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."

THE RADIO AND ELECTRONIC ENGINEER

The Journal of the Institution of Electronic and Radio Engineers

VOLUME 39 No. 1 JANUARY 1970

'The Engineer in State and Private Enterprise'

The Presidential Address of

Harvey F. Schwarz, B.Sc., C.Eng., F.I.E.R.E.*

Delivered at the Annual General Meeting of the Institution in London on 22nd October 1969

I am most appreciative of the honour given me by our members in electing me President of the Institution. I am even more conscious of that honour when considering the list of my distinguished predecessors in this office, all of whom are well known in the world of electronics.

The standards set by my predecessors make my difficulty this evening something akin to sadism since a new President is required to justify himself from the very moment of his election. I am all in favour of the proposal of Colonel Raby that the President of the Institution should be allowed to give an account of his stewardship at the end of his term of office, rather than at the beginning.

The argument in favour of the present procedure is, of course, to enable members to hear and read something of the views of their new President. Even this approach, however, made my heart sink for my predecessors have covered almost every facet of electronics—research; development; the achievements of the profession; the role of the engineer; the increasing need for specialization—all these aspects, and many others, have been covered and, as well, their forecasts for the future have frequently been proved brilliantly prophetic.

It is when the full awareness of all this sinks home that the incoming President begins to wonder wildly what he can possibly do to escape this task which will undoubtedly reveal his inadequacy. It may even result in his establishing a highly unfortunate record—that of being elected President and being asked to resign all in the same evening!

However, as Doctor Johnson pointed out, we can be sure that when a man knows he is about to be hanged, it concentrates his mind wonderfully. And in my own case, that concentration has taken the form of a decision to break completely with the usual substance of Presidential Addresses. Instead I wish to discuss a subject which is, I believe, of paramount importance to our profession and to the growth of our industry. That subject is 'The engineer in state and private enterprise'.

How State and Private Enterprise Differ

You may recall that almost four hundred years ago, Francis Bacon said, in an essay:

> 'He that hath wife and children hath given hostages to fortune, for they are impediments to great enterprises'.

I think today one could paraphrase Bacon to the extent that:

> 'He who relies on governmental support hath given a hostage to fortune for it may well prove an impediment to a great enterprise'.

Now in so paraphrasing, I have no intention of denigrating in any way the assistance we have all had —some to a greater, some to a lesser extent—from Government. And I am well aware that a number of accomplishments could not have been achieved without such help.

* Managing Director, The Decca Navigator Co. Ltd.; Director of Decca Ltd. and The Decca Record Co. Ltd.

Let me also make it clear that the development of the radio and electronic engineering profession and industry owes much to the engineering contribution of the Armed Services, Government Research Establishments and various Ministries. Such corporations as the Post Office, Cable and Wireless, the B.B.C., etc., have also given invaluable help to industry, although in principle their efforts are primarily to help their own business.

I also acknowledge, of course, the great part played by many Universities in co-operating with industry—not necessarily, however, involving Government machinery.

Because of the different approaches of governmental and commercial enterprises to their particular business there is an incompatibility which can be, and often is, detrimental to the advancement of a project. However sympathetic or enthusiastic a government department may be for a project, it is, for instance, less concerned with the time factor than is private enterprise. This is because the very structure of the civil service business quite rightly starts by ensuring that public funds are properly spent and used to the best advantage. All too frequently this pre-requisite breeds a caution which often becomes so exaggerated that by the time the project has been surveyed from every possible aspect by a series of committees, the chance may well have been lost. Success in private enterprise can only be achieved by not only securing value for money, but also meeting competition by minimal delay in decision taking.

The essence of survival in an increasingly competitive world, and surely few worlds can be more competitive than that of electronics, is the assessment of an opportunity, the decision to grasp it and the drive to push it forward to its logical conclusion. This, of course, is much easier said than done and particularly in the face of opposition from those to whom any element of risk is anathema. But it is abundantly true, as you all know, that there are very few stone certainties in business; some element of risk is one of the facts of life and, indeed, the old saw 'No accumulation without speculation' is not too far off the mark. Equally, however, there is at times a very fine dividing line between what I might term the reasonable risk and the outright gamble, and I suppose that it is the recognition of this which distinguishes the true captains of industry.

At the risk of sounding parochial I must, in this context, illustrate my point by referring to the Decca Company and, in particular, to Sir Edward Lewis, the chairman. He has written a book called 'No C.I.C.'[1]† which, although its main purpose was to show how the steady increase of government restrictions in the post-war years would have prevented a company like Decca from surviving and growing, the book is also a wonderful account of the determination of one man not to let a business he had started fail, even when the odds against him seemed insurmountable.

If we examine such enterprise on a credit and debit basis we will see that it brings with it many more benefits than those which may be immediately apparent. For not only does its exercise permit the company concerned to operate on an unrestricted basis in an ambience in which, as I have said, time is often of the essence, but from this there is engendered a spirit of what, for want of a better term, I might describe as 'press on', which permeates the entire organization. In the laboratories and the workshops and the factories there is scope for initiative and free thought, for inventiveness and for the pursuit of the best possible end by the best possible means.

By contrast, the cushioning effect of reliance on governmental support can—human nature being what it is—stultify initiative and enterprise. Understandably, delays in obtaining state support are very often protracted and engender apathy in engineers to the detriment of the final project.

All of us have had experience of really good concepts, which have been disregarded, slowed down or allowed to atrophy. For example, consider the dollar expenditure that could have been saved and the revenue that might have been earned from exports by proceeding with the *TSR* 2. Certainly if it had been possible to have handled it as a free enterprise project it would not have been abandoned when success was virtually within grasp.

The Electronics Industry and Private Enterprise

Starting with wireless and developing into radio and electronics, the industry with which we are all associated was developed as a free enterprise. In the book 'A 20th Century Professional Institution'[2] there is also evidence of the enterprise of our founder members in pressing the then President of the Board of Education to recognize radio engineering as a discipline which should be taught in technical colleges and universities. But even in 1937 Government departments were still doubtful as to whether radio was a sufficiently important subject to warrant special faculties and departments in our educational system!

Electronics continues to offer vast scope for creating wealth. Indeed, the application of electronics to every other field of human endeavour is almost beyond comprehension. If we in Britain are to get our fair share of this huge potential, we must be given

† C.I.C. = Capital Issues Committee—set up by the British Treasury in the 'fifties to regulate company expansion.

every opportunity and seize every chance quickly and resolutely. Perhaps we may make mistakes, but there is considerable truth in another old saying: 'The man who never made a mistake never made anything'.

My view is that the engineer must be given freedom to develop his talents. I say this not simply to secure commercial glorification of an industry. As individuals we must first ensure the wellbeing of our country, and this can only be done if there is freedom to seize opportunities. If we are even to maintain, let alone improve our standard of living, we must secure a transfusion of funds from export; our ingenuity, our know-how, our manufacturing expertise must be harnessed to that end.

Britain—and I need hardly remind you that I speak as an American and can therefore be absolved of any chauvinism in what I am about to say—Britain has led the world in inventiveness in electronics, as in so many other spheres. It is, in fact, a salutary exercise for any non-British electronics engineer such as myself to survey the span of developments in this field and to see how many of them originated in, or from, this country. We do not obtain, from such inventiveness, anything like the measure of return which is our due; too often, it seems to me, others are allowed to cash in—literally—on these ideas either because they haven't been pushed hard enough by those who produced them, or because the governmental support for which they sought has taken much too long to obtain.

There are, of course, many projects which no company, unaided, can implement without governmental help. But Government and industry together must hammer out some *modus operandi* whereby we can short-circuit the delays which are seemingly inherent in attaining such support. Even if we can at least obtain a firm negative quickly it will enable us to turn our attention and endeavour to other more productive causes. Vacillation and procrastination are not only thieves of time, initiative and enthusiasm, but are export 'killers'.

Enthusiasm and Enthusiasts

Which brings me to another of the essential offshoots of the free enterprise philosophy. Enthusiasm. Now I realize that in an increasingly cynical world, enthusiasm is almost a dirty word. It is, today, generally used in the pejorative sense. To be branded as an enthusiast means so often to be categorized as a sort of latter-day social pariah. To the extent even that when one produces a dialectically unanswerable argument it may be dismissed with a light laugh and the seemingly definitive statement: 'Well, of course, he's an enthusiast'—as if that statement were a substitute for logic!

I do not know of any major achievement in man's history that was accomplished without the enthusiasm of individuals for a particular cause. By contrast, there are many, many examples of the enthusiast having to fight for his cause—through thickets of obstructionists; denigrators; cold-water pourers and the like. It is also seen that, when he had proved his point the hard way, the noise of those same erstwhile detractors leaping on the band wagon was almost deafening. Even worse, however, is the list of enthusiasts whom history has proved to have been right but who failed, despite every effort, to prevail against the forces of apathy or indifference.

Enthusiasm is the basic quality which has enabled man to conquer so many enemies; surmount so many barriers; expand so many horizons. From it has come the ability to reach the stars and, in that regard, I think that all of us can take some measure of pride in the part which electronics have played in landing men on the moon. Who could possibly have thought that less than seventy years after Marconi transmitted those first, faltering signals across the Atlantic, countless millions of people on earth would be watching, and listening to, astronauts on another planet?

And in the achievement of that end, could anyone possibly discount the part which enthusiasm played in over-riding all the myriad objections to its attempted accomplishment? Those of us concerned with the Institution's 1961 Convention on 'Radio Techniques and Space Research' recall only too well the cold charity we received from Government and other sources in organizing our programme. The Institution was described to me as 'a bunch of enthusiasts'!

A Case History: The Decca Navigator Story

One of the most outstanding examples that I have met of enthusiasm was the fundamental work of my colleague, W. J. O'Brien in his basic work of developing what is now known as the Decca Navigator System. Basically, O'Brien invented in 1937—and I would like to repeat 1937—a system for measuring the speed of aircraft flying between two synchronized radio transmitters. There is no need for me to repeat tonight the technical data as this was given by Mr. O'Brien in his paper at the Institution's first post war Convention—in May 1947—for which he was awarded the senior Premium of the Institution.[3]

On the day war was declared O'Brien started a write up of a position-fixing system based on his earlier work, and sent it to me in London as he thought it would help Britain. I very much believed in the idea and submitted a detailed proposal to the Air Ministry in October 1939. It was assessed by

Robert Watson-Watt—who certainly gave a very quick decision—which was that the system just wouldn't work! There was at that time, of course, a number of promising navigational aids in process of being developed and one of them, the Gee system, attracted Watson-Watt because it used pulses as does radar.

This turn-down meant that we were unlikely to get sufficient official support for even a transmitting licence. It was also very difficult for me to get out to America—but this is another story. Suffice to say I did manage it and with Bill O'Brien started an experimental laboratory in Los Angeles. We produced two low-powered transmitters which we set up on a farm nearby and one receiver which was installed in a motor car. We were able to demonstrate the existence of the hyperbolic lines and the very high sensitivity of the receiver indicator to any deviation from one of them.

After showing the equipment to a representative of the British Embassy, I returned to England in September 1941, and armed with proof that the system would work I again tried to interest the Air Ministry and the Navy.

It was perhaps not coincidence that the Royal Navy, which undoubtedly was the pioneer of the original wireless effort in Great Britain, gave us our first official encouragement. I cannot express adequate appreciation for the help then given by the old Admiralty Signals Establishment. In July 1942, our set-up in Britain was joined by Bill O'Brien who originally came over from America to spend three months helping us, but who has stayed with our Company ever since.

The next developments have been graphically described by Sir Edward Lewis:[1]

'In September transmitters were set up in Anglesey, a receiver installed in a Dutch trawler, and the first full-scale trial took place, the target being a lighthouse off the Isle of Man at about 50 miles range. Navigating solely with the aid of this new and revolutionary Decca equipment, the trawler moved accurately to its destination in a heavy swell and strong currents. The Air Ministry representatives commented that the system was *too* accurate for them. The Post Office representatives stated that they had something in the development stage still better. Fortunately the Admiralty decided that we must press on with further development work with all possible speed, and that we did. Six months later, in the Spring of 1943, the Admiralty carried out further trials with H.M.S. *Saltburn* up to a range of 200 miles off the west coast of Scotland. These trials were again entirely successful.

'Finally, after a full-scale invasion exercise in January 1944, the Decca Navigator had proved so successful that the vital decision was made by the Admiralty to use the system in critical operations connected with the landings of the British Army in Normandy. Special transmitting gear had to be designed and manufactured and new receivers made and tested. Time was short, but the stations were set up and the ship installations produced on time. When "D" Day arrived the Decca Navigator was there, first guiding the leaders of the minesweeping flotillas, and then guiding the first landing craft through the narrow mineswept lanes to their exact locations on the Normandy beaches with the phenomenal accuracy of some twenty yards. D. H. Toller-Bond at A.S.E., Haslemere, had worked closely with us and had been responsible for setting up the stations for "D" Day. It was an exciting time, for as soon as we picked up the continuous transmissions we knew that the invasion was under way.'

Obviously, within the limitations of this Presidential Address I have not gone into technical detail, nor paid sufficient tribute to the very closely knit team which O'Brien and I set up to develop the Decca Navigator System. I have already been invited by the Institution to give a separate discourse on the history of this venture and I hope to accept the invitation at a later date. For the present, I will confine myself to referring to two other papers which have been published in the Institution's *Journal*.[4,5] I repeat that this whole concept, which in the end played such a great part in the invasion of Europe and ending the European war, was only successful because of the determination and enthusiasm of individuals, coupled with private enterprise which managed to overcome official discouragement.

This very personal instance—not the only one I have experienced, is the reason for my putting forward this thesis which urges the need for the enthusiastic engineer to be given freedom to exploit his natural talents. Discovery cannot be made by order or regulation, nor can the engineer's ingenuity be trammelled by limitation of opportunity.

The Fallacy of Mergers

I deeply regret that my generation has also seen the birth of another phenomena likely to hamstring the enthusiastic engineer or indeed, deter the recruitment of young people into our profession. I describe this phenomena as 'Gadarene Swine-manship'. This manifests itself in a seemingly almost frenetic compulsion—not to dash over a cliff as did those unfortunate animals—but to huddle together in a larger and larger conglomerates in a positive fury of gregariousness. And, by a curious paradox, the more entities such conglomerates absorb, the more amorphous they become.

However satisfying such amalgamations may be in terms of high finance, I cannot believe that they always lead to greater efficiency or to the ultimate good of the entire industry. It really saddens me to see such famous names as de Havilland and Vickers disappearing like so many Excaliburs in a sea of take-over bids. What is even more serious is that the competitive spirit may also be submerged in the welter of mergers. The cynical maxim: 'If you can't beat them, join them' brings us, like most cynicism, to a dead end, for it leaves unanswered the highly pertinent question: 'What happens after you HAVE joined them?'

Elimination of competition creates fertile ground for the procreation of time delay, decision making committees, rejection of the enthusiast, and all the dangers inherent in governmental control. There is, moreover, an added danger; that of discouraging, if not suppressing, the acorn of an enthusiastic idea because mammon is too big to worry about cultivating the oaks of the future.

And yet, was it not the acorn of broadcasting and simple wireless receivers made by small competitive firms which led to the international oak forest of the twentieth century electronics industry? Have we already forgotten the pioneer work of the British Broadcasting Company, the danger of monopoly when it became the British Broadcasting Corporation, and the public demand for and political recognition of the Independent Television Authority? What was deemed to be in the public interest is surely applicable to giving the qualified engineer wide and competitive scope in which to prove his ability and ideas.

Obviously there are areas of engineering development or research where a combination of effort can be, and is, to the overall good, and I will be referring in a moment to the means by which this can be further improved. I hope, however, that I have indicated the problems which will arise if there is a continuous process of mergers and 'take-overs'. It will result in feather-bedding which may stifle initiative and enthusiasm, and the pride in belonging to a close knit-team, as against being merely one of a vast throng which is a breeding ground for private pique and a human rat race.

The profession of radio and electronic engineering has grown and thereby thrived on inventiveness, initiative and enthusiasm. Like liberty, the price of these, it seems to me, is unceasing vigilance. We must ensure that we preserve the ambience, the milieu, in which these qualities may continue to flourish, for with them we can achieve almost anything, and without them very little of real progressive worth.

Some Solutions to the Problems

Having recognized the problems facing electronic engineers in State service and in private enterprise, I would like to adopt engineering practice by suggesting solutions to the problem. First, that we must continue to persuade Government actively to support any engineering venture which either helps the defence services or is of benefit to the nation as a whole, but which is of a financial magnitude which cannot fairly be borne entirely by industry.

Secondly, that we should ourselves be aware of the future requirements of such Corporations as broadcasting, Post Office, etc., so that industry can meet their needs.

Other than these two basic facts, there must be encouragement for competition in private industry so that the incentive is there to develop new ideas and to encourage the enthusiast. In other words, within private industry there should be the type of organization that can foster the acorn to which I have already referred.

Last—perhaps it should have been put first—is the need to understand and help the requirements of industry at large and the public—in other words, the consumer. In the end it is always public demand that wins and private enterprise must be left free to meet public demand.

The Role of N.E.C.

Now all of these requisites that I have mentioned have, in fact, found expression in the formation of the National Electronics Council. Since the Institution has given so much support to N.E.C., and it is moreover headed by one of our most illustrious Past Presidents, I hope I may be forgiven in saying that my admiration of N.E.C. is only disturbed by my own feeling that N.E.C. itself is in danger of being hamstrung.

I would first remind you that N.E.C. was not the creation of Government or Corporation. It was the expressed desire of private enterprise to meet with Government and Universities to discuss common problems in the electronics field. Very wisely, the scope of N.E.C. has been widened to incorporate major users of electronic equipment and even extends to helping industry and Government in the problems of recruitment to the engineering profession.

But is this enough? Here we have an independent organization, assisted by Government and supported by all the representatives of industry, universities, etc., that I have already mentioned. It is never in possession of the complete picture of electronics economics and development because there are numerous other councils or committees, all Government originated, which are supposed to have some concern with the wellbeing of industry and therefore of the engineers employed in it.

The National Electronics Council, under the Chairmanship of Lord Mountbatten, has been the broker to arrange a 'marriage' between Government, universities and industry on such things as selective dissemination of information, colour television, computer usage etc. Is it, however, a complete 'marriage'? Or will it not be so until it also gathers within its counsel all the things which will help—such as education, manpower, economic viability, exports, etc. Surely all of these things must come within the framework of N.E.C. so that there is continuity within both Government and industry on policy matters which affect the same interests sitting around the same table to argue out problems which not only affect the wellbeing of the electronics industry, but the wellbeing of Great Britain.

I can only hope that with the power of N.E.C. and all the other Government Committees concerned with electronics, we shall make better progress. Electronics is a young and very specialized business; neither in Government nor in the conglomerations to which I have referred should we put back the clock by regarding electronics as a tiny little off-shoot of some other major interest. Electronics is a major faculty, a major industry, and a vital necessity to the country's economy. It is imperative that we do not dampen the enthusiasm which is the specialization of the radio and electronic engineer.

The Engineer and his Contribution

I believe that the Chartered Engineer recognizes his ability to contribute to the wellbeing of the community. His contribution is, however, hindered by economic resources which are so easily found in war but apparently so hard to find in peace. If therefore, we engineers are to live up to our Charter to advance science for the benefit of mankind, we must wrestle with the problem of peace-time economics. Otherwise we shall remain in our present apathetic state with an inability to encourage initiative and end by being only assemblers of the ideas of other nations.

It is perhaps in this spirit more than any other that I believe we must not be afraid to evaluate the problem of the engineer in Government and the engineer in private enterprise. Without the manufactured products of the engineer, without the produce of the farmer, or other indigenous effort, there cannot be the profit by which Government obtains its necessary revenue. I argue, therefore, that Government must do more to encourage engineers to reap reward for their efforts in private industry and thereby provide the basis of a happy marriage between the nationalized or neo-nationalized effort and the enthusiasm of the private sector. Only in this spirit can we accomplish as much in peace as we do in war.

It is my constant hope that together, irrespective of race, colour or creed, engineers will find a way to give every man full opportunity to use his abilities for the benefit of all people.

References

1. Lewis, E. R., 'No C.I.C.' (Universal Royalties, London, 1956).
2. 'A Twentieth Century Institution: The Story of the Brit. I.R.E.', (Institution of Electronic and Radio Engineers, London, 1960).
3. O'Brien, W. J., 'Radio navigational aids', *J. Brit. Instn Radio Engrs*, **7**, No. 5, pp. 215–46, October 1947.
4. Roberts, G. E., 'The design and development of the Decca flight log', *J. Brit. Instn Radio Engrs*, **12**, No. 2, pp. 117–31, February 1952.
5. Powell, C., 'Dectra: a long-range radio-navigational aid', *J. Brit. Instn Radio Engrs*, **18**, No. 5, pp. 277–92, May 1958.

(*Address No. 42*).

Harvey Fisher Schwarz was born in Edwardsville, Illinois, U.S.A., in 1905, and studied at Washington University, St. Louis, graduating with a B.Sc. degree in electrical engineering in 1926. He then joined the General Electric Company, Schenectady, as an engineer, and in 1928 went to the Brunswick Radio Corporation as Assistant Chief Engineer, later becoming Chief Engineer. When Warner Brunswick Ltd. was acquired by the Decca Record Company Mr. Schwarz was made Chief Engineer of the new Brunswick Ltd. He was made Technical Director of Decca Radio and Television Ltd. on its formation in 1938.

In 1939 William O'Brien (Fellow), an old friend of Mr. Schwarz, invented a c.w. hyperbolic navigational system which was taken up by Decca, and Mr. Schwarz was closely concerned with its development and application for use by the Royal Navy. After the war the Decca Navigator Company was formed and Mr. Schwarz became its Managing Director in 1950, a position he still holds. In May Mr. Schwarz and Mr. O'Brien received the Pioneer Award of the Institute of Electrical and Electronics Engineers (U.S.A.) in recognition of their pioneer work on the Decca Navigational System. Mr. Schwarz is one of the representatives of the Conference of the Electronics Industry on the General Committee of the National Electronics Council.

Mr. Schwarz was elected a Fellow of the Institution in 1953 and was appointed to the Finance Committee in 1957. He was elected to Council in 1959 and first elected a Vice President in 1965.

U.D.C. 621.391:681.327.64

Possibilities of a Sinusoidal Memory for an Extendable Cybernetic Machine

By

J. F. YOUNG,
C.G.I.A., B.Sc., C.Eng., M.I.E.E., M.I.E.R.E.†

The possibility of using sinusoidal signals in the memory circuits of a cybernetic associating-machine is examined. Possible methods and the associated difficulties are discussed, and the future possibilities of such machines are considered.

List of Principal Symbols

S_a, S_b	stimulus signals
s_a, s_b	effect signals
$R(S_a, S_b)$	recorded signals
$P(A)$	probability of occurrence of A
$P_A(B)$	probability of occurrence of B, given A
$P(A \& B)$	probability of joint occurrence of A and B
N	number of inputs
f_a	frequency a
ω_a	angular frequency $2\pi f_a$
$V \sin \omega t$	sinusoidal voltage, amplitude V
K	constant
V_1	input signal
V_0	output signal
A	forward gain
B	feedback gain
T	delay time.

1. Introduction

When examined closely, much of the work on trainable learning machines[1] is seen to belong to a single class. This might be called the 'adjustable-weight, majority logic' approach. Such systems are not well adapted to operation under changing conditions once the training period is over.[2]

The machines envisaged in the present research use a different approach, based on a simulation of Pavlovian conditioning by association.

2. Basic Requirements

The initial basic requirements of the associatory learning machines envisaged here are:

(1) Some means of detecting coincidences between the occurrences of a number of input stimuli. The machines should preferably be capable of extension to operation with large numbers of inputs.

(2) Means of recording these coincidences. The means adopted should ensure that there cannot be confusion between the recordings corresponding to coincidences between different pairs of stimuli.

(3) Means of making use of any recorded coincidence if any of the input stimuli occur in the future.

(4) A machine should operate on a probabilistic basis. It is not required that, once stimulus signals S_a and S_b have coincided, stimulus S_a should always inevitably produce effect s_b corresponding to stimulus S_b. Rather it is required that the probability of production of effect s_b by stimulus S_a should increase with the past frequency of occurrence of the coincidence (S_a, S_b).

(5) An additional desirable factor would be that older recordings should decay slowly, so that they are of less importance than are more recent recordings. In effect, such a process would help to avoid overloading of the memory. It is not essential to incorporate such a 'forgetting' process into initial work, though the eventual need must be kept in mind.

(6) In a similar way, it might be required later to introduce an inhibitory process.

It should be noted that freedom from error after training is not taken as a requirement. In the biological case, the occurrence of occasional small errors of operation appears to be an essential feature of the continuous re-training which is required in order to cater for slow changes of environment.[3]

3. Detailed Statement of Requirements

In an associatory learning machine, if input signals S_a and S_b occur simultaneously, then a record $R(S_a, S_b)$ must be made. This record must be kept available so that if signal S_a occurs alone at a future time, then the effect s_b of signal S_b can be produced even though signal S_b is not actually occurring.

The recorded signal should be characteristic only of the associated input signals. For example, if the input

† Department of Electrical Engineering, The University of Aston in Birmingham, Birmingham 4.

signals are $S_a, S_b, S_c, S_d \ldots S_z$, then it must be true that the recorded signals

$$R(S_a, \ldots S_b) \neq R(S_x, \ldots S_y)$$

where $S_x, \ldots S_y$ are any signals other than $S_a, \ldots S_b$. It should be noted that this relationship must be fulfilled even if, to take a simple two-member example $x = a$ but $y \neq b$. In this case again, we must have:

$$R(S_a, S_b) \neq (R(S_a, S_y).$$

Thus any particular recorded coincidence $R(S_a \ldots S_b)$ must be exclusive to a particular set of signals $S_a, \ldots S_b$ and must never correspond to any other set, including any set which differs from the original set by only one member.

It is an additional requirement that each and every subset coincidence should be separately detected and recorded, since each of these provides information which might be required by the machine at some future time. Indeed in the case of an organism, subset coincidence might be of vital future interest. Consequently it is necessary to arrange that

$$\begin{aligned} R(S_a, S_b, S_c, S_d, \ldots S_z) = {} & R(S_a, S_b) + R(S_a, S_c) + \ldots \\ & \ldots + R(S_a, S_z) + R(S_b, S_c) + \\ & + R(S_b, S_d) + \ldots R(S_b, S_z) + \\ & + R(S_c, S_d) + \ldots R(S_c, S_z) + \\ & + \ldots + R(S_y, S_z). \end{aligned}$$

It is sufficient to ensure that associations are recorded in pairs. This will include associations in threes, fours, etc., provided that there is a high probability that each and every pair association is always recorded. One effect of this requirement is that the device is then capable of pattern completion or of operating on incomplete patterns of input stimulation.

An example will help to illustrate the requirements and difficulties. Suppose that there are four possible input signals S_a, S_b, S_c, S_d, and that the following associations occur:

$$R(S_a, S_b), \quad R(S_a, S_c), \quad R(S_a, S_c, S_d).$$

Now if all subset pairs are separately recorded, then

$$R(S_a, S_c, S_d) = R(S_a, S_c) + R(S_a, S_d) + R(S_c, S_d).$$

If this condition is met, then at some point in the future input signal S_a appearing without input signal S_c will be capable of producing association S_d, so that wrong information is stored. Signal S_a should not produce signal S_d unless signal S_c is also present. Thus it is necessary to ensure that

$$R(S_a, S_b, S_c, \ldots) \neq R(S_a, S_b) + R(S_a, S_c) + \ldots.$$

It should be noted, however, that if this requirement is rigidly enforced then the machine will inherently only be capable of a minimal amount of pattern-completion.

It can be seen that there appear to be conflicting requirements:

(1) All subset pairs should be separately recorded, since every pair can provide useful information for pattern-completion.

(2) Subset pairs should not be separately recorded, since it is possible for a pair to provide incorrect information.

In the animal, it is possible that the conflict is resolved by the three features of probability, inhibition and forgetting which were mentioned in the previous section.

Since the best method to resolve the conflict in an associatory learning machine is not known at this stage, it is desirable that all three features should be capable where necessary of separate introduction, modification and elimination at a later date. Comparative tests can then be carried out.

4. Conditional Probability

It is necessary to consider cases such as:

```
A ******
B *********************
  ——S——————————T——
```

where the * indicates occurrence. It is sometimes suggested that, since at time T

$$P_A(B) \neq P_B(A),$$

then these two probabilities should be recorded separately. No doubt from a simple and purely mathematical point of view this is correct, but such a view ignores the physiological phenomenon of 'forgetting', which seems to be a vital action in animal systems. If the point T is far enough removed in time from point S, the early association between A and B would be almost completely forgotten by any animal system. We have no information on the exact law of forgetting, but it is tempting to postulate an exponential decay of memory traces.

Now at point S, $P_A(B) = P_B(A)$

but at point T, $P_A(B) < P_B(A)$.

Thus any scheme involving the use of such conditional probabilities necessitates the provision of separate storage facilities for $P_A(B)$ and $P_B(A)$, and it gives a unidirectional form of storage.

To avoid such a preferential form of memory, and to give an improved economy of storage, it is suggested that the memory system ought simply to store $P(A \& B)$. Then at point S above, $P(A \& B) = 1$

but at point T above, $P(A \& B) = 6/21$,

i.e. $P(A \& B)$ has decayed because of non-occurrence of the coincidence. Thus it is proposed that storage facilities are provided not for conditional (or unidirectional) probabilities, but rather for joint (or bi-directional) probabilities.

If there are N inputs $A, B, C, \ldots N$, then the conditional or unidirectional probabilities are

$$P_A(B), P_A(C), P_A(D), \ldots P_A(N),$$
$$P_B(A), P_B(C), P_B(D), \ldots P_B(N),$$
$$P_C(A), \ldots \text{etc.}$$

There are therefore $N(N-1)$ unidirectional probabilities to be stored.

However, the bidirectional probabilities are

$$P(A \& B), P(A \& C), P(A \& D), \ldots P(A \& N),$$
$$P(B \& C), P(B \& D), \ldots P(B \& N),$$
$$P(C \& D), \ldots \text{etc.}$$

There are therefore $N(N-1)/2$ of these bidirectional probabilities for which storage provision must be made. The storage capacity ratio between bidirectional and unidirectional methods is therefore one-half.

5. Learning and Forgetting Curves

If it is assumed that the quantity to be stored is $P(A \& B)$, then it is possible to draw learning and forgetting curves for different circumstances by plotting the value of $P(A \& B)$ against time. This is done in Fig. 1. It is assumed here that the stimulus B is repeated continuously while stimulus A appears as a short train of stimuli. The shape of the resulting probability curve depends on the instant of occurrence and the length of pulse train A. Transfer from a learning curve to a forgetting curve takes place at the end of the A pulse train.

It should be noted that the relationship between the probabilities $P_A(B)$ and $P(A \& B)$ is given by

$$P_A(B) = \frac{P(A \& B).}{P(A)}$$

Now the quantity $P(A)$ is always fractional or equal to unity, so it follows that

$$P_A(B) \geqslant P(A \& B).$$

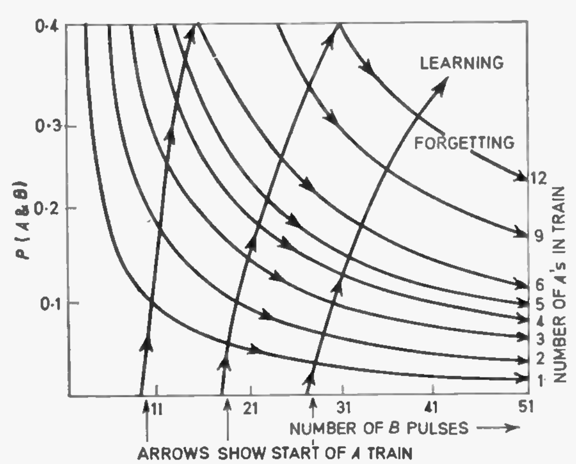


Fig. 1. Learning and forgetting curves.

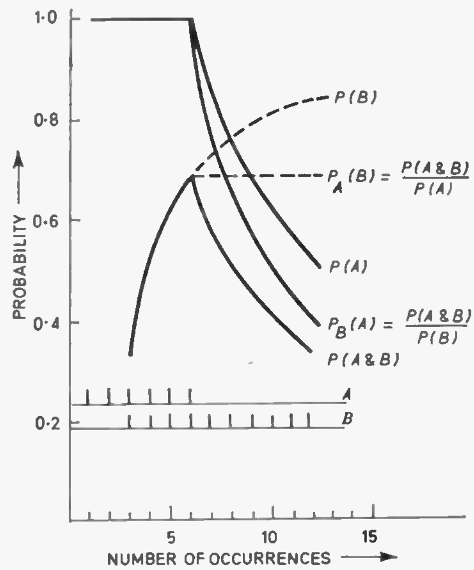


Fig. 2. Probability relationships for a particular case.

The curves of Fig. 2 illustrate the relationships for a practical case.

It will be desirable to introduce such probability relationships into the operation of the machines envisaged here. While it is not absolutely necessary to incorporate these arrangements at the commencement of the work, the methods adopted must not prevent such incorporation at a later stage.

6. Use of Sinusoidal Recording

In early attempts to produce an associatory learning machine, the writer made use of the properties of non-linear electrical circuits. For example, if two sinusoidal voltages having different angular frequencies are added together and applied to a squaring element such as a Hall crystal,[4] then only one of the output terms has a low angular frequency. This term can easily be separated out by filtering. The fact that such filtering is necessary if the low-frequency term is to be separated out might make it unnecessary to use expensive squaring devices such as Hall plates and make it possible instead to use any simple non-linearity which will introduce cross-modulation.

Now consider a number of sinusoidal oscillators each producing a different angular frequency ω_x, whenever it is energized by the occurrence of an input. If the outputs at frequencies ω_a and ω_b of any two oscillators are applied simultaneously to a non-linear modulator, then a modulator output containing the angular frequencies $2\omega_a$, $2\omega_b$, $\omega_a+\omega_b$

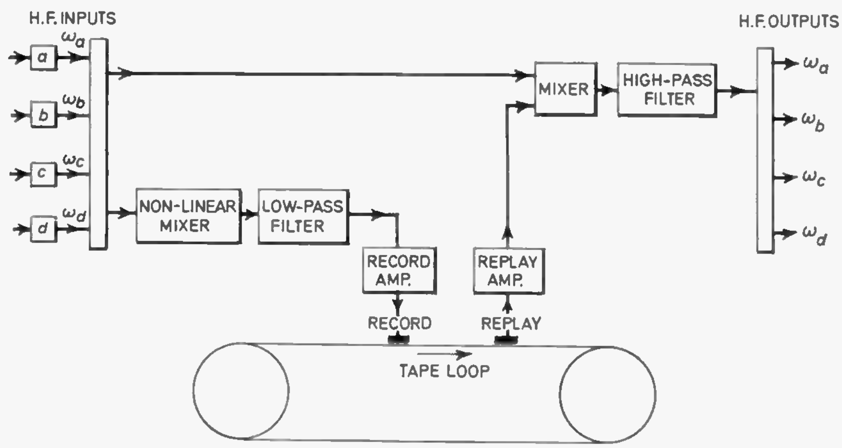


Fig. 3. Block diagram of association-recording machine.

and $\omega_a - \omega_b$ is obtained. The latter, low-frequency term can be filtered out from the rest.

6.1. Possible Association-recording Machine

The basis of a possible association-recording machine based on the use of this simple non-linear modulator principle is shown in Fig. 3. If, for example, oscillators *a* and *b* are energized simultaneously, producing signals at angular frequencies of ω_a and ω_b, then a signal of frequency $\omega_a - \omega_b$ is recorded on the magnetic tape loop. The signal $\omega_a - \omega_b$ can then be picked up from the tape via the replay arrangements at any time in the future to indicate that the two oscillators *a* and *b* have been simultaneously energized at some time in the past.

Before considering further how the recorded information might be used, it will be as well to consider obvious difficulties with the system described. There are two main problems:

(1) The recorded information can be ambiguous if the frequencies $\omega_a \ldots \omega_z$ are not carefully chosen. For example, if $\omega_c - \omega_b = \omega_y - \omega_x$, then there is no way of distinguishing between recorded signals indicating an association of *c* with *b* and those indicating an association of *x* with *y*. In order to avoid such ambiguity, a careful choice of oscillator frequencies is necessary, and this choice is not easy if a large number of inputs has to be handled.

(2) In the system as shown in the diagram, it is necessary to record newly-acquired information on to the magnetic medium directly over older recorded information, without first erasing the latter. This procedure introduces problems of partial erasure of older information and of loss of information due to tape saturation, and it is therefore not a desirable mode of operation.

In order to avoid the first difficulty, it is necessary to construct tables of numbers having exclusive differences (or possibly exclusive sums). To illustrate the problem, a simple difference table, ignoring harmonic effects, is given below.

f	Differences						
10							
11	1						
13	2	3					
17	4	6	7				
22	5	9	11	12			
30	8	13	17	19	20		
40	10	18	23	27	*29*	30	
54	14	24	32	37	41	43	44
69	15	*29*					

The above difference table has been constructed by inserting at the beginning of each line of differences the first difference integer which has not previously been used in the table. Inevitably, eventually one encounters a number which has previously appeared and which is not therefore available. The final entry (29) in the above table is such a case, 29 having been used two lines previously.

In the above example, there are only eight integral values of f which can be used from the 44 equally spaced frequency channels nominally available. There are different ways of constructing such tables. For example, the second line of differences could have started with 3 instead of 2, giving:

f	Differences						
10							
11	1						
14	3	4					
19	5	8	9				
25	6	11	14	15			
32	7	13	18	21	22		
42	10	17	23	28	31	32	
44	2	12	19	25	30	33	34

Here the difference 2 has been introduced near to the end of the table. Eight channels are now available from only 34 instead of the 44 channels required by the previous table.

It is of interest to consider the effective channel utilization in the cases considered. The utilization can be defined as

$$\text{channel utilization} = \frac{\text{number of usable channels}}{\text{number of available channels}}$$

Then for the first case considered above:

No. of channels	1	3	7	12	20	30	44
Usable channels	1	3	6	10	15	21	28
Utilization %	100	100	86	83	75	70	63

while for the second case above:

No. of channels	1	4	9	15	22	32	34
Usable channels	1	3	6	10	15	21	28
Utilization %	100	75	66	66	68	66	82

Consider yet another possible difference table:

2					
3	5				
4	7	9			
6	10	13	15		
8	14	18	21	23	
11	19	25	29	32	34

For this case:

No. of channels	2	5	9	15	23	34
Usable channels	1	3	6	10	15	21
Utilization %	50	60	66	66	65	39

6.2. Discussion of Difference Tables

The difference tables given have been constructed empirically, and such a process could be programmed on to a digital computer. However, it was felt that such a fundamental problem of channel selection should be avoided at this stage of the work. One requirement is the eventual extension of the number of available channels at least to twenty and preferably beyond, while limiting the necessary total number of channels. It would also be desirable if possible to limit the rapid increase of the values in the first difference column in order to equalize the spacing of the channels as far as possible. If the original frequencies f could always be equal in value to the logarithms of prime numbers, then they would always have exclusive sums. Unfortunately however, such frequencies would be non-integral. In addition, the channel spacing would decrease as the numbers increased. These facts make the log (prime) approach of little value even though published tables of such numbers exist.[5]

The absolute values of frequencies f would of course be of little importance, since the table of differences is the real basis of construction. In order to use other frequencies, the whole table can be multiplied by a constant, or a constant can be added to the f column.

It should perhaps be mentioned that a fairly complex modulation method is required for direct utilization of the simple difference tables given above.

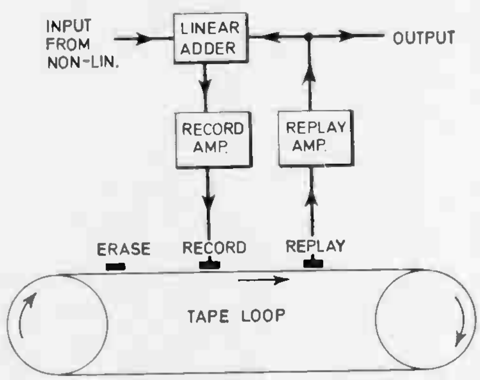


Fig. 4. Basis of more complex association-recorder.

7. Continuous Re-recording

The recording of newly-acquired information over previously-recorded information introduces difficulties as mentioned earlier. In order to avoid such problems, the more complex arrangement shown in Fig. 4 was considered and tried.

Here, the information recorded on the tape is picked up by a replay head. In addition to being taken to the output for external use, the replayed information is added linearly to new inputs and the sum is recorded. An erase head ensures that the tape is magnetically 'clean' before it reaches the recording head. With this system, the entire information about the past action of the machine is stored on the short length of tape between the record and replay heads. Such an arrangement can be thought of as being analogous to the neural recirculating loops believed to form part of the short-term memory system in animals.

7.1. Electronic Recirculating Storage

Consider a positive feedback system having an input signal V_1, a forward transfer function $A \exp(-sT)$, and output signal V_0, and a positive feedback function B. For such a system, the real frequency response is given by

$$\left|\frac{V_0}{V_1}\right| = \sqrt{\frac{1+A^{2(1+N)}B^{2N}-2A^{(1+N)}B^N \cos \omega NT}{1+A^2B^2-2AB\cos \omega T}}$$

Here, because of the finite bandwidth of any practical system, it has been assumed that terms of the expansion above $\exp(-(N-1)sT)$ can be ignored.

In the special case where $A = 1$ and $B = 1$, the relationship simplifies to

$$\left|\frac{V_0}{V_1}\right| = \left|\frac{\sin \omega NT/2}{\sin \omega T/2}\right|$$

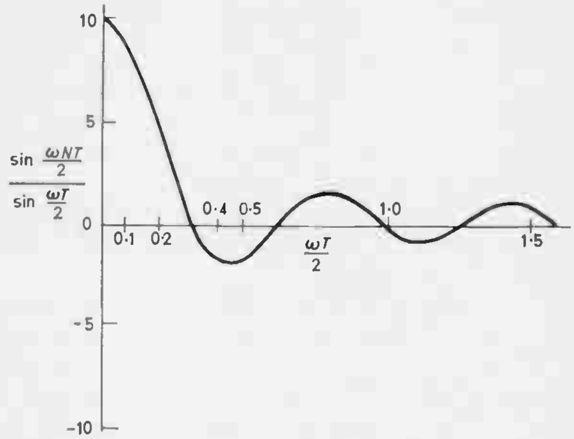


Fig. 5. Response of limited-bandwidth recirculating loop.

To illustrate this relationship, a curve showing the function for the special case where $N = 10$ is plotted in Fig. 5.

Unavoidable non-linearities cause signal degradation in a closed-loop recirculating system. Saturation must cause eventual distortion of stored signals if the overall loop gain exceeds unity. If the loop gain is less than unity, then the recirculating information is subject to decay. Although the decay is a desirable feature, if it is not to be excessively rapid, then it is necessary to maintain the value of the loop gain only very slightly less than unity.

Additional difficulties can be introduced by the fact that, as shown above, the response of a sinusoidal recirculating scheme is not at all constant with frequency.

7.2. The Effect of Noise

In a recirculating system of the type discussed above, it is of interest to consider the ratio:

$$\frac{\text{r.m.s. output noise}}{\text{r.m.s. input noise}} = (1-A^2B^2)^{-\frac{1}{2}}$$

This ratio is plotted against the value of AB in Fig. 6.

It can be seen from this curve that if the loop gain AB approaches unity, then the r.m.s. noise is increased excessively by the positive feedback. Since the loop gain is required to be only very slightly less than unity in this system, the method of direct application of positive feedback to a linear system cannot be regarded as a desirable way of remembering occurrences.

8. Practical Work with Sinusoidal Re-recording

Preliminary investigation of the possibilities of the continuous re-recording method was carried out using two different forms of commercial magnetic recorder. One was a magnetic disk recorder which was intended originally as an early 'portable' dictation machine. An additional recording head was fitted to one of these machines, together with additional electronic amplification. This machine had limitations in that the driving motor was clockwork and the amplifiers were battery operated; nevertheless, with this machine preliminary experience of the difficulties encountered with the continuous re-recording method was obtained.

Some further work was carried out using a low-cost commercial magnetic tape deck and amplifiers. With this equipment, which was operated from an a.c. supply, the severe problem which can be caused by interference from external fields was encountered. Even small amounts of interference can be greatly enhanced in a linear recirculating scheme such as this.

The experience gained with these early systems has had a large influence on the form of machines now being investigated by the writer.

9. Conclusions on Continuous Re-recording

The experience gained with attempts to use the method of continuous re-recording of linearly added sine waves produced the following conclusions.

(1) It is impracticable simply to recirculate information in sinusoidal form using magnetic recording, since the unavoidable distortion is cumulative. However, the sine waves could possibly be reshaped by the insertion of filters into the loop. This possibility has not yet been fully investigated.

(2) Great care is necessary with a recirculating scheme to avoid the cumulative effects of stray induced voltages. Once again, there are possibilities of minimizing such effects by use of filters in the loop.

(3) It would be necessary to develop some form of fast-acting automatic loop gain control to hold the loop gain very slightly less than unity. There is the possibility here of using a master pilot tone in an automatic gain control system.

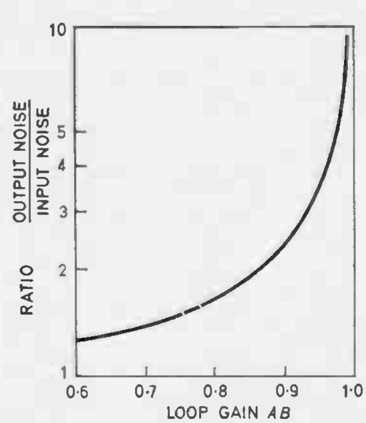


Fig. 6. Variation of noise with loop gain.

(4) The selection of frequencies for the primary sources is not easy, and this limits the possible amount of storage in any one system.

10. Acknowledgments

The writer wishes to thank Professor J. E. Flood for facilities for the present research, the Science Research Council and the National Research Development Corporation for their support, and Mr. N. Hollingshead for his valuable assistance with the experimental work. The work forms part of the 'ASTRA' research programme being carried out in the cybernetics research laboratory of the University of Aston.

11. References

1. Nagy, G. J., 'Adaptive devices and systems', *Electro-Technology*, 77, pp. 40–9, 1966.
2. Young, J. F., 'Cybernetics', *Brit. Commun. Electronics*, **12**, pp. 302–5, 1965.
3. Young, J. F., 'Cybernetics' (Iliffe, London, 1969).
4. Young, J. F., 'Electronic relays for power system protection', *Elect. Rev.* **179**, pp. 352–4, 1966.
5. Carmichael, R. D. and Smith, E. R., 'Mathematical Tables and Formulae', p. 210 (Dover, New York, 1962).

Manuscript first received by the Institution on 26th March 1969 and in final form on 10th July 1969 (Paper No. 1295/IC70).

ASTRA—A Cybernetic Machine

The first part of a programme of research in cybernetics has just been completed in the Cybernetics Laboratory at the University of Aston. The research is concerned with methods of constructing cybernetic machines which can learn in a similar way to that adopted by animals and humans. The Russian scientist Pavlov demonstrated that, in the animal nervous system, one external stimulus could be associated with another. Much, if not all, of human and animal learning activity can be explained on the basis of this simple associatory activity.

The object of the research programme at Aston is the construction of machines which learn by association and overcome the limitations of earlier machines by introducing methods of association which can be extended for use with large numbers of input 'nerve cells'. The prototype machine now completed, known as ASTRA, has 100 inputs from its 'nerve cells' and ten outputs to its 'muscles'. The nerve cells can take the form, for example, of touch sensors, of artificial retinas for visual sensing, or of artificial ears for sound sensing. The muscles can take the form of, for example, electrically operated actuators. Such peripheral devices are under investigation by students in the Cybernetics Laboratory at Aston.

The successful operation of the prototype machine ASTRA Mk. 3 is now to be followed, finance permitting, by the construction of a more extensive machine capable of handling the inputs from several thousands of nerve cells and of controlling the outputs to several hundreds of muscles. At the same time as this new ASTRA Mk. 4 machine is being constructed, it is intended to continue investigations into the operation of the present ASTRA Mk. 3 machine.

Features which can be added in varying degrees to these machines are those of probabilistic memory, of inhibition, and of forgetting. These features all appear in animal and human nervous activity, and one aim of the present work has been to ensure that the engineering methods adopted do not prevent the easy introduction of the additional features experimentally at a later stage.

Eventually, this sort of work is expected to reveal possible methods by which a given sound or word is recognized by animals or by humans. The word 'five', for example, can be recognized even when it is spoken in different accents or when it is degraded in various ways, such as by transmission over the telephone. It is known that severe degradation of a verbal signal does not reduce its intelligibility excessively, but it is not known why this should be so.

Investigation of such a problem using animals and humans as subjects is difficult, because by their very nature such subjects are adaptable, and they can learn to operate in difficult circumstances. With the machines envisaged, it will be possible to eliminate this facility for adaptation whenever required, so that repeatable experiments can be performed.

Because of the introduction of probability, these machines will not necessarily be infallible in action. It appears to be one basic principle of biology that 'to err is human'. Only by making occasional mistakes is it possible for an organism or for a machine to learn and to re-learn from experience.

New techniques of engineering are expected to appear as natural by-products of the research. Examples are speech-operated machine tools and speech-operated typewriters. Investigation has also started into the possibilities and difficulties with mobile learning machines of the 'housemaid' type.

INSTITUTION NOTICES

New Year Honours

The Council of the Institution has congratulated the following members whose names appear in Her Majesty's New Year Honours List:

To be Companion of the Most Honourable Order of the Bath (C.B.):

Henry Edward Drew (Fellow)
(Mr. Drew is Director General of Quality Assurance, Ministry of Technology; he is a past member of the Council of the Institution and was chairman from 1960 to 1964 of the Membership Committee, to which he was first appointed in 1946.)

To be an Officer of the Most Excellent Order of the British Empire (O.B.E.):

Leonard Frederick Mathews (Fellow)
(Mr. Mathews is Director and General Manager (Midlands) of Associated Television Network Ltd.; he was for several years examiner for Television in Part 5 of the Graduateship Examination.

To be a member of the Most Excellent Order of the British Empire (M.B.E.):

Thomas Elwin Allon (Member)
(Mr. Allon is Engineer-in-Charge of the B.B.C. Monitoring Station at Caversham.)

Management and Economics in the Electronics Industry

The provisional programme of the international symposium on management and economics in the electronics industry to be held in the University of Edinburgh from 17th to 20th March 1970 has now been published.

Keynote addresses will be given in the main sessions at the symposium on: marketing, A. L. Humphreys, managing director, International Computers Ltd.; management of innovation, S. Shima, managing director, Sony Corporation, Japan; management services, A. Chargueraud, president, Diebold Europe SA, France; manpower and training, F. Metcalfe, chief education and training officer, E.I.T.B.; the role of governments, Lord Beeching; economics and scale of manufacture, C. H. Villiers, managing director, Industrial Reorganisation Corporation. Some ninety papers will be presented in three parallel sessions at the symposium in addition to the keynote addresses, and will include contributions from France, Israel, Norway, Switzerland and the U.S.A. The final session will be on management/unions/shop floor relationships.

The symposium is being organized by the Institution of Electrical Engineers in association with the Institution of Electronic and Radio Engineers, the Institute of Physics and the Physical Society and the Institute of Mathematics and its Applications; thirteen other professional and government organizations and learned societies are supporting the symposium.

Further details of the symposium and registration forms are now available from the Manager, Conference Department, I.E.E., Savoy Place, London, W.C.2.

Conference on Laboratory Automation—Preliminary Announcement

The Institution, with the association of other institutions and societies, is planning a Conference on Laboratory Automation, to be held in London on 10th–12th November 1970. The subject will embrace the two main themes of Automatic Analysis and Computer Controlled Experiments and the definition of a 'laboratory' is being deliberately interpreted widely to include research and development in any science or technology, as well as routine measurements of an analytical character, for instance. Production and process control and measurement are excluded.

Further details of the Conference, including a 'Call for Papers' will be published shortly. Meanwhile offers of papers falling within the scope indicated may be submitted, preferably accompanied by a synopsis, to the Conference Registrar, I.E.R.E., 9 Bedford Square, London, W.C.1.

Change of Date of Meeting

The London meeting of the Management Techniques Group announced for 11th February at which a lecture on 'Management Effectiveness for Engineers' is to be given by M. H. Makepeace, has been postponed. The meeting will now take place on Tuesday, 10th March, at 6 p.m. in the I.E.R.E. Lecture Room at 9 Bedford Square, London, W.C.1.

Director's Visit to India and Middle East

The Director of the Institution, Mr. Graham D. Clifford, C.M.G., is visiting the Institution's offices in Bangalore and the various Zones of the Indian Division—Delhi, Calcutta, Bombay and Madras—during February. Details of his itinerary are being sent to members in India. Mr. Clifford will break his journey to India in Bahrain where he has been invited to address the Bahrain Group of Professional Engineers.

Postponement of Symposium

Due to circumstances beyond the control of the Organizing Committee, the symposium on 'The Canadian Aspects of Satellite Communication,' scheduled to be held in Montreal in May, 1970 has been postponed. The new date, expected to be within the next few months, will be announced as soon as possible. An outline of the Quebec Section's plans was given in the September, 1969 issue of the *Journal*.

Conference on Automatic Test Systems

University of Birmingham — 14th to 17th April, 1970

This Conference is being organized by the I.E.R.E. with the association of the Institutions of Mechanical, Electrical and Production Engineers, the Royal Aeronautical Society and the Institute of Electrical and Electronics Engineers, and it is under the aegis of the United Kingdom Automation Council.

A provisional programme is given below but it must be stressed that this may be subject to additions and amendments. The final programme will be published shortly before the Conference.

PROVISIONAL PROGRAMME

Tuesday, 14th April.

Introduction by the Chairman of the Organizing Committee, BRIGADIER R. KNOWLES, C.B.E.

DETERMINATION OF REQUIREMENTS AND PLANNING A TEST SYSTEM

'Automatic Test Equipment—Gilt Edge or Gamble?' SQDN. LEADER D. R. HITCHINS, *Ministry of Defence.*

'A Philosophy for A.T.E. Procurement,' SQDN. LEADER C. G. H. FRANK, *Central Servicing Development Establishment, R.A.F.*

'Specifying the Interface Between Design and Test Organizations,' COMMANDER M. W. YOUNG, *Admiralty Surface Weapons Establishment.*

'Management Planning for Multi-System Automatic Testing—A Means of Improving the Engineering Support for the Army,' LIEUT. COLONEL R. W. A. LONSDALE, R.E.M.E., *School of Electronic Engineering, Arborfield.*

'The Impact of Automatic Testing on Product Design and on the Production Testing Organization,' R. KITCHEN, *The Marconi Company, Chelmsford.*

'Automatic Production Testing of Electronic Equipments—A Realistic Approach,' S. T. WILSON, *The Plessey Company Limited, Ilford.*

'The Interdependence of "Built In," "On Board" and "Ground Based" Test Facilities for Aircraft,' A. S. DELAHUNTY and A. G. HAYES, *British Aircraft Corporation, Warton, Lancashire.*

'Automatic Testing—Quality Assurance Aspects,' A. E. HAWLEY, *A.Q.A.D., Ministry of Technology.*

'Configuration Management of Software,' F. LIGUROI, *Emerson Electric Company, St. Louis, Missouri, U.S.A.*

Wednesday, 15th April.

CHOICE OF SYSTEM AND DESIGN CONSIDERATIONS

'A General Purpose Computer Interface for Automatic Test Equipment,' A. C. BURLEY, *Hawker Siddeley Dynamics, Hatfield, Herts.*

'A Computer Controlled Digital Equipment Test System,' C. CLARKE, *The Marconi Company, Chelmsford.*

'Melvin, or, Why buy racks of conventional test equipment when you've got a Computer in the Loop?' D. MACTAGGART, *Canadian Marconi Company, Montreal.*

'A Multi-User Laboratory Computer Facility,' DR. F. FALLSIDE, R. D. JACKSON, C. P. PRICE, P. D. RICE and J. M. THOMPSON, *Engineering Laboratory, University of Cambridge.*

'Synthesis of Test Patterns for Logic Networks,' MISS J. BLYTHIN and K. J. CROOK, *International Computers Limited, Stevenage, Hertfordshire.*

'The Use of F.R.A.M. to achieve Burn-In in G.W. Electronic Equipment,' K. E. HEAD and P. A. W. JONES, *Sperry Gyroscope, Bracknell.*

HARDWARE, INTERFACES AND APPLICATIONS

'The Man-Machine Interface in Automated Measurement,' K. BREWSTER, *Elliott Automation Limited, Rochester.*

'Interface Problems in the Automatic Test of a Doppler Radar (AN/APN–172) using the VAST (AN/USM–247) System,' R. J. BACANSKAS, M. N. GRANIERI, T. J. CARMODY and F. X. GRIBBON, *P.R.D. Electronics Inc., Long Island, N.Y., U.S.A.*

'A Multi-Station Electromechanical Automatic Test System,' P. J. REYNER, *Ether Engineering Limited, Watford.*

'Development and Application of a University Facility for Automatic Testing and Dynamic Analysis,' PROFESSOR J. L. DOUCE, *University of Warwick.*

'Test Equipment Engineering at Texas Instruments Limited,' P. Lingley, M. Gardner, J. W. Adams, G. Kermez and W. G. Peters, *Texas Instruments, Bedford.*

'Measurement of Turbojet Engine Performance under Simulated Flight Conditions,' K. F. A. Walles, *National Gas Turbine Establishment, Pyestock, Farnborough, Hampshire.*

Thursday, 16th April.

SOFTWARE PROGRAMMING AND LANGUAGE

'Data Structures for Generalized Automatic Test Systems,' J. A. O'Brien, *Advanced System Engineering Laboratory, Raytheon Company, Sudbury, Massachusetts, U.S.A.*

'The Management of Test Data Validation with Integral Computer Controlled A.T.E.' J. A. Hill and L. C. Painter, *Elliott Flight Automation, Rochester, Kent.*

'Program Preparation for Automatic Test Equipment,' A. M. Greenspan, *R.C.A. Corporation, Burlington, U.S.A.*

'Computer Generated Diagnosing Procedures for Logic Circuits,' A. H. Boyce, *Marconi Research Laboratories, Great Baddow.*

'The Effective Use of Computers in Automatic Test Systems,' R. J. Patrick, *Elliott Flight Automation, Rochester, Kent.*

'The Use of Atlas by an Equipment Manufacturer,' J. W. Anstead and M. T. Challenger, *Smiths Industries Ltd., Bishops Cleeve, Gloucestershire.*

'An Automatic Test System for Software Checkout of Naval Tactical Data Systems,' T. B. Brereton and W. J. Lukowski, *Fleet Computer Programming Center, Pacific, San Diego, California.*

MANAGEMENT OF A.T.S. AND USER ORGANIZATION ASPECTS. EFFECTS ON EQUIPMENT DESIGN. FEEDBACK OF TEST DATA

'Management Aspects and User Organizational Problems,' D. J. Bloomer, *Hawker Siddeley Dynamics, Hatfield, Herts.*

'Operator Participation and Man/Machine Interfaces,' R. Crosher, *Honeywell Limited, Hemel Hempstead, Herts.*

'Production Acceptance Testing and Automation,' T. K. White, *Henry Wiggin and Co. Ltd., Hereford.*

'The Impact of A.T.E. on Equipment Design,' A. G. Studd, *Smiths Industries Ltd., Bishops Cleeve, Gloucestershire.*

'A Programmable Digital System for Testing Automatic Gearboxes,' K. W. P. Drysdale, *Advance Industrial Electronics, Bishops Stortford, Herts.*

'Automatic Measurement of Television Waveforms,' G. A. McKenzie and R. Vivian, *Independent Television Authority, London, S.W.3.*

'A Comprehensive Automatic Dynamic Response System for Testing Semiconductors, L.S.I. and Modules,' G. C. Warburton, *E-H Research Laboratories U.K. Limited, West Drayton, Middlesex.*

Friday, 17th April.

FUTURE TRENDS

'Designing for Automatic Testing: the Concept and General Approach,' W. R. Ogden and C. R. Thomas, *Ferranti Limited, Edinburgh.*

'Direct Frequency Response Determination Exploiting the Deterministic Characteristics of Pseudo-Random Binary Sequences,' J. D. Lamb, *University of Wales Institute of Science and Technology.*

'Frequency Domain Approach to Automatic Testing of Control Systems,' Professor D. R. Towill and P. A. Payne, *University of Wales Institute of Science and Technology.*

'Failure Prediction for Preventive Maintenance,' R. F. Barry, *R.C.A., Burlington, Massachusetts, U.S.A.*

'A Cost Effective Airborne Test System for the Seventies,' R. C. Rogers, *Honeywell Limited, Brentford, Middlesex.*

'Automatic Test Equipment in the Assembly Shop of the Future,' C. H. Collingwood, *Hawker Siddeley Dynamics, Hatfield, Herts.*

Registration Forms and Further Information

Registration Forms which include details of charges etc. are now available from The Conference Registrar, Institution of Electronic and Radio Engineers, 8–9 Bedford Square, London W.C.1. (Telephone 01–580 8443, ext. 3.). A form requesting information etc. is given at the end of this issue.

U.D.C. 621.372.43

A Contribution to Element Reduction in Transformerless Driving-point Impedance Synthesis

By

JIŘÍ GREGOR,

Doc.Ing., C.Sc.†

A theorem formulating the necessary and sufficient conditions for the decomposition of a real positive function into a sum of real positive functions is presented. It is shown that every such decomposition depends on an arbitrary real positive function and on a positive number only. The appropriate choice of this function gives decompositions with desired properties.

This theorem enables us to treat the known methods of transformerless synthesis, which do not use bridge networks, as special cases of a more general approach. Resistance extraction, Bott and Duffin's method and Miyata's method are shown as special cases. It is possible to compare these methods and derive some new procedures.

The discussion and examples show that a reduction of the number of elements required in the synthesis procedure can be obtained. This reduction is due to the fact that the decomposition can be done without common poles between the two summands in some cases; in this way it is possible to avoid the exponential growth of the number of required elements which is found in other methods of transformerless driving-point impedance synthesis.

1. Series Connexion of Networks

It is well known that the necessary and sufficient conditions for a rational function F of the complex variable

$$s = \sigma + j\omega \quad \text{where } \sigma,\ \omega \text{ are real, } j^2 = -1,$$

to be a driving-point impedance of linear, passive, time-invariant, continuous and finite network are that F should be analytic in the open right half-plane and

$$\operatorname{Re} F(s) > 0 \text{ for every } \operatorname{Re} s > 0$$

(See for instance, Weinberg,[1] p. 144.)

Functions with these properties will be called Brune's functions and we shall write $F \in \mathscr{B}$. The class of functions denoted by $\mathscr{B}$ forms a subset of the set $\mathscr{R}$ of positive real functions.

The majority of the known methods of driving-point impedance synthesis use the decomposition of a given Brune's function into a sum of two (or more) Brune's functions. Such decomposition into a sum of the 'lossless part' (with poles on the imaginary axis only) and the remainder, which is often called a minimum reactive or minimum susceptive function if we hold it as an impedance or admittance, forms the first step of the synthesis procedure. It is the decomposition of this remainder which claims our attention. Several such decompositions are known. Resistance extraction could be mentioned here as the simplest example. Bott and Duffin's method uses, in fact, the same basic idea, as well as Miyata's, Unbehauen's and many other methods (see Weinberg,[1] pp. 455–484). Some of them require a smaller number of elements in comparison with Bott and Duffin's method, but we have no unified approach which enables us to explain such distinctions. Theorem 1 and its consequences offer an approach to this problem.

Let us denote the sub-class of the set of Brune's functions analytic in the closed right half-plane (including the point ∞) by $\mathscr{B}^*$. In a similar manner $\mathscr{R}^*$ means the set of positive real functions which are analytic in the closed right half-plane ($\operatorname{Re} s \geqq 0$) including the point ∞.

Hereafter, the following notation will be used

$$\operatorname{Ev} f(z) = \tfrac{1}{2}(f(z) + f(-z)).$$

If

$$f(z) = \frac{P(z)}{Q(z)},$$

where P, Q are polynomials, then

$$\operatorname{Ev} f(z) = \frac{M_1 M_2 - N_1 N_2}{M_2^2 - N_2^2},$$

where M and N respectively stands for the sum of even and odd powers and the suffixes denote the numerator (1) and denominator (2).

The following theorem, which has been proved elsewhere[2] (see also the Appendix) can therefore be derived.

† Formerly at the Technical University of Prague, Czechoslovakia; now with the University of Khartoum, Faculty of Engineering and Architecture, P.O.B. 487, Khartoum, Sudan.

Theorem 1. Let $f \in \mathscr{B}^*$, then for every function $\varphi \in \mathscr{B}^*$ the functions g, h satisfying the equations

$$Mg(s) = f(s)\,\mathrm{Ev}\,\varphi(s) - r(s)$$
$$h(s) = f(s) - g(s) \qquad \ldots\ldots(1)$$

both belong to the class $\mathscr{B}^*$. Here

$$M \geqq \max_{s=j\omega} \mathrm{Re}\,\varphi(s),$$

$$r(s) = \sum_{j=1}^{p} \sum_{k=1}^{q_j} [B_{jk}(s) + (-1)^k C_{jk}(s)] \frac{f^{(k-1)}(-\beta_j)}{(k-1)!},$$

$$B_{jk}(s) = \frac{1}{2\pi j} \int_{\Gamma+} \frac{\mathrm{Ev}\,\varphi(\zeta)}{s+\zeta} (\zeta+\beta_j)^{k-1}\,\mathrm{d}\zeta,$$

$$C_{jk}(s) = \frac{1}{2\pi j} \int_{\Gamma-} \frac{\mathrm{Ev}\,\varphi(\zeta)}{s+\zeta} (\zeta-\beta_j)^{k-1}\mathrm{d}\zeta, \qquad \ldots\ldots(2)$$

β_j is a q_j-tuple pole of φ and the circles $\Gamma_{\pm}: |\zeta \pm_j| \beta = \varepsilon_j$ have sufficiently small radii and are positively oriented. Conversely, if $f \in \mathscr{B}^*$, then for every pair of functions g, h, satisfying the conditions $f = g+h$ and $g, h \in \mathscr{B}^*$, there exists a function $\varphi \in \mathscr{B}^*$ and a real number M, satisfying (1).

Although we shall use Theorem 1 in the formulated form, it has to be pointed out that it remains true under more general assumptions: $f, g, h \in \mathscr{R}^*$.

The function r in (1) has the following properties:

(i) r is an odd function, i.e. $r(s) + r(-s) = 0$;

(ii) each pole of the function $\mathrm{Ev}\,\varphi(s)$ is a pole of the same order of the function r and r has no other poles.

The first property can be easily proved, as

$$C_{jk}(-s) = (-1)^{k-1} B_{jk}(s)$$

and

$$C_{jk}(s) = (-1)^{k-1} B_{jk}(-s).$$

The second one is a simple consequence of Cauchy's integral formula. We can now conclude:

Corollary: Suppose $f, g, h \in \mathscr{B}^*$. The equation $f = g+h$ holds if and only if there exists a function $\varphi \in \mathscr{B}^*$ and a real number

$$M \geqq \max_{s=j\omega} \mathrm{Re}\,\varphi(s)$$

such that

$$M\,\mathrm{Ev}\,g(s) = \mathrm{Ev}\,f(s)\,\mathrm{Ev}\,\varphi(s)$$

$$\mathrm{Ev}\,h(s) = \left(1 - \frac{\mathrm{Ev}\,\varphi(s)}{M}\right) \mathrm{Ev}\,f(s). \qquad \ldots\ldots(3)$$

This means: Every decomposition of a minimum reactive or minimum susceptive Brune's function into a sum of two Brune's functions depends only on the choice of a function $\varphi \in \mathscr{B}^*$ and a real number M (we are supposing here that the uniqueness of the above-mentioned extraction of the Brune's function with poles on the imaginary axis only—the reactance function—is obvious). In other words, every method of the series–parallel transformerless impedance synthesis can be characterized by the adequate choice of a special Brune's function φ and a positive number M.

Let us deal with some special cases:

1. Let

$$f \in \mathscr{B}^* \quad \text{and} \quad \min_{s=j\omega} \mathrm{Re}\,f(s) = \mu \neq 0.$$

Under these assumptions there exists a function $\varphi \in \mathscr{B}^*$ satisfying the condition $\mathrm{Ev}\,\varphi(s) = 1/\mathrm{Ev}\,f(s)$ for almost every s. The maximum of its real part on the imaginary axis is evidently equal to μ^{-1}. If we use Theorem 1, with $M = \mu^{-1}$, we get $\mathrm{Ev}\,g(s) = \mu = g(s)$, which gives the well-known *resistance extraction.* We get essentially the same result with the function φ satisfying the following condition:

$$\mathrm{Ev}\,\varphi(s) = 1 - \frac{\mu}{\mathrm{Ev}\,f(s)}.$$

2. Let $f \in \mathscr{B}^*$, $\varphi(s) = a/(a+s)$, $a > 0$.

Then

$$\max_{s=j\omega} \mathrm{Re}\,f(s) = 1 \quad \text{and} \quad \mathrm{Ev}\,\varphi(s) = \frac{a^2}{a^2-s^2}.$$

After some calculations using the above-mentioned properties of the function r we get

$$g(s) = a\,\frac{af(s)-sf(a)}{a^2-s^2}, \quad h(s) = s\,\frac{af(a)-sf(s)}{a^2-s^2}, \qquad \ldots\ldots(4)$$

$$g, h \in \mathscr{B}^*, \quad g+h = f.$$

Now let the number a satisfy the condition $f(a) + f(-a) = 0$. This means: There exists a positive number λ with the property $f(a) = \lambda a$. The functions g and h are exactly those we get applying *Bott and Duffin's procedure.* At the same time we have here some explanation for the large number of elements required using this method: it does not utilize the possibilities of cancellation in the formulas (3) and that is why the degree of the two resultant components is unnecessarily high.

3. Let us suppose that $f \in \mathscr{B}^*$ and

$$\mathrm{Ev}\,f(s) = \frac{P(s)+Q(s)}{R(s)}$$

where P, Q, R are even polynomials with real coefficients. Let P and Q be non-negative and not both zero on the imaginary axis. We can choose a function φ satisfying the following condition:

$$\mathrm{Ev}\,\varphi(s) = \frac{P(s)}{P(s)+Q(s)}.$$

It is obvious that $\varphi \in \mathscr{B}^*$. Using Theorem 1, or its corollary, we get the decomposition used by Miyata in his method, provided the coefficients of P and Q are positive. The assumption that the polynomials P, Q, and R have no common factors is superfluous. As is known, we can extend the applicability of Miyata's method multiplying the numerator and the denominator of the function f by a common factor to get the coefficients of the numerator of the function $\mathrm{Re} f(j\omega)$ positive. We can here consider *Miyata's method of synthesis* as a special case of Theorem 1 (see Weinberg,[1] pp. 464–482).

4. There are many other ways of choosing the function φ. We shall demonstrate the previous methods and some of these possibilities with the following example (see Weinberg,[1] pp. 469–478). Let the impedance

$$f(s) = \frac{s^3+7s^2+6s+2}{s^3+3s^2+8s+4} \qquad \ldots\ldots(5)$$

be given. Its even part is

$$\mathrm{Ev}\, f(s) = \frac{s^6-7s^4+14s^2-8}{s^6+7s^4+40s^2-16}$$

and it has the following zeros $s_{1,2} = \pm 1$, $s_{3,4} = \pm\sqrt{2}$, $s_{5,6} = \pm 2$.

Let us start by choosing $\varphi(s) = 1/(s+1)$ and applying, in fact, Bott and Duffin's method. We get the decomposition

$$(s) = \frac{s^3+6s^2+4s}{s^3+3s^2+8s+4} + \frac{s^2+2s+2}{s^3+3s^2+8s+4} \qquad \ldots\ldots(6)$$

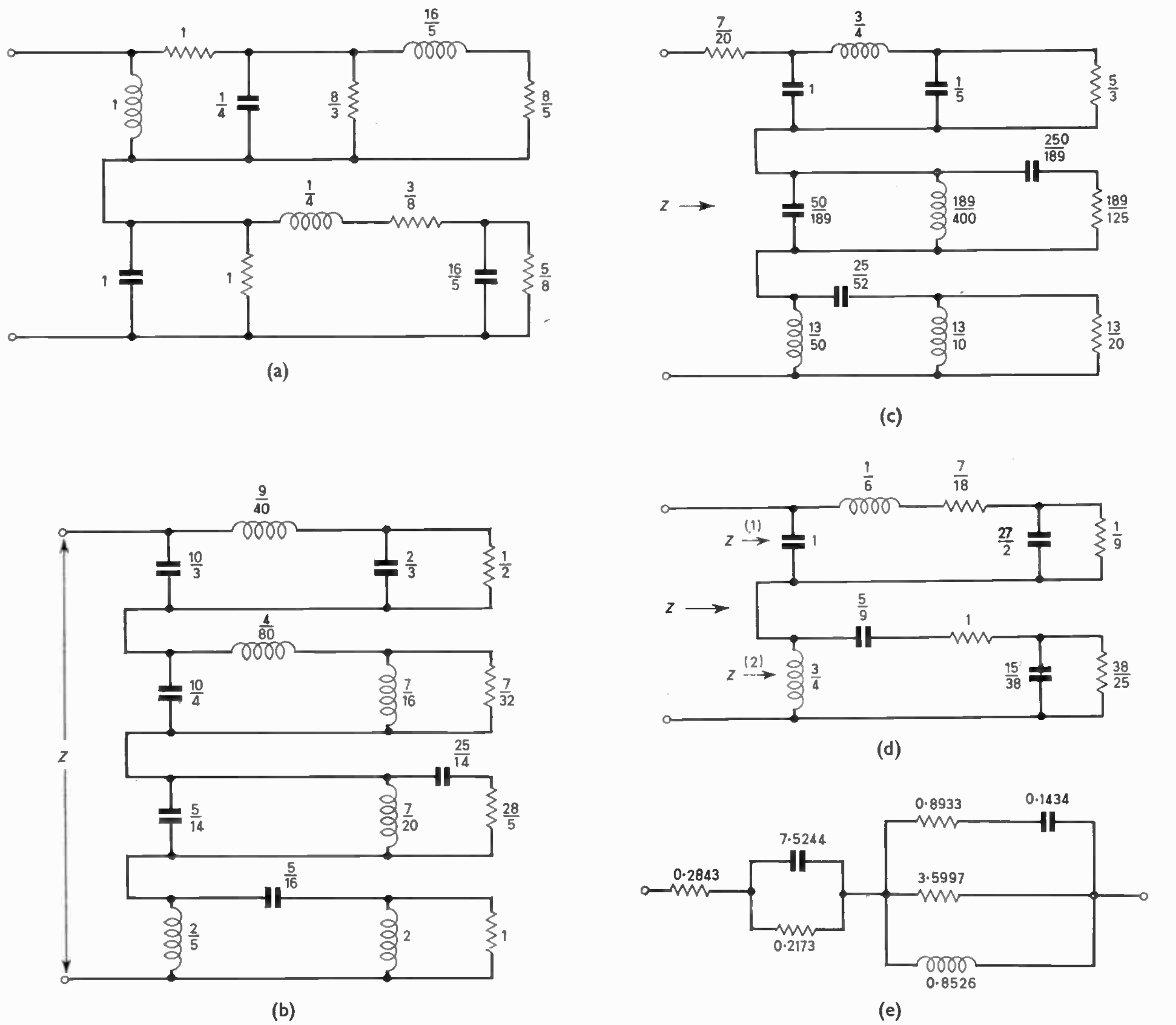


Fig. 1. Realizations of the impedance function given by equation (5) using various methods. (a) Bott and Duffin's method, (b) Miyata's method, (c) modified Miyata method, (d) 'half-*n* method' of Miyata, (e) method described in the paper. Values are in ohms, henrys and farads.

The impedance can be realized as a series connexion of two ladder networks with no more than 12 elements (see Fig. 1(a)). Using now Miyata's procedure in its first form, we get a realization requiring 16 elements (Fig. 1(b)), or, in a more convenient form, only 13 elements (Fig. 1(c)). The so-called 'half-n method' of Miyata gives a realization with only 10 elements (Fig. 1(d)) with the aid of the following decomposition

$$f(s) = \frac{s^2+3s+2}{s^3+3s^2+8s+4} + \frac{s^3+6s^2+3s}{s^3+3s^2+8s+4}. \quad \ldots\ldots(7)$$

These results are given in Weinberg's book. Let us deal with the other possibilities given by the formulated theorem. We can write the even part of the given impedance in the following form

$$\mathrm{Ev}\, f(s) = \frac{(s^2-1)(s^2-2)(s^2-4)}{(s^4+7{\cdot}3742s^2+42{\cdot}7594)(s^2-0{\cdot}3742)}.$$

Let us choose a function $\varphi \in \mathscr{B}^*$ satisfying the condition

$$\mathrm{Ev}\, \varphi(s) = \frac{s^4+7{\cdot}3742s^2+42{\cdot}7594}{(s^2-1)(s^2-4)}.$$

Such a function evidently exists. Using the corollary of Theorem 1, with $M = 20$, we get after some calculations the following decomposition:

$$f(s) = \frac{1}{20}\cdot\frac{s+3{\cdot}2696}{s+0{\cdot}6117} + \frac{0{\cdot}95s^2+6{\cdot}135s+1{\cdot}532}{s^2+2{\cdot}3883s+6{\cdot}5391}. \quad \ldots\ldots(8)$$

According to this result a realization of the form as in Fig. 1(e) may be given. This contains only 7 elements, which is exactly the number of independent coefficients in the given impedance.

2. Extraction of Poles

Theorem 1 and its corollary, as well as the example given, show that the different methods of transformerless synthesis together with the properties of the resulting networks are, in fact, determined by the function φ.

As the sum of two impedance functions is involved, the basic topological arrangement may comprise the series connexion of two networks. However, within these limits, we can discuss the degrees of resulting impedances, the number of required elements and many other problems, keeping in mind the properties of the function φ only. We cannot ignore the fact that, using Theorem 1, we are able to prove in an analytical way that no series realization does exist for some impedances. In addition to these theoretical results some practical methods can be derived from Theorem 1.

We can discuss all the possible cases of a series realization. But which are the simplest components of an impedance for which we are looking? It is known that the decomposition of an impedance into a sum of an RC and an RL part is, in general, not possible. The necessary and sufficient conditions for an impedance to be realizable as a ladder structure with all the three kinds of elements but without mutual coupling are not known. That is why almost the only possibility is to take the degree of an impedance† as the measure of its complexity. In the last mentioned case the following problem arises: Let a function $f \in \mathscr{B}^*$ be given. Under which conditions does there exist a decomposition of the form $f = g+h$, $g, h \in \mathscr{B}^*$ satisfying the requirement $\deg f = \deg g + \deg h$? How can this decomposition be constructed? As shown by the given example we can solve this problem by discussing the possibilities of cancellation in the product $\mathrm{Ev} f(s)$. $\mathrm{Ev}\,\varphi(s)$ (see eqn. (3)). In this way we get some important special cases of transformerless synthesis with the smallest number of elements.

In answer to the formulated problem, let the even part of the given function $f \in \mathscr{B}^*$ be in the form

$$\mathrm{Ev}\, f(s) = \frac{P(p)}{R(p)\,Q(p)},$$

where $p = s^2$ and P, Q, R are relative prime polynomials with real coefficients, $\deg P \leqq \deg RQ$. The following theorem is almost obvious as a consequence of Theorem 1:

Theorem 2. Supposing $f \in \mathscr{B}^*$, then $f = g+h$, where $g, h \in \mathscr{B}^*$ and $\deg f = \deg g + \deg h$ if and only if there exists a polynomial S and a positive number m so that

$$m \geqq \frac{S(p)Q(p)}{P(p)} \geqq 0 \quad \text{for every } p \leqq 0,$$

and R divides the polynomial $mP - QS$.

To prove the sufficiency of the conditions we put $\mathrm{Ev}\,\varphi(s) = (Q.S)/P$ in the corollary of Theorem 1. In a similar manner (using Theorem 1) we can prove that the conditions are necessary. Let us note that the function g in this case can be determined from the equation

$$\mathrm{Ev}\, g = \frac{S}{mR}.$$

Some special cases of this theorem could be of practical interest.

Theorem 3. Let a function $f \in \mathscr{B}^*$ be given and let

$$\mathrm{Ev}\, f(s) = \frac{P(p)}{(p_0-p)Q(p)},$$

$$P(p_0) \neq 0, \quad Q(p_0) \neq 0, \quad p_0 > 0.$$

† The number of poles (or zeros) counted with their proper multiplicities is called the degree of an impedance.

Then

$$f(s) = \frac{as+b}{s+\sqrt{p_0}} + h(s), \qquad \ldots\ldots(9)$$

with $a, b \geqq 0$, $h \in \mathscr{B}^*$ if and only if there exist non-negative numbers α, β, $\alpha^2+\beta^2 \neq 0$ so that

$$\alpha[Q(p)-\mu P(p)] \leqq \beta[pQ(p)-\mu p_0 P(p)], \quad \mu = \frac{Q(p_0)}{P(p_0)} \qquad \ldots\ldots(10)$$

for every $p \leqq 0$. Here,

$$a = \frac{1}{\mu}\frac{\beta}{\alpha-\beta p_0}, \quad b = \frac{1}{\mu\sqrt{p_0}}\frac{\alpha}{\alpha-\beta p_0}. \qquad \ldots\ldots(11)$$

The proof can be omitted. In our example (see (5), (8)) the conditions were fulfilled; in the example $\alpha = 2$, $\beta = 1$, $p_0 = 0{\cdot}3742$. It has to be mentioned that the decomposition according to Theorem 3 is in no case unique if it exists. The Theorem gives, in fact, the necessary and sufficient conditions for the removal of a single real pole from a given impedance function.

For practical purposes it is useful to have some easily verifiable necessary condition. If the decomposition mentioned in Theorem 3 exists, then there exist non-negative numbers α, β so that

$$\frac{q_1}{q_2}(\alpha-\beta p_0) \geqq \alpha, \quad \alpha^2+\beta^2 \neq 0,$$

where

$$q_1 = \frac{Q(p_0)}{Q(0)}; \quad q_2 = \frac{P(p_0)}{P(0)}.$$

Considering the function

$$f(s) = \frac{12s^2+10s+3}{s^2+4s+3},$$

with $p_0 = s_0^2 = 1$ we get $q_1 = \frac{8}{9}$, $q_2 = \frac{20}{9}$ and the necessary condition cannot be satisfied with any non-negative α, β. It is therefore impossible to realize the given impedance (see Ladenheim[3]) as a series interconnexion of two impedances of first degree.

The necessary conditions are evidently not satisfied, for example, when

$$0 < \frac{P(0)Q(p_0)}{Q(0)P(p_0)} < 1.$$

Very simple sufficient conditions can be derived from (10) and expressed in terms of the coefficients of polynomials P, Q. Considering

$$P(p) = a_0 p^n + a_1 p^{n-1} + \ldots + a_n,$$

$$Q(p) = b_0 p^n + b_1 p^{n-1} + \ldots + b_n,$$

we can easily prove: If non-negative numbers α, β, $\alpha^2+\beta^2 \neq 0$ fulfil a system of $(n+1)$ inequalities

$$(-1)^{n-k}(\alpha(a_k\mu - b_{k-1}) + \beta(b_k - a_k\mu p_0)) \geqq 0,$$
$$k = 0, 1 \ldots n \qquad \ldots\ldots(12)$$

then the decomposition as in Theorem 3 is possible. Here, $b_{-1} = b_n = 0$. In the general case it is possible to verify the conditions (10) using Sturm's theorem.[1]

Essentially the same method can be used when discussing the removal of a pair of complex poles from a given impedance. The following lemma will be useful:

Lemma: For any angle α, $-\pi < \alpha \leqq \pi$ and any complex number $p_0 = a+jb$ with $b \neq 0$ there exists a polynomial S of arbitrary degree $n \geqq 1$ with real coefficients so that

$$S(p_0) = \lambda\, e^{j\alpha}, \qquad \lambda > 0.$$

For $n > 1$ the modulus λ can always be determined so that $S(p) \geqq 0$ for every $p \leqq 0$; the same holds for $n = 1$ if and only if $\sin\alpha < 0$ and $b\cos\alpha - a\sin\alpha \geqq 0$.

The proof is very simple. It is sufficient to give the explicit form of the polynomials for $n = 1, 2$. The polynomials of higher degree can be constructed with arbitrarily given $(n-2)$ roots. It is easy to verify that

$$S_1(p) = \frac{\lambda}{b}(p\sin\alpha + b\cos\alpha - a\sin\alpha),$$

$$S_2(p) = p^2 + \left(\frac{\lambda\sin\alpha}{b} - 2a\right)p + a^2 + b^2 + {}$$
$$+ \lambda\frac{b\cos\alpha - a\sin\alpha}{b} \qquad \ldots\ldots(13)$$

have the required properties.

Let us use now the same notation as in Theorem 2 and suppose the polynomial R to be in the form $(p-p_0)(p-p_0^*)$, where $\operatorname{Im} p_0 > 0$ and $Q(p_0) \neq 0$. The polynomial $mP(p) - Q(p)S(p)$ will be divided by R if and only if

$$mP(p_0) - Q(p_0)S(p_0) = 0$$

and the necessary and sufficient conditions of Theorem 2 will be satisfied if, at the same time,

$$m \geqq \frac{Q(p)S(p)}{P(p)} \quad \text{for every } p \leqq 0.$$

Denoting now

$$\alpha = -\arg\frac{Q(p_0)}{P(p_0)}$$

we can find the polynomials S_1, S_2 mentioned in the lemma so that

$$S_1(p_0) = S_2(p_0) = \lambda\, e^{j\alpha}.$$

With these polynomials the equation

$$\frac{Q(p_0)}{P(p_0)}S_i(p_0) = m_i, \qquad i = 1, 2$$

determines two real numbers

$$m_i = \lambda \left|\frac{Q(p_0)}{P(p_0)}\right|.$$

The removal of an impedance, whose even part has the polynomial R as its denominator, will now be possible if and only if for some $\lambda > 0$ and $i = 1$ or $i = 2$ the following two conditions are satisfied:

$$\lambda \left|\frac{Q(p_0)}{P(p_0)}\right| \geqq \frac{Q(p)}{P(p)} S_i(p), \quad S_i(p) \geqq 0 \text{ for every } p \leqq 0. \qquad \ldots\ldots(14)$$

For $i = 1$ there is $S_1(p) \geqq 0$ at the non-negative half-axis if and only if

$$\sin\alpha < 0, \quad b\cos\alpha \geqq a\sin\alpha \qquad \ldots\ldots(15)$$

and under these assumptions the necessary and sufficient conditions can be written in the form:

$$\left|\frac{Q(p_0)}{P(p_0)}\right| \frac{b}{p\sin\alpha + b\cos\alpha - a\sin\alpha} \geqq \frac{Q(p)}{P(p)} \quad \text{for every } p \leqq 0. \qquad \ldots\ldots(16)$$

This result can be summarized in the following

Theorem 4. Let a function $f \in \mathscr{B}^*$ be given and let

$$\operatorname{Ev} f(s) = \frac{P(p)}{Q(p)(p-p_0)(p-p_0^*)},$$

$$P(p_0) \neq 0, \quad Q(p_0) \neq 0, \quad \operatorname{Im} p_0 > 0;$$

then $f = g + h$, where $g(s) = (as+b)/(s^2+cs+d)$,

$$d = p_0 p_0^*, \quad c = (p_0 p_0 - \operatorname{Re} p_0)\sqrt{2},$$

$g, h \in \mathscr{B}^*$ and g, h have no common poles if and only if the conditions (15), (16) are satisfied.

Let us now consider the polynomial $S_2(p)$. There is $S_2(p) \geqq 0$ for every $p \leqq 0$ if and only if

$$\lambda^2 \sin^2\alpha - 4b\lambda\cos\alpha - 4b^4 \neq 0 \qquad \ldots\ldots(17)$$

and

$$\lambda(b\cos\alpha - a\sin\alpha) + b(a^2+b^2) \geqq 0. \qquad \ldots\ldots(18)$$

The following theorem has just been proved:

Theorem 5. Under the same assumption as in Theorem 4, there is $f(s) = g(s) + h(s)$,

where $g(s) = (ks^2+as+b)/(s^2+cs+d)$,

$g, h \in \mathscr{B}^*$ and g and h have no common poles if and only if for some $\lambda > 0$ the conditions (17), (18) and the inequality

$$\frac{\lambda}{S(p)}\left|\frac{Q(p_0)}{P(p_0)}\right| \geqq \frac{Q(p)}{P(p)} \qquad \ldots\ldots(19)$$

for every $p \leqq 0$ is satisfied.

3. Examples

Let us deal with some examples:

Example 1. Let us consider the impedance function

$$f(s) = \frac{7s^4 + 20s^3 + 38s^2 + 38s + 25}{s^4 + 4s^3 + 8s^2 + 7s + 4}.$$

Its even part can be written in the form

$$\operatorname{Ev} f(s) = \frac{7s^8 + 14s^6 + 65s^4 + 86s^2 + 100}{(s^4+s^2+1)(s^4-s^2+16)}.$$

The polynomials mentioned in Theorem 2 are the following:

$$P(p) = 7p^4 + 14p^3 + 65p^2 + 86p + 100$$

$$Q(p) = p^2 + p + 1, \qquad R(p) = p^2 - p + 16.$$

The zero point of the polynomial R is

$$p_0 = \tfrac{1}{2}(1 + \mathrm{j}\sqrt{63}) \quad \text{and} \quad \frac{Q(p_0)}{P(p_0)} = -\frac{3+\mathrm{j}\sqrt{7}}{288}.$$

Further

$$\left|\frac{Q(p_0)}{P(p_0)}\right| = \frac{1}{72}, \quad \cos\alpha = -\frac{3}{4}, \quad \sin\alpha = \frac{\sqrt{7}}{4}.$$

Therefore

$$S_2(p) = p^2 + \left(\frac{\lambda}{6} - 1\right)p + 16 - \frac{5\lambda}{6}.$$

Here $\sin\alpha > 0$, the assumptions of Theorem 4 are not satisfied and therefore it is impossible to remove an impedance of second degree with linear numerator. Using Theorem 5, we have $m = \lambda/72$ and after subtraction of the function

$$\operatorname{Ev} g(s) = \frac{72}{\lambda} \frac{p^2 + \left(\frac{\lambda}{6} - 1\right)p + 16 - \frac{5\lambda}{6}}{p^2 - p + 16}$$

from the given $\operatorname{Ev} f(s)$ we get

$$\operatorname{Ev} h(s) = \frac{(7\lambda - 72)p^2 + (9\lambda - 72)p + 10\lambda - 72}{\lambda(p^2+p+1)}.$$

Both these functions are even parts of impedances for some value of λ if and only if the inequality in Theorem 5 is satisfied. But here it is more convenient to find directly, using the known test for biquadratic impedances, the appropriate interval for the parameter λ. We get

$$g(s) = \frac{\frac{72}{\lambda}s^2 + \left(\frac{216}{\lambda} - 9\right)s + \frac{288}{\lambda} - 15}{s^2 + 3s + 4},$$

$$h(s) = \frac{\left(7 - \frac{72}{\lambda}\right)s^2 + \left(8 - \frac{72}{\lambda}\right)s + 10 - \frac{72}{\lambda}}{s^2 + s + 1}$$

and $g, h \in \mathscr{B}^*$ if and only if

$$\frac{36}{199}(50 + \sqrt{112}) \leqq \lambda \leqq 18.$$

When realizing now the given impedance by the modified Bott-Duffin method we need 22 elements, while after the given decomposition we need only 16. The half-n method of Miyata can be used only with a surplus factor of second degree in the variable s and therefore it needs 24 elements (see Weinberg,[1] p. 476).

It is well known that the number of elements grows exponentially for every known method of transformerless impedance synthesis. Therefore a decomposition into terms with no common poles always gives a reduction of the total number of elements.

Example 2. Let the impedance

$$f(s) = \frac{s^4+22s^3+16s^2+35s+1}{s^4+s^3+5s^2+2s+4}.$$

Its even part can be written in the form

$$\mathrm{Ev}\, f(s) = \frac{p^4-p^3+6p^2-p+4}{(p^2+2{\cdot}04191p+1{\cdot}17243)(p^2+6{\cdot}96151p+13{\cdot}62289)}.$$

Let $Q(p) = p^2+2{\cdot}04191p+1{\cdot}17243$. After some calculations we realize that, for example, if $p_0 = 0$ then

$$\frac{\lambda\left|\frac{Q(p_0)}{P(p_0)}\right|}{S_2(p} \leqq \frac{Q(p)}{P(p)}$$

for every positive λ and, therefore, the conditions of Theorem 5 cannot be satisfied. However, considering the admittance $F(s) = 1/f(s)$, the decomposition becomes possible. Carrying out the calculations in this case we will meet subtractions of large numbers with small differences, which has to be done with high accuracy (e.g. with a computer using double-precision programs with 18–20 decimals). This is because of the necessary cancellation between two polynomials as required in Theorem 2.

4. Conclusions

Theorem 1 and its consequences yield a new approach to the driving-point impedance synthesis. This includes the known methods based on the decomposition of an impedance into a sum of impedances and, moreover, gives all such decompositions. Among all these decompositions it is possible to choose the most advantageous one. Because the total number of required elements in transformerless synthesis grows exponentially with the degree of the impedance the most advantageous decomposition is that with no common poles between the components. This way we can decrease the number of required elements in the transformerless synthesis or, if we are allowed to use series–parallel connexions only, we can prove that no such reduction is possible.

5. References

1. Weinberg, L., 'Network Analysis and Synthesis' (McGraw-Hill, New York, 1962).
2. Gregor, J., 'On the decomposition of a positive real function into positive real summands', *Apl. Mat.* (*Czechoslovakia*), 6, 1969 (to be published).
3. Ladenheim, E. L., 'Three-reactive five-element biquadratic structures', *I.E.E.E. Trans. on Circuit Theory*, **CT-11**, pp. 88–97, March 1964.
4. Nevanlinna, R., 'Asymptotische Entwicklungen beschränkter Functionen und das Stieltjessche Momentenproblem', *Ann. Acad. Sci. Fennicae*, **A18**, 1922.
5. Herglotz, G., 'Ueber Potenzreihen mit positivem reellem Teil im Einheitskreis', *Leipz. Ber.*, **63**, 1911.
6. Seshu, S. and Reed, M. B., 'Linear Graphs and Electrical Networks' (Addison-Wesley, London, 1961).

6. Appendix

The proof of Theorem 1 can be based on a slightly modified result of Nevanlinna[4] and Herglotz[5]:

Theorem 6. Let f be a complex function finite in the open right half-plane and let

$$f(s) = \mu s + \left|\int_{-\infty}^{+\infty}\right| \frac{\mathrm{d}\tau(s)}{s+\mathrm{j}t} \quad \text{for Re } s > 0,\dagger \qquad \ldots\ldots(20)$$

where $\mu \geqq 0$ and τ is a non-decreasing function with its even part equal to zero almost everywhere and

$$\int_{-\infty}^{+\infty} \frac{\mathrm{d}\tau(t)}{1+t^2} < \infty.$$

Then $f \in \mathscr{R}$. Conversely: if $f \in \mathscr{R}$ then there exists a real non-negative number μ and a function τ with the mentioned properties so that (20) holds in the open right half-plane.

It is not difficult to prove that

$$\lim \frac{f(s)}{s} = \mu$$

when s tends to infinity and at the same time

$$|\arg s| \leqq k < \frac{\pi}{2}.$$

The formula (20) can be written in another form

$$f(s) = s\left(\mu + 2\int_0^{\infty} \frac{\mathrm{d}\tau(t)}{s^2+t^2}\right). \qquad \ldots\ldots(21)$$

If the function f is a Brune function, analytic in the closed right half-plane, then the function τ becomes absolutely continuous and its derivative is a rational function. Moreover, if f is analytic at the point ∞ then $\mu = 0$.

† Hereafter, $\left|\int_{-\infty}^{+\infty}\right|$ means the '*valeur principal*',

i.e. $\left|\int_{-\infty}^{+\infty}\right| \varphi(t)\,\mathrm{d}\tau(t) = \lim_{A\to\infty} \int_{-A}^{A} \varphi(t)\,\mathrm{d}\tau(t).$

We can give the proof of Theorem 1 now:

Proof of Theorem 1: According to Theorem 6 we have

$$f(s) = 2s \int_0^\infty \frac{\mathrm{d}\tau(t)}{s^2+t^2} \qquad \text{for } \operatorname{Re} s > 0,$$

where τ has the properties as in Theorem 6. Let us choose an arbitrary function $\varphi \in \mathscr{B}^*$ and let us denote

$$r(t) = \operatorname{Re} \varphi(\mathrm{j}t).$$

The function φ being analytic in the closed right half-plane we can conclude that

$$0 \leqq r(t) \leqq M = \max_t \operatorname{Re} \varphi(\mathrm{j}t).$$

Let us consider now a function

$$\tau_1(t) = \int_0^t r(\theta)\, \mathrm{d}\tau(\theta);$$

it is evidently non-negative, non-decreasing and

$$\int_0^\infty \frac{\mathrm{d}\tau_1(t)}{1+t^2} < \infty.$$

All the requirements of Theorem 6 are satisfied and therefore the function

$$g(s) = \frac{2s}{M} \int_0^\infty \frac{\mathrm{d}\tau_1(t)}{s^2+t^2}$$

is a positive-real function. The same holds for

$$\tau_2(t) = \int_0^t \left(1 - \frac{r(\theta)}{M}\right) \mathrm{d}\tau(\theta)$$

and

$$h(s) = 2s \int_0^\infty \frac{\mathrm{d}\tau_2(t)}{s^2+t^2}$$

because

$$1 - \frac{r(\theta)}{M} \geqq 0.$$

At the same time $\tau_1(t)+\tau_2(t) = \tau(t)$ and therefore

$$g(s)+h(s) = f(s).$$

Let us denote now

$$\phi(\xi, s) = \frac{\operatorname{Ev} \varphi(\xi)}{s+\xi}.$$

The function φ being a rational function, the only poles of $\operatorname{Ev} \varphi(\xi)$ are $\pm\beta_i$, $i = 1, 2, \ldots p$. We can write

$$\begin{aligned}\phi(\xi, s) = \frac{A(s)}{s+\xi} &+ \sum_{k=1}^{q_1} \left[\frac{B_{1k}(s)}{(\xi+\beta_1)^k} + \frac{C_{1k}(s)}{(\xi-\beta_1)^k}\right] + \\ &+ \sum_{k=1}^{q_2} \left[\frac{B_{2k}(s)}{(\xi+\beta_2)^k} + \frac{C_{2k}(s)}{(\xi-\beta_2)^k}\right] + \ldots + \\ &+ \sum_{k=1}^{q_p} \left[\frac{B_{pk}(s)}{(\xi+\beta_p)^k} + \frac{C_{pk}(s)}{(\xi-\beta_p)^k}\right] \qquad \ldots\ldots(22)\end{aligned}$$

for any s with $\operatorname{Re} s > 0$. Consider the positively oriented circles $|\xi+\beta_i| = \varepsilon_i$, $|\xi-\beta_i| = \varepsilon_i$ respectively, with sufficiently small ε_i and similarly the circle $|s+\xi| = \varepsilon_0$. Multiplying (22) by $(\xi+\beta_i)^{k-1}$ and $(\xi-\beta_i)^{k-1}$ respectively, and integrating along the mentioned circles we get

$$B_{ik} = \frac{1}{2\pi\mathrm{j}} \int_{K+} \phi(\xi, s)(\xi+\beta_i)^{k-1}\, \mathrm{d}\xi,$$

$$C_{ik} = \frac{1}{2\pi\mathrm{j}} \int_{K-} \phi(\xi, s)(\xi-\beta_i)^{k-1}\, \mathrm{d}\xi,$$

$$i = 1, 2, \ldots p, \quad k = 1, 2, \ldots q_i,$$

$$A = \frac{1}{2\pi\mathrm{j}} \int_K \phi(\xi, s)\, \mathrm{d}\xi = \operatorname{Ev} \varphi(s), \qquad \ldots\ldots(23)$$

where obviously the functions A, B_{jk}, C_{jk} depend only on the function φ. Let us consider now the formula

$$f(s) = \left|\int_{-\infty}^{+\infty}\right| \frac{\mathrm{d}\tau(t)}{s+\mathrm{j}t}.$$

The following differentiation of the integral for any s in the open right half-plane can be easily justified:

$$(-1)^{k-1} \frac{f^{(k-1)}(s)}{(k-1)!} = \left|\int_{-\infty}^{+\infty}\right| \frac{\mathrm{d}\tau(t)}{(s+\mathrm{j}t)^k}, \quad k = 1, 2 \ldots.$$

Expecially, for $s = -\beta_i$ there is $\operatorname{Re}(-\beta_i) > 0$ and therefore

$$I_{ki}^- = \left|\int_{-\infty}^{+\infty}\right| \frac{\mathrm{d}\tau(t)}{(-\beta_i+\mathrm{j}t)^k} = \frac{f^{(k-1)}(-\beta_i)}{(k-1)!}(-1)^{k-1}.$$

Furthermore

$$I_{ki}^+ = \left|\int_{-\infty}^{+\infty}\right| \frac{\mathrm{d}\tau(t)}{(\beta_i+\mathrm{j}t)^k} = (-1)^k \left|\int_{-\infty}^{+\infty}\right| \frac{\mathrm{d}\tau(t)}{(-\beta_i-\mathrm{j}t)^k}$$

$$= (-1)^k \int_{-\infty}^{+\infty} \frac{\mathrm{d}\tau(t)}{(-\beta_i+\mathrm{j}t)^k} = (-1)^k I_{ki}^-.$$

But

$$Mg(s) = \left|\int_{-\infty}^{+\infty}\right| \phi(\mathrm{j}t, s)\, \mathrm{d}\tau(t),$$

hence

$$Mg(s) = f(s)\ \mathrm{Ev}\ \varphi(s) + \sum_{j=1}^{p} \sum_{k=1}^{q_j} (B_{jk}(s) I_{kj}^{+} + C_{jk}(s) I_{kj}^{-}),$$

which completes the first part of the proof.

We shall prove now the converse statement. Let $f = g+h$, $g \in \mathscr{B}^*$, $h \in \mathscr{B}^*$ and let us consider the ratio

$$\frac{M\ \mathrm{Ev}\ g(s)}{\mathrm{Ev}\ f(s)} \quad \text{for } s = jt,\ t \text{ real.} \qquad \ldots\ldots(24)$$

If the function $\mathrm{Ev}\, f$ has a zero on the pure imaginary axis then at the same point the function $\mathrm{Ev}\, g$ has a zero (remember that the real part of the function h cannot be negative at any point of the pure imaginary axis). Therefore the ratio (24) is finite for any $s = jt$. Let us choose the number M so that (24) will be $\leqq 1$. This is always possible and we have

$$0 \leqq \frac{M\ \mathrm{Ev}\ g(s)}{\mathrm{Ev}\ f(s)} \leqq 1 \qquad \text{for } s = jt.$$

Moreover, the ratio (24) is a rational function. Therefore, there always exists a certain function φ so that $\varphi \in \mathscr{B}^*$ and

$$\mathrm{Ev}\ \varphi(s) = \frac{M\ \mathrm{Ev}\ g(s)}{\mathrm{Ev}\ f(s)}.$$

We can simply verify now that all the formulae in Theorem 1 remain true with the above defined function φ. This completes the proof.

Manuscript first received by the Institution on 30th December 1968 and in final form on 15th August 1969. (Paper No. 1296/CC62).

STANDARD FREQUENCY TRANSMISSIONS—December 1969

(Communication from the National Physical Laboratory)

Dec. 1969	Deviation from nominal frequency in parts in 10^{10} (24-hour mean centred on 0300 UT) GBR 16 kHz	MSF 60 kHz	Droitwich 200 kHz	Relative phase readings in microseconds N.P.L.—Station (Readings at 1500 UT) *GBR 16 kHz	†MSF 60 kHz	Dec. 1969	Deviation from nominal frequency in parts in 10^{10} (24-hour mean centred on 0300 UT) GBR 16 kHz	MSF 60 kHz	Droitwich 200 kHz	Relative phase readings in microseconds N.P.L.—Station (Readings at 1500 UT) *GBR 16 kHz	†MSF 60 kHz
1	− 300·2	0	+ 0·1	634	528·5	17	− 300·1		+ 0·1	641	
2	− 300·1	0	0	635	528·5	18	− 300·0		0	641	
3	− 300·1	−0·2	0	636	530·5	19	− 299·9		+ 0·1	640	
4	− 300·1	−0·1	0	637	531·0	20	− 300·0		+ 0·1	640	
5	− 300·1	0	+ 0·1	638	531·0	21	− 300·1		0	641	
6	− 300·0		0	638		22	− 300·1		0	642	
7	− 300·2		+ 0·1	640		23	− 300·0		0	642	
8	− 300·0		0	640		24	− 300·0		+ 0·1	642	
9	− 300·0		+ 0·1	640		25	− 300·1		+ 0·1	643	
10	− 300·0		+ 0·1	640		26	− 300·1		+ 0·1	644	
11	− 300·1		+ 0·1	639		27	− 300·0		+ 0·1	644	
12	− 300·0		+ 0·1	639		28	− 300·1		+ 0·1	645	
13	—		+ 0·1	—		29	− 300·1		0	646	
14	—		0	—		30	− 300·1		+ 0·1	647	
15	—		0	—		31	− 300·1	−0·1	0	648	531·6
16	− 300·1		0	640							

All measurements in terms of H.P. Caesium Standard No. 334, which agrees with the N.P.L. Caesium Standard to 1 part in 10^{11}.
Note: The frequency offset of GBR for 1970 will be -300×10^{-10}
* Relative to UTC Scale; ($\mathrm{UTC_{NPL}}$ − Station) = + 500 at 1500 UT 31st December 1968.
†Relative to AT Scale; ($\mathrm{AT_{NPL}}$ − Station) = + 468·6 at 1500 UT 31st December 1968.

Contributors to this Issue

J. F. Young (M.1956) served a student apprenticeship with the General Electric Company, and then worked on closed-loop control systems at their Witton Laboratories. After a period with the laboratories of W. & T. Avery and of Thorn Automation, he returned to G.E.C. to start the Electronics Laboratory of the Switchgear Division. In 1959 he became manager of the Electronics Division of Donovan Electrical. After several years experience of industrial administration, he became a lecturer in Applied Electronics at the University of Aston. He is the author of several textbooks, in addition to about fifty published papers in the fields of electrical engineering and electronics. He has served on a number of Institution Committees, being currently a member of the Instrumentation and Control Group Committee.

Charles J. Drane, Jr., has been a research physicist with the microwave physics laboratory of the U.S. Air Force Cambridge Research Laboratories since 1955. He obtained his B.S. (Phys.) with honors in 1950 from Boston College, Newton, Massachusetts, and his S.M. (Phys.) in 1953 and his S.M. (Math.) in 1966 from the Massachusetts Institute of Technology. Since 1966 Mr. Drane has been studying applied physics at Harvard University, where he is at present a candidate for the Ph.D. At A.F.C.R.L. he has investigated the application of correlation techniques to the design of high-resolution aerials and his current fields of interest continue to be the analysis and synthesis of aerial arrays.

John F. McIlvenna received the B.S.E.E. degree from Massachusetts Institute of Technology in January 1962 and the M.S.E.E. from Northeastern University in June 1969. He is employed at the U.S. Air Force Cambridge Research Laboratories, Bedford, Massachusetts, as an electronic engineer, specializing in electromagnetics. He began his career in the microwave physics laboratory as a cooperative student in 1959 and became a permanent employee upon his graduation in 1962. His main interest has been in array and reflector aerial theory and its applications such as radar detection and ground mapping.

Robert R. Hoge has held his present position of vice-president, research and engineering, with Computing Devices of Canada Ltd., since June 1966. He has been employed by Bendix Corporation, the major stockholder in Computing Devices, since 1957. At the Bendix Research Laboratories his responsibilities included development of control instrumentation for nuclear propulsion power plants, development of numerical controls for machine tools, and early application of micro-integrated circuits. He was director, advanced technology applications in the Bendix Executive Offices from 1964 until his appointment at Computing Devices. During the period 1951–56 Mr. Hoge was employed by Battelle Memorial Institute, the U.S. Army Signal Corps Engineering Laboratories, and the Electronics Systems Laboratory at M.I.T. In 1955 he gained the M.S. degree of Massachusetts Institute of Technology and in 1957, following a research studentship at the University of Birmingham, received the M.Sc. degree; his thesis formed the basis of a paper on 'A decimal product accumulator' published in the *Journal* in February 1958.

Jiri Gregor graduated in 1953 at the Czech Technical University in Prague and was then appointed as a lecturer in the Department of Mathematics at the Faculty of Electrical Engineering, becoming a senior lecturer in 1957. In 1962, after postgraduate studies and submission of a thesis he obtained the degree of the Candidate of Science (C.Sc.) in applied mathematics. The main field of his research work has been the theory of analytic functions and its application to electric network theory, and following theoretical work in this subject he was appointed Associate Professor (Docent) in 1964. In 1968 he joined the Department of Engineering Mathematics at the University of Khartoum.

P. A. Vaughan graduated from the University of Swansea in 1960 with an honours degree in electrical engineering and joined Vickers Armstrong Aircraft Limited (now B.A.C.) at Weybridge, where he spent 18 months working on avionic systems. He was subsequently employed for five years by E.M.I. Electronics Limited at Feltham on the design and assessment of systems for the detection and measurement of infra-red energy. His present appointment is with Elliott's Space and Weapons Research Laboratory at Frimley.

U.D.C. 535.374:551.510.5

Application of Dye Lasers to Probe the Upper Atmosphere by Resonance Scattering

By

M. R. BOWMAN, M.A., D.Phil.,†

A. J. GIBSON, B.A.,†

M. C. W. SANDFORD, M.A.†

Reprinted from the Proceedings of the I.E.R.E. Conference on 'Lasers and Opto-Electronics' held at the University of Southampton on 25th to 28th March 1969.

A dye laser, tuned to one of the sodium D lines, has been used to measure atomic sodium which occurs in the atmosphere at a height of about 90 km. Details of the laser, tuning arrangements and other equipment are given, and the principle of the measurement and method of calibration are described. The results obtained and their relevance to atmospheric theory are indicated. The possibility of measuring other atmospheric constituents by the same technique is considered, and the laser requirements are indicated.

1. Introduction

Hitherto lasers have been used to study atmospheric density up to about 100 km by measuring Rayleigh scattering of the light pulse by molecules.[1,2] Scattering from aerosol layers has also been observed. Use of Raman scattering to measure number densities of individual major constituents in the troposphere has been demonstrated.[3,4] In the work reported here the laser has been tuned to a resonance frequency of a minor atmospheric constituent: a D line of atomic sodium. The cross-section for the resulting resonance scattering is many orders of magnitude greater than the Rayleigh cross-section and makes possible the measurement of the height profile of the small quantities of sodium which occur in a layer at about 90 km.

For such resonance-scattering experiments a tunable pulsed laser is required. Early suggestions included a neodymium-in-glass laser, a pulsed gas laser and a Raman laser. However, much greater flexibility of tuning is now available from optical parametric oscillators and dye lasers. For the generation of wavelengths within the visible spectrum, flashlamp-pumped dye lasers[5] are the obvious choice on account of their simplicity and economy.

2. Dye Laser

A flashlamp-pumped rhodamine 6G-in-ethanol laser[6] was used for the measurement of atmospheric sodium.

2.1. Flashlamp

A coaxial flashlamp, of the type used by Sorokin *et al.*,[5] was used as a pump (Fig. 1). A disadvantage of the original design was that since the electrical discharge does not fill completely the space between the silica tubes, non-uniform pumping of the dye solution and large beam divergences resulted. To obtain more uniform pumping, magnesium oxide powder is packed around the outer tube and the inner tube has a ground finish. This also prevents total internal reflection and 'whispering' modes. The use of rubber 'O' rings for all seals facilitates the replacement of the silica tubes which is necessary about every thousand shots. The lamp is connected by low-inductance connectors to a 0·5 μF rapid-discharge capacitor, normally charged to 14 kV, and the lamp is fired at intervals of about 5 seconds by reducing the air pressure between the silica tubes.

2.2. Basic Laser Characteristics

The dye solution flows continuously through the inner tube of the flashlamp. Optical feedback is provided by one of the windows of the dye cell, which carries a dielectric coating of 50% reflectivity, and an external 99% reflector. The flashlamp pulse, having a rise time of about 300 ns, generates a laser pulse of approximately 400 ns duration and 10 mJ output energy. The overall beam divergence is about 5 mrad and the direction of the beam does not change by more than 1 mrad from shot to shot. This stability of beam direction is most important in enabling the fine tuning described in the next section to be carried out.

2.3. Tuning

Spectral narrowing and tuning was performed by means of tilted Fabry–Perot interferometers inserted between the flashlamp assembly and the external reflector (Fig. 2). Three Fabry–Perot interferometers having plates with dielectric coatings of 50% reflectivity and free spectral ranges of 10 nm, 1·5 nm and 0·2 nm at 589 nm were employed; tuning was carried out by varying the angles of tilt. The output spectrum was examined on Polaroid film with a

† Radio and Space Research Station, Ditton Park, Slough, Bucks.

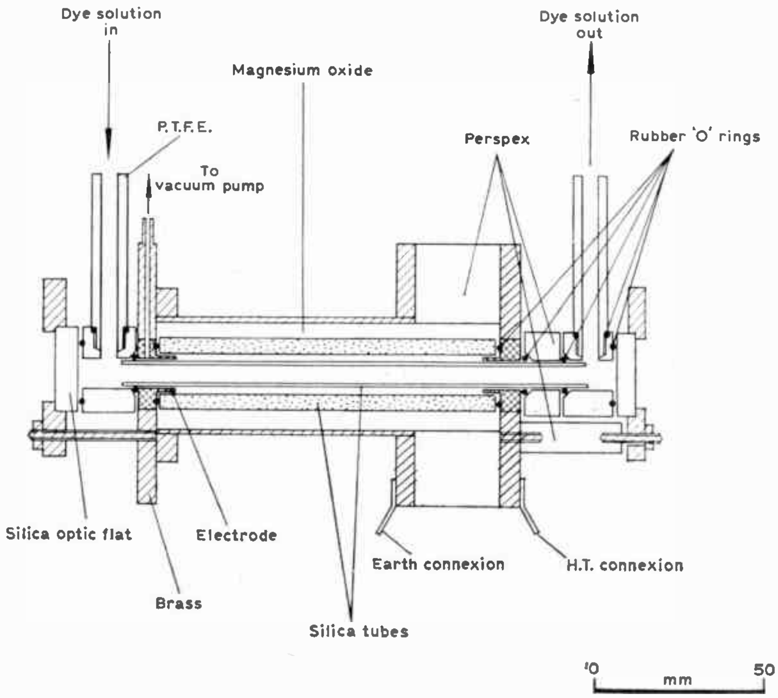


Fig. 1. Diagrammatic view of coaxial flashlamp.

grating spectrometer producing a dispersion of 0·8 nm per mm, crossed by a Fabry–Perot interferometer having 75% reflectivities and 56 pm free spectral range. The laser spectrum was narrowed from 10 nm to 0·005 nm, whilst the output energy was decreased from 10 mJ to about 3 mJ. The loss of energy is believed to be mainly due to reflexions from the outer faces of the Fabry–Perot plates.

It was found, however, that the laser could not always be tuned exactly to the required wavelength merely by adjusting the angles of tilt of the Fabry–Perot interferometers, because the windows of the dye cell, which were fused silica flats of 6·5 mm thickness, produced resonances at 20 pm intervals. For example, the front reflector of the laser consists of two parallel surfaces, the dielectric coating of 50% reflectivity and the silica–ethanol surface of 0·13% reflectivity. This seemingly low reflectivity of 0·13% cannot be neglected because the maximum and minimum reflectivities of the combined reflector, calculated from the sum and difference of the amplitude reflexion coefficients, are about 55% and 45%. Since the fractional change in double-pass transmission of the Fabry–Perot interferometers for a wavelength shift of 10 pm from the wavelength of peak transmission can be smaller than the fractional change in reflectivity of the front reflector, the final wavelength becomes locked to a resonance of the front reflector. Hence the Fabry–Perot interferometers set the wavelength to ± 10 pm and the final tuning must be carried out by varying the temperature of the windows. The 20 pm spaced resonances were found to shift by about 5 pm per deg C which is in agreement with the temperature coefficient of optical thickness of silica.

3. Measurement of Sodium

3.1. Equipment

The transmitter consisted of the dye laser followed by an inverted telescope with a 0·3 m diameter objective, which reduced the beam divergence to less than 0·2 mrad. Scattered photons were collected by a 1 m diameter mirror with a surface sufficiently accurate to define an overall beamwidth of 0·5 mrad. The received photons, detected by a photomultiplier, were recorded digitally as a function of range and the counts summed over many laser shots. The background count was measured between laser shots by means of a ratemeter. Further details have been given in previous papers.[1,2,7]

3.2. The 'Radar' Equation

The mean number of photo-electrons, C, counted in excess of the mean background in a height channel of thickness δh at height h, where $\delta h \ll h$, is given by

$$C = \frac{NAT^2\eta\delta h n(h)\sigma}{4\pi h^2} \qquad \ldots\ldots(1)$$

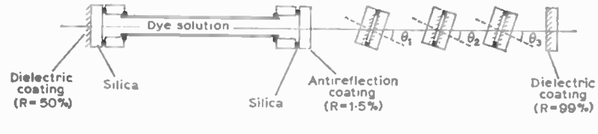


Fig. 2. Diagram of complete laser resonator showing tuning interferometers.

where N = number of transmitted photons
A = receiver area
T = atmospheric transmission coefficient
η = efficiency of the receiver in counting photons
$n(h)$ = number density of scattering atoms
σ = effective total scattering cross-section of an atom.

The resonance radiation is assumed to be isotropic. When the laser linewidth, $\Delta\nu$, is greater than the width of the resonance line, $\Delta\nu_R$, the effective cross-section is given by an equation of the form,

$$\sigma = \frac{e^2 f}{4\varepsilon_0 m_e c} \frac{1}{\Delta\nu} \qquad \ldots\ldots(2)$$

where e = electronic charge
f = oscillator strength of the resonance transition
ε_0 = permittivity of free space
$m_e c$ = electronic mass
c = velocity of light.

As the laser line width is reduced below the resonance line width, σ increases towards a limiting value,

$$\sigma = \frac{e^2 f}{4\varepsilon_0 m_e c} \cdot \frac{1}{\Delta\nu_R} \qquad \ldots\ldots(3)$$

The exact definitions of $\Delta\nu$ and $\Delta\nu_R$ depend upon the frequency profiles of the laser line and resonance line.

It is not necessary to measure N, A, T and η individually in order to determine $n(h)$ because the system can be calibrated by observing the signal scattered from an altitude of 30–35 km at the same time as the signal from the 90 km region is being measured. A value of $NAT^2\eta$ can be calculated from the 30–35 km signal which is believed to be entirely Rayleigh scattering by the molecules of the air whose density and cross-section are known and constant to a few per cent at this height. This method is used because N, T and η are difficult to measure. A spectrograph measurement of the transmitted line profile and knowledge of the resonance line profile gives the value of σ according to eqn. (2) or eqn. (3), and hence $n(h)$ can be obtained from eqn. (1).

3.3. Results

The first results have been given elsewhere.[7] More recently, using a narrower transmitted line, better measurements, shown in Fig. 3, have been made. However, since a large number of shots were used for the purpose of tuning the laser and aligning the transmitter, the output energy during the 450 shots of the measurement was considerably less than the maximum described in Section 2. The horizontal bars indicate the range of height over which the signal was averaged. The vertical error bars indicate $\pm$ one standard deviation as derived from the photon counts in each height range. Based on the errors involved in the measurement of the laser spectrum, there is 95% confidence that the absolute value of the density scale is within $\pm 50\%$ of that indicated in Fig. 3.

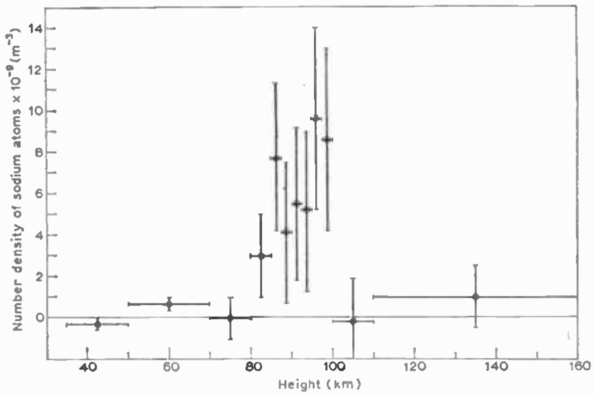


Fig. 3. Measurement of atmospheric sodium. (24 Nov. 1968, 2110–2253 U.T.)

3.4. Relevance of Sodium Measurements

Previously it has been possible to study the sodium only in daytime or twilight when sunlight excites resonance radiation. The laser method can provide profiles at night to complete a 24-hour picture of the layer. Already the measurements have shown that night-time heights and densities are similar to those in daytime. As the laser is improved and more accurate measurements are made, it is expected that a study of the changes of the night-time profile will throw light on the balance of the reactions that are believed to produce sodium from its oxide, namely

$$Na + O_3 \rightarrow NaO + O_2, \quad NaO + O \rightarrow Na + O_2$$

and thus on the ratio O/O_3 which is thought to be quite variable. It may also be possible to discern the effects of meteors and dust layers that may provide sources or sinks for sodium.

4. Measurement of Other Constituents

Table 1 lists some atmospheric constituents which it may be possible to measure by laser radar. The laser energy necessary to give one photo-electron per shot has been calculated; this is a rough value of the minimum useful signal. When the signal count, C, is appreciably greater than the background count, the r.m.s. error in C is $C^{\frac{1}{2}}$ since the number of photo-electrons obeys a Poisson distribution. Thus the total column number density will be measured with 10% error from 100 laser shots of the energy in Table 1. In our sodium measurements, the ratio of signal counts to background counts was about $\frac{1}{3}$, the background being mainly due to sodium street illumination. Hence four times as many shots were required com-

Table 1. Laser requirements for the measurement of atmospheric constituents

Species	Wavelength (nm)	Maximum Effective Cross-section [a] $\sigma(m^2)$	Column Number Density [b] $\int n(h)\,dh(m^{-2})$	Height h (km)	Efficiency Factor [c] $T^2\eta$	Energy Required[d] (J)
Na	589·0	$1{\cdot}6\times10^{-15}$	6×10^{13}	80–100	1×10^{-2}	5×10^{-5}
K	769·9	$1{\cdot}3\times10^{-15}$	6×10^{11}	80–100	$1{\cdot}5\times10^{-3}$	6×10^{-2}
Li	670·8	$1{\cdot}2\times10^{-15}$	6×10^{10}	70–90	7×10^{-3}	5×10^{-2}
Ca^+	393·4	8×10^{-16}	3×10^{12}	100–200	$1{\cdot}5\times10^{-2}$	2×10^{-3}
Ca	422·6	$1{\cdot}8\times10^{-15}$	?	?	2×10^{-2}	?
Al	396·1	$1{\cdot}3\times10^{-16}$	?	?	$1{\cdot}5\times10^{-2}$	?
$He^*(2^3P-2^3S_1)$	1083·0	2×10^{-16}	5×10^{11}	>200	7×10^{-5}	40
$(3^3P-2^3S_1)$	388·9	1×10^{-17}			$1{\cdot}5\times10^{-2}$	10
$NO^*(b^4\Sigma^- - a^4\Pi_1)$	870(1–0) 790(2–0)	5×10^{-19}?	2×10^{14}?	70–120	5×10^{-4}	1
$N_2^+(A^2\Pi_u - X^2\Sigma_g)$	391·4(0–0)	2×10^{-19}	10^{13}	>200	$1{\cdot}5\times10^{-2}$	20
$N_2(B^3\Pi - A^3\Sigma_u^+)$	885(1–0)	4×10^{-20}	?	?	5×10^{-4}	?

[a] Cross-section per molecule for laser linewidth $\leqslant 0{\cdot}5$ pm, ignoring hyperfine structure (which reduces Na cross-section by 30%), and taking the diatomic molecules thermally spread over 20 rotational states of the zero vibrational level.

[b] Column number densities are deduced from twilight measurements. NO* assumes 10^{-4} of NO molecules measured by daytime rockets are in the metastable $a^4\Pi_1$ state.

[c] $T^2\eta$ has been calculated assuming attenuation for a very clear atmosphere, a S1 photo-cathode for $\lambda > 0{\cdot}8$ µm, S20 for $\lambda < 0{\cdot}8$ µm and an optical transmission of 0·25 for the receiver.

[d] Energy required for one photo-electron per laser shot, calculated from (1), putting $A = 0{\cdot}6$ m².

pared with a dark site. At a suitable site, the background could probably be neglected for night-time measurements of all the constituents in Table 1. The night-time densities of He* and N_2^+, however, may be very low, and a study of the twilight densities given in the table may be prevented by the background of scattered sunlight.

Clearly there are two main requirements: a high energy, near ultra-violet, laser to measure Ca^+ and possibly Ca, He* and N_2, and a high energy laser in the 760–800 nm region (possibly a ruby-laser-pumped dye-laser) to measure K, NO* and N_2^*.

5. Conclusion

It has been shown that a dye laser can be sufficiently well stabilized and tuned to make possible the excitation of resonance radiation from atmospheric sodium, and useful night-time height profiles have been obtained. Although sodium is the easiest to measure, with improved dye lasers a number of other interesting constituents should be measurable.

6. Acknowledgment

The work described was carried out at the Radio and Space Research Station of the Science Research Council and this paper is published with permission of the Director.

7. References

1. Bain, W. C. and Sandford, M. C. W., 'Light scatter from a laser beam at heights above 40 km', *J. Atmos. Terrestr. Phys.* **28**, pp. 543–52, June–July 1966.
2. Sandford, M. C. W., 'Laser scatter measurements in the mesophere and above', *J. Atmos. Terrestr. Phys.* **29**, pp. 1657–62, December 1967.
3. Leonard, D. A., 'Observation of Raman scattering from the atmosphere using a pulsed nitrogen ultraviolet laser', *Nature*, **216**, pp. 142–3, 14th October 1967.
4. Cooney, J. A., 'Measurements on the Raman component of laser atmospheric backscatter', *Appl. Phys. Letters*, **12**, No. 2, pp. 40–42, 15th January, 1968.
5. Sorokin, P. P., Lankard, J. R., Moruzzi, V. L. and Hammond, E. C., 'Flashlamp-pumped organic-dye lasers', *J. Chem. Phys.* **48**, 10, pp. 4726–41, May 1968).
6. Gibson, A. J., 'A flashlamp-pumped dye laser for resonance scattering studies of the upper atmosphere', *J. Sci. Instrum. (J. Phys. E.)* Ser. 2, **2**, pp. 802-6, September 1969.
7. Bowman, M. R., Gibson, A. J. and Sandford, M. C. W., 'Atmospheric sodium measured by a tuned laser radar', *Nature*, **221**, pp. 456–7, 1st February 1969).

Manuscript received by the Institution on 10th April 1969 (Short Contribution No. 129/Com.25).

Research and Development Project Management

Techniques for Guiding Technical Programmes towards Corporate Objectives

By

R. R. HOGE,
B.S.(E.E.), S.M., M.Sc.†

Presented at a meeting of the Ottawa Section of the Canadian Division of the Institution on 5th February, 1969.

The paper first discusses the evaluation of proposed projects and the selection of projects. Techniques for planning and managing projects are then described and finally some of the aspects of transferring results to production are dealt with.

1. Introduction

We are involved in an industrial research revolution that is affecting the dynamics of corporations to an extent comparable with that of the industrial revolution of 150 years ago. Today corporations view the R & D process as being one of continuous operation to provide the organization with a steady flow of new and improved products and processes. Accordingly, in the past 20 years there has been a rise in the organizational responsibility of the R & D manager, so that now there are a continual production process and a production manager for Research and Development.

As an illustration, the significant element in President Kennedy's 1961 decision to launch the United States on the lunar programme was not that the work would involve thousands of individuals but that a specified completion date of 1970—a decade later—could be announced for so complex an R & D task. Thus, at both national and corporate levels the point has been reached where R & D is regarded as a production process subject to planning, control, and optimization.

2. Evaluation of Proposed Products and Selection of Projects

2.1. *Corporate Objectives*

Definition of company objectives is the initial and most important step in the general management process, largely because the objectives of an organization are its sole reasons for existence. Collectively, the objectives of a company are the yardstick against which all requirements and accomplishments are evaluated.

As a prerequisite to evaluating new product ideas and to selecting development projects, the company operating objectives should be established without regard to specific projects. (See Appendix.)

The general goals are largely the kind which will be decided by the Board of Directors. But they may not be explicit on the subject of product lines. In that case, the top Marketing and Engineering executives should jointly select new product lines and, with evaluation help from financial analysts, show that the proposed new products will ensure achievement of the desired growth and financial objectives. General management approval can then be obtained.

With new product-line objectives established, the R & D manager can formulate general R & D goals and plans to achieve them—and it is important that the top marketing executive understands these goals and supports them.

2.2. *Guidelines for New Products*

The passing down of general R & D objectives to the lower tiers of R & D management sets the framework for preparing new product development proposals. The importance of clearly communicating these goals downward, both within Marketing and within Engineering cannot be overemphasized.

Marketing staff ignorance of or disregard for product guidelines can also cause waste. For example, we receive many Requests for Proposal from the Government to perform paid Research and/or Development. Sometimes a salesman tries to initiate a proposal exercise on RFP's which are obviously outside the company's interests. Attention to clearly-stated product guidelines and R & D objectives should discourage a 'buck-shot' approach to new business.

An important related topic for companies producing military equipment has been stated by Professor E. B. Roberts of the Management School at M.I.T. He says that, 'contrary to superficial appearance, the vast majority of R & D contracts go to "a preferred sole-source", even though there has actually been a formal competition.'[1] The companies that get the big jobs are those who were carrying on constructive technical discussions with the sponsor's technical specialists long before the competitive tender was published. Experience in our own company bears this out. So, for companies in the military hardware business such as ours, it is essential for the

† Computing Devices of Canada Ltd., P.O. Box 508, Ottawa, 4, Ontario.

Marketing staff to develop new business through technical interchange with customers on subjects and products which are compatible with company objectives.

2.3. *Product Ideas and Project Proposals*

Any company can concentrate with advantage on the early stages of determining 'what should be developed'. On the whole, it takes just as much money to develop a new-product failure as it does to create a spectacular winner in the market place. Successful project management begins with tackling problems which have solutions for which there is a marketable demand.[2]

Where do new product ideas come from? As you are probably aware, the largest single source in the typical quality-product company is the sales or market research staff. These people are in frequent contact with customers, hence they can best determine what will sell. Their ideas, however, usually deal with modifications or improvements in present products; the less frequent ideas for radically new or different products usually come from research or engineering.[1]

If our company is typical, once a year financial and operating plans are reviewed and updated, being especially particular about the next fiscal year. For us this includes preparation and comparative evaluation of proposals for projects to be undertaken as company engineering expense. Key managers are advised on general product and engineering goals, together with the approximate amount of money which has been budgeted for each area, and they are asked to prepare technical and cost proposals for development projects within this framework.

Corresponding product specialists on the Marketing staff get the same initial information. These product specialists work with the engineering proposers, formulating:

(i) desirable performance specifications
(ii) target selling prices
(iii) required operating environments
(iv) reliability and maintainability goals.

These are in the order that the salesman and engineer think of and convey them. There is a tendency for both to overlook operating environments and reliability goals in the first instance—but these are important because they affect the cost of both the development programme and the resulting product.

The R & D Manager then produces a development project proposal, which includes the following:

Specifications and objectives (from above).
Block-diagram outline of the proposed technical approach.
An Engineering Job Analysis.
A time-phased labour estimate (broken down into work orders).
A schedule by activity and/or work order.

Usually, the above is actually generated in two cycles. In the first cycle, the engineer proposing makes a very rough estimate of development and production costs which is discussed with Research and Engineering and with Marketing before proceeding with the full proposal.

Concurrently with the R & D proposal effort, the Marketing Product Specialist should prepare a market analysis, delineating

Size of the probable market.
Probable share of the market.
Specific known customer projects and programmes which require this type of product.
Projected volume of business to be expected.

The above should be coupled with a life-cycle analysis by financial and production cost-estimating staff of the expected profitability of the proposed new product, which is every bit as important as the development project proposal. Companies experienced in R & D say, in effect, 'Our men can develop anything. Show us enough sales and enough profit and we'll spend the necessary time and money to develop it'. This is because most development work should and does take place within technology largely understood at the outset of the project.

2.4. *Project Selection*

When the proposals are all in, their cost probably adds up to more than the budget will support. So, assuming all proposed products satisfy the general product-line criteria it is necessary to eliminate the least promising projects on the basis of financial performance or low urgency.

With regard to financial performance, the key factor is to use cash flow and Return on Investment analysis for at least the next five years as measurement criteria. Obviously, the accuracy of these techniques for evaluating financial performance is much dependent upon the validity of underlying market and engineering forecasting. Nevertheless, if all proposed projects are evaluated by the same method, a useful comparative rating of proposals is obtained, indicating those which cannot qualify for funds.

Most new product failures are not the result of technical incompetence; generally products 'flop' because the idea or concept was wrong for the market in the first place. Thus, management attention to project selection is an excellent source of high returns.

3. Planning Proposed Projects

It is pointed out in Section 2.3 that development project proposals include product specifications, a technical approach in block diagram form, an Engineering Job Analysis, a time-phased labour estimate, and a project schedule. The last four of these actually make up a development project plan for achieving the first. We require this plan, regardless of whether the proposal is for approval by our own management or for competitive submission to a customer.

3.1. *Assign a Project Manager*

We assign a senior development engineer with appropriate experience to coordinate the project planning, labour estimating and scheduling. If possible, we assign the same man we expect to be Project Engineer or Engineering Project Manager if the project is approved. This involvement is necessary because:

> 'Participation tends to increase commitment; commitment tends to heighten motivation; motivation which is job-oriented tends to make managers work harder and more productively; and harder and more productive work by managers tends to enhance the company's prosperity; therefore, participation is good.'[3]

The planning starts with a Quotation Request which

ENGINEERING JOB ANALYSIS FORM 1347	PROGRAM NAME VOTE 5 DOPPLER COUPLED MAP	PROGRAM NO. 0250	ISS 5	DATE OF ISSUE 14 June 1968
PROJECT ENGINEER C.R. Robinson	MANAGER R.J. Patton	DIRECTOR I.J. Irvine		

1. SCOPE

This EJA is re-issued as a result of a clearer definition of the requirements of the program brought about by discussion and meeting between Engineering and Marketing. The scope of this work will be to design and build circuitry additional to that existing in the Solid State PMDS in order to couple such a system to Doppler or Air Data Inputs. This extends to breadboarding and proving feasibility only.

2. RELATED DOCUMENTS

Memo dated 17 May 1968, J.C. Alexander to K.G. Thorne.

WG 13 Moving Map Display Proposal T060/909D/TP1.

T060/906C Q.R.

3. INTRODUCTION

The original approach reflected in previous issues of the EJA was to develop digital modules capable of performing a variety of analog functions, usually performed electromechanically, and to subsequently interconnect these to form a solid state equivalent of the Analog Map System. This approach has now been slightly modified and the completely modular approach has now been superseded by an approach which will produce the most economical circuit design to meet the complete requirement. The following constraints are placed on this program:

a) Maximum commonalty, consistent with economical design, with the existing design of the 0360/A7E PMDS,

b) WG 13 program requirements to be considered as growth items and the design to reflect this,

c) The potential application of the design to the requirements of a solid state PHI to be considered, i.e., interface directly with BDH Indicator etc.

d) The system to incorporate the present Display Unit (0360/A7E) with minimum modification.

e) Production hardware resulting from this program should be available by early 1970, therefore design considerations should be limited to the most advanced currently available state of the art components and techniques.

The system design will proceed to a solid state pictorial PHI having 1) Heading Card, 2) Bearing Pointer, 3) Range Indicator (numeric), together with the Moving Map Display.

After breadboarding of new designs, it is assumed that at least a breadboard Display Unit will be available together with some breadboards of the A7E/G360 EAU. Initially Engineering sketches will be produced and circuits will be built up on either plug boards or I.C. master PCB's. Some circuits may be transferred to experimental PCB's to test effect of configuration. Inter board connection to be made via PCB connectors. The circuit card assemblies should be assembled in a veripak or similar structure for system breadboard testing.

Fig. 1. Typical engineering job analysis.

4. DETAIL

Project No.	Work Order	Cost Centre	Activity
G 251	102		System Design
		1140	The system concept will be defined evolving out of the documentation listed and the requirement as stated by marketing for a "low cost navigation system". The system design will be continued to approach an optimum configuration in terms of state of the art technology and compatible requirements of military MMDS and PHI's.
G 251	103		Circuit Design
		1130	A hardware design philosophy will bo formulated concurrent with the production of the system specification.
			The design effort will be concentrated in four main areas:
			1) Implementation of the 'CORDIC' algorithm. This module will perform rotation, vectoring and binary to BCD conversion and will form the heart of the computer sub-system.
			2) Non-volatile memory. Presently a small 256 bit linear select core memory is predicated. Investigation of this and other approaches will be made and a breadboard constructed with associated read/write logic.
			3) Velocity Sensor Interface. This module will accept Doppler pulse trains VH and VD and effectively resolve these into N-S and E-W components. Use will probably be mado of the 'CORDIC' coordinate rotation capability.
			4) Map Electronics Interface. This area will require extensive study and breadboarding. A trade-off will be made between retaining the present servo error generating circuitry and combining this in the arithmetic section of the nav. computer.
			The design of power supplies will not be considered - bench supplies will be used as required.
G 251	103		Mechanical Design
		1740	A packaging study will be made for the core memory with particular attention to ease of assembly and vibration requirements. Some PCB layouts may be roquired and schematics will require drawing.
G 251	103		Fabrication
		1860	A number of panels and chassis may be required from rough engineering sketches. Estimated initially as one of each. The panel to contain push buttons and rotary switches. The chassis to hold say 10 A7E PCB's.
G 251	103		Assembly
		1850	A number of PCB's may require assembling and a chassis wired, assume 10 A7E PCB's.
G 251	103		R & M
		1870	Examination of engineering schematics and discussion with engineers. Formulation of realistic R & M targets and high lighting potential problem areas. RFI diagnostic appraisal to be performed at intervals.
G 251	103		Specifications
		1870	Assistance in specification and location of new components.
G 251	104		System Assembly and Test
		1130	Most of the initial design and development will be at the module level however further design of additional circuitry will be required to interface them at the system level. The majority of breadboard testing including module testing will be at the system level. Design modifications will be incorporated as determined and all testing will be at room temperatu[re] The system should be assembled in a Veripak unit and tested with either an available Map Head or the existing breadboard film transport. This work order will include the preparation of preliminary test specifications and procedures to apply to breadboard testing and the production of test results.
G 251	105		Project Engineering
		1130	Control of technical and cost considerations and responsibility for preparation of final documentation test procedures, test results, specifications and circuit and system descriptions.

Fig. 1 (cont.)

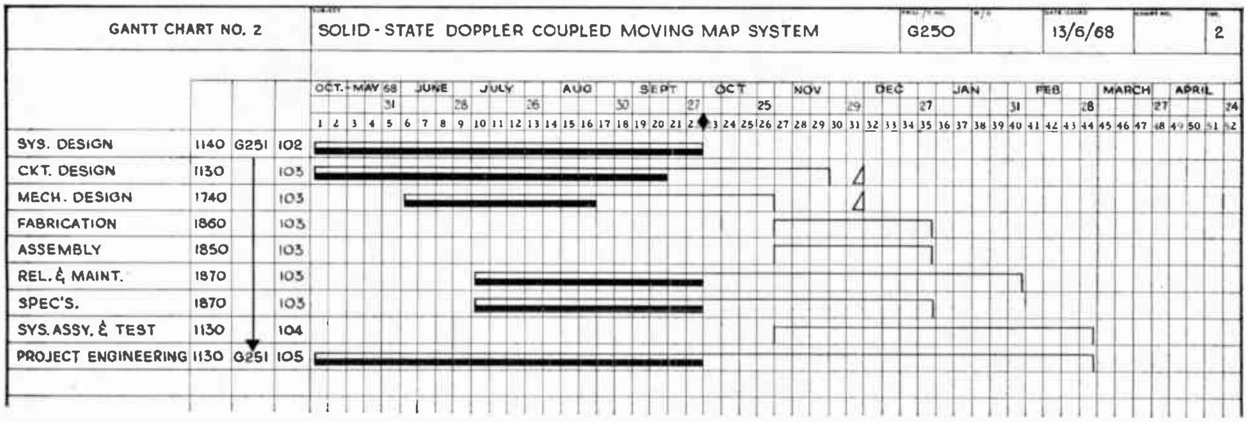


Fig. 2. Sample Gantt chart.

includes a work statement. It should convey, directly or by reference, all essentials of what is to be achieved by the project.

Our prospective project engineer first works out the technical approach to be followed. Frequently he enlists the help of one or more other experienced engineers—perhaps systems specialists—to formulate the basic product concepts. Block diagrams and preliminary layout drawings are prepared to convey the product concepts to specialists in supporting departments who will be involved in the programme and will therefore participate in the project planning process.

3.2. *Formulate an Engineering Job Analysis*

Next the technical proposal coordinator or project engineer will prepare an Engineering Job Analysis.[4,5] This is a word description of the project, broken down into clearly definable project phases and/or tasks. Figure 1 is a sample of a typical EJA.

The EJA begins with a paragraph on the Scope of the project, followed by a section identifying Related Documents, specifications procedures, etc. Next an Introduction expands on the statement of Scope, detailing limits and conditions of the project, relationship with other preceding or concurrent projects, special constraints and Security restrictions. The balance (and bulk) of the EJA deals with the tasks which make up the Project Work Order structure. Typical tasks are as follows:

- (a) System study and analysis.
- (b) Preparation of system and unit specifications.
- (c) Preliminary electrical design.
- (d) Preliminary mechanical design.
- (e) Prepare reliability and maintainability plan.
- (f) Design and build breadboard(s).
- (g) Electrical specifications and parts lists.
- (h) Mechanical design and detail drafting.
- (j) Construction of engineering models.
- (k) Engineering design proving procedures.
- (m) Design proving tests.
- (n) Production drawings and documentation.
- (o) Instruction and maintenance manuals.

Dates for Key Events:

- (1) Design reviews
- (2) Progressed reports
- (3) Scheduled completion, delivery dates.

A paragraph is written about each project task (work order) describing what is to be accomplished and also outlining as specifically as possible the responsibility of each department on that task. In connection with this, the project engineer will also usually prepare a preliminary Gantt schedule chart such as Fig. 2 showing his desired or suggested timing for all events. The technical approach and the EJA describe HOW the project will achieve WHAT the 'quotation request' work statement requested. The Gantt chart says WHEN.

3.3. *Compile Cost Estimates*

This resulting EJA is used for communicating the project requirements and duties to the managers of all departments, so that they can provide their own estimates of labour required (more participation). The labour estimate of every department is recorded in a time-phased manner, as indicated in Fig. 3. (For authorized projects, this later becomes 'Forecast Project Data.')

We don't expect the prospective Project Engineer to accept the labour estimates from supporting departments without question. We want him to understand them, and to challenge each estimate if necessary, so that he is eventually convinced that it is justifiable. Here the proposal coordinator is in the 'General Contractor' role. Incidentally, in the process of estimating and negotiating with supporting departments, some adjustments to the Gantt schedule are usually made.

ENGINEERING ESTIMATE FORM   **COST CENTRE** 1130   PAGE 1 OF 1

T. T060-906C   E.P.P. ISSUE 2 (13/6/68)   E.J.A. ISSUE 5 (14/6/68)   TITLE DOPPLER COUPLED MAP   DATE 16/6/68   PREPARED BY C. R. ROBINSON (PROJECT ENGINEER)   APPROVED BY R. J. PATTON (MANAGER)

FORM 1396A

MONTHS ARO

PROJ. 1–4	W.O. 5–8	DESCRIPTION	* 9–12	* 13–16	LAB. CAT.	1 17–20	2 21–24	3 25–28	4 29–32	5 33–36	6 37–40	7 41–44	8 45–48	9 49–52	10 53–56	11 57–60	12 61–64	13 65–68	14 69–72	15 73–76	UNSHED 77–80	TOTAL HOURS	TOTAL £	REMARKS
1	3	CIRCUIT DESIGN			SR ENG'R	5.60	5.60	5.90	5.80	5.40	5.20													
					x TOTAL DL	5.60	5.00	5.90	5.80	5.40	5.20													
					MATERIAL				1.50														150	
		* POSTED TO PAGE ___ LINE ___			T L	2.50	2.50																500	
					D.C.		3.00																300	
1	4	SYSTEM ASSEMBLY AND TEST			ENG'R					.30	1.80	2.80	1.30	1.00										
					TECH IV					.40	.60	.40	1.90	.60										
					x TOTAL DL					.80	2.40	3.20	3.20	1.60										
					MATERIAL	30.00	30.00																6000	
		* POSTED TO PAGE ___ LINE ___			T L																		–	
					D.C.																		–	
1	5	PROJECT MANAGEMENT			SR ENG'R	1.60	1.60	1.20	1.20	1.20	1.20	1.20	1.60	1.60										
					x TOTAL DL	1.60	1.60	1.20	1.20	1.20	1.20	1.20	1.60	1.60										
					MATERIAL																		–	
		* POSTED TO PAGE ___ LINE ___			T L																		–	
					D.C.																		–	
					x TOTAL DL																			
					MATERIAL																			
		* POSTED TO PAGE ___ LINE ___			T L																			
					D.C.																			

Fig. 3. Engineering estimate form.

At this point, the Engineering Administration staff provides help to compile and integrate the time-phased labour estimates into a consolidated package showing:

time-phased labour by category and cost-centre for each work order,

total labour hours for each cost centre.

The consolidated project plan and estimate is then returned to the prospective Project Engineer to sign-off, and for approval by his supervisor and the supervisor's superior.

The consolidated engineering estimate is turned over to Central Estimating and Analysis, who convert it to a dollar estimate by application of the appropriate labour rates, overheads, burdens, G & A factors, fee, etc.

3.4. *Proposal and Summary Plan*

While all this planning, estimating, scheduling, etc., is going on, the proposal coordinator is also supervising an *ad hoc* staff who write a technical proposal. Typically this document describes the proposed product, the customer-satisfying features it embodies, and it outlines the project organization and plan for developing the product. This technical proposal is often itself a major production.

Getting back to the cost estimate and plans—when the financial estimators are finished, the whole package is reviewed at two levels:

First by Finance, Marketing and (sometimes) Engineering to consider whether all bases are covered.

Then by Finance and the President/General Manager for approval of the dollar value of the project.

Sometimes the project plan and Engineering estimate is sent back to Engineering for 'fine-tuning'—because somebody (the General Manager or the customer) wants it modified. In that case, a revised Gantt schedule and consolidated engineering estimate eventually emerges—and is duly signed off by all levels of engineering management—and it then becomes Engineering's revised commitment. An important point is that no one, except the General Manager, is allowed arbitrarily to cut an engineering project labour estimate.

An aid to manpower planning is a by-product of the proposal and cost estimating phase whereby we key-punch the time-phased labour estimates for each department (cost centre). As soon as a project plan and labour estimate passes 'bid review' we decide on the most probable calendar month for starting and add the manpower load to our existing project time-phased load. At least once a month we make a computer print-out histogram as shown in Fig. 4. Each manager gets a copy for his department. The heads of R & D, Marketing and sometimes Industrial Relations also get copies for all departments. This keeps Marketing aware of availability or dearth of manpower in future months. In our company, there is much more paid R & D than 'company expense' engineering—so it is Marketing's job to keep the Engineering backlog up.

4. Managing the Project

Some of the proposed Project Plans will get approval, providing the R & D organization with a commitment to fulfil.

The first problem is always one of properly staffing the programme. Each group who will participate now has acquired some additional 'backlog'. Can they cope with it using existing staff?

The latest set of manpower loading histograms will give the Project Engineer/Manager a 'feel' for which department will have manpower trouble. Each department head usually knows his manpower situation intuitively. With luck, each department has a range of jobs and schedule commitments which permit some juggling to satisfy the commitments on all jobs. If not, we try to hire or rent people to make up discrepancies.

4.1. *Project Structure*

A functionally-organized R & D staff is associated with project and product-oriented work planning. We develop products through well-planned, well-managed projects which are astride the functional R & D departments 'matrix-wise'. (See Fig. 5.) The traditional pyramid organization of functional departments has the following advantages:

(i) It helps individual researchers or specialists to keep in touch with others at work in the same subject field and encourages him to keep up to date in his technology.

(ii) It fosters teamwork within a specialization.

(iii) It enables a specialist to match his thoughts, ideas, hypotheses and conclusions with others skilled in his field.

(iv) It assists in avoiding duplication of technical manpower and physical facilities for a given service or skill.

(v) It facilitates starting a new study within a discipline.

(vi) While individuals may periodically assume project responsibilities, functional departments serve as permanent homes. The specialist needs less often to adjust to new chiefs and radically different working relationships. This gives him a greater feeling of stability and security.

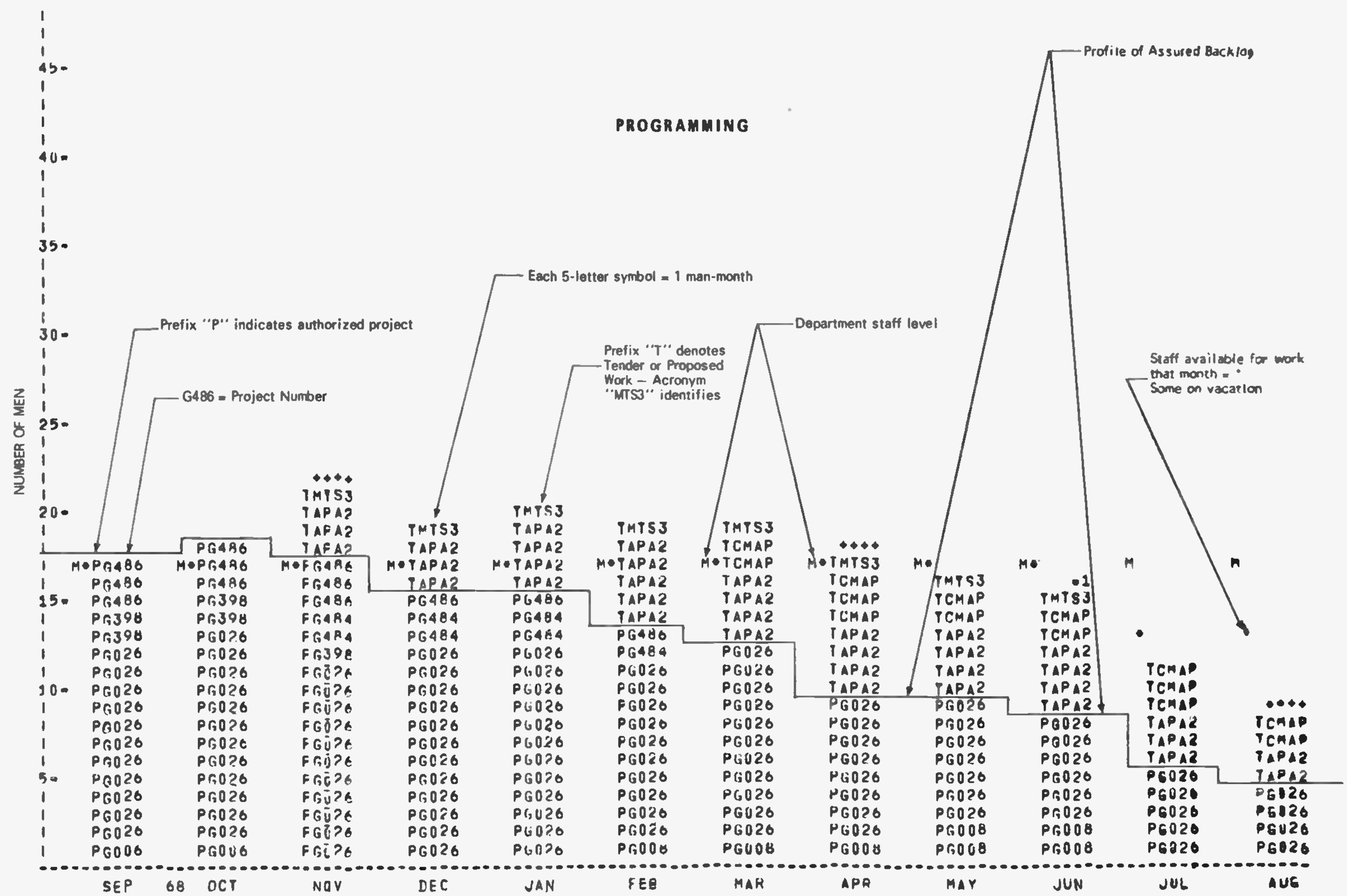


Fig. 4. Manpower loading histogram.

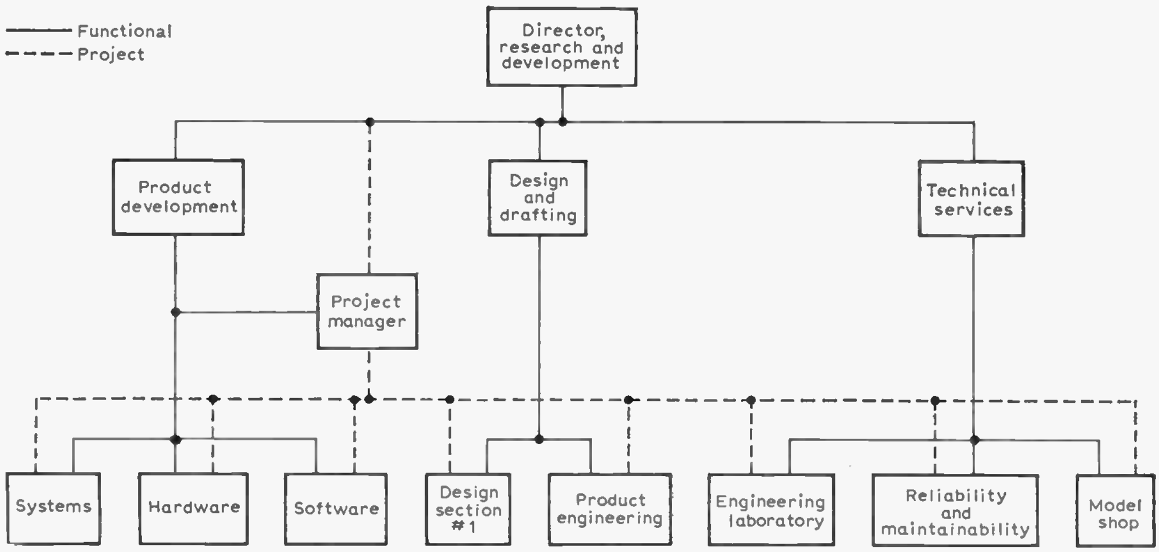


Fig. 5. Functional R & D organization.

But in today's economy, any going concern is constantly confronted with a stream of projects that supply the work for the members of the organization. Each project is in a different stage of completion; one may be merely a concept undergoing feasibility study, another in development, some in production, and some being phased out of the product line in favour of new models. This stream of projects, each with its own problems and peculiarities, requires that an individual be designated as Project Manager, with the responsibility for keeping abreast of all the company's work on that project. Project management, therefore, is a general management activity and includes such functions as planning, organizing, motivating, integrating, directing, and controlling efforts to obtain a specific goal. In many ways, project management is similar to functional or traditional management. The Project Manager, however, may have to accomplish his ends through the efforts of individuals who are paid and promoted by someone else in the chain of command.

4.2. *Characteristics of Project Management*

Project management has several characteristics that do not exist in traditional management.

First, project management requires that functional lines (parent company) and organizational lines (outside organizations) be crossed to accomplish project tasks. Project management is more concerned with the flow of work in horizontal and diagonal relationships than with the vertical scalar chain of authority. But project management cannot exist alone. The Functional Managers and the Project Manager share the authority and the responsibility for the project activities as follows:

The Project Manager:

(1) Unifies the project affairs such that the objectives are satisfied.
(2) Establishes the funding, scheduling, and performance standards for the project on non-contractual matters.
(3) Acts as the focal point for customer technical contact with the company.
(4) Resolves any conflict that threatens to disrupt the project activities.

The Functional Manager:

(1) Provides functional facilitation to this as well as the other projects in the organization.
(2) Prescribes how the day-to-day tasks will be performed.
(3) Maintains an existing capability up to the state-of-the-art in the department's speciality.

The Project Engineer/Manager needs authority to accomplish this work, but his authority is usually far removed from the chief executive of R & D. Yet the Project Manager is, in effect, the 'general manager' or 'general contractor' of the project as far as the company is concerned, and exercises two types of authority:

Legal authority—which is granted by a specific job charter or 'terms-of-reference' document.

Personal authority—which is his influence, accepted by his associates. This means that the project manager's authority has no functional or organizational constraints. Instead, it pervades the environment, seeking out the people and ideas it wishes to control.

PROJECT STATUS REPORT

Computing Devices of Canada Limited — Bendix

1 Oct 1968 (STATUS DATE) Title: MALLARD - MESSAGE COMPOSITION APPLIQUE Proj. No. G 450

Date Prepared 9 Oct 1968 RECORD Project Engineer Manager: D. G. MOULDING

START DATE	COMPLETE DATE	WORK SCHEDULE: TASK	YEAR 1968 MONTH: J F M A M J J A S O N D
1 JAN	29 MAR	FUNCTIONAL DESCRIP. OF SYSTEM	
29 "	30 SEPT	INTERMOD. SYSTEM DEFINITION	
		FINAL TRADE-OFF MATRIX	
1 MAR.	25 SEPT	SOFTWARE FLOW CHARTS	
28 SEPT		SOFTWARE ASSESSMENT	
10 MAY		SIMULATION ON ADC-2	
15 MAR	25 SEPT	HARDWARE STRAW SYSTEMS	
29 JULY		COMPONENT EVALUATION	
3 SEPT		RELIABILITY & MAINTAINABILITY	
24 MAY		MECHANICAL DESIGN	
		COST ESTIMATING	
		FINAL REPORT	

FISCAL DATA: (Direct Labour)

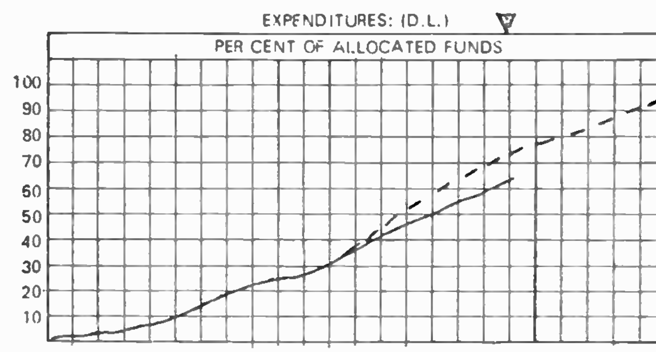
Allocated Funds .... $ 24,288
Actual Expend through last month ... 15,468
EST Expenditure this month 3,642
Cost on Completion 24,288

PROGRESS:

☐ On Schedule

☒ Behind Schedule

Anticipate ... 1/2 .. Month(s) Slippage



TECHNICAL STATUS: (Use 2nd sheet if required)

1. Summarize progress for the month. 2. List planned accomplishments for next month. 3. Estimate expenditures for next month.

NOTES:

1. Discussions were held with technical representatives from 8 manufacturers of "read-only" memories. The suitability of these companies' products for use in the M.C.A. was discussed and projected cost data obtained for the memories and associated addressing logic.

   An MOS pulse pattern generator has been designed and will be used in a demonstration of addressing and read-out from two memories; (Philco-Ford-1024 bit, General Instrument - 2048 bit).

   Fabrication of a tray and circuit board to hold 200 dummy memory modules has commenced and moulds to produce the modules have been made.

   A first listing of the required program instruction set has been made together with counts of the number of program instructions to be stored and executed to perform the applique's composition and editing functions.

2. Breadboarding to demonstrate memory addressing and read-out. Environmental testing including bounce, vibration and temperature of the read-only memory tray. An assessment of applique software.

3. EST expenditure next month (Nov) $3,151.00

DISTRIBUTION:

FORM 1406B

Fig. 6. Project status report.

Reference 6 covers this subject in more detail.

4.3. *Implement the Plan*

The manager implements the plan by having appropriate people assigned and by opening the appropriate accounts for work-order charges. The Project Engineer provides the technical leadership for the development effort in his own department.

Depending on the size of the project, the Project Engineer and Project Manager may be separate men or may be the same person. In any case, the personality wearing the Project Manager's hat has the responsibility to assure that all supporting departments (a) get their inputs at the right time, (b) start and finish their job on time, (c) satisfy—but only just satisfy—their requirements.

Since the supporting departments probably have an open work order to charge to, it is good business for the Project Manager to prevent having them 'sell him' more of their support or service than the project really needs. The converse is also true, but less likely.

Summarizing the Project Manager's role—he Plans the work. . .then he Works the plan.[4]

4.4. *Monitoring Progress*

But no plan is perfect and one cannot anticipate all the snags. Periodic project reviews and reports are essential to monitor the degree of success in following the plan. Our Project Engineers are required to

(a) publish a monthly status report, and
(b) review their Forecast Project Data, i.e. time-phased anticipated manhours on a monthly basis.

Figure 6 is an example of the monthly status report. The fiscal portion is a summary of detailed expenditure data provided to the Project Engineer by Engineering Administration about the 10th of each month. The Gantt-chart section is copied or adopted from the original project plan, by work orders. To facilitate assessing whether tasks are being completed according to the original plan, we use the symbols in Fig. 7 for updating the Gantt bars.

In the balance of the report the engineer summarizes progress for the past month, including problems encountered, attempted solutions and their effect. He also explains schedule slippage, if any, and what is being done to make it up. Lastly, he indicates what accomplishments are planned for the current month.

In conjunction with the monthly status reviews the Project Engineer also reviews and reaffirms or updates manpower loading charts—the Forecast Project Data sheets (as shown in Fig. 8). These are time-phased charts derived from the original project plan and labour estimate. Monthly Forecast Project Data sheets are used together with Forecast Proposal Data to update the manpower loading histograms. The

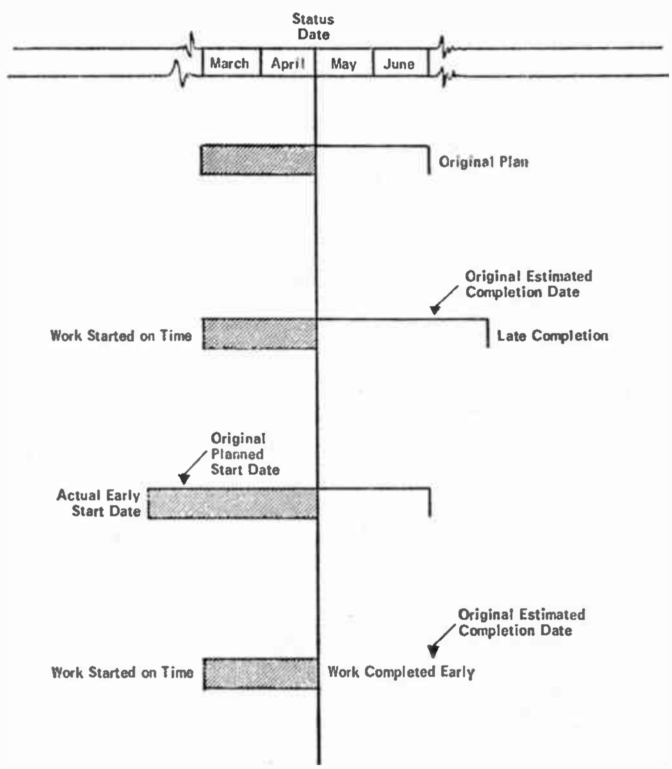


Fig. 7. Gantt chart symbols.

PROJECT ENGINEERING DIVISION

FORECAST
PROJECT DATA

FORECAST PERIOD FROM 1 OCT 1968 TO 1 FEB 1969

FORM 1201B

PROJECT – W.O. (1–4)	PROJECT MALLARD DESIGN (5–7)	RESP. CODE	COST CENTRE 9–12	OCT 13–16	NOV 17–20	DEC 21–24	JAN 25–28	FEB 29–32	33–36	37–40	41–44	45–48	49–52	53–56	57–60	61–64	65–68	69–72	UNSCHD 73–76	TOTAL HOURS	RATE			TOTAL $
P G 4 5 1 (F A N G)	W.O. 1 0 1 — TITLE: PROJECT MANAGEMENT	D	1130	3	3	3	3													12				69
			TOTAL																	12				69
			MAT'L																		1000			
			TRAVEL																		1000			
P G 4 5 1 (F A N G)	W.O. 1 0 2 — TITLE: SYSTEM INVESTIGATION	D	1130	92	1 28	52														2 72				1569
			1120	61	1 33	49														2 43				1402
			TOTAL																	5 15				2971
			MAT'L																		1000			
			TRAVEL	3 00																	1000			300
P G 4 5 1 (F A N G)	W.O. 1 0 3 — TITLE: SOFTWARE SIMULATION	D	1120	1 20																1 20				692
			TOTAL																	1 20				692
			MAT'L																		1000			
			TRAVEL																		1000			
P G 4 5 1 (F A N G)	W.O. 1 0 4 — TITLE: HARDWARE INVESTIGATION	D	1130	1 32	80															2 12				1201
			1740	2 40	80															3 20				1269
			1870	80																80				351
			TOTAL																	6 12				2821
			MAT'L																		1000			
			TRAVEL																		1000			

PREPARED BY D. MOULDING (PROJ. ENGR.) DATE 9 OCT 1968 APPROVED BY R. J. PATTON PAGE 1 OF 2

Fig. 8. Forecast project data.

FORM 1281B

FORECAST PROJECT DATA

PROJECT ENGINEERING DIVISION

FORECAST PERIOD FROM 1 OCT 1968 TO 1 FEB 1968

PROJECT – W.O. (1–4) (5–7)	RESP. CODE	COST CENTRE 8–12	OCT 13–16	NOV 17–20	DEC 21–24	JAN 25–28	FEB 29–32	33–36	37–40	41–44	45–48	49–52	53–56	57–60	61–64	65–68	69–72	UNSCHD 73–76	TOTAL HOURS	RATE			TOTAL $
P G 4 5 1 W.O. 1 0 5																							
F A H G TITLE	D	1130		40															40				231
		1120		32															32				185
		1740		30															30				118
ESTIMATING		1850		30															30				132
		1860		30															30				111
		1870		30															30				109
		TOTAL																	1,92				887
		MAT'L																		1000			
		TRAVEL																		1000			
P G 4 5 4 W.O. 1 0 1																							
F A H G TITLE	D	1130	1,20	1,04															2,24				1292
		1120	1,20	1,01															2,21				1275
		TOTAL																	4,45				2567
		MAT'L																		1000			
		TRAVEL																		1000			
P W.O.																							
F A H G TITLE													PROGRAM TOTAL				D.L		18,96				10007
																	T.L						308
		MAT'L																		1000			
		TRAVEL																		1000			
P W.O.																							
F A H G TITLE																							
		MAT'L																		1000			
		TRAVEL																		1000			

PREPARED BY D. MOULDING (PROJ. ENGR.) DATE 9 OCT 1968 APPROVED BY R.J. PATTON PAGE 2 OF 2

Fig. 8 (contd.)

cost estimating staff also apply labour rates to the manpower totals by cost centre and calculate revised total project costs.

When this monthly status review is done conscientiously and objectively, it is possible early in the project to identify problems which jeopardize achievement of the objectives. Then additional manpower, specialist consultants or a change in approach can be adopted to correct the problem.

5. Transferring Results to Production

For complete success, a key and often difficult phase of guiding an R & D project is transferring the results to production. This involves a definite transition for the product, wherein a variety of functional specialists skilled in fabrication, assembly, test and calibration techniques take over to make quantities of exact replicas of the development prototype product.

5.1. *Development Project Output*

Development projects are expected to do a comprehensive job, with the following typical output:

One or more prototypes which satisfy the functional and cost specifications.

Procedures for 'design proving' tests, alignment and calibration.

Reports on the design proving tests at room and extreme environments, demonstrating that the specifications are met. The Quality Assurance Department monitors these tests and (hopefully) certifies that the product is ready for production.

All necessary design documentation including layout drawings, assembly drawings, detail drawings, specification control drawings for purchased components, and basic alignment and test procedures.

But the R & D project does not include the industrial engineering function of methods and process planning. And development engineers never succeed in documenting all necessary information and they never succeed in anticipating all production problems. Once the product is released to the functionally-organized factory—who will undertake an 'anxious mother' role for the product? One possibility would be the Project Engineer for the original development, because he obviously has the best understanding of how the original idea and specifications were translated into this product. But in practice few companies probably use this route because:

(a) The Project Engineer's combination of inventive engineering skills and management flair are in demand for the next development project.

(b) The engineering skills most useful in successfully launching production are seldom found in development engineers.

5.2. *The Product Engineer*

Many companies handle the problem by assigning product responsibility to a Product Engineer who resides organizationally at the Engineering-Production interface. In some companies this Product Engineer is part of Industrial Engineering. In our company he belongs to the Design and Drafting Division.

Our Product Engineers, who incidentally are usually graduate engineers, are responsible for the conduct of engineering aspects of a new product, from production release through to the time when the scheduled production rate is achieved. But how does he acquire the necessary know-how?

To fulfil his responsibility, a Product Engineer acquires his knowledge of the product during the development phase—by assisting the Project Engineer in preparing the design, drafting, and testing documentation required for the 'Release to Production'. He participates in or monitors the design proving tests during the development phase. He coordinates between the design and drafting function and the manufacturing planning functions to ensure that the configuration† of the product is commensurate with the scale of intended production, as indicated on the project work statement. Our Product Engineers are key men who perform a 'support' function. During the development phase they absorb specific product information while they support the Project Engineer. During the production phase they speak for the R & D organization in providing engineering support to manufacturing.

Figure 9 shows the relationship between various activities and responsibilities during the development/production transition.

6. Presentations on Project Aims, Progress and Results

Somewhere during the history of the project the Project Manager will probably have to give a presentation on his work. An effective presentation is part of good project management because a project cannot be classified as a success unless the man who paid for the work supports him in that conclusion. Also, this sponsor (who may be inside the company, or the customer) will probably have to make a decision regarding further action based on the results of the project. So it is vital that he understands the Project Manager's results and recommendations.

There must be a comprehensive final report—and the oral presentation is not designed to replace it. The presentation is often used to present the major findings to the top decision makers, who lack the

† Definition of 'configuration': the complete technical description required to fabricate, test, accept, operate, maintain and logistically support systems/equipment.

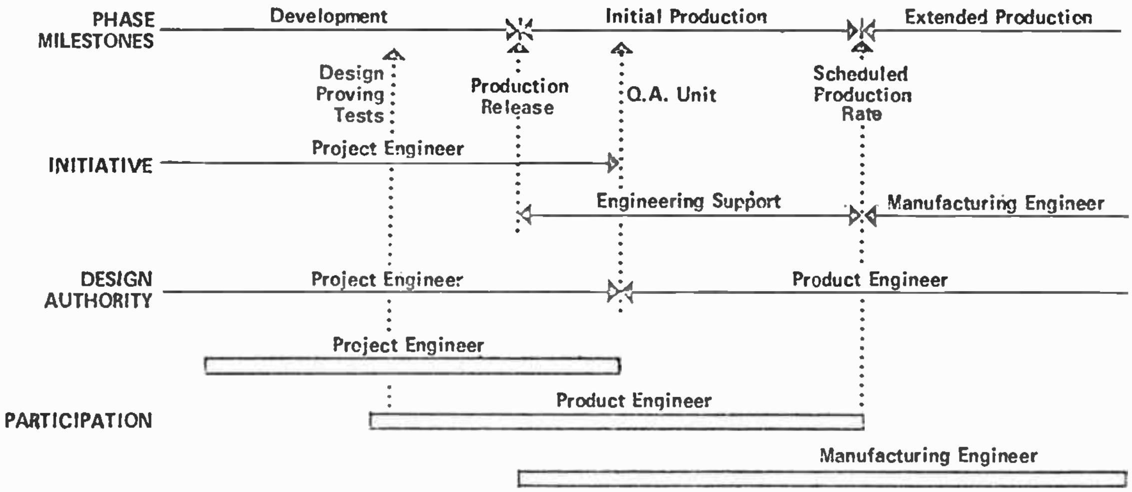


Fig. 9. Responsibilities during the development–production transition.

time to read and assimilate detailed reports and working papers. Even if they do have time, a presentation on results and recommendations makes better use of our communication senses. For example, a University of Michigan Study found that people spend their time communicating as follows:

Writing	11%
Reading	15%
Talking	31·9%
Listening	42·1%

Another study by Socony Vacuum Oil suggests that in acquiring data, we retain:

10% of what we read
20% of what we hear
30% of what we see and hear simultaneously
and 70% of what we say as we talk.

Both sides probably retain more of what we say as we talk because a briefing provides a face-to-face, give-and-take communication situation in which misunderstandings, misinterpretations and differences of opinion may be exposed and an attempt made to resolve them.

7. Conclusions

Reduced to bare essentials, the techniques for guiding technical programmes are:

Starting with general company objectives, derive R & D goals and product guidelines. Prepare product development proposals together with marketing and profitability analysis. Select projects for authorization on the basis of comparative financial performance.

Plan proposed development projects thoroughly, whether for company or customer approval. Acquire manager motivation and commitment by assigning the prospective Project Engineer to coordinate project planning. Ensure participation of supporting skills in project planning.

Use a Project Manager to coordinate project tasks within a functional organization structure. Use him as a 'General Contractor' within your organization. Feed back interim results and correct the approach.

Ensure an effective transfer to production by providing knowledgeable engineering support during early production phases. Train the product support engineer on the development project.

Use oral presentations to convey the key results and recommendations to management or customer.

8. References and Bibliography

1. Roberts, E. B., 'How the U.S. buys research', *International Science and Technology*, September 1964, p. 70.
2. Dean, Burton V., 'Evaluating, Selecting and Controlling R & D Projects', American Management Association Research Study 89, 1968.
3. Miller, Ernest C., 'Objectives and Standards—an Approach to Planning and Control', American Management Association Research Study 74, 1966.
4. King, W. J., 'The Unwritten Laws of Engineering', American Society of Mechanical Engineers, 1944.
5. 'PERT Guide for Management Use', U.S. Government Printing Office, Washington, D.C., 1963.
6. Cleland, David I., 'Understanding project management', *Manage Magazine*, 19, No. 9, 1967.

9. Appendix

Background Information on the Company that provided Experience and Examples for this Paper

The basic principles for managing technical programmes, as outlined in this paper, are fairly universal. Many of the illustrative examples are drawn from procedures within Computing Devices of Canada. In order to better understand the viewpoint, it may be helpful to give a brief background of the company.

Computing Devices of Canada and its U.S. parent, The Bendix Corporation, are high-technology companies. The various Bendix Divisions and its affiliated companies are heavily engineering oriented. At Computing Devices the business is about 70% military, and it is characterized by relatively short production runs of very complex and expensive products. Most production programmes are preceded by some development engineering. A very substantial part of our total research and engineering activity is government-funded, and is often obtained through competitive tender. Projects are usually undertaken by us on a fixed-price and fixed-schedule basis with well-defined objectives. Our leaning toward government work and toward paid engineering in no way diminishes the need for basically sound techniques for guiding technical programmes. If anything, failure to practise sound research and development project selection, planning and management can lead to financial disaster very quickly in the military fixed-price paid R & D business.

9.1. *Synopsis of Objectives*

In September 1967 the Directors and Officers of the Company laid down the following objectives:

Sales volume. Growth by 10% per year.

Profitability. By 1970, 5·5% on sales and return on investment of 11·5%, before taxes.

Market posture. Increase exports to 55% of sales by 1970.

Products. Greater diversity and at least one new product per year. Push for production of multi-sensor data processing systems.

Engineering. Conduct Independent Research and Development at 2·3% of sales. Develop commercial moving map display and solid-state position and homing indicator (PHI).

Seek paid R & D leading to production follow-on.

Orient current product engineering at meeting or beating unit manufacturing-cost goals.

Production. Improve efficiency through better planning, control, methods and plant modernization.

Capability and productivity. Acquire ability to do new things. Improve performance on current jobs.

Inter-company relations. Earn the respect and co-operation of Bendix.

Utilize our Bendix identity to facilitate our participation in the U.S. market.

9.2. *New Product Criteria*

In recent years, our own company has seen the consequence of poorly communicated product goals. Well-meaning engineers have often submitted new product ideas, sometimes after considerable personal effort—only to be disappointed because we have rejected their apparently excellent product idea. To a large extent the fault was the company's because specific guidelines for new products were not generally known. We now have such guidelines, of which the following are the salient points:

1. *General.* Must contribute to company goals for profit, stability and growth.

   Should be basically electronic or electro-mechanical.

2. *Marketing.* Should be of interest to our present customers:

   (a) Canadian Armed Forces.

   (b) U.S. Department of Defense.

   (c) North American air-frame and airline companies.

   (d) Export customers that can be reached through Bendix.

   Preference for items which are related to our existing products, including those which can obsolete our existing products.

3. *Engineering.* Technology involved should be related to our present R & D capability.

   Should take advantage of our military specification know-how.

   Development time should not exceed 20% of the expected time during which sales volume would increase.

   Development cost should be between $30,000 and $300,000.

4. *Manufacturing.* Preference for items which could be manufactured economically with present facilities.

   Cost of new equipment must be amortized and recovered during product life.

5. *Financial.* Preference for products eligible for customer or Government participation in development costs.

Manuscript received by the Institution on 18th March 1969 (Paper No. 1297/MT1).

U.D.C. 621.396.677

Gain Maximization and Controlled Null Placement Simultaneously Achieved in Aerial Array Patterns

By

C. DRANE, Jr.,
B.S.(PHYS.), M.S.(PHYS.), M.S.(MATH.)‡

and

J. McILVENNA,
B.S.(E.E.), M.S.(E.E.)‡

This paper presents a matrix method for maximizing aerial directive gain, while simultaneously placing nulls in the far-field radiation pattern of the array. The technique is applicable to aerial arrays of N elements, arbitrarily positioned, and allows the designer to specify the directions of up to (N—1) independent pattern nulls and/or side-lobes, while providing maximum gain in some prespecified direction. This control is achieved by varying only the amplitude and phase of the element currents. The method can be used even when non-isotropic aerial elements and inter-element mutual coupling effects are included in the expression for directive gain.

List of Principal Symbols

N	number of array elements
M	number of pattern constraints
Superscript *	complex conjugate
Superscript †	complex conjugate transpose of a vector or a matrix
Superscript T	transpose of a vector or matrix
d_n	location of nth element in array
μ	$\cos\theta$ where θ is measured from line of array
λ	wavelength
γ	eigenvalue
$E(\mu)$	far-field amplitude radiation pattern
G	directive gain of the array
$\bar{A}$ and $\bar{B}$	$(N\times N)$ square matrices occurring in numerator and denominator, respectively, of the gain formula
I	an N-element column vector with the element currents as entries
p	one of the M prespecified constraint vectors
$\bar{P}$	$(N\times N)$ constraint matrix constructed from p vectors
$\mathbf{I}$	an N-element transformed current vector
$\bar{A}_c$ and $\bar{B}_c$	$(N\times N)$ numerator and denominator matrices of the transformed gain formula
$\bar{A}_a$ and $\bar{B}_a$	$(N-M)\times(N-M)$ abridged matrices formed by deleting the first M rows and M columns of $\bar{A}_c$ and $\bar{B}_c$
Γ	constrained gain.

‡ Air Force Cambridge Research Laboratories, Office of Aerospace Research, L. G. Hanscom Field, Bedford, Massachusetts 01730, U.S.A.

1. Introduction

The frequent need to discriminate between closely-spaced radar targets has led to considerable interest in the development of high-resolution, high-gain radar aerials. Both resolution and gain are describable in terms of the radiation pattern of the aerial. High resolving power implies that the pattern has a narrow main beam and low side-lobes, while high directive gain implies a large amount of energy radiated by the aerial in the direction towards which the beam points, relative to the total energy radiated. This paper considers a technique for achieving maximum gain for aerial arrays whose radiation patterns must satisfy constraints on the location of null values that control the beamwidth and/or side-lobe levels of the pattern.

These two problems, gain maximization and radiation pattern beam-shaping, are usually considered separately. In the former,[1-11] generally little attention is given to the radiation pattern structure that results from the maximization procedure. For example, for linear arrays (in which the elements are located along a straight line), associated with the maximum gain solution is a pattern with a relatively high side-lobe structure that makes this solution somewhat unattractive; such high side-lobes could cause undesirable interference in target discrimination. In beam-shaping operations on the other hand,[12,13] control over the pattern structure is the prime objective, but one often obtains such control only at a significant sacrifice in gain.

It is then the purpose of this paper to consider jointly gain maximization and beam shaping. Applications calling for such techniques include the following problems. Firstly, there are often in radio communications links the simultaneous design goals of maximum gain in the direction of some distant transmitter and/or receiver and the reduction or elimination of interference or jamming from other directions. Secondly, in radio astronomy it is often desirable to use an aerial that can reduce the signal from a strong source and at the same time produce maximum gain in the direction of a nearby weaker source of interest. Finally, in a real life tactical environment, aerial siting is seldom optimum, and reflexions from natural or man-made objects can interfere with operations (as, for example, the interference experienced by an aerial from other aerials on board ship). In all these applications one would like to minimize the deleterious effects perhaps by placing pattern nulls in the directions of the interference while at the same time radiating and/or receiving, as the case may be, maximum energy in the direction of the desired target or signal.

The technique proposed here is applicable to an array of N aerial elements that are arbitrarily located in three-dimensional space. In addition, it will be shown that the designer can specify arbitrary directions for up to a total of $(N-1)$ pattern nulls and/or side-lobe positions, while guaranteeing maximum gain in some prespecified direction. Only the element excitation coefficients need be varied (in amplitude and phase), making the method attractive for adaptive operations in which the pattern must be rapidly reconfigured, often without the possibility of a change in element locations, in order to adjust continuously the null locations to coincide with new directions of the interfering signals or targets.

In this method, the mathematical technique consists of a rather general matrix procedure for finding the extremal values of the ratio of two quadratic forms that are subject to constraints. We shall next consider the basic elements of this procedure.

2. Array Matrix Theory

2.1 General Theory

Although the discussion below is to be limited to aerial directive gain, the techniques outlined are generally applicable to a much wider class of problems.[14] Other quality criteria for aerial arrays, such as power gain, signal/noise ratio, gain-bandwidth product and aerial Q-factor can be similarly handled. Even in circuit problems in electrical engineering, for example the maximization of the ratio of the power dissipated in an N-port load to that dissipated in the internal impedances of the N feeding generators, these techniques can be applied.

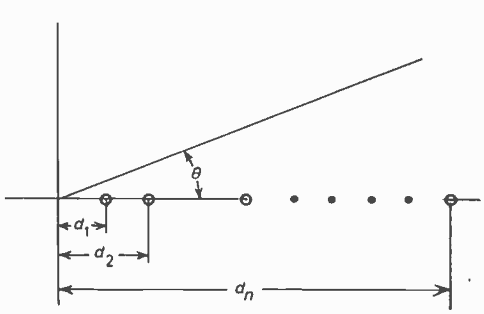


Fig. 1. Linear array geometry.

We summarize here the results of a recent matrix approach to aerial array optimization.[3–9,14,15] It serves as the basis for the constraint technique outlined in a later section.

The directive gain, often used as an indicator of overall array performance, is defined as

$$G = 4\pi \frac{\text{power radiated in a particular direction}}{\text{total power radiated}} \tag{1}$$

For a linear array, such as that shown in Fig. 1 with N isotropic elements arbitrarily spaced along a line, the far-field amplitude radiation pattern is (to within a constant factor) given by

$$E(\mu) = \sum_{n=1}^{N} I_n \, e^{-jD_n\mu},$$

where $D_n = (2\pi/\lambda)\, d_n$, $\mu = \cos\theta$ and the I_n are the complex element excitation coefficients, amplitude and phase. The power radiated in a particular direction, μ_0, is

$$P(\mu_0) = |E(\mu_0)|^2 = \sum_{n=1}^{N} \sum_{m=1}^{N} I_n I_m^* \, e^{j(D_m - D_n)\mu_0},$$

which the reader will recognize as a quadratic form; the symbol * denotes complex conjugate. Arranging the currents in an N-element column vector I,

$$I = \begin{bmatrix} I_1 \\ I_2 \\ \cdot \\ \cdot \\ \cdot \\ I_n \end{bmatrix}$$

and defining a square $(N \times N)$ matrix $\bar{A}$, with elements

$$A_{mn} = e^{\,j(D_m - D_n)\mu_0}, \qquad 1 \le m, n \le N,$$

leads to the more familiar matrix formulation

$$P(\mu_0) = I^\dagger \bar{A} I,$$

with † denoting the combined operations of complex conjugation and transposition. Note that $\bar{A}$ is a Hermitian matrix, i.e. $A_{mn} = A_{nm}^*$ and that, in addition, it is a one-term dyad, i.e. it can be expressed as the outer product of two column vectors

$$\bar{A} = aa^\dagger,$$

where

$$a = \begin{bmatrix} e^{jD_1\mu_0} \\ e^{jD_2\mu_0} \\ \cdot \\ \cdot \\ \cdot \\ e^{jD_n\mu_0} \end{bmatrix}.$$

(This property of $\bar{A}$ will lead to a significant simplification in finding the maximum directive gain.)

The total power radiated by the aerial is that contained in the power pattern and is given by

$$P_{\mathrm{T}} = 2\pi \int_{-1}^{+1} |E(\mu)|^2 \, d\mu = \sum_{n=1}^{N} \sum_{m=1}^{N} I_n I_m^* \int_{-1}^{+1} e^{j(D_m - D_n)\mu} \, d\mu,$$

or

$$P_{\mathrm{T}} = 4\pi \sum_{n=1}^{N} \sum_{m=1}^{N} I_n I_m^* \frac{\sin (D_m - D_n)}{(D_m - D_n)}.$$

We define the $(N \times N)$ matrix $\bar{B}$ with elements

$$B_{mn} = \frac{\sin (D_m - D_n)}{(D_m - D_n)}.$$

(Note that $\bar{B}$ is real, symmetric, and positive definite.) Then,

$$P_{\mathrm{T}} = 4\pi (I^{\dagger} \bar{B} I),$$

equation (1) for aerial directive gain becomes a ratio of quadratic forms

$$G = \frac{I^{\dagger} \bar{A} I}{I^{\dagger} \bar{B} I}, \qquad \text{......(2)}$$

and the usual goal is to find the currents which maximize G.

There are other equivalent forms for directive gain but the representation in eqn. (2) has at least two distinct advantages. First, necessary calculations are readily performed on a digital computer using standard library routines for the matrix operations. Second, and most important, is that the extremal values, of quality criteria cast in this form, can be determined quite simply. This result is based on a theorem[4,5,16,17] which states that with $\bar{A}$ and $\bar{B}$ Hermitian and $\bar{B}$ positive definite, all the relative maxima and minima of the ratio are given by the eigenvalues determined from

$$\bar{A} I = \gamma \bar{B} I,$$

where γ denotes the eigenvalues which must satisfy

$$\det(\bar{A} - \gamma \bar{B}) = 0.$$

The vectors $\{I\}$ which produce these extrema, are the corresponding eigenvectors.

The case of directive gain (and other quality criteria) is especially simple. Since $\bar{A}$ is a one-term dyad, all but one of the eigenvalues of the gain ratio are zero. This one non-zero eigenvalue is, in fact, the maximum gain and is given by[4,5,9]

$$G_{\max} = \gamma_1 = a^{\dagger} \bar{B}^{-1} a.$$

The corresponding eigenvector, representing the currents which produce this maximum gain, is

$$I = \bar{B}^{-1} a$$

These two results provide an elegant and compact solution for the maximization of directive gain and other quality criteria. This formulation has been extensively used to study maximum gain arrays with, however, no constraints on the aerial radiation pattern behaviour. The following section will show how the notion of pattern constraints can be introduced in a way which preserves the compactness and simplicity of the matrix approach.

2.2. Pattern Constraints

The general problem of the maximization or minimization of a quantity that is subject to constraints is quite often approached by the method of Lagrange multipliers. This traditional approach has in fact been used in several studies of aerial optimization including the maximization of gain subject to constraints on the supergain ratio and the beam efficiency[10,11,18,19] and, more recently, the maximization of gain with constraints on the pattern behaviour.[20,21] The Lagrange technique is characterized by the fact that as the number of applied constraints increases, so also does the complexity of the manipulations required for a solution. The constraint technique outlined below, however, actually reduces the matrix dimensions by exactly the number of constraints applied and can provide, therefore, an important simplification as far as computer operations are concerned.

The constraints take the form of setting either the pattern value and/or the pattern derivative value to zero in certain directions. The former locates nulls, the latter side-lobes in the resulting constrained pattern. Other authors have developed techniques for placing nulls in the radiation pattern[12,22,23], although some of these methods are applicable only to uniformly spaced arrays and none of them is coupled to a simultaneous maximization of the gain in the constrained pattern. Let us now outline a maximum gain-constraint technique.

For the situation in Fig. 1, the array factor, which for isotropic elements is the radiation pattern, is given by

$$E(\mu) = \sum_{n=1}^{N} I_n \, e^{-jD_n \mu}.$$

Setting the pattern equal to zero in the directions μ_i where $i = 1, 2, \ldots M$, $0 \leq M \leq (N-1)$, produces a set of M linear, homogeneous equations in the currents I_n:

$$\left\{ \sum_{n=1}^{N} I_n \, e^{-jD_n \mu_i} = 0 \right\}, \qquad i = 1, 2, \ldots M.$$

The left-hand sides of each of these constraint equations can be expressed as the inner product of a constraint vector p_i

$$p_i^{\mathrm{T}} = \{e^{-jD_1 \mu_i}, \; e^{-jD_2 \mu_i}, \ldots e^{-jD_N \mu_i}\}, \quad i = 1, 2, \ldots M$$

and the current vector I. Or, setting the pattern derivative equal to zero in certain directions gives rise to the equations

$$\left\{\sum_{n=1}^{N}(-jD_n)I_n\,e^{-jD_n\mu_i}=0\right\},\quad i=1,2,\ldots M$$

and thus locates side-lobe positions. The overall restrictions on the constraint technique are that the constraint equations must be linear, independent, and above all, homogeneous. The last restriction will be shown to be important so far as retention of the form common to the unconstrained problem is concerned.

Following a procedure outlined by Guillemin,[24] we first form a constraint matrix $\bar{C}$, with N rows and N columns. Each of the first M rows in $\bar{C}$ consists of one of the constraint vectors p_i; the remaining $(N-M)$ rows are filled with any arbitrary collection of N element, independent vectors. A new normalized constraint matrix $\bar{P}$, orthogonal by rows and columns, i.e. $\bar{P}^\dagger=\bar{P}^{-1}$, is generated by using the Gram-Schmidt or equivalent procedure on the matrix $\bar{C}$.

One can now define a transformation

$$I=\bar{P}^\dagger\mathbf{I}\quad\text{or}\quad\mathbf{I}=\bar{P}I,\qquad\ldots\ldots(3)$$

which when applied to eqn. (2) gives:

$$G=\frac{\mathbf{I}^\dagger(\bar{P}\bar{A}\bar{P}^\dagger)\mathbf{I}}{\mathbf{I}^\dagger(\bar{P}\bar{B}\bar{P}^\dagger)\mathbf{I}}=\frac{\mathbf{I}^\dagger\bar{A}_c\mathbf{I}}{\mathbf{I}^\dagger\bar{B}_c\mathbf{I}},\qquad\ldots\ldots(4)$$

where $\mathbf{I}$ is an N element column vector and $\bar{A}_c$ and $\bar{B}_c$ are, of course, square $(N\times N)$ matrices. We will presently show that this form permits several significant simplifications.

The orthogonalization procedure is of such a nature that each of the first M rows in $\bar{P}$ turns out to be a linear combination of the original constraint vectors $\{p_i\}$. The definition of $\mathbf{I}$ in eqn. (3) shows that each of the first M entries in $\mathbf{I}$ is therefore some linear combination of the original constraint equations, all of which however were homogeneous. It follows that the first M entries in $\mathbf{I}$ are zeros. This means that in eqn. (4) one can discard the first M entries in $\mathbf{I}$ (we call this abridged vector $\mathbf{I}_a$) and therefore the first M rows and M columns of $\bar{A}_c$ and $\bar{B}_c$. The resulting abridged form for gain, incorporating all the constraint effects, is

$$\Gamma=\frac{\mathbf{I}_a^\dagger\bar{A}_a\mathbf{I}_a}{\mathbf{I}_a^\dagger\bar{B}_a\mathbf{I}_a}\qquad\ldots\ldots(5)$$

where now $\bar{A}_a$ and $\bar{B}_a$ are square matrices, reduced in dimensions to $(N-M)$ by $(N-M)$. This reduction in dimensions is consistent with intuitive notions that constraints always reduce the number of degrees of freedom of the variable involved in the system. It is important to note that the formula for gain subject to these homogeneous constraints remains the ratio of two quadratic forms.

It can be shown that a typical element in $\bar{A}_c$ is

$$A_c^{mn}=\sum_{k=1}^{N}\sum_{j=1}^{N}P_{mk}A_{kj}P_{nj}^*,\quad 1\le m,n\le N.\qquad\ldots\ldots(6)$$

Note that

$$A_c^{mn}=(A_c^{nm})^*$$

(where * denotes complex conjugate); hence $\bar{A}_c$ is Hermitian. $\bar{B}_c$ has the same property. Since $\bar{A}_a$ and $\bar{B}_a$ are formed by deleting the first M rows and M columns of $\bar{A}_c$ and $\bar{B}_c$, $\bar{A}_a$ and $\bar{B}_a$ will likewise be Hermitian.

We have seen that the quadratic form appearing in the denominator of eqn. (2), representing as it does the total power in the radiation pattern, must be positive definite. So also must be the quadratic form associated with the matrix $\bar{B}_c$, obtained from the other quadratic form merely by the transformation given in eqn. (3). The matrices $\bar{B}$, $\bar{B}_c$, and therefore also the abridged matrix $\bar{B}_a$ must be positive definite. Finally note that

$$\bar{A}_c=\bar{P}\bar{A}\bar{P}^\dagger=\bar{P}(aa^\dagger)\bar{P}^\dagger=(\bar{P}a)(\bar{P}a)^\dagger=a_c a_c^\dagger,$$

so that like $\bar{A}$, $\bar{A}_c$ is a one-term dyad. A little thought shows that deletion of M rows and M columns of $\bar{A}_c$ to form $\bar{A}_a$ is equivalent to deleting the first M entries in the vector a_c (we call this abridged vector a_a). Thus

$$A_a=a_a a_a^\dagger$$

where it can be shown that

$$a_a^{\mathrm{T}}=\left\{\sum_{k=1}^{N}P_{M+1,k}\,e^{j\mu_0 D_k},\ \sum_{k=1}^{N}P_{M+2,k}\,e^{j\mu_0 D_k}\ldots,\ \sum_{k=1}^{N}P_{N,k}\,e^{j\mu_0 D_k}\right\}.$$

Thus the transformation from eqn. (2) to eqn. (5) has introduced the pattern constraints in a manner which leaves intact the form of the expressions for gain, and reduces the dimensions of the matrices involved. In addition, all the simple matrix operations used to compute the maximum gain and the corresponding current distribution remain valid here. For there is again but one non-zero eigenvalue; it is now the maximum‡ gain of the array subject to the pattern constraints and, referring to the result quoted in the preceding section, is given by:

$$\Gamma^{\max}=\gamma_1^c=a_a^\dagger\bar{B}_a^{-1}a_a.\qquad\ldots\ldots(7)$$

The corresponding eigenvector is found from

$$\mathbf{I}_a^{\max}=\bar{B}_a^{-1}a_a,$$

and the current distribution, $I_c^{\max}$, associated with this maximum constrained gain is obtained by eqn. (3).

‡ Throughout this discussion, the assumption has been made that element spacings, though arbitrary, are fixed. There are various schemes for perturbing[15,25] the element spacings to produce higher gains.

But, if one recalls that the N-element column vector **I** is always formed from the $(N-M)$-element column vector $\mathbf{I}_a$ by inserting zeros for the first M elements of **I** and following these with the $(N-M)$ elements of $\mathbf{I}_a$, eqn. (3) in this case simplifies to

$$I_c^{max} = \bar{P}_d^{\dagger}\,\mathbf{I}_a^{max},$$

where the $(N-M)\times N$ matrix $\bar{P}_d$ is obtained from the $(N\times N)$ matrix $\bar{P}$ by deletion of the first M rows.

When the eigenvalues (positive semidefinite) are ordered according to their numerical size, it can be shown that in any generally constrained system the eigenvalues have an 'interleaved' relationship[24] with the unconstrained eigenvalues. That is, if a general form such as eqn. (1) is subjected to, say, a single constraint, $(M = 1)$, the set of unconstrained eigenvalues $\{\gamma_i\}$, $i = 1, 2, \ldots N$, and the set of constrained eigenvalues $\{\gamma_i^c\}$, $i = 1, 2, \ldots (N-1)$, are related as

$$\gamma_1 \geq \gamma_1^c \geq \gamma_2 \geq \gamma_2^c \geq \ldots \geq \gamma_{N-1}^c \geq \gamma_N.$$

In the case of gain, only γ_1 and γ_1^c are non-zero, implying the important result that

$$\Gamma^{max} \leq G^{max}.$$

It is of interest to the array designer to minimize the amount of gain he must sacrifice to achieve some desired degree of pattern control. Some examples of this technique will be discussed in Section 3.

2.3. *Mutual Coupling Effects*

It is important to state here in passing that mutual coupling and element factor effects can be included in the gain formulation in a variety of equivalent ways. For the purposes of this paper, we need only note that with mutual effects and even element losses accounted for, gain can still be represented as the ratio of two quadratic forms, possessing all the necessary Hermitian, positive definite and dyadic properties described up to now.[14,22,26-28] The elements in the numerator and denominator matrices are otherwise different, but in a non-essential way.

It is well known that maximum gain is associated with a uniform current distribution across the elements.[2] In this sense, the maximum gain constraint

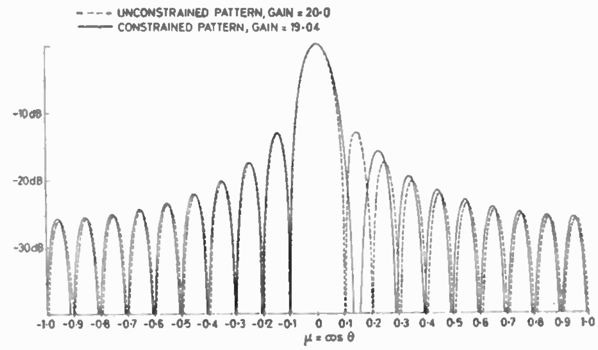


Fig. 2. Use of a single pattern constraint to replace a peak with a null.

technique represents the smallest possible deviation from uniform current illumination that will place the pattern nulls as desired. This trend toward uniform illumination means that the excitation currents cannot oscillate wildly (a situation encountered in superdirective aerials) and insures that the maximum gain current distributions can be realized in practice. Keeping the illumination as uniform as possible is also an aid in calculating and accounting for mutual coupling effects.[12]

3. Results of the Constraint Method

To provide a valid basis for comparison of constrained and unconstrained maximum gain patterns, we shall use the maximum gain (uniformly illuminated) broadside pattern of a twenty-element array, spaced uniformly at $0{\cdot}5\lambda$, with gain of 20·0, as a reference pattern.[2] The selection of this spacing, number of elements and broadside operation is solely for demonstration and does not imply any restrictions on the constraint technique itself. All patterns presented are normalized by their largest value.

As a first simple demonstration one might consider eliminating a secondary peak of the unconstrained pattern by setting a null in that direction while requiring maximum broadside gain. The result is shown in Fig. 2. Note that the side-lobe structure is

Table 1

Number of elements	Element spacing	Maximum gain unconstrained	Maximum gain with first side-lobe peak set to a null	Maximum gain with two peaks set to nulls	Gain for a Chebyshev with −30 dB side-lobes
8	$0{\cdot}90\lambda$	13·35	12·62	11·92	10·91
8	$0{\cdot}50\lambda$	8·0	7·57	7·20	6·76
20	$0{\cdot}955\lambda$	35·7	34·12	32·5	31·09
20	$0{\cdot}50\lambda$	20·0	19·04	18·15	17·38

nearly unchanged over most of the pattern, and the gain has decreased only slightly to 19·04.

If additional peaks were to be set to nulls, the gain would drop correspondingly. Some quantitative results, demonstrating typical gain losses, are shown in Table 1 for broadside arrays of eight and twenty elements having various spacings. The Chebyshev gains[29,30] are included to show gain losses that one might encounter by simply depressing all side-lobes to say, -30 dB. The purpose of this comparison is to ascertain whether or not, as an alternative to the technique proposed in this paper, it would be preferable to choose the element excitation coefficients that result in all side-lobes being below some value at which their interference in the operation of the main beam is assuredly minimal. We see that such an alternative leads to a very definite reduction in gain relative to the optimum solution. It is also important to realize that the dynamic range of the required Chebyshev excitation coefficients (or the coefficients of any other design whose side-lobes are as low) can be excessive and perhaps even unattainable in practice, whereas the maximum gain excitations on the other hand are more nearly uniform. The Chebyshev solution represents complete pattern control over the whole side-lobe region, and as such it is an over-specification of the problem that can be costly in terms of gain. The optimum solution, on the other hand, restricts the pattern only where necessary (those directions corresponding to unwanted interference or reflections), so that in other regions, where it is constraint-free, the pattern will assume a reasonable shape consistent with the attainment of maximum gain.

Also, the proper placement of two adjacent nulls can be used to control the side-lobe level in some angular sector. (This objective becomes especially attractive for those applications in which, say, the source of interference is of some angular extent, or fairly broadband operation is envisioned.) Figure 3 shows the results of moving closer together the first two nulls to the right of the broadside direction of the unconstrained pattern. This side-lobe suppression (to -31 dB from $-13·5$ dB unconstrained) is achieved with only a slight decrease in gain, i.e. 20·0 to 19·2.

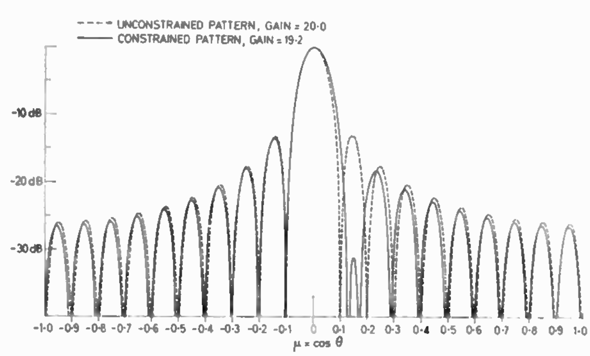


Fig. 3. Use of two pattern constraints to reduce the radiation level in an angular sector.

Note that in Figs. 2 and 3, there have been slight increases in the null-to-null beamwidth. If operating conditions are such that one cannot tolerate even this small amount of beam broadening, he can preserve the desired beamwidth through the constraint process simply by using the known first nulls as additional constraints. Of course, there is some price to pay in gain! Figure 4 shows a comparison of patterns obtained by depressing the side-lobe level to about -24 dB, with the beamwidth constrained in one case and unconstrained in another. The technique used here was to move the second null closer to the origin while holding the beamwidth between first nulls fixed. (Other variations are of course possible!) The constrained beamwidth pattern has a gain of 18·4 compared to a gain of 19·4 when the beamwidth is unconstrained. One could, of course, completely eliminate the peak while maintaining the beamwidth. The gain for this situation is 18·06 compared to 19·04 with the beamwidth unconstrained.

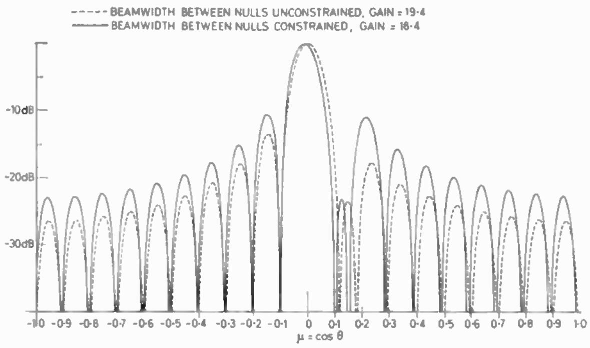


Fig. 4. Use of constraints to control the radiation in an angular sector either with or without the beamwidth fixed.

The variations of this approach are endless and these few results demonstrate some of the gain maximizing corrective techniques that can be applied to a given pattern. Of note is the fact that the starting or unconstrained pattern remains relatively unaffected over most of its range and that gain losses, at least with a few constraints applied, are not very severe. Although the use of constraints will, in general, decrease the directive gain, one should not conclude that gain losses always increase with the number of constraints applied. Depending on the particular case, two constraints may well reduce the gain by an amount less than one constraint does. For example, a single constraint used as in Fig. 2 decreased the gain to 19·04 while the two constraints used in Fig. 3 caused the gain to drop only to 19·2.

One feels intuitively that as the number of array elements increases, the negative effect of a few constraints on gain should become less pronounced. That is, the relocation of a few of these nulls should affect to a lesser extent patterns with a relatively large number of peaks and nulls than patterns with but a few nulls. This feeling is supported by the results shown in Table 1, although the differences between eight and twenty elements are expectedly slight.

As we have seen, maximum gain in uniformly spaced broadside arrays occurs with interelement spacings that are an appreciable fraction of a wavelength.[2] This results in radiation patterns with high end-values at $\mu = \pm 1{\cdot}0$, higher in fact than any interior sidelobes. Even though such patterns provide maximum gain, such pattern behaviour is usually undesirable. Some other spacing could perhaps be used, and if for example the usual $\lambda/2$ spacing is selected, the corresponding gain loss (13·35 at $0{\cdot}9\lambda$ spacing for an eight-element array down to 8·0 at $0{\cdot}5\lambda$) can be appreciable. A use for the constraint method then could be to depress the high endfire radiation, to some acceptable level while obtaining a broadside gain higher than that at $0{\cdot}5\lambda$ spacing. Figure 5 shows the effects of moving the last nulls of the unconstrained, eight-element pattern (elements spaced at $0{\cdot}9\lambda$), closer to $\mu = \pm 1{\cdot}0$. Note that the end value has been reduced to a level which is about equal to the highest interior side-lobe and the corresponding gain is still 13·07. Further movement of the nulls in the direction of endfire will serve to further reduce the end values but the inner side-lobes will increase somewhat and the gain will decrease. In passing, one should perhaps note that for spacings of $\lambda/2$ or λ the two directions $\mu = \pm 1{\cdot}0$ are the only directions in which one cannot exactly place a null. For these spacings the constraint vectors $\{p_i\}$ become a non-independent set and the null-placing technique fails. For other spacings, placing nulls precisely at $\mu = \pm 1{\cdot}0$ involves one in problems of maintaining computational accuracy.

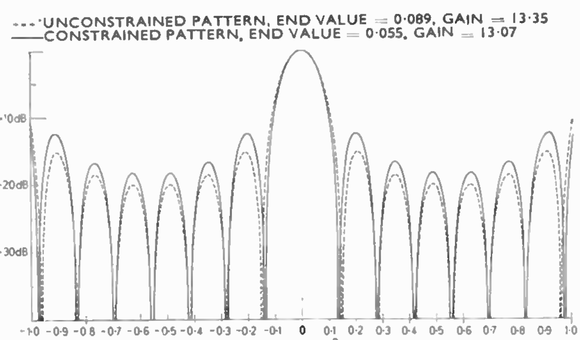


Fig. 5. Use of constraints to modify the endfire radiation of the maximum gain, eight-element array with $0{\cdot}9\lambda$ interelement spacing.

In all of this discussion, only one or two nulls have been relocated to improve pattern characteristics. It is just as simple to use several nulls simultaneously: for some applications this will result in considerably improved patterns.

These cases demonstrate the advantages of a technique which couples null placement with maximum gain, especially for situations in which aerial element relocation is difficult or impossible to arrange.

4. Some Additional Uses for the Constraint Method

The constraint method has uses other than the corrective type alluded to above. With it, one can investigate the basic properties of maximum gain patterns. For example, one customarily links maximum gain to narrow beam widths with the feeling that the narrower the beam, the higher the gain. Actually, the energy squeezed from the main beam redistributes itself in the side-lobe structure. Recalling that directive gain involves an integral over all of the radiation pattern, it is not obvious that the gain will increase as the beamwidth decreases. To investigate this point further, the interelement spacing was held fixed and the constraint method was used to move the first two nulls of the maximum gain patterns closer together, thus narrowing the beamwidth. But the gain, in fact, decreases as the beam gets narrower implying that to improve the beamwidth of any maximum gain pattern, one must sacrifice gain. (As the nulls are moved further into the main beam itself, the cost in gain is, expectedly, very high. For example, in an eight-element array, with a beam one-third of the unconstrained beamwidth, gain drops from 13·35 to 0·32 and the pattern is a many-lobed structure. Such results indicate, however, that with this technique, pattern nulls can be placed in almost any direction.) From the constraint method point of view, the decrease in gain is predictable, since, as seen earlier, any constraints on the unconstrained, maximum gain pattern will reduce the gain. Ideally then, maximum gain will be obtained if one chooses that number of elements, their spacings and a uniform excitation such that he obtains a pattern with its nulls occurring naturally in the directions of the interfering signals. As pointed out earlier however, conditions may not allow the complete freedom necessary to accomplish this goal and the constraint technique becomes a very attractive alternative.

Some calculations for other than broadside situations were also performed. For an eight-element array, spaced uniformly at $\lambda/2$ and with no constraints placed on the pattern, the maximum off-broadside gain pattern is simply the maximum broadside gain pattern, shifted to point in the desired direction by means of a linear phase taper modification of the

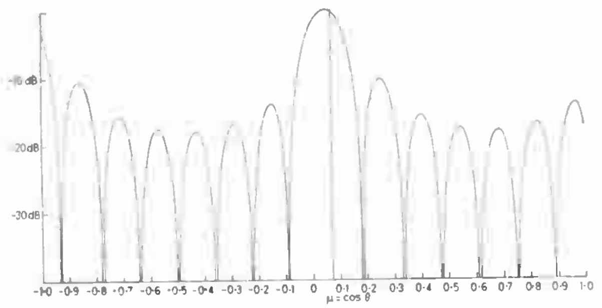


Fig. 6. Beam direction not always coincident with direction chosen for maximization of gain.

element currents. That is, the gain does not change with scanning. If $0{\cdot}9\lambda$ spacing is used, however, the gain drops rapidly with scan angle, because the large endfire radiation associated with this spacing begins to intrude into the visible region. Neither of these results is unexpected! An interesting case however occurs if one plots the unconstrained pattern for maximum gain at 86° as shown in Fig. 6. Contrary to intuition, the beam is not pointing exactly at 86°! (This same phenomenon has been noted recently in four-element circular arrays.[26]) The explanation here is simply that at this spacing, pointing the beam exactly in the maximum gain direction brings the large endfire peak too far into the visible range, and the overall effect is a decrease in gain. Gain is of course the ratio of power in a given direction to the total power radiated. One thinks of maximizing this ratio by making the numerator of eqn. (2) as large as possible, i.e. by pointing the beam in the desired direction. The phenomenon above demonstrates that the denominator factor can in fact dominate the numerator and decrease the gain even though the numerator is as large as possible. This denominator behaviour will only occur for scanned patterns with significant endfire radiation. It is a point of interest, rather than of practical significance since such wide element spacings are seldom used for conventional electronically scanned arrays.

5. Conclusions

A general matrix method for maximizing the ratio of two quadratic forms that are subject to constraints has been outlined and applied to the specific case of aerial directive gain. (The very same method can be used, without modification, in the optimization of other quantities such as signal/noise ratio or aerial power gain.) The constraint technique, applicable even when mutual coupling effects are accounted for, is an attractive way of making corrective alterations to the radiation pattern structure in a manner which, at the same time, assures maximum gain. Alternatively, the method can be used to synthesize the maximum gain pattern corresponding to a given placement of desired pattern nulls and/or side-lobe locations. It can be shown that the excitation currents are unique whether the aerial pattern for maximum gain is constrained or not.[31]

Computed radiation patterns were used to demonstrate that the constraint method can locally modify a pattern, usually with only minor losses in gain. The technique is applicable to any number of arbitrarily located aerial elements, but large arrays will require large-scale matrix operations in the computer. This can be a practical limitation on the method. But, all matrices have a large degree of symmetry, a characteristic that can be used advantageously to reduce the computational effort involved for large arrays.

6. References

1. Uzkov, A. I., 'An approach to the problem of optimum directive antenna design', *C.R. (Doklady) Acad. Sci. U.R.S.S.*, **53**, No. 6, pp. 35–38, 1946.
2. Tai, C. T., 'The optimum directivity of uniformly spaced broadside arrays of dipoles', *I.E.E.E. Trans. on Antennas and Propagation*, **AP-12**, pp. 447–454, July 1964.
3. Krupitskii, E. I., 'On the maximum directivity of antennas consisting of discrete radiators', *Soviet Physics (Doklady)*, 7, pp. 257–259, September 1962.
4. Harrington, R. F., 'Antenna excitation for maximum gain', *I.E.E.E. Trans.*, **AP-13**, pp. 896–903, Nov. 1965.
5. Cheng, D. K. and Tseng, F. I., 'Gain optimization for arbitrary antenna arrays', *I.E.E.E. Trans.*, **AP-13**, pp. 973–974, November 1965.
6. Villeneuve, A. T., 'Optimization of gain of arbitrary array antennas', AD 628–943†, December 1965.
7. Shih, S., 'Optimization and synthesis of antenna arrays', AD 638–034,† July 1966.
8. Tseng, F. I. and Cheng, D. K., 'Gain optimization for arbitrary antenna arrays subject to random fluctuations', *I.E.E.E. Trans.*, **AP-15**, pp. 356–366, May 1967.
9. Cheng, D. K. and Tseng, F. I., 'Maximization of directive gain for circular and elliptical arrays', *Proc. Instn Elect. Engrs*, **114**, pp. 589–594, May 1967.
10. Lo, Y. T., Lee, S. W. and Lee, Q. H., 'Optimization of directivity and signal-to-noise ratio of an arbitrary antenna array', *Proc. I.E.E.E.*, **54**, pp. 1033–1045, August 1966.
11. Butler, J. K. and Unz, H., 'Beam efficiency and gain optimization of antenna arrays with non-uniform spacings', *Radio Science*, **2** (New Series), pp. 711–720, July 1967.
12. Strait, B. J., 'Antenna arrays with partially tapered amplitudes', *I.E.E.E. Trans.*, **AP-15**, pp. 611–617, September 1967.
13. Hyneman, R. F., 'A technique for the synthesis of line-source antenna patterns having specified sidelobe behavior', *I.E.E.E. Trans.*, **AP-16**, pp. 430–435, July 1968.

† Reports identified by an AD number are available from the Defense Documentation Center (DDC), Cameron Station, Alexandria, Va. 22314. To the best of the authors' knowledge, this material has not been published in a form other than as a scientific report on a U.S. Air Force contract.

14. Harrington, R. F., 'Field Computation by Moment Methods', Chapter 10, (Collier-Macmillan Limited, London, 1968.)

15. Drane, C., Jr., 'Some optimum antenna arrays', AD 648–633†, October 1966.

16. Hildebrand, F. B., 'Methods of Applied Mathematics', pp. 1–119. (Prentice-Hall, New Jersey, 1952.)

17. Gantmacher, F. R., 'The Theory of Matrices', Vol. I. (Chelsea Publishing Co., New York, 1959.)

18. Butler, J. K. and Unz, H., 'Optimization of beam efficiency and synthesis of non-uniformly spaced arrays', *Proc. I.E.E.E.*, **54**, pp. 2007–2008, December 1966.

19. Uzsoky, M. and Solymar, L., 'Theory of superdirective linear arrays', *Acta Physica* (*Hung*.), **5**, No. 2, pp. 185–204, 1956.

20. Zheleznyak, M. M. and Kashin, V. A., 'Optimization of the directive gain of an antenna array with an auxiliary condition', *Radio Engng Electronic Physics*, **13**, No. 11, pp. 1816–1818, November 1968.

21. Yevstropov, B. A., *et al*., 'Synthesis of linear grids with maximum directive gain', *Radio Engng Electronic Physics*, **13**, No. 4, pp. 513–521, April 1968.

22. Strait, B. J. and Hirasawa, K., 'On radiation and scattering from arrays of wire antennas', Proc. Nat. Electronics Conf., December 1969.

23. Davies, D. E. N., 'Independent angular steering of each zero of the directional pattern for a linear array', *I.E.E.E. Trans*., **AP-15**, pp. 296–298, March 1967.

24. Guillemin, E. A., 'The Mathematics of Circuit Analysis', pp. 132–182. (Technology Press, Cambridge, Mass., 1949.)

25. Tseng, F. I. and Cheng, D. K., 'Spacing perturbation techniques for array optimization', *Radio Science*, **3** (New Series), pp. 451–457, May 1968.

26. Cummins, J. A., 'Analysis of a circular array of antennas by matrix methods', Ph.D. Thesis, Syracuse University, New York, December 1968.

27. Albin, D. H., Nakatani, D. T. and Villeneuve, A. T., 'On the use of experimental and computational techniques in analyzing the performance of scanned arrays', AD 834–954†, March 1968.

28. Gately, A. C., Stock, D. J. R. and Ru-Shoo Cheo, B. 'A network description for antenna problems', *Proc. I.E.E.E.*, **56**, pp. 1181–1193, July 1968.

29. Brown, L. and Scharp, G., 'Tchebyscheff antenna distribution, beamwidth, and gain tables', NAVORD Report 4629 (NOLC Report 383), Naval Ordnance Laboratory, Corona, Calif., February 1958.

30. Drane, C., 'Useful approximations for the directivity and beamwidth of large, scanning Dolph–Chebyshev arrays', *Proc. I.E.E.E.*, **56**, No. 11, pp. 1779–1787, November 1968.

31. Moyatskiy, V. I., 'Maximum power gain of antennas consisting of discrete radiators', *Telecomm. Radio Engng*, **22**, No. 7, pp. 79–82, July 1967.

Manuscript first received by the Institution on 15th May 1969 and in final form on 25th September 1969 (Paper No. 1298/AMMS26).

Recent Contributors to the *Journal*

Aleksander Woroncow (M.1967) obtained the degree of Dipl. Ing. in electrical engineering from the Warsaw Polytechnic. He then joined the Research Department of the Polish Broadcasting Corporation, working on acoustics, studies of transmitter coverage, and reception techniques. During the war he served with a Polish Air Force unit of the R.A.F., attached to the Admiralty Signal Establishment (now A.S.W.E.), where he worked on the early development of radar display systems. After the war he joined the Royal Naval Scientific Service, serving at A.S.W.E. first on the development of pulse and display circuits for radar, and latterly on receiver circuits for radar and wideband low noise amplifier circuits for communications antennas. Mr. Woroncow is the holder of a number of patents in display and receiver circuits—the most recent being for a new type of true i.f. logarithmic amplifier circuit. His joint paper with Dr. Croney on this circuit was published in the *Journal* and was awarded the Clerk Maxwell Premium for 1966.

Dr. Joseph Croney graduated from London University in electrical engineering, and before the war he was on the staff of the Northampton Polytechnic, London (now the City University). In 1939 he joined H.M. Signal School (now A.S.W.E.) and worked on the early developments of naval radar receivers for metre waves. In 1960 he was one of the small team that built the first naval 10 cm radar. He developed the first sea-clutter rejection circuits, and, after the war, the first noise-limiting successive-detection logarithmic amplifiers for which he holds the patent. In 1950 he became head of the Civil Navigational Aids Division of A.S.W.E., and in 1954 of the Antenna Techniques Division where he carried out work both on clutter decorrelation techniques using high-speed antenna scanning, and on electronic scanning antennas. He is the author of numbers of published papers in these fields including two in the Institution's *Journal* for which he has received Premiums. He received his Ph. D. from the University of London in 1966 for a thesis on Radar Clutter Reduction Techniques, and an 'Individual Merit' promotion to Deputy Chief Scientific Officer in 1967 for his contributions to radar technology; he is author of the chapter on Civil Marine Radar in the forthcoming McGraw-Hill 'Handbook of Radar'.

David R. Noaks is currently a lecturer in the Electronic and Electrical Engineering Department of the University of Birmingham. He received his engineering training as a student apprentice with the British Thomson Houston Company, Rugby, between 1950 and 1955, graduating with an external London B.Sc. in electrical engineering. He accepted a Short Service Commission in the Instructor Branch of the Royal Navy and spent five years at the Royal Naval College, Greenwich, lecturing on digital systems. His main research interests are in hybrid computer systems and the application of finite automata theory to the design of digital systems.

S. Marjanovic obtained his B.Eng. in the Faculty of Electrical Engineering at the University of Belgrade in 1962. He was then engaged as an assistant lecturer in the Electronics and Telecommunications Department until 1964 when he came to England to study for a higher degree. He worked on hybrid computer circuitry at the University of Birmingham and on being awarded a Ph.D. returned to Yugoslavia where he is currently engaged on the design of a hybrid computing system at the Institute for Automation and Telecommunications, Belgrade.

The paper by Dr. Croney and Mr. Woroncow, entitled 'Radar Polarization Comparisons in Sea-clutter Suppression by Decorrelation and Constant False Alarm Rate Receivers' was published in the October 1969 issue of *The Radio and Electronic Engineer*.

Dr. Marjanovic's and Mr. Noaks's paper on 'A High Speed, High Accuracy, Digitally-set Potentiometer' appeared in the December 1969 issue.

U.D.C. 535.374:629.13.052

A Pulsed Laser Altimeter

By

P. A. VAUGHAN,

B.Sc.†

Reprinted from the Proceedings of the I.E.R.E. Conference on 'Lasers and Opto-Electronics' held at the University of Southampton on 25th to 28th March 1969.

The paper discusses the development of an equipment which determines the height of an aircraft by measurement of the transit time of a pulse of laser light travelling from the transmitter to the ground, and back to the receiver. The transmitter source is a GaAs laser diode mounted in a Dewar, and cooled to 80° K by a self-regulating Joule–Thomson cooler. A cold cathode modulator is employed to drive the laser, the emitted energy being collimated by a multi-element transmitter lens. The receiver is mounted coaxially with the transmitter and takes the form of a Cassegrainian telescope. Detection is provided by a silicon photodiode which drives a high-gain video amplifier via a low-noise f.e.t. pre-amplifier. Transit time measurements are made with a modified high speed counter, and ancillary circuitry is incorporated to allow height information to be fed to an aircraft tape recorder system.

The system design parameters are presented together with a discussion of the major component design features and problems. This includes design of the Dewar, laser modulator, transmitter and receiver optics, and the receiver detector/amplifier assembly. The Dewar design problems result from the requirement for a low inductance feed into the Dewar for the laser drive pulse, and the necessity of retaining a vacuum over long periods of time under adverse environmental conditions. The laser modulator design presents the problem of driving the laser with fast high current pulses (40 ns pulse width, 200 A). The detection system design requires the optimization of the detector/pre-amplifier combination for fast, sensitive, low noise performance. The engineering, environmental testing, and functional testing of this equipment is mentioned, and the performance figures given.

1. Introduction

In 1964 the Services Electronics Research Laboratory, Baldock, demonstrated the feasibility of using an airborne pulsed laser system for the measurement of aircraft altitude.‡ The present paper discusses the subsequent design, development and test of a fully engineered equipment for permanent use in a combat aircraft. It will be employed as a height datum for trials work on radio altimeters which suffer from inaccuracies due to 'penetration effects'. The system employs a gallium arsenide laser emitting at 0·85 μm, and determines the height by measurement of the transit time of a pulse of laser energy travelling from the transmitter to the ground and back to the receiver.

2. System Parameters

The analysis given below shows the relationship between the system parameters and the achievable range and accuracy, thus indicating how range and accuracy may be maximized.

† Elliott Space and Weapons Research Laboratory, Frimley, Camberley, Surrey.

‡ Hambleton, K. G., 'Results of Gallium Arsenide Laser Altimeter Trials (Flights 1 and 2)', Technical Report No. M226, August 1965.

2.1. *Range*

The received power is given by

$$P_r = \frac{K_t K_r P_t A f(\theta)\, e^{-2\alpha R}}{2\pi R^2} \text{ watts} \qquad \ldots\ldots(1)$$

assuming a terrain which scatters the laser energy isotropically into a hemisphere and a receiver field of view which overlaps that of the transmitter. K_t and K_r are the transmitter and receiver optical transmissions, P_t the laser output power, A the receiver aperture area, $f(\theta)$ the terrain scattering coefficient, R the range and α the one-way atmospheric attenuation coefficient.

The noise which competes with this signal consists of two components: noise from the detector/amplifier assembly and background noise produced by the solar energy scattered from the terrain. Now the background power entering the receiver is given by:

$$\frac{1}{2\pi}\beta\phi K_r B A W f(\theta) \text{ watts} \qquad \ldots\ldots(2)$$

again assuming the terrain to behave as an isotropic scatterer. β and ϕ give the receiver field of view, B is the spectral pass-band of the receiver filter and W is the solar spectral irradiance at sea level. The background power will generate a detector current I which will in turn produce an r.m.s. noise current given by $\sqrt{(2eI\Delta f)}$ where e is electron charge and Δf the electrical post detector bandwidth. Therefore the background noise is given by

$$\sqrt{\frac{e\Delta f\beta\phi K_r BAW(f\theta)S}{\pi}} \text{ amperes r.m.s.} \quad \ldots\ldots(3)$$

where S is the detector sensitivity.

Now the peak signal to r.m.s. noise current ratio for a detector/amplifier noise limited system is given from eqn. (1) as

$$\text{s.n.r.} = \frac{K_t K_r P_t A f(\theta)(e^{-2\alpha R})S}{2\pi R^2 i_n} \quad \ldots\ldots(4)$$

where i_n is the r.m.s. noise current generated by the detector/amplifier assembly, being independent of the background. This equation gives the expected result that maximum range will be achieved by maximizing K_t, K_r, P_t, A, S and minimizing i_n.

For a background-noise-limited system eqns. (1) and (3) give

$$\text{s.n.r.} = \frac{K_t P_t \sqrt{K_r A f(\theta) S}\, e^{-2\alpha R}}{2R^2\sqrt{\pi e\Delta f\beta\phi BW}} \quad \ldots\ldots(5)$$

Equation (5) is further complicated as K_t, A, β and ϕ are inter-related due to their common dependence on the focal length, aperture and F number of the transmitter optics. The general requirements for optimization of s.n.r. for a background-noise-limited system can be established from eqn. (5) but a detailed optimization is complex and will not be considered here.

The system parameters of the equipment described below were chosen on the basis of a compromise between the requirements of eqns. (4) and (5) and the requirements of engineering considerations.

2.2. *Accuracy of Height Measurement*

This is basically dictated by the accuracy with which the transit time can be measured with the high-speed counter employed (see 3.1.5). A further error is introduced by the rise-time of the receiver signal pulse and the noise present with the signal. The inaccuracy due to pulse rise-time is given by $\pm t/2 \times c/2$ where c is the velocity of light and t is the rise-time, i.e. $\sim \pm t'/4$ feet where t' is the rise-time in nanoseconds. It is therefore apparent that the rise-time of the transmitter light pulse and the receiver detector/amplifier assembly must be kept at a minimum in order to minimize t'. The effect of noise on accuracy

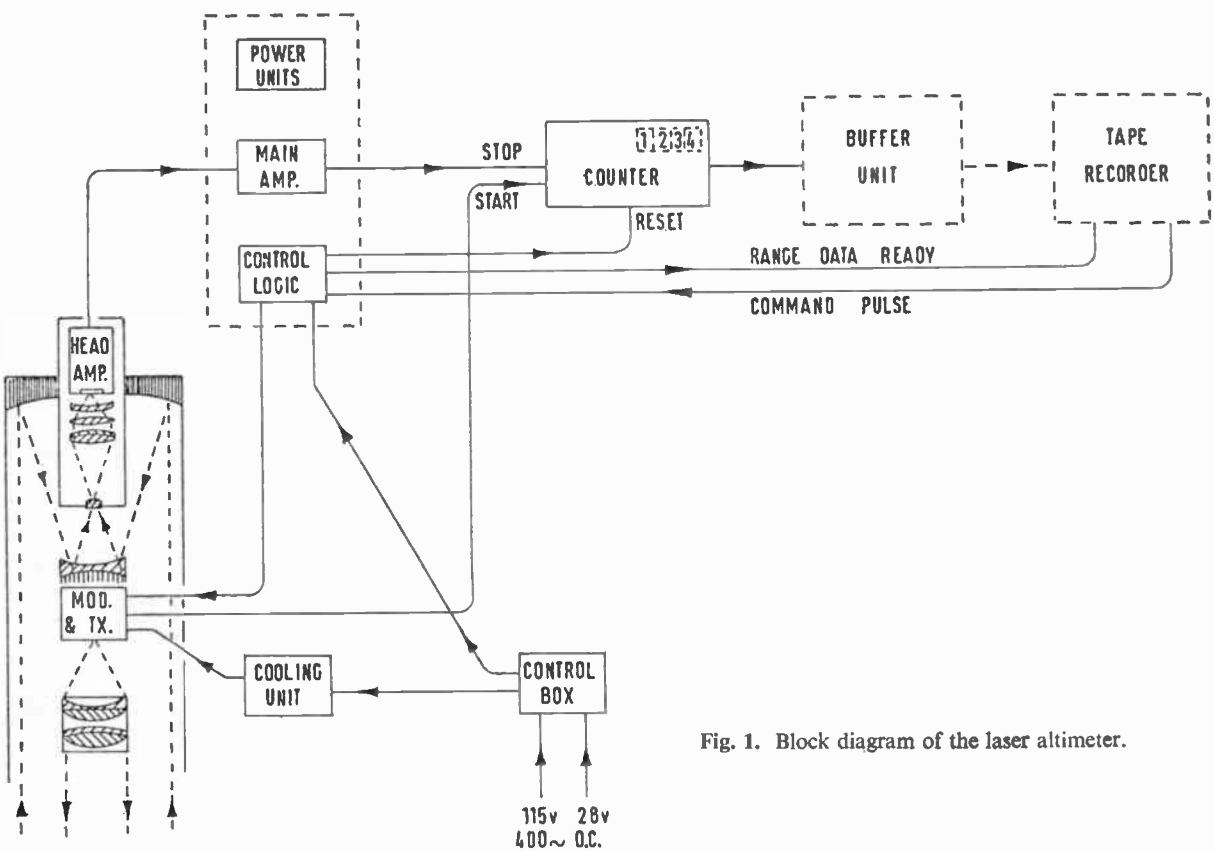


Fig. 1. Block diagram of the laser altimeter.

is complex but it is unlikely to be significant in comparison with the inaccuracies due to the counter and rise time.

3. Description of Equipment

A block diagram of the system is shown in Fig. 1. The transmitter is mounted coaxially with the receiver in order to overcome parallax problems and make best use of the available circular aircraft window.

3.1. *Transmitter*

3.1.1. Laser Dewar assembly

The transmitter source is a gallium arsenide laser diode mounted in a Dewar assembly allowing cooling to approximately 80°K, the diode efficiency increasing with decrease in temperature. High efficiency is necessary in order to achieve a high peak laser output power at a reasonably low input current level which can be generated by a practical drive circuit (see 3.1.2). Cooling is provided by a Joule–Thomson system employing a self-regulating minicooler fed from a high pressure pure air pack. For a given size of air storage bottle, the self-regulation gives a much greater operating time over conventional minicoolers as the flow rate is automatically controlled to maintain just sufficient liquid air in the Dewar.

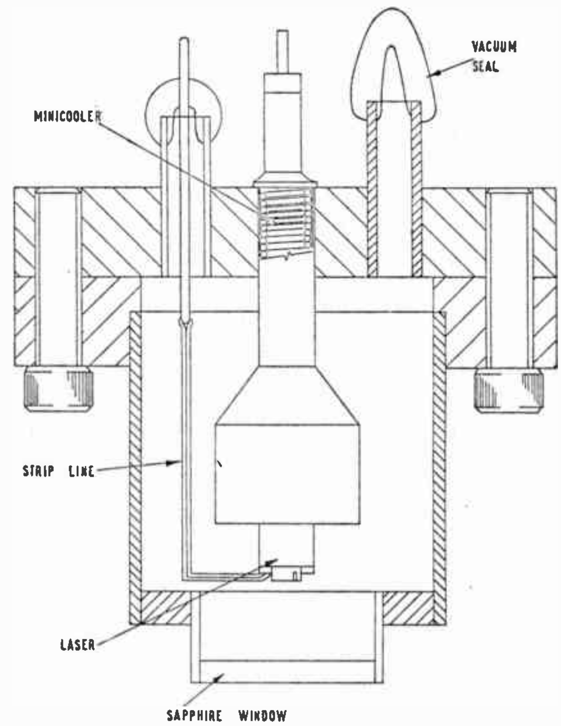


Fig. 2. Arrangement of the Dewar.

The Dewar (see Fig. 2) is made of stainless steel to give robustness and ease of manufacture. It consists of a 3·8 cm (1·5 in) diameter cylindrical outer wall which is welded at its upper end to the inside diameter of a heavy gauge annulus, the lower end accepting a sapphire window. The upper end of the inner cylindrical wall is similarly welded to a second heavy gauge annulus. The lower end of the inner cylinder is flared outwards to accommodate a 1·9 cm (0·75 in) diameter disk which serves as a heat sink for the laser. The two annular sections are bolted together and the vacuum seal provided by a gold ring. The minicooler fits down inside the inner cylinder thus placing precision tolerances ($\pm$0·0076 mm, 0·0003 in) on the bore of this tube (7·2 mm, 0·283 in), the efficiency of the minicooler being dependent on a good fit. The tube must also have very thin walls (0·127 mm, 0·005 in thick) in order to minimize heat transfer from the outside of the Dewar to the cold laser.

The inner cylindrical wall assembly described in the preceding paragraph was designed with considerable care. It is necessary to avoid fabrication techniques which entail machining across the grain of the stainless steel, as this can result in the formation of leaks through the thin sections of titanium-loaded stainless steel due to 'stringers'. These stringers may be eliminated by the use of niobium-loaded, vacuum-smelted stainless steel but this material is difficult to acquire in small quantities. The solution chosen in the present application was to use a construction which avoided the requirement for machining across the grain. This consists of a narrow bore precision tube joined to a larger bore tube by an annulus. The two tubes are machined with their axes along the length and thus along the grain of a stainless steel rod. The annulus is manufactured with suitable grain orientation from a stainless forged blank, as is the laser heat sink. Ideally, copper should be used for the heat sink in order to give good thermal conductivity. However, the use of stainless steel results in a better manufacturing procedure due to the ease of welding the heat sink to the large bore tube.

The pressure in the Dewar envelope must be sufficiently low to produce the required low heat transfer between the cold laser and the outside of the Dewar. This is necessary to achieve high cooler efficiency and prevent misting of the Dewar window. The desired pressure level was computed as less than 10^{-2} torr, it being further considered necessary that this should be maintained for a period of one year in order to achieve a reasonable service life. This pressure and time fixes the maximum outgas rate allowable for the materials within the Dewar, these being stainless steel, kovar, copper, glass, mylar and nylon. The mylar is used as an insulating medium for the strip line (see 3.1.2) and nylon screws are used to secure the laser diode to the heat sink via copper jaws. The outgas rates for the metals and glass will all lie in the range 10^{-15} to 10^{-14} torr litres s^{-1} after outgassing for 24 hours at 200°C. This is much lower than the outgas rate for mylar and nylon which is in the range 10^{-11} to 10^{-10} torr litres s^{-1} after outgassing for 60 hours at 120°C (maximum allowable temperature). It is thus obvious that the outgassing from the mylar and nylon will predominate in spite of the metal area being greater by a factor of 10. It is for this reason that the mylar and nylon areas are kept to a minimum when it is possible to achieve the desired pressure level given above, provided great care is taken to properly outgas all the Dewar materials. The laser diode also resides inside the vacuum envelope and will outgas. However, it consists of small areas of materials with reasonably low outgas rates, i.e. copper, gallium arsenide, indium and glass. It therefore presents no problem provided it is kept clean during manufacture and handling.

3.1.2. Laser modulator

The cooled laser diode used in this equipment requires a current drive of 200 A in order to achieve the maximum peak power of 50 W. It is also necessary to generate a drive pulse with short rise-time in order to achieve good accuracy (see 2.2). The basic form of this drive circuit is a fast switch which discharges a highly-charged capacitor stack through the laser diode. The circuit thus consists of a capacitor stack in series with a switch, a laser diode, and stray inductance, i.e. an RLC series circuit. The current which flows in this circuit when the switch is closed is given by

$$i = \frac{V}{\sqrt{(L/C)-(R^2/4)}} \exp\left(\frac{-Rt}{2L}\right) \sin\left(\frac{1}{LC} - \frac{R^2}{4L^2}\right)^{\frac{1}{2}} t \qquad \ldots\ldots(6)$$

assuming that the switch behaves perfectly and the circuit is under-damped, this being the general case. R is the diode plus circuit resistance, L the stray inductance, C the storage capacitor and V the voltage to which the capacitor is charged. This current will reach a maximum value of

$$i_{\max} = V\left(\frac{C}{L}\right)^{\frac{1}{2}} \exp\left(\frac{-R\pi}{4}\right)\left(\frac{C}{L}\right)^{\frac{1}{2}} \qquad \ldots\ldots(7)$$

at a time $t = \frac{\pi}{2}(LC)^{\frac{1}{2}}$ ......(8)

assuming $\frac{1}{LC} \gg \frac{R^2}{4L^2}$

Equation (7) indicates the general circuit requirements of large V and small R in order to achieve a high maximum current. The general requirements for good rise-time are given from eqn. (8) as small C and L.

The switch chosen for this modulator circuit is a cold-cathode trigger tube (E.G.&G. type KN 2). Three low-inductance high-voltage capacitors are connected in parallel to form the capacitor stack which is charged to 1 kV, this being the maximum voltage at which reliable switch operation can be achieved. The value of capacitance is chosen to give adequate peak current whilst not unduly degrading rise-time. The circuit is constructed with heavy-gauge connectors of minimum length in order to minimize circuit series resistance and inductance. The modulator circuit is mounted on top of the Dewar assembly and the low impedance connexion to the laser diode represents a severe problem. This connexion is approximately 7·6 cm (3 in) long and must pass through the vacuum envelope. A parallel-plate kovar strip line is used, the inductance being inversely proportional to the plate separation. This separation is kept very small by the use of 0·0127 mm (0·0005 in) thick mylar as the dielectric medium. The strip line was initially taken through the vacuum envelope with an araldite seal. The first seals cracked readily on temperature cycling due to the differential expansion between the stainless steel and the araldite. This problem was improved by the use of silica flour fillers to decrease the expansion coefficient of araldite, and flexible resins to accommodate the differential expansion. The resultant seals were good up to the equipment maximum survival temperature of +70°C but unfortunately failed at the lower survival temperature of −65°C. The araldite seals were therefore discontinued and replaced with a glass-to-metal seal providing a twin wire feed through, the strip line being reformed inside the Dewar. It was of course necessary to accept a degradation in performance due to the increased inductance.

It is desirable to make the characteristic impedance of the strip line equal to that of the laser diode in order to deliver maximum drive current. The impedance of the diode is very low (approx. 20 mΩ) and in practice the line impedance is made as small as possible (i.e. 0·15 Ω).

The modulator circuit described above can be triggered via the cold cathode tube grid at rates from 0–100 pulses/second. It delivers 200 A with a rise-time of 25 ns (10–90%). The main disadvantage of this circuit is that the cold cathode trigger tube is limited to a life of approximately 10^7 pulses.

3.1.3. Optics

The radiation emitted from the line source laser is collimated by a transmitter lens of 16·5 cm focal length and 8·25 cm aperture producing a beam divergence of 6 × 2 milliradians. The power reflected from the ground is collected by the receiver which takes the form of a Cassegrainian telescope with a

30·4 cm (12 in) aperture. After reflexion by the primary and secondary mirror, the received light is brought to a focus at the field stop which defines the receiver field of view at 10×6 mrad. The energy from the stop is then refocused on to a silicon detector of 2·5 mm diameter.

The aberrations in the above optical system must be made small in comparison with the field of view, i.e. they should be kept less than 0·5 mrad on both the transmitter and receiver. Aberrations on the transmitter optics are kept to a minimum by the use of a 4-element lens with optimized radii of curvature. The receiver optics are designed with spherical surfaces throughout in order to ease manufacturing problems. The secondary mirror has a concave refracting surface to allow for correction of aberration which is minimized once again by optimizing the radii of curvature of the primary and secondary mirrors. The optical chain employed for refocusing the light from the receiver stop employs proprietary lenses, the degree of aberration that is introduced being tolerable.

The optimization of radii of curvature was carried out using a computer ray-tracing program. The manufacturing radii were then chosen as those available test plates closest to the optimum. The theoretical aberration for this system is 0·05 mrad for the transmitter and 0·003 mrad for the receiver, allowing degradation by a factor of 10 in manufacture. Measurements carried out on the completed optics showed the aberrations to be less than 0·5 mrad.

3.1.4. Detector/amplifier assembly

The detector choice lies between a photomultiplier and a silicon detector.

3.1.4.1. *Photomultiplier Tube.* The photomultiplier has high internal gain and the noise generated in the subsequent amplifier may be ignored in comparison with the noise produced by the tube. The noise is therefore a combination of the components given below. (It should be noted that peak rather than r.m.s. noise must be considered due to the peak threshold detector used in the altimeter):

(i) Dark current noise
(ii) Background noise
(iii) Signal induced noise.

(i) The magnitude of the r.m.s. dark noise current is given by

$$i_n = \sqrt{2eI_d\Delta f} \qquad \ldots\ldots(10)$$

where e is electron charge, I_d the dark current, and Δf the post-detector electrical bandwidth. Now the value of I_d depends on the ambient temperature and the type of photocathode employed. For the present application an S1 photocathode is required in spite of its high dark current, as it exhibits high quantum efficiency at 0·85 μm. The probability distribution of the dark noise is non-Gaussian and therefore the normal relationship between peak and r.m.s. noise cannot be applied. Experiments designed for the system under consideration were carried out to find this ratio. It was established that for a probability of false alarm of 1% and a receiver 'on' time of 20 μs, the peak to r.m.s. noise ratio is 15.

(ii) The problem of background noise has already been discussed in 2.1.

(iii) The current generated by the signal will generate noise given by

$$\text{r.m.s. noise current} = \sqrt{PS2e\Delta f} \qquad \ldots\ldots(11)$$

where P is the signal power falling on the photocathode and S the cathode sensitivity. The signal/noise ratio is thus given by

$$\frac{PS}{\sqrt{2PSe\Delta f}} = \sqrt{\frac{PS}{2e\Delta F}} \qquad \ldots\ldots(12)$$

Of the three noise mechanisms discussed above, the dark noise will normally predominate giving a minimum detectable signal at the photocathode of 7×10^{-7}W at 55°C for a Δf of 25 MHz and a 1% probability of false alarm. The equivalent figure for a silicon detector is $1{\cdot}3 \times 10^{-7}$ (see 3.1.4.2). This makes the silicon detector the obvious choice especially in view of the attendant advantages of ruggedness, reliability and consistency of performance.

3.1.4.2. *Silicon Detector/Amplifier Assembly.* The avalanche silicon detectors were not considered for this equipment as they were not available at a sufficiently early point in the contract.

The conventional silicon detector has no internal gain and the thermal noise in the subsequent amplifier normally predominates. The detector behaves as a current source and thus the pre-amplifier input impedance must be low in order to achieve fast detection. A virtual-earth amplifier therefore represents an ideal choice, an operational amplifier being the most convenient form of such an amplifier. The optimum choice of input device is a f.e.t. which will provide the best noise performance and adequate speed of operation. The arrangement employed is a low-noise f.e.t. (2N3823) input operating as a source follower. This is followed by a grounded base n-p-n device which provides efficient current coupling into a p-n-p common emitter stage. The output stage is a bootstrapped n-p-n emitter follower which is capable of driving a 50Ω load. The bootstrapping is provided in order not to load the previous emitter follower. Overall negative feedback is applied from the output emitter to the f.e.t. gate. The silicon detector is a.c.-coupled to the front of the amplifier, a bias of 90 V being supplied through a large bias resistor. The noise of this system comes from three sources:

(i) Feedback resistor
(ii) Field-effect transistor
(iii) Source impedance, i.e. detector impedance.

(i) The thermal current noise from the feedback resistor can be reduced by increasing the resistor value but this increases the amplifier gain and reduces the bandwidth. A compromise must therefore be chosen.

(ii) The f.e.t. noise may be minimized by using a low-noise device (i.e. 2N3823) and selecting for low I_{DSS}.

(iii) It is necessary to maintain a high source impedance in order to minimize the current noise from this component. This is achieved by operating the diode at high bias voltage (90 V) which results in a low junction capacitance and high reactance. Low junction capacitance also leads to fast detector performance. Noise may be further reduced by the use of a series padding resistor but this will reduce the detector speed and a compromise must be reached.

A detailed analysis was made of the three noise mechanisms and a computer program used to calculate noise performance over a wide range of conditions. The amplifier was optimized as a result of this work and the noise measurements made on this amplifier agree well with the calculated values. It was further shown that the noise from the source impedance predominates although that from the feedback resistor and f.e.t. is still significant.

The detector/pre-amplifier assembly achieves a noise level of 25 nA r.m.s. and a bandwidth of 25 MHz.

3.1.5. Measurement and control electronics

A video amplifier is used to raise the output from the pre-amplifier to a level of about 5 V for driving the counter. This amplifier gives a voltage gain of 10^5 and will deliver 8 V with a rise-time of 10 ns. (10–90%) into a 50 Ω load.

A modified Marconi high-speed counter is used to measure the transit time, the count being started by a sample from the transmitter current pulse and stopped by the returned signal pulse from the video amplifier. The clock frequency is 98·3312 MHz which results in a read-out in increments of 5 ft (1·52 m). The accuracy is ± 1 count resulting in a height accuracy of ± 5 ft. Altitude is displayed on a 4-digit display using gas-filled numerical indicator tubes ('Nixie' type) and a binary coded decimal output is also provided for feeding height information to the aircraft tape recorder.

The altimeter can be triggered externally at rates from 0–100 pulse/s or run off its own internal oscillator at 10 pulse/s, this normally only being used for ground testing. Control logic is provided for resetting the counter after each measurement and opening the input gates on the aircraft tape recorder.

The equipment is operated remotely from a control panel and has been designed to operate from 400 Hz 115 V and 28 V d.c. aircraft supplies.

4. Engineering

The general standard of engineering for the altimeter must be high as the equipment must meet a severe environmental specification (see Section 5). The main engineering feature is the transmitter/receiver assembly which houses the optics. It consists of five annular castings and two tubes, the whole assembly mounting together to form a cylinder some three feet in length. The assembly must be sufficiently rigid to provide optical stability under vibration and machined to a high accuracy to allow automatic positioning and alignment of the optics on assembly. This overcomes the requirement for complex adjustment mechanisms and extended setting-up procedures.

5. Testing and Performance

The equipment was subjected to environmental testing. It was shown to function correctly over the temperature range -30°C to $+55$°C and survive the range -55°C to $+70$°C, the lower survival temperature being accompanied by a pressure equivalent to an altitude of 50 000 ft (12 700 m). Correct functioning was demonstrated during vibration testing conducted over the frequency range 5 to 150 Hz at a vibration level which reached a maximum value of 8 g at 150 Hz. Acceleration survival was demonstrated at 6·5 g and the equipment was shown to function during a damp heat test.

Limited functional tests were carried out in the laboratory and the system is now undergoing acceptance testing in a *Comet 4C*. It is expected that the equipment will work over the altitude range 15–2000 ft and achieve an accuracy of ± 5 ft or ± 1% whichever is the greater. The running time, set by the air storage bottle size, is approximately 10 hours.

6. Possible Improvements

Since the design and development of this equipment considerable advances have been made in silicon avalanche diodes and room temperature gallium arsenide laser diodes. These advances, if incorporated, would not improve performance greatly but should give an appreciable reduction in size and complexity.

7. Acknowledgments

The development of the laser altimeter was sponsored and funded by the Ministry of Technology. The laser was provided by S.E.R.L., Baldock.

Manuscript received by the Institution on 13th January 1969. (Paper No. 1299/AMMS27).