but the very smallest size, high production volume systems, it would seem sensible to use 16-bit processors in nearly all circumstances.

7 Conclusion

It has been shown that there are marked differences between microprocessors in respect of numbers of source statements, sizes of compiled programs and execution speeds.

Several architectural factors contribute to these differences, and these differences are mostly characteristic of the differences that arise from the word length (i.e. the a.l.u. word-length), with 16-bit processors being superior in all three respects to the 8-bit processors and both types being far superior to the single 4-bit processor examined.

The lower costs of the shorter word-length c.p.u.s may, then, be outweighed by the extra memory costs, higher programming costs and lower performance when compared to the longer word length c.p.u.s.

Manuscript first received by the Institution on 12th January 1977 and in final form on 25th July 1977.
(Paper No. 1793/Comp. 186)

† 'Microprocessor/Minicomputer Trends and West European Markets', Mackintosh Consultants Co. Ltd., London, 1976.

UDC 621.391.81.029.63

Degradation in the efficiency of medium-size parabolic antennas with propagation conditions

W. G. BURROWS,

Ph.D., D.I.C., C.Eng., M.I.E.E., M.I.E.R.E.*

SUMMARY

The results of a simple, short-term experiment suggest that significant errors may arise in the overall design and costing of a communication system where parabolic antennas are employed at the receiver terminal. Under high signal conditions the relative gains of two parabolic antennas of different diameters are in close agreement with the theoretically predicted value. Under low signal conditions, when a larger diameter antenna might be preferred in practice, there appears to be a considerable reduction in the proportion of power collected by a large antenna in relation to the proportion collected by an antenna of significantly smaller diameter.

* *Formerly of the Department of Electrical Engineering, Imperial College, London; now Head of Department of Marine Electronics, Brunel Technical College, Ashley Down, Bristol BS7 9BU.*

1 Introduction

In recent years, there has been a significant increase in the establishment of tropospheric communication systems operating over medium to long-haul paths (200–500 km). Whilst the majority of the older systems were designed and built in times of considerably less financial and physical constraints than abound at the present time, the modern system poses many new problems often by virtue of its application. In the earlier system, designed and planned for high reliability on the basis of 'worst month' path loss statistics, considerable use was made of the larger diameter antennas offering narrow beam widths and high gain. Frequently the delicate balance of costing associated with terminal equipment, geographical location, system performance etc., was maintained by the use of these antennas which in many cases are of diameters extending to some 20 metres.

In some of the more modern applications of tropospheric communication systems there are severe environmental constraints which impose firm limits on the entire physical dimensions of a complete terminal, including the antenna. The recent developments in the oil industry, with the need to provide communication links between marine oil rigs and the shore, are a suitable example. In the type of application where there are specific requirements of very high reliability and where the adverse effects of the propagation medium are to be reduced to a minimum, the antenna becomes the subject of additional interest in the processes of optimizing system performance in relation to capital costs and available physical space.

The larger antenna, with its theoretically high gain, is a very costly component in any system. Its sheer physical dimensions present serious mechanical design problems associated with its use on any site. Furthermore, once it has been installed there is very little latitude for modification. In many instances, the larger antenna is employed only to anticipate 'worst month' conditions and provide a suitable gain margin for all possible 'predictable' eventualities.

2 Theoretical Background

In the practical assessment of the plane-wave gain G for an antenna of effective aperture A_{eff} the simple theoretical relationship $G=4A_{\text{eff}}/\lambda^2$ is subject to some modification by factors such as the 'illumination factor', and the antenna–medium coupling loss. Whilst the 'illumination factor', which is largely a function of the physical dimensions and properties of an antenna, may be determined by theory or by measurement, the antenna–medium coupling loss is the subject of both conjecture and speculation.

The received power degradation of an antenna, frequently associated with antenna–medium coupling loss, has been the subject of a number of studies.[1–4,6] There remains, however, very little agreement amongst workers in this area as to the exact or real reasons for this loss.

Norton *et al.*[5] and Staras[1] give a refined account of the difference between the theoretical and actual gains of very narrow beam antenna by considering the relationship between a rapidly reducing scatter volume and the corresponding reduction in antenna beam width. These arguments assume, of course, that the primary mechanism of propagation in the direction of the antenna is one of 'scattering'. Also, Boithias and Battesti[3] suggest that the apparent gain degradation of antennas used on trans-horizon systems increases with diameter and may become significant in the case of the very large antenna.

In many respects this evidence does not appear to favour the large antenna for receiving purposes because of the loss of power which is basically attributed to amplitude and phase perturbations in the incident wavefront. There are however, possible explanations of effects whose causes must be attributed to, and directly associated with, the mechanism of trans-horizon propagation of the incident wave which is not necessarily 'scattered'. This immediately leads to the simple suggestion that, for a given system, this loss in the power collected by the larger antenna may vary with propagation conditions.

In the truly practical situations where a system is to be designed to meet high reliability performance criteria against severe constraints of costs and available space, antenna dimensions become significant. Under these conditions a large antenna, with all its assumed virtues, may be rejected in favour of an array of smaller antennas. The array does have the virtue of offering the performance of either a 'simple' or an 'adaptive' type. Such a design decision however would be influenced by the relative effective gains offered by antennas of differing dimensions operating under similar circumstances and on the same site.

3 Experimental Arrangements

A very simple investigation was recently carried out to compare the relative powers collected by two parabolic receiving antennas of different diameters, in close proximity, operating under a wide range of signal levels from a distant transmitter. The investigation was made possible by utilizing the facilities offered by a 900 MHz tropospheric radio link, operating over a trans-horizon path of length 140 km, between Christchurch, Dorset, and Imperial College, London, which had been used primarily for tropospheric propagation research.

At the transmitter, in Christchurch, a 3·66 m (12 ft) parabolic antenna was sited at a height of 6 m (20 ft) above ground level and provided with a horizontally polarized dipole feed. An identical 3·66 m antenna, sited at Imperial College and at a height of 44 m (145 ft) above ground level, was used for the reception of c.w. transmissions made for measurement purposes. With this basic geometry the total path of 140 km was essentially trans-horizon for meteorological conditions where $K \leqslant 2$. There were occasions however, when this K value for the path was exceeded to give near 'line-of-sight' conditions, over an otherwise 'smooth-earth' for considerable periods of time.

For the purpose of this investigation, an additional 1·22 m (4 ft) diameter antenna was sited adjacent to the 3·66 m receiver antenna with a horizontal spacing of 3 m between antenna centres. The lengths of the feeder cables from the two antennas to the recording receiver were adjusted to give identical loss values when connected via a coaxial changeover switch. The coaxial switch, arranged for remote operation, was used to provide a rapid change-over from one aerial to the other as required. The receiver was equipped to record long-term path loss variations, over a dynamic range of 25 dB, by means of a conventional electro-mechanical pen recorder at a recording speed of 7·6 cm/h. Facilities were available in the receiver to introduce attenuation steps of 5×10 dB in the received signal level if required. The receiver was also equipped with modest data-logging apparatus for the purpose of recording short-term ($3\frac{1}{2}$ min) path loss variations in a form suitable for subsequent analysis by computer. Feeding this raw sampled data, obtained at a rate of 5 samples per second, into a simple plotting routine provides a means of obtaining the equivalent of a short-term high-speed pen recording which can serve to highlight the finer detail in a fading pattern. This same process also allows the 'data' to be checked for undesirable artefacts prior to subsequent analysis.

The relative gains of the two receiving antennas were compared *in situ* by siting a portable transmitter approximately 2·5 km down the path, towards the main transmitter, and at an elevation similar to that of the site at Imperial College. Under these conditions, with an output power of 1 mW fed into a simple $\lambda/2$ dipole at the portable transmitter, the received signal was completely devoid of any significant fluctuations due to atmospheric changes etc. Changing the receiver antenna from one of 1·22 m diameter to one of 3·66 m diameter, under this near ideal 'free-space' condition, gave a signal enhancement of 8·5 dB. Without making any allowance for illumination factors, antenna–medium coupling loss etc., on either antenna, the theoretical signal power enhancement would be given $G_{3\cdot6}/G_{1\cdot2} = D_{3\cdot6}{}^2/D_{1\cdot2}{}^2 = 9{\cdot}54$ dB. Since the measured enhancement of 8·5 dB was a totally repeatable figure, obtained under near ideal conditions, this value was used as a reference in subsequent measurements made with respect to the two receiver antennas.

4 Results Obtained

Over the period 10th November 1976 to 15th January 1977 continuous recordings were made of the received signal level in London with frequent changes of the receiver antenna. Every attempt was made to compare the relative powers collected by the two antennas under a variety of propagation conditions for the total Christchurch–London path. It was purely coincidental, but nevertheless convenient, that during this period the conditions of the atmosphere were such that the overall

path loss for the system was subject to the greatest variations in hourly median values compared with those that had been experienced over a period of the previous five years. During the course of one day, 26th November 1976, the progress of a cold front across, and approximately parallel to, the path caused an enhancement of the median level which was some 30 dB above the highest value obtained during the previous four years. This condition prevailed for a period of approximately seven hours. For the remainder of the period of this investigation the median levels were reduced to values which were normal for the time of the year. The normal seasonal variations for this path are in the region of 25–30 dB on a total path loss of approximately 200 dB.

Figure 1(a) shows an example of the typical received signal variations for this time of the year under what might be described as 'good average' conditions. The total period shown is approximately 1½ hours. For the purpose of assessing the extent of the signal variations in question, the system was calibrated from a local source to give only a reference scale on the recordings which does not represent the actual path loss.

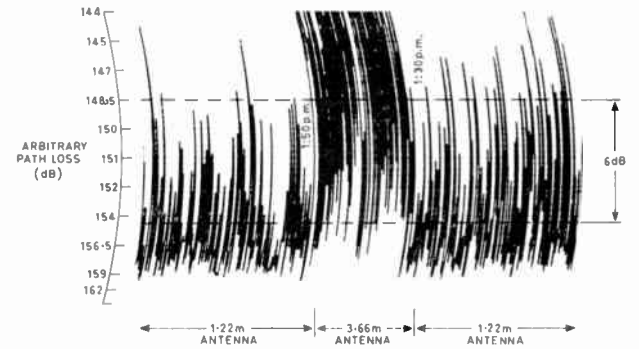
A reasonably steady average level for the signal received on the 1·22 m antenna before and after the short period (approximately 20 minutes) when the antenna was changed to one of 3·66 m diameter is indicated in Fig. 1(a). Whilst the general nature of the fading is typical for this link, it can be seen that the average enhancement of the signal due to the change of antenna is in the region of 6 dB.

Figure 1(b) shows an example of the signal variations under propagation conditions which are considerably improved relative to those applying in Fig. 1(a). During the period of 1½ hours in the sample shown, the 3·66 m antenna was used for a period of approximately 15 minutes to give an average enhancement of 8·4 dB. Unfortunately, the somewhat limited 'writing range' of the actual recording pen inhibits the visual representation of the variations in the higher level signal.

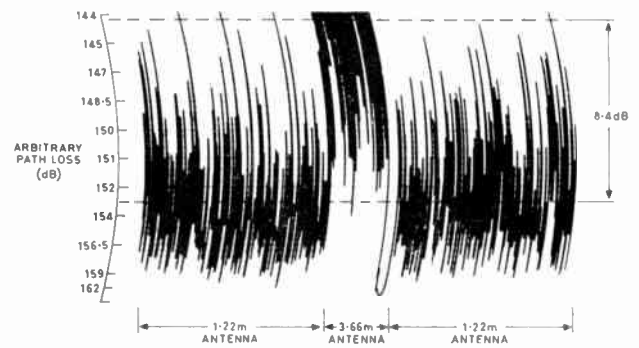
Figure 1(c) shows an example of the signal variations under propagation conditions which were extremely poor but provided a signal from the 1·22 m antenna at a significant level above the receiver noise. For these conditions, the recording amplifier gain was changed to give an expanded loss scale. Under these conditions, it can be seen that a change of approximately 2 dB results from the change of antenna diameter.

Whilst only three examples are given of specific epochs from the recordings made during the whole period of the investigation they are each representative of the typical effects which occur as the result of changing the antenna diameter under different overall path-loss conditions.

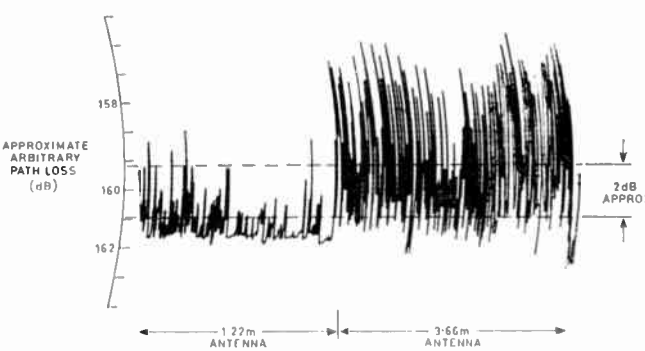
For each period of use of the two antennas short-term samples, of 3½ minutes duration, of the received signal level variations were recorded with the data-logging equipment. For the most part, these samples were obtained at times close to the instants when a change of



(a) 'Average';



b) 'Good';

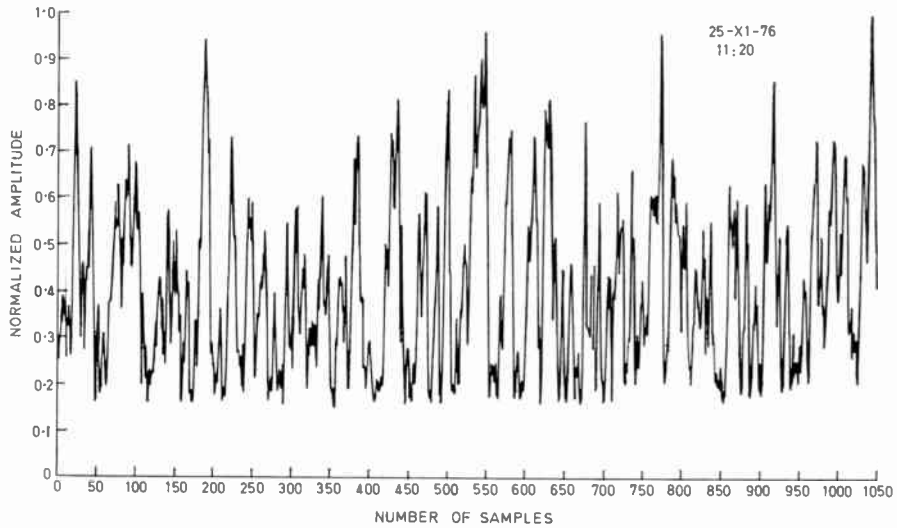


(c) 'Poor'.

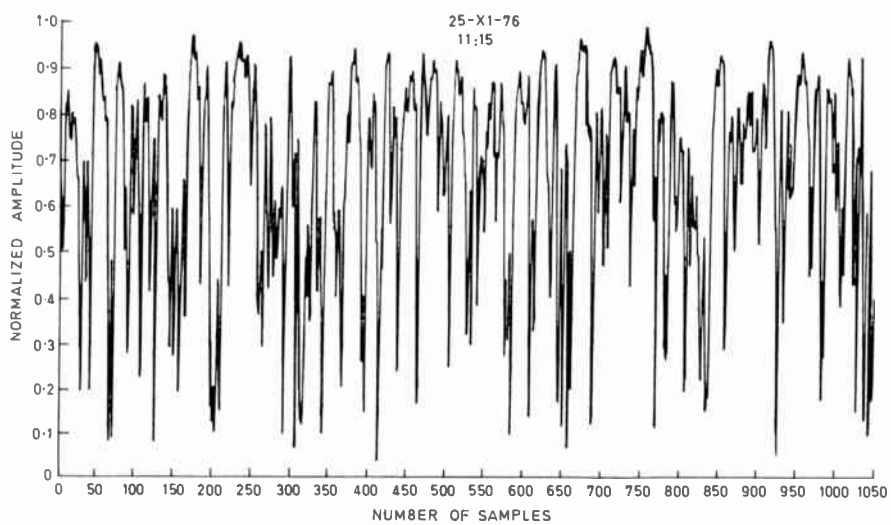
Fig. 1. Variations in received signal level with propagation conditions.

antenna occurred. This procedure reduced the possibility of significant changes in the propagation mechanisms along the path adversely affecting any comparisons that might be made between any two recorded samples.

With the aid of a computer, these samples were processed to give normalized fading characteristics, variance values, and spectral density (signal voltage) distributions. The Fast Fourier Transform program for the spectral density distribution was constrained to a maximum frequency of 2·0 Hz instead of the normal 2·5 Hz determined by the sampling rate. Figure 2(a) and (b) are examples of the raw data plotted to give the equivalent of a high-speed pen recording of the signal level variations obtained from the 1·22 m and 3·66 m antennas respectively. These two samples, which are representative of



(a) 1·22m antenna; normalized variance = 0·033.



(b) 3·66 m antenna; normalized variance = 0·049.

Fig. 2. Data samples of received signal variations.
Sampling rate = 5 per second. Epoch duration = 3·5 min.

those obtained during the investigation, were recorded immediately before and after a change from the 3·66 m antenna to the 1·22 m antenna during a settled period of 'good propagation' conditions. The spectral density distributions for these two fading characteristics are given in Fig. 3. The difference between the two distributions is again representative of all of those obtained during the investigation. The difference did not appear to suffer any significant change with variations in the general path-loss conditions.

5 Discussion

By virtue of somewhat limited practical and financial circumstances the scope of this investigation has been severely restricted. Although basically very simple, experiments of this type are very expensive in terms of both equipment and manpower. In many cases the commercial advantages that might be gained from such ventures are not always obvious.

However, in spite of the extensive literature which has hitherto been published on the properties of antennas, the relationships between these properties and the propagation medium are, as yet, neither understood or established. As a result, in the planning and design of a modern tropospheric radio communication system, experience and empirical information still play an important role. The overall costing of such systems is therefore not without a certain degree of speculation.

From the results obtained in this simple investigation, changing the receiver antenna diameter in a tropospheric

system appears to produce at least two effects which could be of considerable significance in the design of a receiver terminal on a site which imposes severe physical limitations on the antenna dimensions.

Figure 3 shows the change in the shape of the spectral density distribution of the fading characteristic typically associated with a change of antenna diameter from 1·22 m to 3·66 m. The basic effect shown by these results may be regarded as one of resolution. By increasing the antenna diameter, any time-varying incoherence in the incident wavefront resulting from the propagation mechanisms will be integrated to give a subsequent reduction in the overall effects of the incoherence. The higher spectral frequency components of the received fading signal will therefore also be reduced. Current practice favours the use of large diameter antennas principally for their gain properties. In the event of smaller diameter receiver antennas being considered for use on a path which might be subject to 'fast-fading', the resultant spectral density distribution of the variations in the received signal is likely to have serious effects on the overall system performance.

A more serious effect is demonstrated by the typical sample results given in Fig. 1.

Whilst the limited scope of the investigation precludes firm conclusions, there is clear evidence of apparent antenna gain degradation, or 'loss of power collected', resulting from adverse propagation conditions. Since there is already evidence to suggest that the smaller diameter antenna is subject only to minimal power collection loss, the results given show that, relative to the 1·22 m antenna, there are considerable variations in the power collected by the 3·66 m antenna as a result of changes in the propagation conditions. By using the 1·22 m antenna as a reference the results given in Figs. 1–3 may be summarized in Table 1.

Table 1

Summary of results

	1·22 m diameter antenna		3·66 m diameter antenna		
Propagation conditions	Arbitrary path-loss $L_{p1\cdot2}$ dB	Change in path-loss median value $\Delta L_{p1\cdot2}$ dB	Arbitrary path-loss median value $L_{p3\cdot6}$ dB	Change in path-loss median value $\Delta L_{p3\cdot6}$ dB	Enhancement $G_{3\cdot6}/G_{1\cdot2}$ dB
Poor	161	−6·5	159·7	−11·2	2
Average	154·5	0	148·5	0	6
Good	153	+1·5	144·2	+4·3	8·4

6 Conclusions and Future Work

These results suggest a number of issues which might be the subject of further, more detailed research.

For the particular path used for this investigation it can be seen that for high signal or 'good' propagation conditions the signal enhancement resulting from a change of antenna diameter is in close agreement with that predictable by theory. However, under 'poor' conditions the effective enhancement resulting from the same change of antenna diameter is considerably less than that value. This result immediately implies that, for practical purposes, there is a need both to understand the underlying reasons for this degradation and to establish an accurate description of the three-dimensional 'surface' relating gain, diameter and 'received signal level' (incident wavefront), as shown in Fig. 4. Such a 'surface' would, of course, only apply to receiving antennas set at a fixed angle of 'shoot' in both the horizontal and vertical planes. There is experimental evidence, which is as yet unpublished, which suggests that under various types of propagation or 'weather' conditions it is necessary to introduce slight variations in both the azimuthal and elevation angles of 'shoot' for a parabolic antenna in

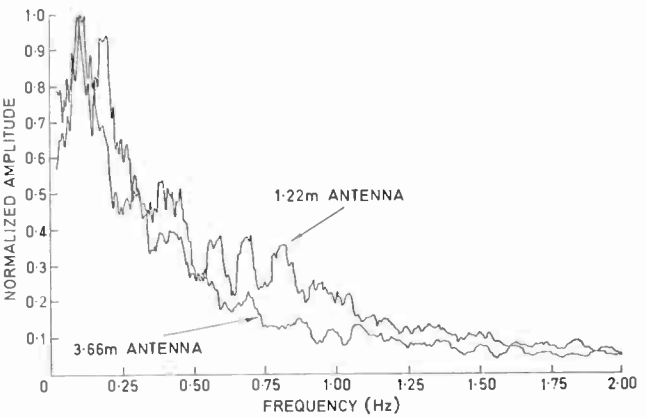


Fig. 3. Comparison of typical spectral density distributions for signal variations received on 1·22m and 3·66m diameter antennas.

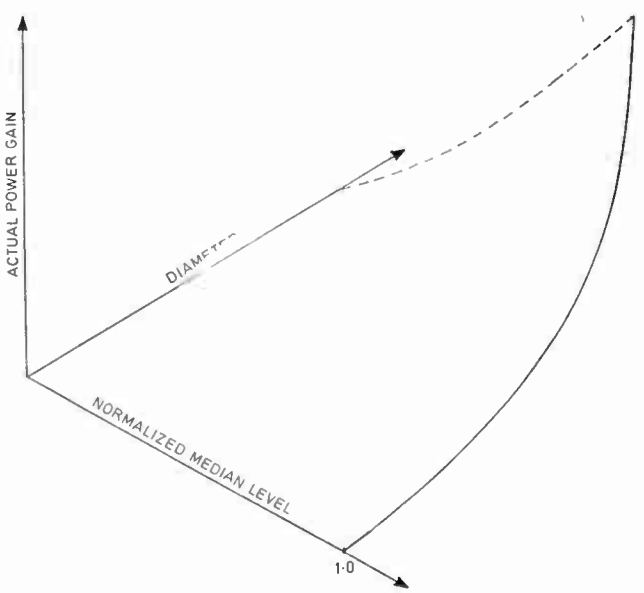


Fig. 4. Suggested 'surface' relating antenna gain, diameter and received signal level.

order to optimize the received signal level from a trans-horizon transmission.

It may also be of interest to note that there is evidence to suggest that, for the three propagation conditions considered, the change ΔL_p for both antennas appears to be related to the simple ratio of their diameters.

In practical terms, these results have very serious implications. In the absence of any assessment of the properties of a wavefront incident on an antenna used for basic path loss measurements, serious errors are likely to be incurred. Such errors are likely to give rise to false conclusions particularly in respect of fading depths, path-loss assessments etc. It would also appear that for a system operating under excessive or severe path-loss conditions, the use of a large antenna to provide additional gain is based on false assumptions and could be needlessly expensive. If the proportion of the 'loss' in the power collected by the smaller diameter antenna is to be neglected, there would appear to be an argument in favour of an array of smaller diameter antennas providing a required gain. For the majority of antenna sites the smaller antenna presents fewer mechanical and structural problems.

7 Acknowledgments

Thanks are due to the help and co-operation provided by the Director of Radio Technology, Home Office and the Director of R.S.R.E. Christchurch. Thanks are also due to those colleagues at Imperial College, who from time to time provided help and useful discussion.

8 References

1 Staras, H., 'Antenna-to-medium coupling loss', *Trans. IRE on Antennas and Propagation*, **AP-5**, No. 2, pp. 228–31, April 1957.

2 Booker, H. E. and de Bettencourt, J. T., 'Theory of radio transmission by tropospheric scattering using very narrow beams', *Proc. IRE*, **43**, No. 3, pp. 281–90, March 1955.

3 Boithias, L. and Battesti, J., 'Experimental study of the reduction in aerial gain in trans-horizon communications', *Annales des Telecommunications (France)*, **19**, No. 9/10, pp. 221–30, September–October 1964; 'New trials of the aerial gain degradation in beyond-the-horizon paths', *loc.cit.*, **22**, No. 11/12, pp. 321–5, November/December 1967; 'New elements on propagation by atmospheric heterogeneity', *loc.cit.*, **26**, No. 1/2, pp. 15–30, January/February 1971.

4 Bullington, K. *et al.*, 'Results of propagation tests at 505 Mc and 4090 Mc on beyond-horizon paths', *Proc. IRE*, **43**, No. 10, pp. 1306–16, October 1955.

5 Norton, K. A. *et al.*, 'The use of angular distance in estimating transmission loss and fading range for propagation through a turbulent atmosphere over irregular terrain', *Proc. IRE*, **43**, No. 10, pp. 1488–1526, October 1955.

6 Levin, E. *et al.*, 'Aperture-to-medium coupling on line of sight paths: Fresnel scattering', *Trans. IRE*, **AP-7**, No. 2, pp. 142–6, April 1959.

Manuscript first received by the Institution on 13th July 1977 and in final form on 5th September 1977.
(Paper No. 1794/Comm. 157)

IERE News and Commentary

The Presidential Address

Professor W. A. Gambling, whose election as President of the Institution for 1977/78 was unanimously confirmed at the Annual General Meeting in London on 13th October, will give his Inaugural Address on Tuesday, 24th January 1978. He has taken as his theme 'Electronics, Universities and Institutions'. The meeting will take place in the Large Theatre of the London School of Hygiene and Tropical Medicine, Keppel Street, Gower Street, London WC1.

Graham Clark Lecture

CEI's 22nd Graham Clark Lecture will be given by Lord Ashby of Brandon, D.Sc., F.R.S., Chancellor of Queen's University, Belfast, on Thursday, 26th January 1978. The lecture will be given at the Institution of Civil Engineers and will start at 6 p.m. Lord Ashby's lecture will be entitled 'Engineers and Politics—a case history'.

Professional Engineers and Trade Union Recognition

The CEI has placed on record its profound concern regarding the wide implications of the failure by ACAS (Advisory Concilliation and Advisory Service) to make a recommendation on trade union recognition at W. H. Allen & Sons Ltd. of Bedford.

The conclusion reached in this case is totally opposed to the overwhelming weight of evidence submitted by the negotiating parties. Despite the expressed wish of 79% of the professional engineers and other professional staff in technical posts at W. H. Allen to join a specific union, ACAS has not recommended that the union should be recognized by their employers. This makes a mockery of the democratic process and poses a serious threat to the interests not only of professional engineers but all those to whom freedom of personal choice remains important.

It is understood that UKAPE, the union involved, is initiating High Court action.

The CEI has recommended publicly that professional engineers should, in their own interest, join a trade union appropriate to their own needs in view of developments in industrial relations legislation in recent years. It has also recommended criteria by which the choice should be made by those to whom the freedom of choice is still available. Paramount in CEI thinking is the need for professional engineers to adhere to their Code of Conduct which has been formulated in the public interest.

Although the CEI cannot take a partisan role in inter-union matters, it has the obligation to protect the broad interests of all professional engineers. It feels most strongly that this pronouncement by ACAS runs counter to those interests and to the need to recognize the right of professional engineers to be represented by unions in which their voice can be effective.

New Advisory Service for the Engineering Industry

A Manufacturing Advisory Service to help small and medium-sized companies in manufacturing industry to increase efficiency by adopting modern but well proven technology and management practices is being set up by the Secretary of State for Industry.

Replying to a Parliament Question, Mr. Eric Varley said that the service is intended mainly for firms employing between about 100 and 1,000 employees in metal working and assembly, which are activities common to many industrial sectors, but a few schemes for specific sectors may be included.

It is understood that industrialists of wide engineering experience accompanied by a suitable team will visit selected companies to join with the management in surveying the firm's manufacturing operations. In suitable cases the management and the team will jointly arrange a first short advisory project, to be undertaken by a team drawn from the most appropriate source of specialist advice, for example industry research associations, Government Research Establishments, industrial consultants and universities. The service will be co-ordinated by the Production Engineering Research Association, who will also provide a technical enquiry and information service and demonstration facilities.

The service will start on a modest scale, aiming to cover some 700 firms in the first two years, at an estimated cost of £1·75M, and after review it may be extended. The Secretary of State intends to appoint a steering committee for this service, including an industrialist chairman, and independent members whose names will be announced later.

Further information on the Service can be obtained from: The Department of Industry, Research Requirements Branch, 3 Abell House, John Islip Street, London SW1P 4LN. (Tel. 01-211 3520).

BIMCAM, CAMA and SIMA Forge Closer Links

Three leading trade associations in the measurement, control and automation industry, which for several years have co-operated in areas of common interest, have now formalized their relationship by setting up a new Group known as 'The Group of Associations of Manufacturers of British Instrumentation, Control and Automation' (GAMBICA).

The three Associations are: The British Industrial Measuring and Control Apparatus Manufacturers' Association, the Control and Automation Manufacturers' Association and the Scientific Instrument Manufacturers' Association, all of which are members of the BEAMA Federation.

The purpose of the new Group will be to combine the resources of the three Trade Associations on issues of common interest while each will retain its individual autonomy.

The combined membership of the three Trade Associations is nearly 240 companies with a total turnover of approximately £1,000M a year of which over 50% is attributable to export.

In a joint statement, the chief office holders of the three Associations (Mr L. R. Pullen—President, SIMA, Mr A. St. Johnston—Chairman, CAMA, and Mr G. S. Kermack—Chairman, BIMCAM) said:—'It is our view that this closer working together will bring about a more powerful voice for this highly important sector of industry, which has a high growth potential at home and overseas and which has an influence on a wide range of industries in the U.K.'

SIMA and BIMCAM will continue to be located at 20 Peel Street, London W8 7PD, (Tel. 01-727 2614) and CAMA at Leicester House, 8 Leicester Street, London WC2H 7BN (Tel. 01-437 0678).

Education, Engineers and Manufacturing Industry

During 1976, following an initiative by the British Association for the Advancement of Science in conjunction with the Royal Society and NEDO, it was decided to commission a report on an important problem facing manufacturing industry, which would be presented at the 1977 Annual Meeting of the British Association.

To this end a Co-ordinating Group was formed, chaired by Sir Montague Finniston, F.R.S. The background to the project was described in the April 1977 issue of the Journal and the terms of reference selected for study were:

'What steps are desirable, with respect to the education (including career advice), recruitment and deployment of professional level engineers, to improve the performance of British manufacturing industry and in particular the effectiveness of production management.'

The project was supervised by a Management Committee, chaired by Sir Ieuan Maddock, C.B., O.B.E., F.R.S. (Past President IERE) and the Report* gives the conclusions of seven months' study by a specially-recruited investigating team based at the University of Aston in Birmingham and directed by the Vice-Chancellor, Dr J. A. Pope.

The study was initiated with the premise that an improvement in the performance of manufacturing industry requires an appropriate supply of able and well-trained engineers, and it accepted that much is already good in British education, British engineering, British manufacturing industry and production management. The continuous criticism levelled at manufacturing industry however damages confidence and may well dissuade many young people from choosing careers in industry.

The investigating team welcomed the current stated policy that economic priority must be given to manufacturing industry to increase its effectiveness. But no single action was likely to achieve this end and to suggest that the supply and deployment of engineers is the main cause of past or present problems, or that improvement in this area alone will cure all problems, was considered to be wholly wrong.

The Report is based on an examination of existing data and studies, and on discussions with persons knowledgeable in the areas covered by the terms of reference. The Management Committee decided at the outset of the study that there was not enough time to allow for any fresh survey work. The effectiveness of production management depends on many interdependent variables and to isolate the effect of a single variable (such as engineering education or the quality of professional engineers) was not possible. Consequently, the Report makes only limited reference to the production management aspect of the terms of reference, and it assumes that an improvement in the supply of high quality engineers must constitute a move in a positive direction.

How the Schools can help

To influence more of the abler young people to choose engineering as a career will require changes in attitudes. School students see engineers, particularly those in manufacturing industry, as having low salaries, low status and low requirements of intelligence compared with other professions. Various initiatives exist to counter this widespread attitude, but they will only have limited success as long as our society has a low opinion of engineering as a profession.

The investigating team emphasized that mathematics is essential for engineering. How it is taught, what is taught, and the number and quality of those teaching it in schools, are issues of widespread concern. There are three distinct problems. Firstly, the dropping of mathematics before 'O' level, or inadequate teaching to this level, can permanently handicap the student and limit his or her future choice of subjects. Moves to establish a common core curriculum may help in this regard. Secondly, there are more than fifty varieties of syllabus at 'A' level, with a range of content such that first-year teaching of university subjects which rely on mathematics cannot assume a common body of knowledge even if all students possess 'A' level mathematics. Moves by the Standing Conference on University Entrance to state a common minimum syllabus content may help in this regard. Thirdly, no action on curriculum or syllabus can succeed unless there is an adequate supply of teachers properly qualified to teach at each level. Government action to train teachers in shortage subjects is welcome; but declining secondary school enrolment, and a policy of maintaining a staff/pupil ratio based on 1976 levels, may preclude some schools from employing the necessary staff. A solution to this problem is urgent and essential and must rest with the Government.

Changes at the Universities and Polytechnics

Specific 'A' level requirements in mathematics *and* physics restrict engineering degree courses to a limited pool of potential candidates, and one which shows no growth. An 'A' level requirement of mathematics and one or more 'A' level passes in any other subjects would increase the number of potential candidates by about 50%, and thus would increase the potential for attracting good quality students. The ability to make this change rests with the universities and polytechnics.

Engineering and science at universities have lost considerable ground in attracting UK men students. The increase of interest in studying engineering in 1976 and 1977 suggests the trend is being reversed. But, whereas science has gained somewhat from the very substantial increase in the number of women studying in universities and polytechnics, engineering has not. Unless more women can be attracted into engineering, the profession has to seek adequate quantity and quality from boys only, whereas competing professions can recruit from both sexes. The investigating team recognized that considerable social adjustments are required, which would take time, but nevertheless they believe that sustained effort to make engineering an attractive career for women is essential.

Significant changes in engineering course curricula have been made during the past ten years and the traditional engineering course which aimed to shape all students in the same mould is rapidly becoming the exception. The most notable developments are the broadening of courses to include economics and management subjects, increased flexibility through providing multiple options in technical and non-technical subjects, and a sharpened interest in design as a vehicle for integrating basic subjects. The investigating team consider these developments are healthy because they enable the universities and polytechnics to cater better for the varied interests, abilities and capacities of students, and for the wide range of industry's needs. Schools need to know of the range of choices available to sixth

* 'Education, Engineers and Manufacturing Industry'. Main Report and Support Papers, price £2·00 each, both volumes together £3·00. Obtainable from The Information Officer, University of Aston in Birmingham, Birmingham, B4 7ET.

formers, and companies need to learn more about the range, and also to work with the universities and polytechnics so that the options are suited to their requirements.

The total output of qualified engineers (all disciplines) is currently little different to the level in the early sixties, but the mix of that output has changed significantly. Higher National Certificate/Diploma (HNC/HND) men dominated the output in earlier years, but nowadays degree graduates account for most of the annual output. One implication of this change in mix is that demand for undergraduate and graduate training places in industry has increased greatly and industrial training places are scarce as is evident from a study considered in the Report which concluded that sandwich courses have reached saturation point.

Universities and polytechnics, industry, the professional Institutions and appropriate Training Boards have parts to play in the education and training of engineers but the present system is ill-defined and performance is mixed. There is an urgent need to improve the integration and co-ordination of the education and training phases. Any new proposals must, however, cater for diversity and a single solution is unlikely to prove satisfactory. In the opinion of the investigating team the responsibility for ensuring that a graduate has received adequate education and training should lie with the educational institutions, and that the professional Institutions should concentrate on monitoring the integrated education and training schemes rather than monitoring the individual.

Industry Attitudes

The Report says that no evidence was found to suggest that engineers of high calibre are held back from reaching the top in a company. Those who achieve general management or functional management positions are paid as well as colleagues in other functions. Those who do not make career progress are not paid well. There is a need for well-defined career structures both in technical functions and in production management, with particular stress on providing satisfactory job opportunity in the early stages of each structure. Companies should also review the *jobs done* by qualified engineers over the age of 30 and earning less than £5,000 p.a., so that opportunities for retraining and redeployment may be identified. If these actions lead to upgrading of the jobs done by engineers, under job evaluation systems of payment, this will improve the average position of engineers in the job hierarchy.

Industry recruits most of its graduate input from first-degree graduates of UK origin, preferably men, and this recruitment base has shown little growth in recent years. The fall in its share of all graduates has been mainly in science, arts and social studies men, and in women.

The 1975 output of mechanical and production engineering manpower—first degree graduates and HNC/HND holders—was a third smaller than it was eight years ago. Admissions data for 1972/75 indicate that output of UK origin graduates in 1976 and 1977 will show a further decline, and no improvement in output will materialize until 1979/80. This suggests that manufacturing industry may have difficulty in recruiting mechanical and production engineering manpower over the next 2-3 years. The difficulty could become acute if the economy recovers rapidly.

The report is mainly directed towards improving the quality of engineers who enter manufacturing industry, but it is clear that the problem involves a number of factors. In the schools, there is a need to secure more able young men and women, who can at least consider engineering as their subject in higher education, and this critically depends on the quality of the teaching of mathematics. The universities and polytechnics need to make engineering available to the widest range of good quality candidates and this can be done by dropping physics as a compulsory requirement. There is a need for integrated education and training, between higher education institutions and industry. There is a need for industry to place graduates in jobs which utilize their full capacity and provide visible career paths in technical or managerial roles. There is a need for both technical and managerial roles to be adequately rewarded, and known to be so. Each of these actions has been considered above, but it is their combined effect—each reinforcing the others—and the recognition by our society of the importance of engineering and of manufacturing industry which these actions would signify, that will ultimately decide the quality of engineers that industry will obtain.

Plan for Action

The Report concludes with an 'action plan' under four headings.

Actions to expand the quality and size of the pool of candidates eligible to study engineering

Government must ensure that an adequate number of properly qualified mathematics teachers are trained and that they are employed by the schools.

Universities and polytechnics should change their admissions requirements for engineering from two specific subjects at 'A' level (generally Mathematics and Physics) to one specified subject (Mathematics) plus one or more good 'A' level passes in any other subject(s).

Efforts should be made to attract more women into the engineering profession. Companies, universities, and schools must initiate action at the local level, and Government, the CBI, the professional Institutions and the Industry Training Boards must assume responsibility at the national level.

Actions to improve the education and training of engineers

Government in conjunction with industry, the educational institutions, the professional Institutions and appropriate Industry Training Boards should carry out a detailed study:

(i) to determine how the full spectrum of engineering education and training (including in-career training) may be properly integrated and co-ordinated.

(ii) to determine the necessary capacities for both the education and training components of the integrated system.

(iii) to determine what additional resources are required, and how they shall be provided in a cost-effective manner.

A solution to this problem is urgent, and in the meantime the educational institutions should take greater responsibility for the total assessment of the integrated education and training of the professional engineer. The professional Institutions should concentrate on monitoring these integrated education and training schemes.

Actions to improve the recruitment and deployment of engineers

Companies should develop well-defined career structures in technical and in management functions, with particular stress on providing satisfactory job opportunity in the early stages of each structure.

Companies should examine the job specifications and career potential of every professional engineer on their staff who over 30 years of age and has a salary under £5,000 per annum.

Actions to improve the status of engineering

Companies should demonstrate in practical terms thier assessment of the importance of industry to the national welfare.

Vacation School Report:

The Practical Use of Control Theory

Reading, 22nd and 23rd March 1977

The Easter Vacation School on 'The Practical Use of Control Theory', organized by the Committee of the Automation and Control Systems Group and held at the University of Reading, was designed to bring participants together with practising experts who are also experienced teachers, and it represented a new venture for the IERE. The School was held in three sessions; the first introduced some basic theory of linear continuous systems, the second illustrated the advantages of using interactive computer graphics in design, and the third and final session demonstrated the use of theory in the design of a number of practical control systems. The discussions following the presentations allowed the participants to question the finer points more closely.

Professor D. R. Wilson of the Polytechnic of Central London chaired the first session and introduced the first lecturer, Mr P. Atkinson of the University of Reading, who gave a lecture entitled 'Review of Control Theory and its Applications'. This covered the natural occurrence of differential equations in practical systems and their solution using the Laplace transform and the significance of these solutions. The ideas of transfer function analysis were introduced together with frequency response representation on Nyquist, inverse Nyquist, and Bode diagrams; the solution of the characteristic equation via the root locus representation was also mentioned. The concepts of mechanistic modelling and model order reduction were discussed with a review of the various methods of compensation.

Dr S. E. Williamson of Surrey University gave the second lecture, which exposed the root locus method in considerable detail. After introducing the connection between the roots of the characteristic equation, the *s*-plane and time domain response of a system, the lecturer explained the geometry of root loci. He then surveyed the shapes of some root loci which are frequently encountered in engineering design. The application of the method in design and compensation was outlined with particular reference to a position control system. Particular problems which occur often in practical systems were then surveyed, including the stabilization of open-loop unstable systems, systems with highly damped complex poles and the effect of pure time delay.

In the next lecture, Lt. Cdr. M. J. Ashworth of the Royal Navy Engineering College, Manadon, explained the necessity for plant modelling in control system design and showed how it is possible to produce a representative low order model of a high order plant which may then be satisfactorily used in the design process. Lt. Cdr. Ashworth emphasized however that plant modelling is only the very beginning of the design process in which the major objectives will include reduction of sensitivity to parameter variations and external disturbances and in the control of the magnitude of the plant-input signal.

The evening session consisted of a dual presentation by Mr A. J. Allen (University of Reading) and Mr P. Atkinson. Mr Allen began by describing a suite of programs which have been developed at Reading for the interactive computer-aided design of closed-loop control systems, for use on a PDP 8 computer with 4 k words of core-store, a floppy disk backing store, c.r.o. display and v.d.u. Mr Atkinson's demonstration of the working programs employed two large display oscilloscopes and four large television monitors to show the computer and operator responses. The flexibility and power of the suite were illustrated as graph after graph of Nyquist diagram, inverse Nyquist diagram, root locus diagram, impulse response, step response, ramp response and disturbance frequency response were flashed on the screens in 'live' interactive designs. After the presentation the chairman asked the audience, who were mainly from industry, how many actually had a comprehensive facility of this kind available within their own organization. Surprisingly enough, there were very few.

The morning session on the second day opened with a lecture by Professor D. R. Wilson in which he showed how quite elementary control theory can be brought to bear on the solution of the radar tracking problem. He emphasized, however, that simple design methods must be supported by in-depth studies involving more detailed structural definition of the hardware components, linearity and operating environments. The use of non-linear systems to improve performance was introduced and practical circuits to achieve the required non-linearities were described.

The second lecture, on 'The Design of Guided Missile Autopilots', was given by Mr P. Garnell of the R.M.C.S. Shrivenham, who showed the structure of some typical autopilots and explained the basic design problems. He gave the linearized transfer functions of associated elements and explained how satisfactory performance could be achieved by use of Towill's coefficient plane techniques in the system design. Mr Garnell demonstrated a design using a special-purpose analogue simulation of a missile autopilot.

Professor D. R. Towill himself next gave a masterly description of his coefficient plane method of design of tracking systems, applied on this occasion to an aircraft automatic landing system. He showed how the low-order modelling techniques described earlier by Lt. Cdr. Ashworth could be combined with the coefficient plane representation to design tracking systems in general and to predict and compare the performances of tracking systems designed using various standard criteria. After his lecture Professor Towill dealt ably with a critic in the audience who compared his method to that described by Whiteley over thirty years ago. The final discussion was really rather too short for an adequate exchange of views but discussions continued during lunch before all the participants left.

The School had been oversubscribed and the enrolment had had to be closed at 66. The lectures were well received and the organizers feel reasonably satisfied that this type of venture can be made an academic success of real value to those who want to learn.

P. ATKINSON

New Books Received

All the books which are described below are available in the Library and may be borrowed by members in the United Kingdom. A postal loan service is available for those who are unable to call personally at the Library.

Handbook for Electronics Engineering Technicians

MILTON KAUFMAN and ARTHUR H. SEIDMAN (*Editors*). McGraw-Hill, London and New York 1976. 16 × 23·5 cm. 520 pages. £16·20 (UK), $19·50 (US).

CONTENTS: Characteristics of resistors. Characteristics of capacitors. Coils. Magnetic circuits. Transformers. Practical circuit analysis. Meters and measurements. Semiconductor devices qad transistors. Integrated circuit technology. Tuned circuits. Filters. Transistor amplifiers and oscillators. Operational amplifiers. Digital circuit fundamentals. Digital integrated circuits. Power supplies. Batteries. Vacuum tubes.

Key Papers on Surface Acoustic Wave Passive Interdigital Devices

DAVID P. MORGAN (*Editor*). Peter Peregrinus, Stevenage, Herts 1976. 21 × 29·5 cm. 372 pages. £9·00.

CONTENTS: Introductory review and survey. of applications. Interdigital transducers. Materials and propagation effects. Bandpass filters. Pulse-compression filters and phase-coded matched filters. Delay lines and special techniques.

A collection of 74 papers from world literature originally published between 1965 and 1975.

Testing Methods and Reliability: Electronics

A. SIMPSON (*Worthing College of Further Education*). Macmillan, London 1976. 23 × 15 cm. 172 pages. £3·95.

CONTENTS: Testing methods and reliability: The need for testing. Reliability. Data presentation, distributions and sampling. Logarithmic units. Instruments. Component measurements. Testing methods —electronics: Electronic voltmeters. Cathode-ray oscilloscope. Sources of test signal. Measurement of Q or magnification factor. Valve and transistor testers. Fault-finding.

This book and the following one are intended mainly for those taking CGLI Technician courses.

Testing Methods and Reliability—Power

A. SIMPSON. Macmillan, London 1976. 23·5 × 15 cm. 176 pages. £3·95.

CONTENTS: The need for testing. Reliability. Data presentation, distributions and sampling. Logarithmic units. Instruments. Component measurements. Measurement of power and energy. Measurement of speed of rotation and temperature rise. Tests on electrical machines. Cathode-ray oscilloscope and its applications. Measurements and fault-finding on valve and transistor amplifiers. Tests on wiring installations and fault location in cables.

Electrical Communications

R. G. MEADOWS (*Polytechnic of North London*). Macmillan, London 1976. 23 × 15 cm. 178 pages. £3·25.

CONTENTS: Signal analysis and system response. Noise. Amplitude and angle modulation systems. Pulse and pcm modulation. Pulse code modulation. Transmission lines and waveguides. Antennas. Short wave, microwave and radar systems.

Computer-aided Design of Surface Acoustic Wave Devices

J. H. COLLINS and L. MASOTTI (*Editors*). Elsevier, Amsterdam 1976. 25 × 16·5 cm. 308 pages. $42·50.

CONTENTS: SAW propagation in piezoelectric solids. Basics of the SAW interdigital transducer. Bulk wave generation by the IDT. Multistrip couplers. Basics of SAW frequency filter design. Interdigital transducer techniques for specialized frequency filters. The design of SAW dispersive filters using interdigital transducers. Design of reflective-array surface wave devices. Surface acoustic wave oscillators. Pattern generation and replication for SAW devices. SAW signal transform techniques. Key signal processing functions performed with surface acoustic wave devices.

This set of papers is published as a special issue of *Wave Electronics*, Vol. 2, issues 1, 2 and 3. They were first read at a European Workshop on Computer Aided Design of SAW Devices in 1975.

Handbook of Operational Amplifier Circuit Design

DAVID F. STOUT (*Aeronutronic Ford Corporation*). McGraw-Hill, New York and London 1976. 16 × 23·5 cm. 434 pages. £20·35 (UK), $24·50 (US).

CONTENTS: Introduction to operational amplifiers. Fundamentals of circuit design using op-amps. Feedback stability. Amplifiers. Comparators. Converters. Demodulators and discriminators. Detectors. Differential amplifiers. Low-pass filters. High-pass filters. Bandpass filters. Bandstop filters. Frequency control. Integrators and differentiators. Limiters and rectifiers. Logarithmic circuits. Modulators. Multipliers and dividers. Multivibrators. Oscillators. Parameter enhancement and simulation. Power circuits. Regulators. Sampling circuits. Time and phase circuits. Waveform generators.

Singular Optimal Control Problems

DAVID J. BELL (*UMIST*) and DAVID H. JACOBSON (*CSIR South Africa*). Academic Press, London and New York 1975. 23 × 15·5 cm. 190 pp. £5·80 (UK), $15·00 (USA).

CONTENTS: Historical survey. Fundamental concepts. Necessary conditions. Control. Sufficient conditions and necessary and sufficient conditions for non-negativity of nonsingular and singular second variations. Computational methods for singular control problems.

Applied Electromagnetics

J. E. PARTON and S. J. T. OWEN (*University of Nottingham*). Macmillan. London 1975. 23·5 × 15·5 cm. 258 pages. £8·95 (hard cover), £4·95 (paperback).

CONTENTS: Vector analysis. The electric field. The electric field and materials. The magnetic field. Magnetic forces and magnetic media. Fields varying in time. Electromagnetic waves. Field problems— non-exact solutions. Field problems— analytical solutions. Some low-frequency applications. High-frequency effects.

Minicomputers and Microprocessors

MARTIN HEALEY (*University College of Wales, Cardiff*). Hodder and Stoughton. London 1976. 23 × 15·5 cm. 353 pages. £9·50 (hard cover), £6·25 (paperback).

CONTENTS: Digital computers and their applications. A rudimentary digital computer. Further c.p.u. features. Input/output. Microprocessors. Peripheral devices. Software. Advanced features. Selecting a computer system.

The aim of this book is to explain how a minicomputer works. It extends the subject beyond the computer itself to include peripheral devices, such as disks, terminals, etc., and explains the significance of the standard software, including compilers, assemblers, operating systems, etc. There is no attempt to teach computer programming.

Convolution and Fourier Transforms for Communications Engineers

R. D. A. MAURICE (*late of BBC Research Department*). Pentech Press. London 1976. 22 × 14·5 cm. 198 pages. £7·50.

CONTENTS: Algebraic convolution and generating functions. An example of algebraic convolution. Mathematical convolution. The delta function. Spectra and characteristic functions. Fourier transforms of the delta function. Response of linear circuits to transient excitations. Some of the rules for playing the Fourier-transform game. Examples of use of the Fourier integral and convolution in broadcasting problems. Convolution division. Convolution square-roots. Differentiation and integration by convolution. Correlation.

Applicants for Election and Transfer

THE MEMBERSHIP COMMITTEE at its meeting on 11th October 1977 recommended to the Council the election and transfer of the following candidates. In accordance with Bye-law 23, the Council has directed that the names of the following candidates shall be published under the grade of membership to which election or transfer is proposed by the Council. Any communication from Corporate Members concerning the proposed elections must be addressed by letter to the Secretary within twenty-eight days after publication of these details.

October Meeting (Membership Approval List No. 239)

GREAT BRITAIN AND IRELAND

CORPORATE MEMBERS

Transfer from Graduate to Member

ATKINS, Ronald. *Manchester.*
LOCKWOOD, Roland Robert. *Reading, Berkshire.*
MORCOM, Christopher John. *Bath, Avon.*
PHILLIPS, Ian. *Malmesbury, Wiltshire.*

Transfer from Associate Member to Member

NICHOL, Brian Christopher. *Middlesbrough, Cleveland.*

Direct Election to Member

ALDHAM, John Paul Edward. *Chelmsford, Essex.*
CARTER, David. *Chesterfield, Derby.*
WHITEMORE, Philip. *Salisbury, Wiltshire.*

NON-CORPORATE MEMBERS

Transfer from Student to Graduate

KAY, Alan John. *London.*
RAHIM, Tehseen. *London.*
ROBERTS, Martin Henry. *Swindon, Wiltshire.*

Direct Election to Graduate

FIRTH, Christopher Nigel. *Woodford Green, Essex.*
LYMER, Anthony. *Burnham-on-Crouch, Essex.*
MILLS, Robert Leybourne. *Nether Poppleton, York.*
ROFFE, Ian Michael. *Plymouth.*

Transfer from Student to Associate Member

PATTERSON, Adrian Gifford. *Yate, Avon.*

Direct Election to Associate Member

ARMSTRONG, Peter Gaynor. *Salisbury, Wiltshire.*
BANHAM, Stephen James. *Salisbury, Wiltshire.*
IRVING, Derek Keith. *Horsham, West Sussex.*

STUDENTS REGISTERED

SMY, Clifford Matitn. *South Benfleet, Essex.*
SPENCER, Andrew John. *Southampton.*
THANKI, Yogeschandra. *Chatham, Kent.*

OVERSEAS

CORPORATE MEMBERS

Direct Election to Fellow

HOLLAND, David. *Kingston, Jamaica.*

Transfer from Graduate to Member

SHEFFIELD, James Hedley. *Scarborough, Ontario.*

Direct Election to Member

AL-ARAJI, Saleh Raoof. *Baghdad, Iraq.*

NON-CORPORATE MEMBERS

Transfer from Student to Graduate

KWAN, Hok-Chi. *Hong Kong.*

Direct Election to Graduate

CHENG, Chun Tung. *Hong Kong.*
KRISHNAMOORTHI, V. *Trivandrum, Kerala, South India.*
SIM, Chin Hong. *Singapore.*

Direct Election to Associate Member

IDIKA, Aaron Chidozie. *Cairo, Egypt.*
KHAN, Saeed Newaz. *Lahore Cantt, Pakistan.*
NJOKU, Evarist Uzodinma. *Kaduna, Nigeria.*
WONG, Tiing Ho William. *Sibu, Sarawak, Malaysia.*

STUDENTS REGISTERED

CHAN, Ng Kun. *Hong Kong.*
CHAN, Yat On. *Hong Kong.*
CHEUNG, Ping Ming. *Hong Kong.*
CHONG, Khiong Lin. *Singapore.*
CHU, Wing Yuen. *Hong Kong.*
ENG, Hock Seah. *Singapore.*
HOE, Lye Soon. *Singapore.*
KO, Yiu Chi. *Hong Kong.*
KONG, Yee Kam. *Kuala Lumpur, Malaysia.*
LAM, Weng Keen. *Singapore.*
WONG, Sang. *Hong Kong.*
YEO, Chor Wah. *Singapore.*
YUEN, Shun Yau. *Hong Kong.*

BBC Engineering Design Information

The BBC research, design and manufacture electronic equipment required for its own use, which cannot be obtained from commercial sources. The whole field of broadcasting technology is covered from d.c. to microwaves—outside broadcasts to studios to transmitters both for radio and television. The majority of BBC-designed equipment is available to British industry for commercial exploitation under a manufacturing licence agreement and many such licences are in force.

Engineering Design Information Sheets are issued describing the designs and the following are among these published in recent months.

Low Noise 15 dB *Video Amplifier* AM5/526

Designed for use at the receiving end of coaxial cables, the low noise of this unit makes it particularly suitable for use as the first amplifier in a system.

U.H.F. Test Equipment EP14M/507

When used in conjunction with a video oscilloscope, this portable test equipment provides comprehensive facilities for the repair and maintenance of u.h.f. television transmitters and transposers. It comprises a tunable modulator and demodulator, a sweep generator and a spectrum analyser with tracking facilities.

I.T.S. Generator and Inserter GE4M/556

This equipment generates National I.T.S. or, with suitable internal reprogramming, International I.T.S. and inserts these on to the correct lines during the field blanking period of a 625-line composite video signal. Several internally or remotely switched full field test signals are additionally provided. Internal programming also allows for selective blanking to black level of lines 7 to 22 and 320 to 335.

Band II F.M. Receiver RC1/12

A high sensitivity v.h.f. f.m. stereo receiver intended for monitoring a f.m. transmission, channel selection is by a pre-tuned crystal oscillator, but the channel can be changed once installed by changing the crystal and making a few adjustments.

Amplitude Measuring Unit UN1/715

Developed to measure video signal amplitudes, this equipment forms part of a measurement system which must include an oscilloscope. It uses the standard method of measurement in which a square wave of variable known amplitude is added to the waveform of the video input signal. A digital display indicates, in decibels, departures from the selected standard.

Pulse Delay Unit UN14L/532

Provides a means of delaying standard television pulses for up to 4μs, the unit contains four 1μs delay lines, each with 50ns steps to allow adjustment to within 25ns.

These design information sheets may be consulted in the Institution's Library. Members who require further information, or would like to be put on the mailing list for future issues should write to Mr. S. A. Snook, Liaison Engineer, Designs Department, British Broadcasting Corporation Broadcasting House, London W1A 1AA. (Telephone: 01-580 4468, extn. 4345/4325).

Forthcoming Institution Meetings

Tuesday, 22nd November

Colloquium on PORTABLE COMMUNICATIONS SYSTEMS MEETING POSTPONED

Tuesday, 29th November

JOINT IERE/IEE COMPUTER GROUP

Colloquium on ELECTRONIC SECURITY AND PERSONAL ACCESS SYSTEMS

Royal Institution, Albemarle Street, London W1, 10 a.m.

Advance registration necessary. For further details and registration forms apply to Meetings Officer, IERE.

Tuesday, 6th December

ELECTRONICS PRODUCTION TECHNOLOGY GROUP

Colloquium on NEW TECHNIQUES AS AIDS TO PRODUCTION

Royal Institution, Albemarle Street, London W1, 10.30 a.m.

Advance registration necessary. For further details and registration forms apply to Meetings Officer, IERE.

Tuesday 10th January

MEASUREMENTS AND INSTRUMENTS GROUP

Colloquium on MEASUREMENT AND POLLUTION

Royal Institution, Albemarle Street, London W1, 2 p.m. Advance registration necessary. For further details and registration forms apply to Meetings Officer, IERE.

Thursday, 19th January

AUTOMATION AND CONTROL SYSTEMS GROUP

Highways and byways of ultrasonics

By A. E. Crawford (*D.D.S. Engineering*)

London School of Hygiene and Tropical Medicine, Keppel Street, London WC1, 6 p.m. (Tea 5.30 p.m.).

Wednesday, 25th January

COMMUNICATIONS GROUP

Colloquium on INTERWORKING BETWEEN P.C.M. AND F.D.M. SYSTEMS

Royal Institution, Albemarle Street, London W1, 10.30 a.m.

Advance registration necessary. For further details and registration forms apply to Meetings Officer, IERE.

Tuesday, 31st January

EDUCATION AND TRAINING GROUP

Colloquium on ENGINEERS CAN COMMUNICATE!

Royal Institution, Albemarle Street, London W1, 10.30 a.m.

Advance registration necessary. For further details and registration forms apply to Meetings Officer, IERE.

Southern Section

Wednesday, 23rd November

Switched mode power supplies

By P. Chapman (*Marconi-Elliott*)

Room ABO11, Portsmouth Polytechnic, King Henry I Street, Portsmouth, 7.30 p.m.

Wednesday, 30th November

JOINT MEETING WITH IEE

The use of microprocessors in numerical control

By Dr. V. Latham (*University of Southampton*)

Lanchester Theatre, University of Southampton, 6.30 p.m.

Wednesday, 7th December

Dolby noise reduction system

By P. Plunkett (*Dolby Laboratories*)

Lecture Theatre F, University of Surrey, Guildford, 7 p.m.

Wednesday, 18th January

JOINT MEETING WITH RTS

The Southampton Hospital broadcasting system

By G. A. Allcock (*University of Southampton*)

Lanchester Theatre, University of Southampton, 7 p.m.

Thames Valley Section

Wednesday, 25th January

Colloquium on THE DESIGN OF HI FI AUDIO POWER AMPLIFIERS

Details of this meeting were given in the October issue (page 481).

J. J. Thomson Physical Laboratory, University of Reading, 2 p.m.

Advance booking is necessary although tickets are free to members: Apply in writing to Mrs. E. R. Atkinson, Department of Cybernetics, University of Reading, 3 Earley Gate, Whiteknights, Reading RG6 2AL, Berks. **Each applicant will be limited to one ticket.**

Kent Section

Monday, 30th January

Why computer aided design?

By R. Fox (*University Computing*)

St. George's Hotel, New Road, Chatham, Kent, 7 p.m. (Tea 6.30 p.m.).

Beds and Herts Section

Thursday, 24th November

Electronic calculators—past and current technology

By C. N. Peart (*Commodore Business Machines*)

Hatfield Polytechnic, Hatfield, Herts. 7.45 p.m. (Tea 7.15 p.m.).

Thursday, 26th January

Quadraphonics

By I. Collins and C. Daubney (IBA)

Synopsis: The paper outlines the development of sound reproduction from the first gramophone systems through stereo to the various current proposals for quadraphony. Both Matrix and so-called discrete systems are considered and suggestions made for the requirements which a broadcast system would need to meet if it were to obtain widespread acceptance in Europe. Following from these suggestions some remarks are made concerning the systems which are preferred for broadcasting purposes.

Extensive use will be made of demonstrations.

Hatfield Polytechnic, 7.45 p.m. (Tea 7.15 p.m.).

East Anglian Section

Thursday, 24th November

JOINT MEETING WITH IEE

Development of miniature TV receivers

By Clive Sinclair (*Sinclair Radionics*)

Engineering Laboratories, Trumpington Street, Cambridge, 6 p.m. (Tea 5.30 p.m.).

Wednesday, 30th November

JOINT MEETING WITH IEE

Switched mode power supplies

By P. Chapman (*Marconi-Elliott*)

Synopsis: With the necessity to reduce the weight, volume and power loss of electronic equipment increasing use is being made of switched mode methods of power control.

The apparent simplicity of the switching mode technique is demonstrated by referring to the design equation of the three basic types. However deeper analysis of the circuits show subtleties that can trap the unwary engineer, and these are highlighted.

A number of less well-known configurations are also presented and the advantages of various power and control circuits are discussed. Finally consideration is given to problems, such as e.m. interference, component selection and filter resonance that inevitably occur in practical situations.

Civic Centre, Chelmsford, 6.30 p.m. (Tea 6 p.m.).

Thursday, 8th December

Tropospheric scatter communications

By B. S. Skingley (*Marconi Communications*)

Synopsis: The recent upsurge of interest in tropospheric scatter systems has been highlighted by the use of the medium to provide broadband reliable communication systems for the Oil and Gas Production Platforms in the North Sea.

The paper briefly surveys the history of the technique, and examines the characteristics of the propagation medium and the equipment features required to maximize the availability of each system. The

influence of major parameters including path length, operating frequency and antenna size on the information-carrying capacity of tropospheric scatter systems, is outlined as are the improvements obtained by diversity and combining systems.
Finally, recent innovations, which maximize the co-ordination distances between systems and reduce the spectrum requirements of networks, are described, together with the author's thoughts on future trends in tropospheric scatter systems.
University of Essex, Colchester, 7 p.m. (Tea 6.30 p.m.)

Tuesday, 17th January
Digital television—a logical choice?
By K. H. Barrett (*IBA*)
Audio Visual Centre, University of East Anglia, Norwich, 7.30 p.m. (Tea 7 p.m.).

South Western Section

Wednesday, 7th December
Underwater acoustic imaging
By Dr. S. O. Harrold (*Portsmouth Polytechnic*)
Chemistry Lecture Theatre No. 4, University of Bristol, 7 p.m. (Tea 6.30 p.m.)

East Midlands Section

Monday, 5th December
Circuit analysis by computer as an aid to electronic circuit design
By Dr. D. Boardman (*Leicester Polytechnic*)
Synopsis: Analysis of electronic circuits has always seemed an ideal subject for programming on a digital computer and indeed many circuit analysis programs have been developed. The paper will describe computer programs for d.c. and a.c. analysis of electronic circuits. The philosophy of the programs will be discussed including the formulation of the admittance matrix, the matrix reduction techniques and the modelling of active devices. Practical examples of the use of the programs for the design of transistor amplifiers and active and passive filters will be described.
Room H.08, Hawthorne Building, Leicester Polytechnic, 7 p.m. (Tea 6.30 p.m.).

Tuesday, 17th January
The engineer—a management/personnel view
By R. Palmer
Lecture Theatre, Physics Department, Leicester University, 7 p.m. (Tea in Charles Wilson Building at 6.30 p.m.).

West Midlands Section

Monday, 5th December
JOINT IEE-IERE-IPOEE MEETING
Automobile electronics
By W. F. Hill (*Joseph Lucas*)
North Staffordshire Polytechnic, Beaconside, Staffs, 7 p.m. (Tea 6.30 p.m.)).

Thursday, 8th December
JOINT MEETING WITH IEE
Multiprocessors for use in telephone systems
By G. Edge and D. Holden (*Plessey*)
GEC Telephone Works, Stoke, Coventry, 6.30 p.m. (Tea 5.45 p.m.).

Wednesday, 11th January
Technology in the service of the police
By T. H. Farr (*Home Office*)
Birmingham Polytechnic, 7 p.m.

South Midlands Section

Wednesday, 7th December
Surround sound reproduction systems
By M. A. Gerzon (*Mathematical Institute, Oxford*)
Foley Arms Hotel, Malvern, 7.30 p.m.

Monday, 9th January
JOINT MEETING WITH IEE
Design of modern radio receivers
By B. Cooke (*Eddystone Radio*)
BBC Engineering Training Centre, Wood Norton Hall, Nr. Evesham, Worcs., 7.30 p.m.

North Eastern Section

Tuesday, 13th December
Microprocessors in railway signalling
By Dr A. H. Cribbens (*British Rail*)
YMCA, Ellison Place, Newcastle upon Tyne, 6 p.m. (Tea 5.30 p.m.).

Tuesday, 10th January
The influence of LSI on digital system design
By Professor D. W. Lewin (*Brunel University*).
YMCA, Ellison Place, Newcastle upon Tyne, 6 p.m. (Tea 5.30 p.m.).

North Western Section

Thursday, 15th December
JOINT MEETING WITH IEE
Charge coupled devices
By D. J. Burt (*GEC Hirst Research Centre*)
Renold Building, UMIST, Sackville Street, Manchester, 6.15 p.m. (Tea 5.45 p.m.)

Thursday, 19th January
Radio paging
By N. W. Brown (*Post Office*)
Synopsis: A general introduction to Radio Paging is given and mention is made of the Post Office Thames Valley trial.
The paper then describes in detail the engineering aspects of the London service which has been designed to serve the Greater London Area from a fully automatic 100 000 code capacity control unit. Full STD access is afforded to persons making the paging calls which are validated and then acknowledged by recorded announcements.
Renold Building, UMIST, Sackville Street, Manchester, 6.15 p.m. (Tea 5.45 p.m.).

Yorkshire Section

Tuesday, 29th November
Advances in railway signalling
By J. W. Birkby (*British Rail*)
Room 3107, Sheffield Polytechnic, 7 p.m. a 6.30 p.m.)

South Wales Section

Thursday, 8th December
JOINT MEETING WITH IEE
New developments in automatic frequency response testing
By Dr. D. Rees (*Polytechnic of Wales*)
Synopsis: The lecture covers automatic testing of dynamic systems using multi-frequency signals and discrete Fourier transforms.
University College, Swansea, 6.30 p.m. (Tea 5.30 p.m.).

Wednesday, 14th December
ANNUAL GENERAL MEETING
Followed by
A digital private telephone exchange
By J. A. Tritton (*GEC Hirst Research Centre*)
Synopsis: This paper gives a brief review of time-pace switching networks and signalling methods, suitable for small telephone systems of up to 500 lines or so. Advances in technology have made an all-digital telephone system a practical possibility.
The paper describes such a prototype private automatic exchange featuring time-division switching and stored program control (s.p.c.) by means of a commercially available microprocessor.
Room 112, Applied Physics Department, UWIST, Cathays Park, Cardiff, 6.30 p.m. (Tea 5.30 p.m.).

Wednesday, 11th January
JOINT MEETING WITH IEE
Analysis for production control in an integrated steel works
By K. E. Morgan (*British Steel*)
Room 112, Applied Physics Department, UWIST, Cathays Park, Cardiff, 6.30 p.m. (Tea 5.30 p.m.).

Merseyside Section

Wednesday, 14th December
Electronic control and instrumentation in the brewing industry
By A. B. Harrison and J. C. McQuinn (*Bass Charrington*)
Bass Charrington, Runcorn (Bus to leave from University of Liverpool at 7 p.m.). Admission by ticket only, obtainable free of charge from Merseyside Secretary: Mr. P. Foster, 5 Berwick Drive, Liverpool, L23 7UH, before 3rd December.

Northern Ireland Section

Thursday, 1st December
Quadraphonic broadcasting
By a speaker from BBC.
Broadcasting House, Ormeau Avenue, Belfast, 7 p.m.

Thursday, 12th January
Microprocessors
By Professor W. D. Ryan (*Queen's University, Belfast*)
Ashby Institute, 6.30 p.m.