ATR

AUSTRALIAN TELECOMMUNICATION RESEARCH

VOL. 17, No. 2, 1983

ATR is published twice a year (in May and November) by the Telecommunication Society of Australia. In addition special issues may be published.
ATR publishes papers relating to research into telecommunications in Australia.

CONTRIBUTIONS: The editors will be pleased to consider papers for publication. Contributions should be addressed to the Secretary, ATR, c/- Telecom Australia Research Laboratories, 770 Blackburn Rd., Clayton, Vic., 3168.

SUBSCRIPTIONS: Subscriptions for ATR may be placed with the General Secretary, Telecommunication Society of Australia, Box 4050, G.P.O., Melbourne, Victoria, Australia, 3001. The subscription rates are detailed below. All rates are post free. Remittances should be made payable to the Telecommunication Society of Australia, in Australian currency and should yield the full amount free of any bank charges.

The Telecommunication Society of Australia publishes the following journals:

1. **The Telecommunication Journal of Australia** (3 issues per year)

   Subscription — To Members of the Society* resident in Australia $8.00
   Non-members of Australia $13.50
   Non-members or Members Overseas $20.00

2. **ATR** (2 issues per year)

   Subscription — To Members of the Society* resident in Australia $10.00
   Non-members in Australia $20.00
   Non-members or Members Overseas $24.00
   Single Copies —To Members of the Society resident in Australia $7.50
   Non-members within Australia $12.50
   Non-members or Members Overseas $15.00

*Membership of the Society $7.00

All overseas copies are sent post-free by surface mail.
Prices are for 1984. Please note the revised rates.

Enquiries and Subscriptions for all publications may be addressed to:
The General Secretary, Telecommunications Society of Australia, Box 4050, G.P.O. Melbourne, Victoria, Australia, 3001.

ATR

AUSTRALIAN TELECOMMUNICATION RESEARCH

ISSN 0001-2777

VOLUME 17, NUMBER 2, 1983

Contents

Challenge . . .

Telecommunication industry is in the throes of revolutionary change, in technique, in organisation, and in societal impact. The world telecommunication system is the most complex coherent system ever designed by humans and it is physically (geometrically) larger than any other. We can well be proud of the quality of the contribution of Australian research in this context. Being part of the thriving, achieving, and needed, telecommunication engineering community is a source of pleasure that enhances our ability to contribute.

However there are challenges to our ability to continue with achievements of which we can be proud, and hence threats to Australia's ability to develop, or even to maintain, its telecommunication systems.

One challenge is the sheer size of the effort needed just to stay abreast of world developments with the strictly limited resources that can be allocated to this area, while the world supply of developments expands. I have been impressed by the efforts my friends in the Telecom Research Laboratories put into being informed. However there is very real difficulty in finding the resources to make significant, innovative, contributions to research, particularly when this is becoming ever more expensive.

A second challenge lies in the change from an expanding economic system to one in which innovations and changes in function have to come from attrition of existing activities, and through restructuring. The expanding system, to which many of us had become accustomed, could not go on for ever, but the change of pattern does detract from the momentum and morale of those who had developed valued specialist expertise.

Thirdly, along with the quantitative constraints, is the consequence that the educational institutions, as well as other laboratories, are now less able to maintain an infusion of new staff from a diversity of backgrounds. The base of academics' experience is not able to expand commensurately with the expansion of knowledge. Thus, graduates of the future may be worse equipped to solve the as yet unrecognised challenges that they will face.

Unless we meet these challenges there will be an increasing trend toward importation of telecommunication technology. There would follow a reduction of the call on our ability to contribute to the country's capability, economic strength, and employment. The effectiveness of engineers is a delicate function of their job satisfaction. This depends to a large extent on their making such contributions, rather than on simply evaluating the sales proposals of foreign suppliers. Thus there is a probability of a serious downward spiral of Australia's ability to provide for future technological needs, including maintenance of existing services, unless the trends are reversed. We are in a position to do something about it.

Let us review some of the potentials of the present situation. Firstly, we have a strength in the unprecedented quality of the young people undertaking electrical engineering degrees, due partly to their perceptions of the relative opportunities of satisfying employment in various vocations.

Secondly, the seniority of academics and engineers mentioned above does go with well developed capability and perspective. Thirdly, some of the academics and institutions that have resisted compromises of their purity are at last learning to be more flexible in relations with other organisations, due to the financial exigencies. An example is the establishment of arrangements for higher degrees to be based on research done off-campus. Overall, the pool of intellectual resource potentially available in the country for tackling telecommunication problems is stronger than ever before.

Now is the time to explore new ways to harness this strength. Enhancing the collaboration between the universities and industrial bodies is one area with some little-tapped potential. Telecom, with before it the PMG Department, has a good reputation for facilitating research in universities, in sharing information, and in collaborating in educational endeavours. However there has been little exchange of personnel. Relatively few academics have spent hard won study leave in any Australian laboratories. And yet how often do we pay lip service to the thought that interaction with people with different, but sympathetic, views can open up new, fruitful ways of dealing with problems?

One opportunity to enhance interaction, with benefit to both sides, would be for some projects in government or industrial research laboratories to involve some academics directly, in planning, supervision, and execution. Such arrangements are well established in some other countries, and many of our academics have contributed to the objectives of the telecommunication laboratories there. Suitable part-time residency would need to be established for on-going effectiveness, and it would be necessary to ensure that the individuals involved did not incur too much personal loss.

Also, there is scope for people from non-university laboratories to work in many universities with academics on problems of interest to, or taken from, those laboratories. The facilities in universities are of course different, and often not as extensive, but the university environment is a very stimulating one and frequently leads to enhanced personal productivity and new skills.

The education of future engineers depends on the awareness of their mentors in universities. If enhanced interaction can bring to the mentors an improved understanding, it can but be of future benefit.

"Anything is impossible if you think about it the right way", and there are always difficulties with changes. However, this challenge is worth facing, and we can do something about it. I ask you to discuss these ideas with some colleagues, and to improve upon them.

Please let me or the Editor-in-Chief have your views on the challenge, on these possibilities, and on other possibilities. We hope readers will develop these themes, and that we can publish some of your ideas or criticisms in future issues.

ROBERT E. BOGNER,
Department of Electrical and Electronic Engineering,
The University of Adelaide, Adelaide 5000.

Telecommunications Paradoxes

L.T.M. BERRY
Applied Mathematics Department
University of Adelaide

An interesting phenomenon called the transportation paradox† *(Ref. 1) is well known: "a paradox arises when a transportation problem admits to a total cost solution which is lower than the optimum and is attainable by shipping* larger *quantities of goods over the same routes that were previously designated as optimal". It is shown that an analogous statement holds for the minimum cost telecommunications network design problem.*

A second "paradox" of greater practical significance is also considered. It is shown that under some circumstances the addition of further circuits to a network may lead to an increase *in total traffic lost due to congestion!*

†*Although called a paradox, it is not strictly a* logical *paradox.*

1. INTRODUCTION

In practice telephone networks are dimensioned with the help of many approximate formulae, the progeny of mathematical models of parts of the real system. If such formulae lead to a sufficiently accurate prediction of the network performance for a given circuit allocation they are useful. The purist may frown if occasionally the approximations give anomalies such as state probability values greater than 1 or negative chain flows or perhaps even hypothetical circuits. Nevertheless, provided the anomalies fall outside the ambit of what occurs normally in practice, the formulae may still be useful.

At times such formulae even give conclusions surprisingly contrary to our intuition which on close investigation prove to be correct. In this paper we examine two commonly held intuitively reasonable related suppositions and show both to be *false*. In our investigation of the second supposition the widely used equivalent random model (Refs. 2,3) plays a prominent part.

Supposition 1. A network having a minimum cost circuit allocation for specified origin-destination grades of service cannot carry more traffic on the same routes at cheaper cost.

Supposition 2. The addition of more circuits to links of a network cannot result in a diminishing of the total flow. Or, equivalently, the removal of circuits cannot lead to an increase in the total carried traffic.

2. THE TRANSPORTATION PARADOX AND SUPPOSITION 1

Consider the balanced transportation problem of Table 1. The entries c_{ij} in the tableau give the cost of transporting one unit from origin i to destination j. The available quantity at origin i is a_i whilst the demand at destination j is b_j. The problem can be formulated as the integer programme:

$$\text{Min } z = \sum_{i=1}^{4} \sum_{j=1}^{5} c_{ij} x_{ij} \qquad (1)$$

$$\sum_{j=1}^{5} x_{ij} = a_i \ , \ i = 1, \ldots 4 \qquad (2)$$

$$\sum_{i=1}^{4} x_{ij} = b_j \ , \ j = 1, \ldots 5 \qquad (3)$$

$$x_{ij} \geq 0 \text{ for all } i,j \qquad (4)$$

$$x_{ij} \quad \text{integer} \qquad (5)$$

TABLE 1 - A 4×5 transportation problem

		DESTINATIONS					
		1	2	3	4	5	a_i
ORIGINS	1	14	15	6	13	14	7
	2	16	9	22	13	16	18
	3	8	5	11	4	5	6
	4	12	4	18	9	10	15
b_j		4	11	12	8	11	

The *optimal* solution with cost 444 units is given by the entries x_{ij} (in corresponding cells i,j) in Table 2. When a_1 is incremented to 10 and b_4 to 11 (i.e. transporting 3 further units) the assignments shown in Table 3 give the minimum cost of 438 units!

TABLE 2 - Optimal solution (Cost 444)

	1	2	3	4	5	a_i
1			7			7*
2	4	6		8		18
3			5		1	6
4		5			10	15
b_j	4	11	12	8*	11	

TABLE 3 - Optimal solution (Cost 438)

	1	2	3	4	5	a_i
1			10			10*
2	4	3		11		18
3			2		4	6
4		8			7	15
b_j	4	11	12	11*	11	

The point to note is that a cheaper solution is found by replacing the *equalities* (2), (3) by the *inequalities*

$$\sum_{j=1}^{5} x_{ij} \geq a_i \ , \ i = 1, \ldots, 4 \qquad (6)$$

$$\sum_{i=1}^{4} x_{ij} \geq b_j \ , \ j = 1, \ldots, 5 \qquad (7)$$

The feasible region has been increased, and under suitable conditions the optimal cost may be decreased by allowing more traffic onto the network.

A partial analogy may be obtained with telecommunications networks. The following example is mentioned in Ref. 4. Two identical Poisson streams with means 6.6 erlangs are offered to the network shown in Fig. 1. The numbers of circuits on the links are represented by n_i , $i = 1, \ldots, 5$. A grade of service 0.002 is specified for each stream.

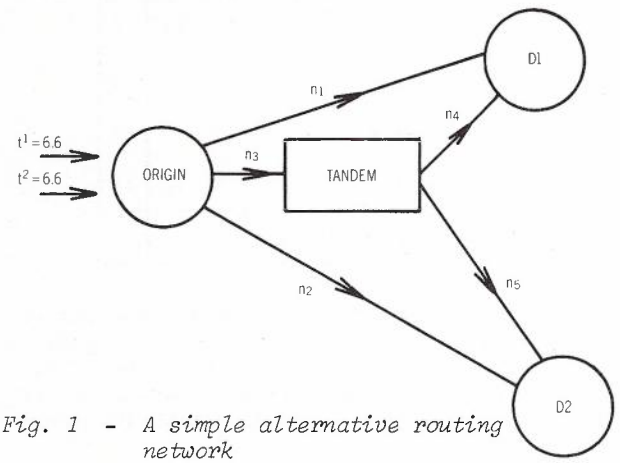


Fig. 1 - A simple alternative routing network

The *only* feasible *integer* solution to the equality constraints

$$\sum_{j=1}^{2} h_j^k = (.998)(6.6),$$

where h_j^k is the mean carried traffic on route j for O-D pair k, is

$$n_1 = n_2 = 15 \qquad (0.998 = E_{15}(6.6))$$

Thus the circuit allocation vector n = (15,15,0,0,0) gives the *optimal* solution regardless of how expensive the circuits on links O-D1 and O-D2 are compared to those on the overflow links. Clearly we can choose link costs to show that replacement of the equality constraints by the inequality constraints

$$\sum_{j=1}^{2} h_j^k \geq (.998)(6.6)$$

can lead to a cheaper solution with the total carried traffic being increased. A similar but slightly more complex example illustrating this point is also given in Ref. 5.

The differences between the above example and that illustrating the transportation paradox are:-

(i) the lower total cost is achieved by "transporting" larger quantities over *different* routes than those previously designated as optimal. It is easy to start with, say,

$$n_1 = n_2 = n_3 = n_4 = n_5 = 5$$

and to choose the equality constraints corresponding to the flows which would occur for such a network. A reduction of n_1 and n_2 to 4 and an appropriate increment of n_3, n_4, n_5 to effect an increase in the total carried traffic for each stream can be made at cheaper cost, provided the costs of the alternate routes are made sufficiently cheap relative to the direct circuit costs.

(ii) integrality conditions are on numbers of circuits rather than flows.

(iii) the objective function is nonlinear rather than linear.

We have illustrated the same phenomenon however, viz., "a total cost solution which is lower than the optimum and is attainable by carrying larger quantities of traffic over the same routes that were previously designated as optimal".

3. SUPPOSITION 2

There is perhaps a widely held belief, based on "commonsense", that whenever more circuits are provided in a telecommunications network there must be a general improvement in the network performance. Consider the network shown in Fig. 2. Exact calculation for the circuit allocation $n_1 = n_2 = n_3 = n_4 = 5$ gives chain flows $h_1^1 = 1.2357$, $h_1^2 = 1.9772$, $h_1^3 = 1.4829$, with a *total* loss of 14.3043 erlangs. Exactly the *same* performance is achieved for all values of $n_2 \geq 5$, $n_3 \geq 5$, $n_4 \geq 5$ when the common link is fixed at 5 circuits. Additional circuits provide no further benefit in performance.

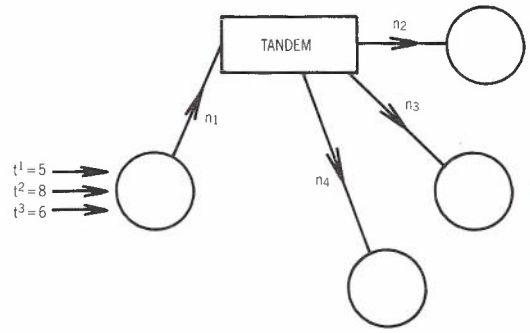


Fig. 2 - A single origin three destination network

Of course, this example is easily understood and is not paradoxical. It does however raise the question of how many *unnecessary* circuits are provided in existing networks. We next seek an example which demonstrates that the addition of further circuits to a network may result in an increase in the total lost traffic, i.e. an actual degradation of performance.

Consider the generalized Erlang system shown in Fig. 3. Each traffic stream has a single chain from the common origin to separate destinations via a tandem exchange.

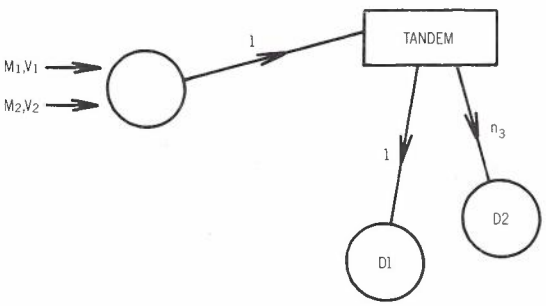


Fig. 3 - Generalised Erlang System

Suppose now that the first stream is Poissonian traffic with mean $M_1 = V_1 = 10$ erlangs and the second stream, obtained by successive overflows, has first two moments $M_2 = 50$, $V_2 = 500$. Clearly with $n_3 = 0$ the *total loss* is $50 + 10E_1(10) = 59.0909$ erlangs. Now, incrementing n_3 to 1 circuit the equivalent random method gives a total loss of 59.1241 erlangs!

The method used to solve the equations of the equivalent system, viz.

$$V = M\left[1 - M + \frac{A}{N - A + M + 1}\right] \qquad (8)$$

$$M = A\,E_N(A) \qquad (9)$$

is described in Ref. 6. Similar examples to illustrate a decrease in overall performance when the extra circuit is added are given in Table 4.

TABLE 4 - Total losses

		TOTAL LOSSES	
M_2	V_2	$n_3 = 0$	$n_3 = 1$
20	200	29.0909	29.1892
60	600	69.0909	69.1107
4	40	13.0909	13.2037
1	6	10.0909	10.1180
1	3	10.0909	10.0975

The differences in total loss for $n_3 = 0$ and $n_3 = 1$ are small in the above cases. To show that these differences are not due to errors in the iterative calculations, in section 4 we establish explicit solutions to the equations of the equivalent system (8), (9) for the network of Fig. 3. It will also become apparent that examples can readily be found for which the increase in lost traffic, due to the addition of the extra circuit, is more spectacular in magnitude. We also defer to section 5 comment on the important question of whether such degradations in performance occur in reality or whether they are due solely to the *approximations* inherent in the equivalent random model.

4. THE EQUIVALENT GROUP

With $n_3 = 1$ the network of Fig. 3 can be analysed by the equivalent random method and the total loss m found directly from (8) and (9) as follows (see Fig. 4).

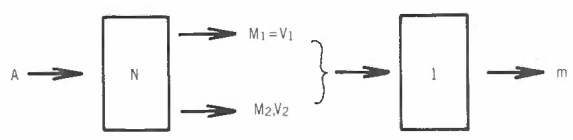


Fig. 4 - The equivalent group

Let M, V be the sums of the means and variances respectively of the streams offered to the common link of the network.

Applying the recurrence formula for the Erlang loss system,

$$m = A E_{N+1}(A) \tag{10}$$

$$= \frac{A \cdot A E_N(A)}{N+1+A E_N(A)} \tag{11}$$

$$= \frac{AM}{N+1+M} \tag{12}$$

From (8),

$$N+1+M = \frac{AM}{V+M^2-M} + A \tag{13}$$

Substitution for $N+1+M$ in (12) gives

$$m = M\left[1 - \frac{1}{M+V/M}\right] \tag{14}$$

Fortunately, both variables A and N are eliminated and the total loss is given as a simple function of M and V. Clearly the factor $(1-1/(M+V/M))$ is the congestion over the *real* circuit. Checking the numerical example of the previous section, with $M_1 = V_1 = 10$ and $M_2 = 50$, $V_2 = 500$ we obtain $M = 60$, $V = 510$ and from (14) $m = 59.124087$ which supports our previous finding.

Also we observe from (14) that the total loss asymptotically tends to M as the combined variance V tends to ∞.

The equivalent random model predicts that supposition 2 is false for the network of Fig. 3 whenever

$$M\left[1 - \frac{1}{M+V/M}\right] > \frac{M_1^2}{1+M_1} + M_2 \tag{15}$$

and clearly this occurs whenever V_2/M_2 is sufficiently large.

5. CONCLUSIONS

It has been shown that a network having a minimum cost circuit allocation, for specified origin-destination grades of service, *can* carry more traffic on the same routes at cheaper cost. This fact is of practical significance when dimensioning telephone networks for minimum cost design. For example, the nonlinear programming formulation (Ref. 7)

$$\text{Min} \quad c(\underset{\sim}{h}) \tag{16}$$

$$\sum_j h_j^k = 0.98 t^k \tag{17}$$

$$h_j^k \geq 0 \tag{18}$$

gives an optimal chain flow pattern $\underset{\sim}{h}^*$ corresponding to a circuit allocation with cost approximately \$2,700,000 for the Adelaide metropolitan network. But replacing (17) by the inequality constraints

$$\sum_j h_j^k \geq 0.98 t^k \tag{19}$$

gives a minimum cost of approx. \$2,400,000†. That is, allowing the total traffic carried for each O-D pair k to exceed the prescribed lower bound $0.98t^k$ gives a different optimal chain flow pattern corresponding to a cheaper circuit allocation (approx. 10% reduction).

Examination of the *equivalent group* equations has revealed that they predict that, under some circumstances, the addition of more circuits to links of a network *can* result in a reduction of the total flow. Or, equivalently, the removal of circuits *can* lead to an increase in the total carried traffic.

It may be thought that this latter "paradoxical" result is due to approximations inherent in the equivalent random method. Even were this so, since many dimensioning models incorporate the equivalent random method, the paradox would nevertheless have to be taken into consideration. Upon reflection, it does seem likely that the introduction of a very rough overflow component on a common link can decrease the overall total carried traffics on chains with this common link. It would be of interest to hear of practical occurrences of such conditions in existing networks. This second paradox is more likely to occur in non-hierarchical networks, where overflow traffic is permitted to subsequently be carried on direct links. Should the effect be of sufficient magnitude an argument could be made for some degree of link access prevention.

6. REFERENCES

1. Swarc, W., "The transportation paradox", Naval Research Logistics Quarterly, Vol.18, No.2, 1971, pp. 185-202.

†This value has been computed by Dr R.J. Harris

2. Wilkinson, R.I., "Theories for Toll Traffic Engineering in the U.S.A.", B.S.T.J., Vol.35, No.2, March 1956, pp. 421-514.

3. Giltay, J., "On Gradings in Automatic Telephony", De Ingenieur - E. Electrotechniek 9 No.23, June 1953, pp. 107-118.

4. Halgreen, C., "Optimizing Telenetworks by Berry's Method", Jydsk Telefon Trafikplanlaegningen, Århus, 1979.

5. Pióro, M., Lubacz, J., "A Modification of Berry's Approach to Hierarchical Telephone Network Optimization Problem", Referaty, Zeszyt 94, Instytut Telekomunikacji Politechniki Warszawskiej, 1980.

6. Olsson, K.M., "Use of computers for analytical and simulation studies when dimensioning telephone plants". Unpublished paper pp. 60-65.

7. Berry, L.T.M., "A mathematical model for optimizing telephone networks", Ph.D. thesis, The University of Adelaide, Dec. 1971.

BIOGRAPHY

LES BERRY was educated at the University of Adelaide and Adelaide Teachers' College. For 10 years from 1959-68 he taught in secondary schools. He then joined the South Australian Institute of Technology as a tutor with the aim of working towards a higher degree. In 1972 he was awarded his Ph.D. for a thesis entitled "A Mathematical Model for Optimizing Telephone Networks". Subsequently he joined the Department of Applied Mathematics at the University of Adelaide as a lecturer.

Reliability Modelling Of Fault-Tolerant Systems

Y.W. YAK
T.S. DILLON
K.E. FORWARD
Monash University

In this paper we describe a Markov model of a fault-tolerant computer system. The model, like others described in the literature, allows systems with variable numbers of both active and spare modules to be represented. This model, however, is unlike other models in that it allows for finite reconfiguration and repair time. A computer program based on this model is also described and examples of the studies made possible are given. As reconfiguration time has previously been assumed to be zero in other general purpose models, these results add to our understanding of its influence on various parameters.

1. NOMENCLATURE

1. N : number of active modules.
2. S : number of spare modules.
3. λ : failure rate of active modules.
4. λ': failure rate of one spare module.
5. μ : repair rate
6. μ': reconfiguration rate.
7. Ca: coverage for recovery from a fault in the active module.
8. Cs: coverage for recovery from a fault in the spare module.
9. Cd(I): coverage vector for degradation.
10. Y(I): sequence of allowed degradation.
11. D : maximum number of degradations allowed.
12. IA: Interval Availability.
13. IR: Interval Reliability.
14. MTFF: Mean Time to First Failure.
15. UDSF: Unavailability Due to System Failure.
16. UDTR: Unavailability Due to Reconfiguration.
17. p.d.f.: probability distribution function.

2. INTRODUCTION

The application of computers to the real time control of industrial processes has not proceeded quite as rapidly as originally envisaged. The main reason for this has been a concern associated with the lack of reliability of the computer based system, as the consequences of failure are likely to be costly and dangerous, extremely undesirable or irreversible. There are however examples of critical applications including control of dynamically unstable aircraft and complex telephone switching systems where computers have been used. The reliability requirement in such applications is extremely high and can only be achieved through the use of fault-tolerant techniques.

The SIFT (Ref. 2) & FTMP (Ref. 3) systems designed for control of dynamically unstable aircraft, for example, have a reliability requirement that the probability of failure should be less than 1×10^{-9} in a flight of 10 hours duration. This is equivalent to it failing less than once every 1,141,550 years of operation. With such a reliability requirement, it is impossible to validate the reliability by testing. It is therefore necessary to assess the reliability by mathematical modelling of these fault-tolerant systems.

During the design stage, it is necessary to estimate the reliability of different types of architecture so as to select the one that is most appropriate, and mathematical modelling is necessary to determine whether or not a particular architecture satisfies a given performance specification. To proceed with the design and construction of a system without first evaluating its reliability could result in an architecture that could be unnecessarily expensive.

These reasons have prompted us to develop a reliability model for fault-tolerant systems that provides a cheap, flexible and effective tool for design and evaluation of fault-tolerant computers.

3. ATTRIBUTES OF FAULT-TOLERANT SYSTEM

The incorporation of fault tolerance consists essentially of adding redundancy to the system. The redundancy may assume several forms namely:

1. hardware redundancy,

2. software redundancy and

3. time redundancy.

In this paper, we are going to look specifically at the hardware redundancy of a fault-tolerant system.

A fault-tolerant system can be viewed as consisting of a set of subsystems, such as memories, processors and buses etc. Each of these subsystems consists of identical modules which can be considered to be active if they are participating in the computation and input/output activities. Alternatively, they can be considered to be spare if they are not currently participating but are to be switched in should one of the active modules fail. It is assumed that every subsystem must survive in order for the system to survive. The reliability of the total system is therefore the product of the individual reliabilities of each subsystem. This is shown in Fig. 1.

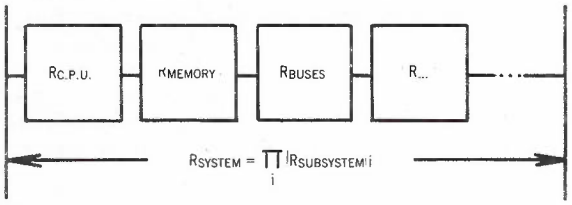


Fig. 1 - Reliability evaluation of a fault-tolerant system

Fault-tolerance is built into each of these subsystems so as to increase the reliability of the whole system but at the same time it is important not to over-protect a particular subsystem. Since this means that it is important to assess the reliability of a particular subsystem, we will model fault tolerance at the subsystem level.

If we examine specific applications of fault-tolerance to systems that have been built, we see that they possess nearly the same attributes, namely N active modules, S spare modules and the possibility of degradation when all the spare modules have been used. On the failure of an active module, one of the spares is switched in to its place.

For specific application, fault-tolerant systems have been built. The SIFT & FTMP, for control of critically stable aircraft, use TMR (Triple-Modular-Redundancy) with spares, i.e., 3 active modules with S spare modules, to achieve its reliability requirement. The ESS (Refs. 7,9) system developed in the Bell Laboratory for telecommunication purposes uses a duplex system, 1 active and 1 spare module, to achieve fault-tolerance. Other systems like the Test and Repair Processor (TARP) in the STAR computer make use of TMR with spares. The Pluribus uses N actives modules with the possibility of degradation.

Individual models have already been developed to evaluate the performance of each of these systems. These models, however, are limited to the system for which they are developed and are not very useful to the designer of a new system at the design stage when he may require answers to the following questions:

1. How many active modules should be used?

2. How many spare modules should be available?

3. How many components should be employed in a module?

4. What should be the quality of fault-detection and fault-recovery?

5. What is the maximum allowable time for reconfiguration?

and several other related questions. The objective of course is to satisfy a given performance specification with the minimum cost.

Some more general models such as those given in Ref. 1 have also been developed, but these are limited in their representation of a fault-tolerant system. Since some of these arise in the aerospace industry, they are confined to closed systems or systems without repair during a given mission time with periodic renewal. Also all previous models do not adequately represent the influence of reconfiguration time. Usually, the assumption is that reconfiguration will be instantaneous.

In the classes of systems of interest to us, e.g. industrial controllers or telecommunication applications, we have systems which have a mixture of repair on demand and periodic maintenance. In addition, since reconfiguration will often involve an interruption to the availability of the service, the influence of the time required for reconfiguration on availability of the service has to be assessed.

The philosophy of a fault-tolerant design is that when a fault occurs in an active module, a spare is switched in to take its place. This process will normally involve an error detection algorithm and a recovery algorithm. Usually, the error detection algorithm will be designed to capture designated classes of faults. Unfortunately, it is possible for undesignated faults not to be detected and hence cause a situation where the system will not recover. For example, if we use a single bit parity check for memory, this method will detect any single bit fault in the memory. If a double bit fault arises, the system will not be able to detect it, and hence will be unable to recover from such faults. Furthermore, it may be possible for certain class of faults to exist which may seriously disrupt the recovery algorithm.

To model the effect that there are certain classes of faults that the system will not be able to recover from, we use coverage factors Ca & Cs where Ca is the conditional probability that the system will recover given that a fault occurs in one of the active modules and Cs is the conditional probability that the system will detect and discriminate the fault occurring in the spare module. Modules which are participating in the computation and voting are Active Modules. Should one of these modules

fail, we have spares to take over the failed active module. There can be any number of spares theoretically ranging from 0 to some number S. When the spares have run out, a TMR system for example may degrade to a duplex or simplex system and a duplex system may degrade to a simplex system. There is therefore a possibility of degradation with or without loss of computation power or functions.

3.1 Active Module and Coverage

As explained above, associated with active module failure there will be a coverage factor Ca to account for the fact that there is a certain class of faults from which this type of module may not be able to recover and these will lead to a system failure. This is described by means of a Markov state transition diagram shown in Fig. 2.

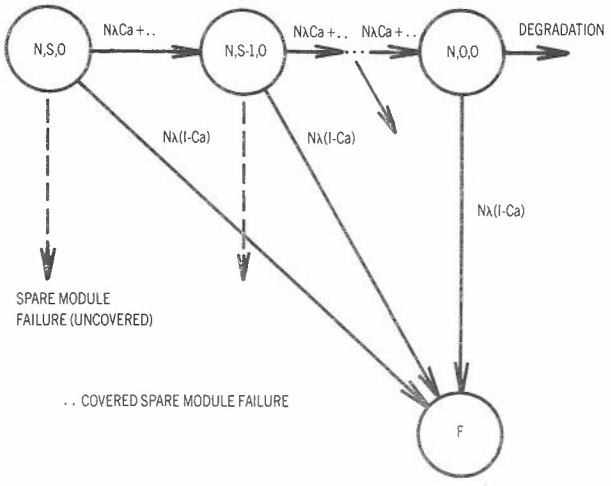


Fig. 2 - Markov state transition diagram for failure of active module only (effect of covered and uncovered failure)

We note from Fig. 2 that any uncovered active module failure will cause a transition to the failure state. If the active module failure is covered, the spare will be switched in to take the place of the failed module. The number of active modules thus remains the same until all the spares are used up. The system then degrades if degradation is possible.

3.1.1 Spare Module and Coverage. Let us next consider what happens when an active module fails. Say we have 3 spare modules X1,X2 and X3. For the first covered active module failure, the system will switch in X1 first. If a subsequent failure occurs in the active module and it is covered, X2 will be switched in and finally X3 will be switched in for a third failure, if the failure is covered. Any subsequent covered failure will lead to a degradation or failure of the system. We see that there is a certain sequence in which the spares will be switched in. Like the active modules, there will be a certain class of faults in the spares that the system is unable to recover from. We earlier introduced the term Cs, the coverage factor for failure of spares, to account for this class of faults. Unlike an active module, an uncovered failure in the spare will not lead to a system failure immediately. This occurs only when the system attempts to switch the spare in. Note also that once an uncovered fault occurs in a spare, the sequence of spare modules that is available is terminated at that spare. Here, we neglect the probability of a spare module developing an uncovered fault and subsequently developing another fault which is covered. The probability of this occurrence is very small, and inclusion of this transition would greatly complicate the model and only improve the accuracy of the model marginally.

Let us consider the case of X1 developing an uncovered fault before it is switched in. The number of spare modules left reduces to 0, since any attempt to switch in X1 in place of a covered failure in an active module results in a system failure. Similarly, if X2 developed an uncovered fault, X1 will be the only spare left. Even though X3 has not developed a fault, the system has lost its capability to use X3 due to the uncovered fault in X2. There will normally be some test program to test that the spare is working correctly before it is switched in. If this test fails to detect that the spare has developed a fault and switches the spare in, it now becomes an active module and its coverage will be that of an active module. Cs will therefore be at least as large as Ca. But it will be larger if the spare testing procedure detects additional faults. A Markov state transition diagram showing transitions due to uncovered and covered spares developing faults is shown in Fig. 3.

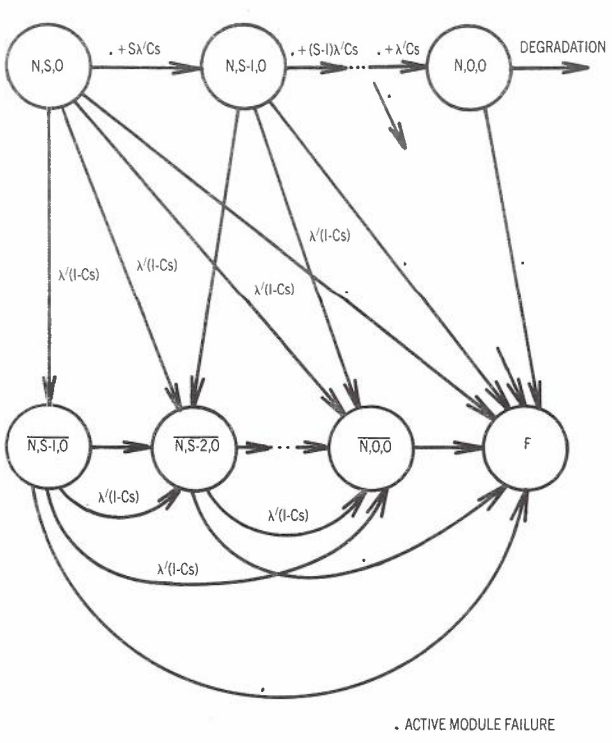


Fig. 3 - Markov state transition diagram for failure of spare module only (effect of covered and uncovered failure)

If any one of the spare modules develops a fault and it is covered, the number of spare modules in the system is reduced by 1. This is denoted by the unbarred transitions. An uncovered fault in a spare module will cause a transition from the state where the fault occurs to a barred state with a lesser number of spare modules. Note also that a fault in a spare (covered or uncovered) will not cause a transition to the failure state. As eventually, in the barred states, the uncovered spare module will be switched in and cause system failure, it is not possible to have degradation from the barred states.

3.1.2 Degradation Mode. After all the spare modules have been used, further failure in an active module will lead to a degradation of the system (considering only the unbarred state transitions). In this mode, the system normally has to give up processing some of its less critical tasks due to a partial loss of computational power. Also the coverage for active module failure in this mode will be different from that before degradation.

For example, consider a TMR system degrading to a duplex configuration. In the TMR configuration, error detection and fault discrimination is done by majority voting. This detection and discrimination method is simple to implement and is effective with a coverage value very close to 1. However, in the duplex configuration, though error detection can still be done by comparing the output of the 2 modules, some other method, probably a diagnostic program, is needed to identify the failed module. This method of discrimination is less effective and will have a lower coverage than that before degradation.

In general, the coverage will be different with every degradation, from TMR to Duplex, Duplex to Simplex, etc. We will therefore need a coverage vector in the degradation mode, Cd(1), Cd(2), ... Cd(D), where D is the maximum number of degradations allowed.

In the degradation mode, a 5-MR system may degrade to a TMR system and finally to a Simplex system. This sequence of degradation depends on the architecture of the system and thus will have to be included as one of the parameters of a fault-tolerant system. We use the parameter Y(I) to denote the sequence of degradation allowed. In the example given above, Y(1) is 3 (3 modules left after the first degradation) and Y(2) is 1. The maximum number of degradations, D, in this case is 2.

Like the active module case, any uncovered fault while operating in the degradation mode will cause a transition to a failure state. The system degrades with each covered module fault until the maximum number of degradations allowed is reached. In the state (Y(D),0,D), the system just has enough modules to carry out all of the critical tasks in an acceptable manner, and any further module failure will result in the system being unable to carry out its tasks in an acceptable manner. There is a possibility of safe-shutdown if this further module failure is covered. This is illustrated in Fig. 4.

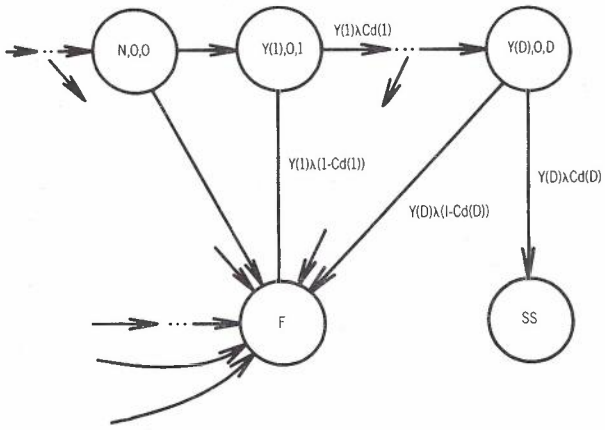


Fig. 4 - Markov transition diagram for fault-tolerant system in the degradation mode

4. GENERAL MARKOV MODEL FOR A CLOSED FAULT-TOLERANT SYSTEM

Having considered the effect of active module failure, spare module failure, where the failure can be either covered or uncovered, and the behaviour of the system in the degraded mode, we can now formulate a general model for a fault-tolerant system. To start with, let us consider a closed fault-tolerant system.

The following assumptions are made in the analysis:

1. The arrival of each module failure is independent and has a constant rate λ for an active module and λ' for a spare module.

2. All spares have the same failure rate. They are either all powered up, i.e., hot standby where $\lambda'=\lambda$, the failure rate of an active module, or unpowered, i.e., cold standby with a failure rate $\lambda'<<\lambda$.

3. Transient faults are neglected in this section. It is assumed that a retry or reinitialization is always successful following a transient fault and has no effect on the performance of the system. It is intended to include transient faults at a later stage of the work.

4. In this section we assume that reconfiguration time is negligible. We later relax this assumption in Section 6.

The state transition diagram for such a closed system is shown in Fig. 5.

From the Markov state transition diagram, we can formulate the state transition matrix A, where

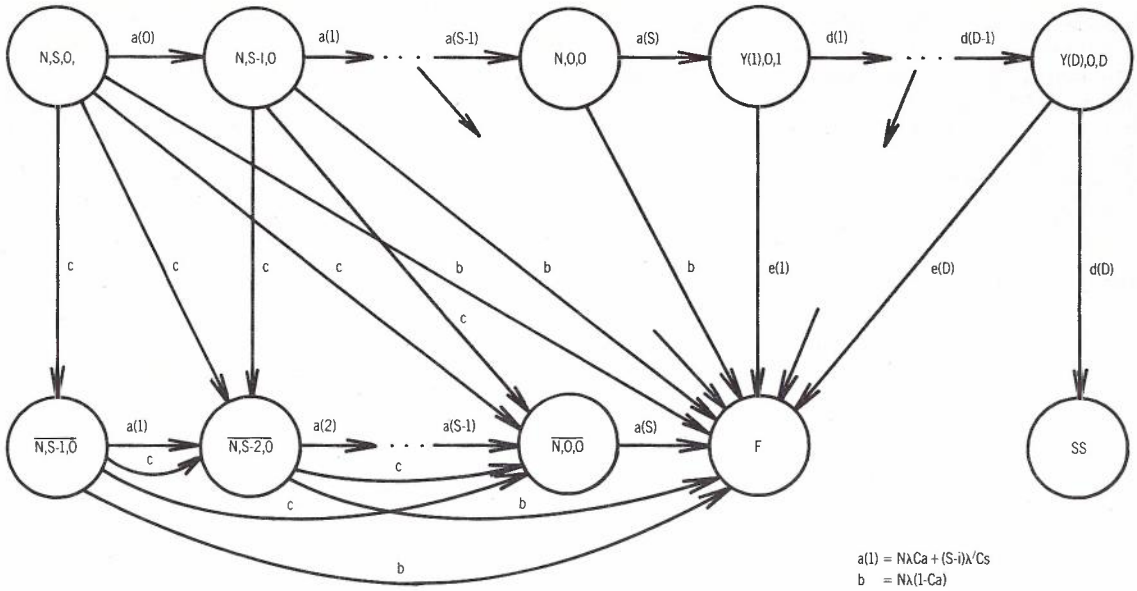


Fig. 5 - Markov model for a closed fault-tolerant system

$$[\dot{P}] = [A]\,[P] \qquad (1)$$

The solution to equation (1) is simply

$$[P] = [EXP(AT)][P(0)] \qquad (2)$$

The probability that the system is in a particular state at time T can therefore be obtained if the probability vector P(0) and EXP(AT) are known.

By Cayley Hamilton's theorem, it can be shown that for a N X N transition matrix [A], EXP(AT) can be reduced to a polynomial of degree N-1, i.e.,

$$EXP(AT) = C_0[I] + C_1[AT] + \ldots + C_{N-1}[AT]^{N-1} \qquad (3)$$

If the transition matrix [A] has N distinct eigenvalues $C_0, C_1, \ldots, C_{N-1}$ can be evaluated from the N equations obtained from

$$EXP(\lambda_i T) = C_0[I] + C_1[\lambda_i T] + \ldots + C_{N-1}[\lambda_i T]^{N-1} \qquad (4)$$

Alternatively, since the eigenvalues are distinct, and EXP(AT) can be expressed as a polynomial, we can evaluate EXP(AT) using the Lagrange Interpolation Formula. The expansion for EXP(AT) in this case is given by

$$EXP(AT) = \sum_{i=1}^{N} e^{\lambda_i T} L_i(A) \qquad (5)$$

where

$$L_i(A) = \prod_{i=1}^{N} \frac{(A) - \lambda_k(I)}{\lambda_i - \lambda_k}$$

λ : Vector of N distinct eigenvalues
I : Identity Matrix

If the eigenvalues are not distinct, i.e.,

$$\lambda_j = \lambda_{j+1} = \ldots \lambda_{j+m}$$

so that λ_j is an eigenvalue of multiplicity (m+1), $C_0, C_1, \ldots C_{N-1}$ can be evaluated from

$$F(\lambda_i) = EXP(\lambda_i T) = C_0(I) + C_1(\lambda_i T) + \ldots + C_{N-1}(\lambda_i T)^{N-1} \qquad (6)$$

with m missing equations. These m equations are given by

$$\begin{aligned} (dF(\lambda)/d\lambda)_{\lambda=\lambda_j} &= C_1 + \ldots (N-1)C_{N-1}(\lambda_j T)^{N-2} \\ &\vdots \\ (d^m F(\lambda)/d\lambda^m)_{\lambda=\lambda_j} &= C_m + (N-m)C_{N-1}(\lambda_j T)^{N-m+1} \end{aligned} \qquad (7)$$

EXP(AT) can then be evaluated from the polynomial as given in equation (3). For closed systems discussed in this section, since the eigenvalues are distinct, we can use equation (5) to obtain EXP(AT).

However, we will find later when using the method of stages that we encounter systems with multiple non-distinct eigenvalues. In this case, we would have to use equations (3),(6) and (7). This can be computationally inconvenient. An alternative approach would be to use a power series expansion method for EXP(AT). This power series expansion is given in equation (8) below.

$$EXP(AT) = 1 + [AT] + [AT]^2/2! + [AT]^3/3! + \ldots \quad (8)$$

By a suitable normalization of [A], we can make the series converge within 10 terms to give us the required accuracy. We then evaluate the state probabilities Pi at time T,2T, ... NT using the following expressions.

$$P(T) = EXP(AT)P(0) \quad (9)$$

$$P(2T) = EXP(2AT)P(0) = EXP(AT)EXP(AT)P(0)$$

$$= EXP(AT)P(T) \quad (10)$$

Likewise,

$$P(3T) = EXP(AT)P(2T) \quad (11)$$

and

$$P(NT) = EXP(AT)P((N-1)T) \quad (12)$$

These equations give the individual state probabilities $P_1, P_2 \ldots P_F$ where $P_F(T)$ is the cumulative probability distribution of system failure in time [0,T]. From this probability distribution, we can obtain several indices which would be useful in characterizing the failure behaviour of a fault-tolerant system. The indices which we have considered to be of interest are as follows:

1. Interval Availability.
2. Interval Reliability.
3. Mean Time to First Failure.
4. Unavailability Due to System Failure.
5. Unavailability Due to Reconfiguration.

The Interval Availability is the expected fraction of time the system is operational during a service interval T and this is given by

$$IA = \frac{\int_0^T \sum_{i \neq F} P_i(t)\, dt}{T}$$

The Interval Reliability is the probability that the system does not fail in time t during the service interval and this is given by

$$IR(t) = \sum_{i \neq F} P_i(t) = 1 - P_F(t)$$

The Mean Time to First Failure is given by

$$MTFF = \frac{\int_0^T \sum_{i \neq F} P_i(t)\, dt}{P_F(T)}$$

for a periodically maintained system with perfect maintenance,

and by

$$MTFF = \int_0^\infty \sum_{i \neq F} P_i(t)\, dt$$

for a system maintained on failure or where maintenance is not possible once put into service. We will see in Section 6 when we consider reconfiguration time that we have to distinguish between the time the system spent in the system failure state and the time the system spent in the reconfiguration states, where the system is unavailable in both cases. We introduce the term UDSF, the Unavailability Due to System Failure and this is given by

$$UDSF = \frac{\int_0^T \sum_{i \varepsilon Rpr} P_i(t)\, dt}{T}$$

and UDTR, the Unavailability Due to Reconfiguration, given by

$$UDTR = \frac{\int_0^T \sum_{i \varepsilon Rcvy} P_i(t)\, dt}{T}$$

The essential difference between IA and IR is that IA is a fraction of time averaged over the service interval T. As long as this fraction is above the value specified, the system is still operating within its specification. It does not matter how many times the system has failed during the service interval. On the other hand, when the system is specified in terms of IR, continued operation during the service interval is to be achieved without the benefits of repair. The system is considered to fall outside the specifications if a failure arises during the service interval. The choice of performance index depends on the application. The SIFT system which is used to control a dynamically unstable aircraft for example would be specified

in terms of IR. IA would be a more appropriate performance measure for the ESS switching system.

5. A PERIODICALLY MAINTAINED FAULT-TOLERANT SYSTEM

In computers used for control of dynamically unstable aircraft for example, where the aircraft flies 100% of the time under computer control, maintenance is not possible during the flight. Maintenance becomes possible only after the plane has landed.

The renewal phase involves locating faults in the modules that have arisen during the operation and removing them. It may not be possible to locate all the faults during this maintenance interval due to the length of time given for maintenance, and to human factors like the expertise of the maintenance crew. This gives rise to imperfect maintenance. Such systems with imperfect maintenance are no longer Markovian and have to be modelled using a Semi-Markov Process. These systems become computationally difficult to manage. Therefore for reasons of mathematical tractability, we will make the following assumptions for the model:

1. The time taken for maintenance can be neglected.

2. Maintenance is perfect, i.e., after the maintenance, all modules are completely renewed.

Such systems can be modelled using a Markov model.

It is reasonable in this type of system to neglect the time taken for renewal. In the aircraft case, after the plane has landed, it does not matter whether the computer subsequently fails during the maintenance period since it is no longer interacting with its environment. As long as all modules are brought back to perfect working order before the next flight, the time taken for maintenance between flights is immaterial, provided maintenance can be completed without disrupting schedules.

The second assumption is not unreasonable if the system is thoroughly checked before the next flight.

The state transition diagram for the mission period is the same as that for a closed system given in the last section. Performance measures for this type of system are Interval Reliability (IR) and Mean Time to First Failure (MTFF). Interval Reliability would normally be specified not to fall below a certain acceptable value k, during the operation period T.

Interval Reliability is given by:

$$IR(t) = \sum_{i \neq F} P_i(t) \geq K$$

We determine the Mean Time to First Failure as follows.

For a periodically maintained system with maintenance period T, the probability that the system is still working at time t, Sn(t) where $nT<t<(n+1)T$, is given by

$$Sn(t) = R(T)^n R(t-nT)$$

where $R(t) = IR(t)$

By definition,

$$MTFF = \int_0^\infty Sn(t)\,dt = \int_0^T R(t)\,dt + \int_T^{2T} R(T)R(t-nT)dt + \ldots + \int_{nT}^{(n+1)T} R(T)^n R(t-nT)dt + \ldots$$

But

$$\int_{nT}^{(n+1)T} R(t-nT)dt = \int_0^T R(t)dt$$

Therefore

$$MTFF = \int_0^T R(t)dt + R(T)\int_0^T R(t)dt + \ldots + R(T)^n \int_0^T R(t)dt + \ldots$$

This is a geometric progression with common ratio $R(T)<1$. Therefore

$$MTFF = \frac{\int_0^T R(t)dt\,[1-R(T)^\infty]}{1 - R(T)} = \frac{\int_0^T R(t)dt}{F(T)}$$

where $F(T) = 1 - R(T)$ is the cumulative probability of system failure in time $[0,T]$.

5.1 A System Maintained at Periodic Intervals or upon Failure

In the case of remote telecommunication exchanges, the adoption of a maintenance policy which was purely maintenance on demand would be extremely wasteful as a full maintenance crew would have to be continually present within reach of the site. Such a crew would only be

active when the system fails. On the other hand, a maintenance policy which relied on periodic maintenance of the system would allow the crew to be better utilized, provided that the system could be assured to be operating at the appropriate level of reliability between maintenance intervals through the use of fault-tolerant techniques. In this sense, it would be similar to the periodically maintained system discussed in the last section. However, in the case of remote exchanges, it will also be necessary to summon a crew to carry out repairs on the rare occasions that the system fails between maintenance periods. Hence, for remote exchanges, it is necessary to represent a maintenance policy that incorporates both periodic maintenance and maintenance on demand.

When the system crashes or fails due to exhaustion of spares, manual repair will be necessary to bring the system back to full working configuration. The repair period will normally have 2 phases:-

1. A passive phase and

2. An active phase.

On the occurrence of a failure, the repair crew will have to be summoned and transported to the site of the system. During this time, though the repair period has begun, no actual repair is carried out on the system and so there is no possibility of the system being repaired. This is referred to as the passive phase.

The time from the instant actual repair is carried out on the system until repair is completed is termed the active phase. From field data, we find that the repair for a particular system will normally take on a mean with a certain variance. The variance of the distribution will depend on the experience of the maintenance crew. The more experienced the crew, the smaller will be the variance. This is because an inexperienced crew will not know where to start looking for the fault. Sometimes, if they are lucky, they will stumble on the fault quickly and hence take a short time to repair the system. At other times they might not be so fortunate, taking a long time to repair the system. The variance of the distribution will therefore be large for such a crew. On the other hand, an experienced crew will know from the nature of the fault where to look for it and will carry out the repair as quickly and efficiently as possible. The variance of the repair time of such a crew will be predominantly due to the variations in the nature of the faults. Thus experience tends to minimize both the mean and variance of the active phase of the repair time. The distribution therefore takes on the following shape shown in Fig. 6.

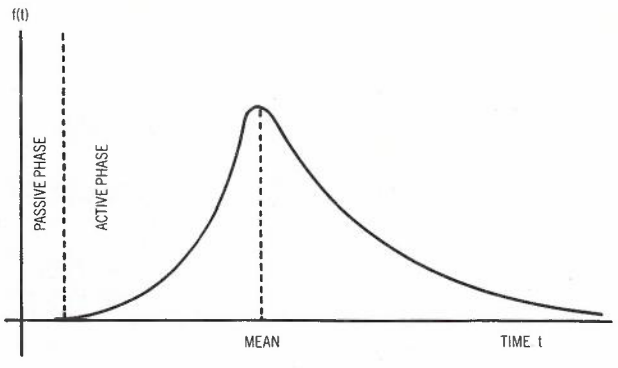


Fig. 6 - Typical p.d.f. for repair time

A good measure for the p.d.f. for the time taken to repair the system will be the coefficient of variation which is defined as follows:

$$\text{coefficient of variation} = \frac{\text{s.d.}}{\text{mean}}$$

As explained above, the value of this measure will depend on the experience of the crew.

The distribution of such a repair time is clearly not exponential. This creates a problem since our basic model relies for its ease of solution on the exponential distribution of failure times.

6. MODELLING RECONFIGURATION TIME

We find a similar problem arises when we consider reconfiguration. The solution of systems with such a non-exponential distribution is discussed in Section 7. All previous models described in the literature assume the time for reconfiguration is negligible. This is not very surprising as most of the models arose from aerospace applications where as long as the system reconfigures fast enough not to cause a system failure, the time spent in this state is immaterial.

In some computer-controlled complex industrial processes, however, frequent interruption in the normal technological process may lead to an irreparable disruption of the process. Telecommunication systems, for example, are normally specified in terms of disruption to services. No distinction is made between the possible causes of such disruption which could be due to system failure or the system's unavailability due to reconfiguration. Both of these contribute to the disruption of user services. In these systems, the influence of reconfiguration time on the availability of the system to the user has to be taken into account.

Before we can proceed to model reconfiguration time, we have first to understand what happens during reconfiguration. Does a covered spare module failure lead to reconfiguration? Is the reconfiguration time the same for all covered active module failures? Basically, all faults detected should lead to reconfiguration in a fault-tolerant system. The reconfiguration may be as simple as updating a

table so that the processor will know which modules have already failed or progressively more complicated to the point where for example, it could even involve having to load the memory from back-up tapes. Thus, depending on the nature of the fault, the time required for each reconfiguration will in general be variable ranging from a few milliseconds to seconds or even minutes.

Take the ESS (Refs. 7,9) processor for telecommunication applications for example. If the main store is corrupted due to a fault in the system, it will try to switch to the other processor using the second processor main store (program data in the main stores are duplicated) hoping that this alternate main store is not corrupted. If the program data is also corrupted in this main store, then the final alternative is to reload memory from the tape unit. Each of these operations will take a certain time and depending on when recovery is successful, the reconfiguration time will vary with the type of fault and the extent of error propagation in the system when recovery is commenced. In general, this results in a distribution which is non-exponential.

Like the p.d.f. for manual repair, the p.d.f. for reconfiguration time can be described using a distribution with the coefficient of variation as a measure to characterise a particular type of module's reconfiguration.

7. METHOD OF STAGES

Basically, we can see that the problem involved is to incorporate the non-exponential distributions that arise in repair-reconfiguration into a model where an easy solution exists because of the exponential distribution assumption.

To exploit the mathematical tractability of the Markov model, we use the method of stages to approximate the non-exponential distribution.

This method uses a number of exponential distributions to approximate a non-exponential distribution. To illustrate this, let us assume in the simplest possible example that the repair or reconfiguration is adequately represented by a distribution that is obtained by two identical exponential repair/reconfiguration stages, each with a transition rate.

Consider a 2 stage model with transition rate μ as shown in Fig. 7.

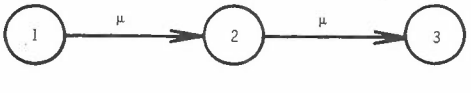


Fig. 7 - A two stage model of repair/reconfiguration time

This may be thought of as a repair or reconfiguration phase where state 1 represents the start of the repair/reconfiguration phase, state 2 is an intermediate stage and state 3 represents the end of the repair/reconfiguration phase.

The Laplace transform of the p.d.f. of the repair/reconfiguration time with transition rate μ in this case is given by

$$f(s) = \frac{\mu}{s+\mu} \cdot \frac{\mu}{s+\mu}$$

Taking the inverse of f(s) gives

$$f(t) = \mu^2 t e^{-\mu t}$$

where f(t) is the p.d.f. of the repair/reconfiguration time.

This is clearly not an exponential distribution although it is approximated by 2 exponential stages. The mean of the resulting distribution is

$$E(t) = 2/\mu$$

with variance

$$\sigma(t) = 2/\mu^2$$

The coefficient of variation is therefore $=1/\sqrt{2}$ which is less than 1 as in the exponential case.

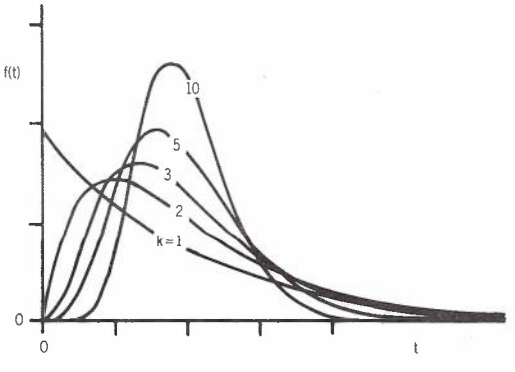


Fig. 8 - p.d.f. for repair/reconfiguration time with 1,2,3,5,10 stages

This idea can be extended to any number of stages. Fig. 8 shows the distributions that are obtained for k=1,2,3,5,10. The distribution corresponding to a particular k is known as the kth order Erlangian distribution. We notice that as the number of exponentials combined to form a distribution increases, it tends to become more "peaky". This method is therefore useful for representing distributions with low dispersion. The lower the dispersion, the larger the number of exponential stages that will have to be used. A measure of dispersion is the coefficient of variation. It can be shown that for k stages, the coefficient

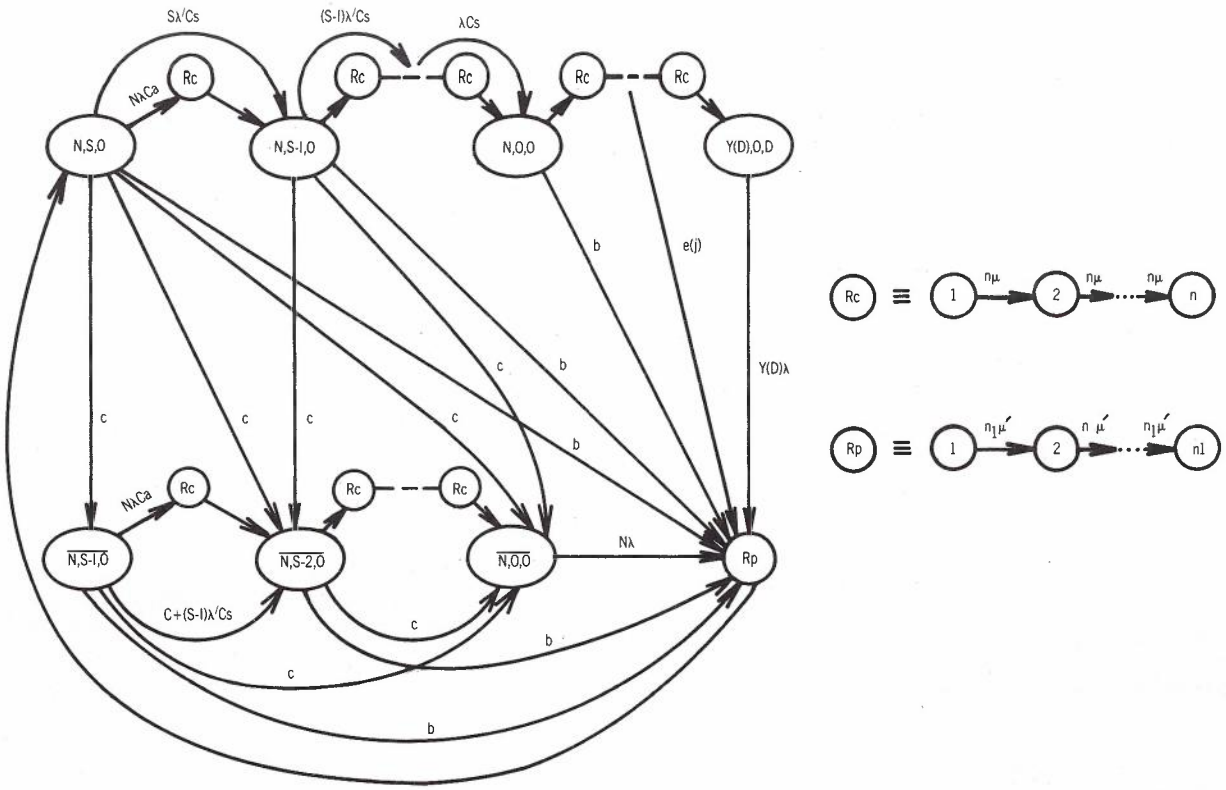


Fig. 9 - Markov transition diagram with repair and recovery states

of variation is $1/\sqrt{k}$ with mean k/μ. Thus by setting up an appropriate number of stages, we can model the repair and reconfiguration phase more accurately.

For a system which is repaired when the system crashes, and if we assume perfect repair, we can model the system by adding in the repair phase as illustrated in Fig. 9.

We neglect the reconfiguration time for covered spare module failures since the system probably has only to update a table without having to initialize and switch modules. We therefore consider only the reconfiguration time due to a covered active module failure and the transition diagram for this is also shown in Fig. 9.

If we assume that either repair or reconfiguration can be represented by a two stage model, the transition matrix for a simple system consisting of 1 active module and 1 spare module is as follows:

$$\left|\begin{array}{cccccc|l} -2\lambda & & & & & \mu & \\ 2\lambda & -\mu' & & & & & \\ & \mu' & -\mu' & & & & \\ & & \mu' & -\lambda & & & \\ & & & \lambda & -\mu & & \\ & & & & \mu & -\mu & \lambda=\lambda' \\ & & & & & & Ca=Cs=1 \end{array}\right|$$

We note that in such a system the eigenvalues are not distinct. We therefore cannot use the Lagrangian Interpolation formula given in equation (5), and we have to use the alternative method for calculating EXP(AT).

For both repair and reconfiguration, having obtained the appropriate transition diagram, we can obtain the transition matrix and solve for the state probabilities using the alternative method of evaluating EXP(AT).

8. DESCRIPTION OF PROGRAM

A computer programme has been developed to compute the state probabilities for a system with repair and recovery states. The program is written in FORTRAN and runs on a VAX 11/780 computer. It is an interactive program which enables the user to enter in the various parameters for a Fault-Tolerant system from the terminal. An example of a session on the terminal, showing typical parameters entered in response to prompts by the program, is shown in Fig. 10. The program proceeds to set up the state transition matrix for the system specified. This could be a closed system, or a system with periodic repair and/or repair on demand, with representation of reconfiguration time, either zero or finite and variable. When the program executes, it evaluates EXP(AT) by a power series expansion method after suitably normalizing [A], to work out the state

```
  RUN RPSYS
NUMBER OF ACTIVE MODULES= 2
NUMBER OF SPARE MODULES=  3
NUMBER OF DEGRADATIONS ALLOWED= 1
SEQUENCE OF ALLOWED DEGRADATION,
Y(1)-Y(D) : 1
FAILURE RATE OF ONE ACTIVE MODULE= 1E-4
FAILURE RATE OF ONE SPARE MODULE= 1E-6
RECONFIGURATION RATE= 60
COVERAGE OF ACTIVE MODULE= .95
COVERAGE OF SPARE MODULE= .95
COVERAGE VECTOR IN DEGRADATION MODE IS
[CD(1),CD(2) ... CD(D)]= .95
NUMBER OF RECOVERY STAGES= 2
REPAIR RATE= 1
NUMBER OF REPAIR STAGES= 2
PLEASE ENTER TIME INTERVAL IN HOURS: 1
```

Fig. 10 - An example of a session on the terminal showing typical parameters entered in response to prompts by the program

probabilities, P(T), P(2T), ... P(NT). The interval T is chosen by the user. From the time distribution of the state probabilities, the MTFF, Unavailabilities Due to System Failure and Unavailabilities Due to Reconfiguration and other parameters are obtained. The flow chart for the program is shown in Fig. 11.

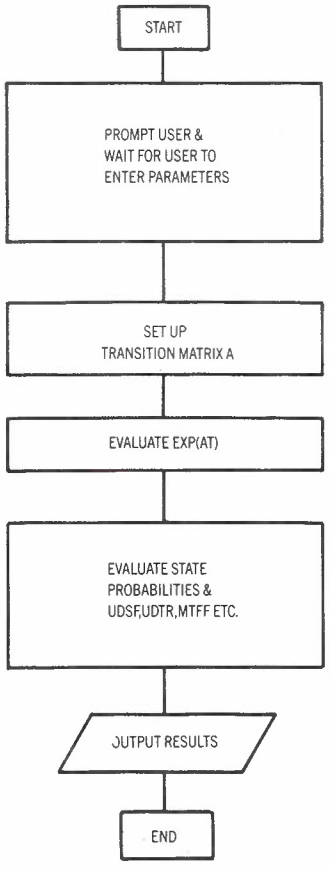


Fig. 11 - Flow chart of the program used to evaluate the reliability of fault-tolerant systems

The program is very flexible and for example, one can study the effect of maintenance interval or study the effect of any one of the parameters on the performance measures of a fault-tolerant system.

9. TYPICAL RESULTS

We have conducted several studies using this program and have included several typical results as shown below. The basic system parameters used for the studies are as follows:

N	λ	Ca	S	λ'	Cs	D	Cd	Y
2	1×10^{-4} /hr	.99	1	1×10^{-4} /hr	.99	1	.99	1

Mean recovery time = 1 min.
Number of recovery stages = 1

Mean repair time = 1 hr.
Number of repair stages = 1

9.1 Study 1: Optimum Maintenance Policy for Duplex System with a Fixed Number of Spares

Consider a duplex architecture with 1 spare, such as might be used for remote telecommunication applications. One might be interested in how often to maintain the system for optimum performance in terms of availability. By varying the maintenance interval keeping all other parameters constant, the unavailability (1 - availability) is obtained and illustrated in Fig. 12. From this graph, it is clear that if we maintain the system at intervals greater than 300 hours, the system performance degrades appreciably.

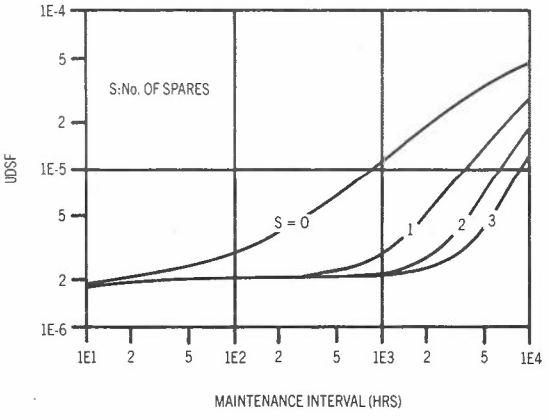


Fig. 12 - Study 1 and 2

9.2 Study 2: Optimum Maintenance Interval as the Number of Spares is Altered

In Fig. 12 we also show the variation of unavailability with maintenance interval when we have 0, 1, 2 and 3 spares. It is clear from the graph that the optimum maintenance interval increases with increased number of spares. The increase, however is not linear and the law of diminishing return applies.

9.3 Optimum Number of Spares with a Fixed Maintenance Interval

For a particular maintenance interval say 3000 hours for example, we may want to know how many spares to put into the system. This is depicted in Fig. 13. We see that there is no reason for us to go beyond the fourth spare as increasing the number beyond that produces very little improvement to the system performance.

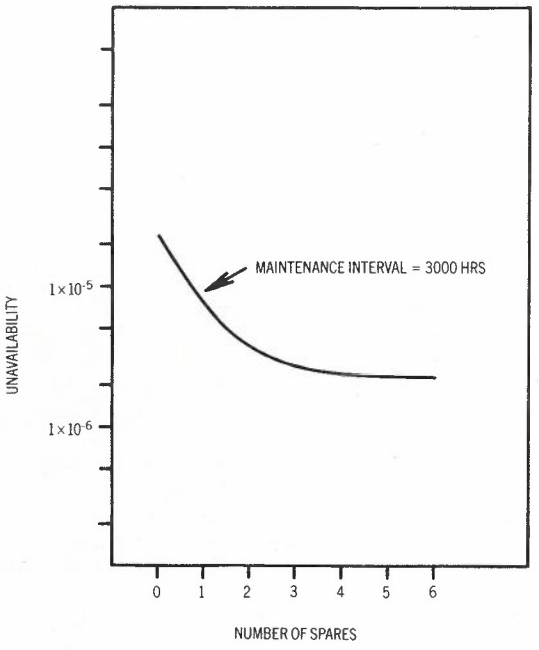


Fig. 13 - Study 3

Using this program, we have also conducted studies on the effect of coverages, reconfiguration time, failure rates and different architectures. The program being interactive enables us to do this easily.

10. SUMMARY AND CONCLUSION

In this paper, we have described a mathematical model for evaluating the performance of fault-tolerant systems. The model is general in scope and allows several classes and types of Fault-Tolerant systems to be modelled including

1. closed systems,

2. systems with periodic repair,

3. systems with periodic repair and repair on demand, and

4. systems with representation of finite reconfiguration time.

We have already developed an interactive program on the VAX 11 and conducted several studies using the program. Some of the results were discussed in Section 9.

11. REFERENCES

1. Y.W. Ng and A. Avizienis, "A Unified Reliability Model for Fault-Tolerant Computers", IEEE Trans on Computers, Vol C-29, No. 11, Nov. 1980, pp. 1002-1011.

2. J.H. Wensley, L. Lamport, J. Goldberg, M.W. Green, L.V. Levitt, P.M. Melliar-Smith, R.E. Shostak, C.B. Weinstock, "SIFT: Design and Analysis of a Fault-Tolerant System for Aircraft Control", Proc IEEE, Vol 66, No. 10, Oct. 1978, pp. 1240-1255.

3. A.L. Hopkins, Jr., T.B. Smith, III, and J.H. Lala, "FTMP - A Highly Reliable Fault-Tolerant Multiprocessor for Aircraft", Proc IEEE, Vol 66, No. 10, Oct. 1978, pp. 1221-1239.

4. B.R. Borgerson and R.F. Frietas, "A Reliability Model for Gracefully Degrading and Standby-Sparing Systems", IEEE Trans on Computers, Vol C-24, No. 5, May 1975, pp. 517-527.

5. R.E. Barlow and F. Proschan, "Mathematical Theory of Reliability", New York: Wiley 1965.

6. T.F. Arnold, "The Concept of Coverage and its effect on the Reliability Model of a Repairable System", IEEE Trans on Reliability, Vol R-22, Mar. 1975, pp. 251-254.

7. G.D. Kraft and W.N. Toy, "Microprogrammed Control and Reliable Design of Small Computers", Prentice Hall 1981.

8. J.M. Cotton, "Processor Reliability Strategies for Telephony", Proc. of National Electronics Conference, Oct. 1978, pp. 83-91.

9. W.N. Toy, "Fault Tolerance Design of Local ESS Processors", Proc IEEE, Vol 66, No. 10, Oct. 1978, pp. 1126-1145.

10. D.R. Cox, "Renewal Theory", New York: John Wiley & Son Inc., 1962.

BIOGRAPHIES

MR Y.W. YAK received his B.E. degree with honours, in electrical engineering from Monash University, Australia, in 1981 where he is currently enrolled as a Ph.D. candidate. His current research interests include fault-tolerant computer systems, reliability modelling of complex systems and microprocessor systems.

DR. T.S. DILLON obtained his B.E. degree with honours, and his Ph.D. from Monash University, Australia, in 1968 and 1974, respectively. He was with the Australian Department of Civil Aviation in 1968. In 1971, he joined the staff of the Department of Electrical Engineering, Monash University, where he currently holds the position of Senior Lecturer. In 1977, he worked for six months on a research project with the Institut fur Electrische Anlagen und Energiewirtschaft at the Technical University, Aachen, West Germany. He was also engaged as a consultant in 1977 by the Swedish State Power Board. In November 1977, he conducted, together with Professor J. Bubenko, an all-Scandinavian symposium on "Economic operation and production planning in hydro-thermal power systems".

His research interests include power-systems, analysis and planning, systems theory, optimal control theory, reliability of complex systems, fault-tolerant computer systems and multi-processor systems.

DR. KEVIN FORWARD is currently a senior lecturer in the Department of Electrical Engineering at Monash University, Melbourne, Australia. He received his B.E. degree with honours in 1961 and his Ph.D. in 1965, both from the University of Western Australia. Until 1967 he was employed as a research scientist by the Australian Government. In 1967, he took a position as a lecturer in electrical engineering at Duntroon, Canberra, Australian Capital Territory. In 1970, he moved to his current position. He spent 1974 at the University of Newcastle-upon-Tyne as a Science Research Council Senior Visiting Fellow in the Department of Electrical Engineering. At Monash, he lectures in computer engineering and his research interests are in the fields of multiprocessor systems, fault-tolerant computing and the design of reliable VLSI systems.

Modulation Techniques For Cable Television Distribution On Optical Fibres

G. NICHOLSON
Telecom Australia Research Laboratories

Modulation techniques are compared for local cable television distribution on optical fibres. The comparison is based on the allowable fibre loss for a given unweighted SNR and transmission bandwidth. A general form of expression is obtained for the SNR of seven modulation techniques.

Pulse width modulation and pulse frequency modulation, both with direct baseband recovery of the analogue video signals, are shown to achieve similar allowable losses as analogue intensity modulation for the same transmission bandwidth. At the expense of increased bandwidth, frequency modulation offers a larger allowable loss, especially with pre-emphasis of the video signal. Pulse code modulation has the best performance, but with a significant increase in circuit complexity compared with the other techniques. The choice of a modulation technique depends on many factors and no one technique is concluded to be a clear candidate.

1. INTRODUCTION

With the progress in optical communications technology and the recent interest within Australia concerning cable television (CTV) services, it is appropriate to review the modulation techniques applicable to CTV distribution on optical fibres. Optical fibre systems for subscriber distribution of CTV and other services are being investigated in several overseas experiments which use a variety of modulation techniques (Ref. 1). Economic considerations are currently against optical fibre distribution compared with alternative metal-based CTV distribution systems, except in particular point-to-point, high capacity applications such as from the CTV head-end to remote distribution points. However, with expected cost reductions and technical improvements in optical components, optical fibre based CTV systems may become a viable proposition.

In this paper a technical comparison of various modulation techniques for a CTV system is developed, in terms of the required transmission bandwidth and the allowable fibre loss for different optical source-detector combinations. A major aim, which is not satisfactorily covered in the literature, is to compare the performance of the modulation techniques (in terms of the same signal-to-noise ratio definition) on a common base for PAL TV signals. The data assumed for the optical components is based on commercially available devices, operating in the 0.85 μm wavelength region. Consideration of the modulation techniques is concentrated on the subscriber distribution application, as it is in this part of a CTV network that costs are largely determined and complex transmission techniques are difficult to justify. The following section outlines modulation techniques which are suitable for an optical CTV system. A transmission system model is introduced in Section 3, including the significant noise sources of the photodetector and receiver pre-amplifier. This is applied in the subsequent Sections 4 and 5 to compare the performance of eight modulation techniques.

2. MODULATION TECHNIQUES

The most direct and simplest approach to transmission of TV channels on an optical fibre is analogue intensity modulation (AIM) of an optical carrier. This approach is analogous to that used in metal-bearer CTV systems and broadcast TV, with frequency division multiplexing (FDM) of the TV channels. However, a significant limitation with AIM for optical transmission is the nonlinearity of light-emitting diode (LED) and laser diode (LD) optical sources. The power-output versus current-input for a LED is approximately linear, over a prescribed range, but the level of total harmonic distortion is usually unacceptable for satisfactory CTV performance, which requires the distortion to be approximately 50 dB or more below the signal level.

The more linear, surface emitting double heterostructure LED's typically have a total harmonic distortion about 45 dB below the signal level for a modulation index of 0.7. The dominant source of nonlinearity is the second harmonic. The modulation index for AIM is defined as the ratio of the peak modulation current to the d.c. bias current of the optical source. The harmonic distortion decreases as the modulation index is reduced, but at the expense of reduced signal power output. The power launched into the fibre can be increased

by using an edge-emitting LED. The combination of a relatively low modulation index to achieve low harmonic distortion and the poor fibre coupling efficiency of LED's, severely limits their application to multichannel CTV distribution (in the context of not wanting to employ line repeaters in the CTV network). Several methods of linearizing LED's have been reported with reductions in the harmonic components of at least 15 dB.

Compared with LED's, LD's enable more power to be coupled into a fibre but some problems are encountered in achieving a linear power-output to current-input characteristic (Refs. 2,3). A stable, low distortion transmission system requires a LD which emits in a single transverse mode. Also, the number of longitudinal modes and the relative power balance between the modes should remain relatively constant over the operating range, including ageing and temperature changes. A typical figure for the second harmonic of a single-transverse, 'single-longitudinal' mode laser is 45 dB below the signal level with a modulation index of 0.7 (Ref. 3). The harmonic levels vary to a degree dependent on the coupled output power. Changes in the distribution of power between longitudinal modes of the LD can give rise to mode partition noise which limits the system signal-to-noise ratio (SNR) (Ref. 2). A LD of narrow spectral width when used with a multimode fibre can result in modal noise, which also limits the SNR (Ref. 2). Low frequency modulation causes LD internal temperature variations, which are a potential cause of mode instability. This effect generally degrades distortion performance for frequencies below 20-50 MHz, depending on the LD type (Ref. 1). Reflections from the fibre end or a nearby joint back into the laser cavity can cause further distortion. Because of the above factors, there is difficulty in achieving low-distortion performance with AIM using a LD source.

The most frequent approach to multichannel transmission by AIM is to choose a FDM arrangement that avoids harmonic distortion from the optical source. The distortion is then determined by the level of intermodulation products, the specifications for which can be phrased in terms of the third-order harmonic distortion of the optical source (Ref. 1). The performance calculations for AIM in Section 4 assume the optical source is sufficiently linear, so that the distortion products are at least 50 dB below the signal level. This assumption is satisfied with a FDM arrangement of the TV channels and selected optical sources. Modal noise and mode partition noise are not considered in the SNR calculation, although they may limit the SNR achieved with particular system configurations.

Pulse width modulation and pulse frequency modulation, both with direct baseband recovery of the analogue video signals (PWM-B and PFM-B), are similar to AIM but overcome the source nonlinearity problem (Ref. 4). Fig. 1 is a block diagram of a general video transmission system on optical fibre. For AIM, direct modulation of the source is employed and there are no modulator and demodulator. The receiver for PWM-B and PFM-B is the same as that for AIM, but the transmitter includes a pulse width or pulse frequency modulator. The transmitted signal, as well as the sampling components at higher frequencies, can be designed to contain a baseband replica of the video input signal. The higher frequency components are removed by the low pass filter and any bandwidth limitation of the fibre to give the desired video output. PWM-B and PFM-B are suited to transmission using a LD (Ref. 4). The suitability of a LED depends on the system design parameters and the modulation bandwidth of the LED. Surface-emitting LED's, with a typical modulation bandwidth (electrical) of about 25 MHz, are of questionable suitability for PWM-B or PFM-B transmission of a 5 MHz-bandwidth video signal. An alternative is a surface-emitting LED with increased modulation bandwidth at the expense of reduced power, or an edge-emitting LED which typically has a wider modulation bandwidth. The PWM-B and PFM-B techniques are considered in Section 4, assuming a LD or a wide modulation-bandwidth LED (electrical bandwidth $\geq$ 80 MHz). A modulation technique not considered is baseband recovery with a delta pulse modulator. Takasaki *et al* (Ref. 4) conclude this technique is poor in performance by comparison with AIM or PFM-B.

Frequency modulation (FM) offers the potential of trading-off increases in transmission bandwidth for increases in the allowable fibre loss (Ref. 5). Furthermore the effect of source nonlinearity is significantly reduced with FM. In Section 4, FM, and FM with pre-emphasis of the transmitted signal in conjunction with de-emphasis of the received signal plus noise, are considered. The frequency dependence of the receiver noise with FM means substantial performance gains can be achieved with emphasis

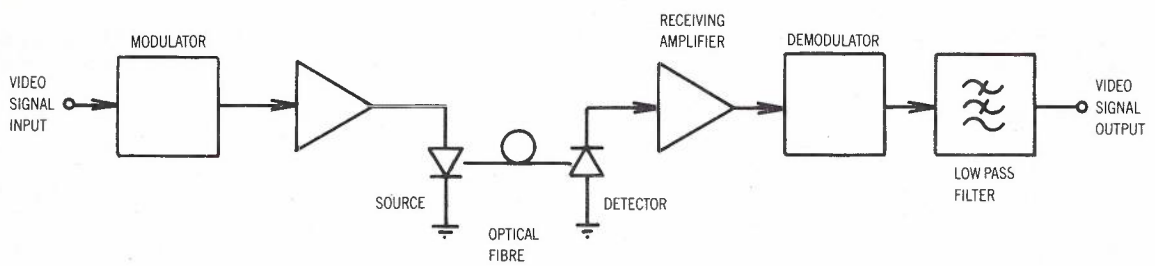


Fig. 1 - Video Transmission System

techniques (Ref. 6). Also considered are pulse interval modulation (PIM) (Ref. 7), pulse position modulation (PPM) (Refs. 5,8) and pulse frequency modulation (PFM) (Ref. 9). In contrast to PWM-B and PFM-B, the above pulse modulation techniques regenerate the pulses at the receiver before demodulation. Other techniques such as pulse width modulation and pulse interval and width modulation (Ref. 10) are not considered, as they do not promise any significant performance improvements compared with the above.

To complement the comparison of modulation techniques, pulse code modulation (PCM) for video transmission is reviewed in Section 4. The advantages of digital transmission include its robustness to transmission deficiencies and its flexibility in terms of compatibility with other services. The disadvantages are the wide transmission bandwidth and significantly increased circuit complexity compared with the other modulation techniques. The bandwidth requirement for PCM precludes the use of a LED, because of the fibre material dispersion in the 0.85 μm wavelength region, to all but one TV channel over relatively short transmission distances. A variety of coding and bandwidth compression techniques are being investigated to reduce the bit rate required for digital transmission of video signals.

3. A SYSTEM MODEL FOR COMPARISON OF THE MODULATION TECHNIQUES

A PAL TV signal is assumed with a baseband bandwidth of 5 MHz. In some instances an FM sound carrier at 5.5 MHz is transmitted with the video signal. The presence of the sound carrier is not included in the calculations, suffice to say that a practical filter rolloff with 5 MHz bandwidth allows adequate reception of the sound. For multichannel TV transmission with AIM, PWM-B, or PFM-B, a FDM arrangement is assumed with a minimum 7 MHz channel spacing (as in Australian broadcast TV). The TV channels are VSB amplitude modulated.

Fig. 2 shows the composite video signal with various signal levels measured by the IEEE scale (Ref. 11). The signal-to-noise ratio (SNR) definition used in this report, for a baseband video signal, is

$$SNR = \left[\frac{\text{current (voltage) between blanking level and white level}}{\text{rms noise current (voltage) in 5 MHz video bandwidth}} \right]^2 = I_{BW}^2 / \langle i_n^2 \rangle , \quad (1)$$

where $\langle i_n^2 \rangle$ denotes the mean-square noise current. The SNR is expressed in dB and the noise is unweighted. For a FDM TV channel with VSB modulation, $\langle i_n^2 \rangle$ is measured in a 6 MHz bandwidth. This SNR definition is taken as an adequate and simple measure of performance, although many different definitions including noise weighting have been advocated. A SNR of 40 dB, based on Canadian CTV specifications (Ref. 12) for the minimum SNR at any subscriber

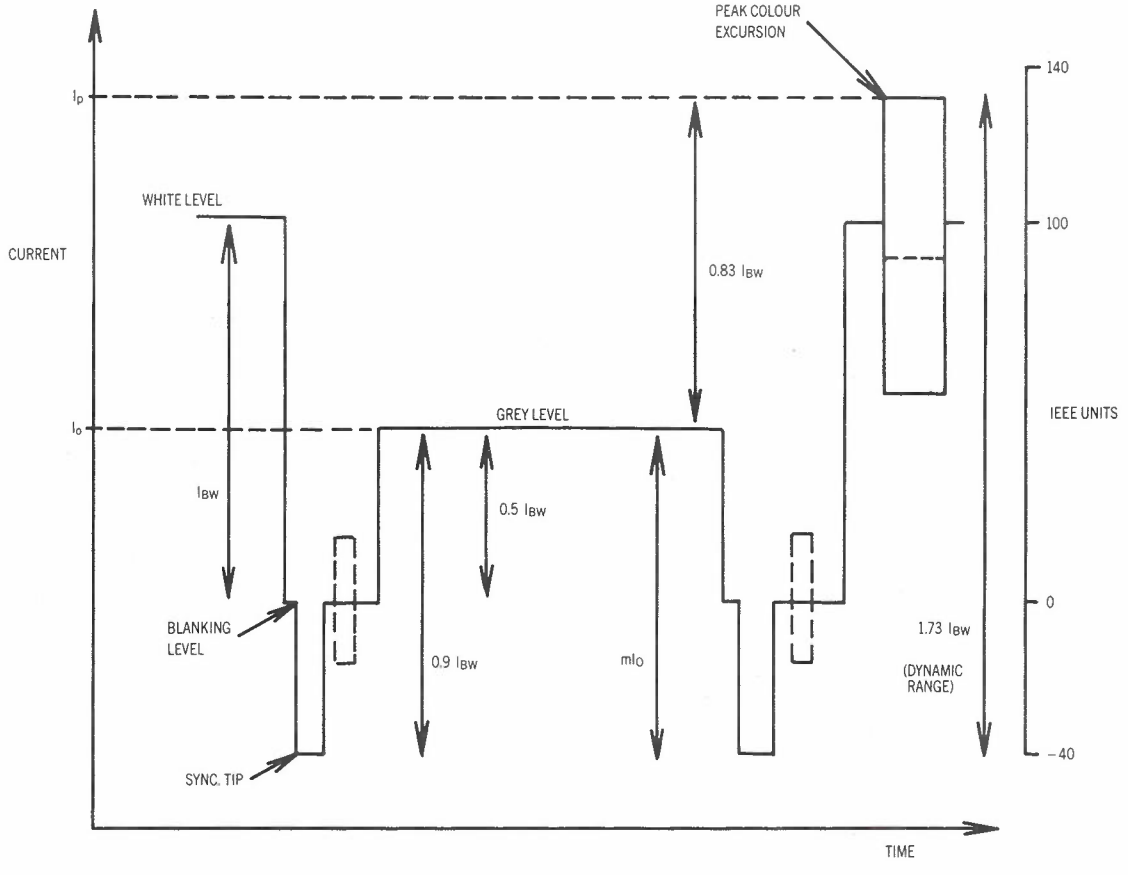


Fig. 2 - Composite Video Signal

terminal, is assumed in the performance calculations. The minimum dynamic range of 1.73 I_{BW}, defined as that between the maximum colour excursion to sync tip levels in Fig. 2, is assumed in the calculations.

The data assumed for optical sources is given in Table 1. The peak power is that coupled into a 50 μm core, 0.2 numerical aperture, graded-index optical fibre. The spectral width of the LD is sufficient so that modal noise should not be a problem. The optical receiver is assumed to consist of a PIN detector or avalanche photodiode (APD), followed by a transimpedance (feedback) amplifier. The transimpedance amplifier is the most common front-end receiver design, because it is capable of wide bandwidths, provides a greater dynamic range than the alternative high-impedance approach and has a noise performance which can approach that of the high-impedance front-end design. A model of the optical receiver with the significant noise sources is given in Fig. 3 (Ref. 5). Data assumed for the PIN detector and APD is given in Table 2.

TABLE 1 - Data Assumed for Optical Sources

Optical Source	Low-Bandwidth LED_1	High-Speed LED_2	Laser Diode
Peak Power Coupled into Fibre in dBm	-11.0	-15.7	+1.8
Spectral Width at Half Intensity in nm	40	45	2.5
Electrical Bandwidth in MHz	25	86	>300
Rise Time of Optical Pulse (10% to 90% points) in ns	14	4	<1

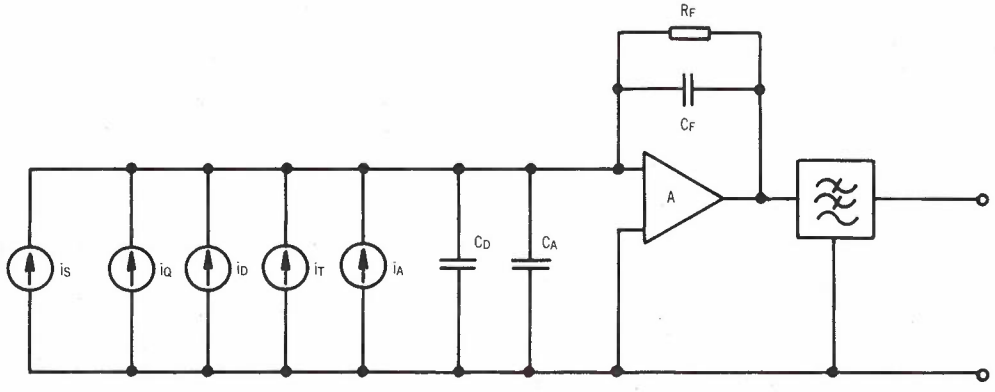


Fig. 3 - Optical Receiver

TABLE 2 - Data Assumed for Optical Detectors

Optical Detector	PIN Detector	APD
Responsivity in A/W	0.6	0.5
Capacitance in pF	2.0	2.0
Dark Current in nA	20	20
Exponent x in Power Law Approximation to Excess Noise Factor	-	0.5
Rise Time (10% to 90% points) in ns	3	0.5

In Fig. 3, i_s is the average detector signal current given by

$$i_s = \frac{\eta q}{h\nu} GP = R_o GP, \qquad (2)$$

where η is the quantum efficiency of the photodiode, q is the electronic charge, hν is the photon energy, R_o is the responsivity, G is the average avalanche gain (G = 1 for the PIN detector) and P is the incident optical power. The significant noise sources of the detector

are the signal-dependent shot noise i_Q and the shot noise i_D of the dark current. The mean-square noise currents in a frequency interval df (single-sided spectra) are respectively,

$$\langle i_Q^2 \rangle = 2q\, I_s G^{1+x}\, df = 2q\, R_o G^{2+x}\, P_o\, df \qquad (3)$$

where $I_s = \langle i_s \rangle$ is the mean avalanche-multiplied signal current corresponding to the average incident power $P_o = \langle P \rangle$ and

$$\langle i_D^2 \rangle = 2q\, I_D\, G^{2+x}\, df, \qquad (4)$$

where I_D is the unmultiplied detector dark current. The random nature of the avalanche gain process gives rise to excess noise, which is approximated by a power law with exponent x.

Consider a model for the noise sources of the transimpedance amplifier, with a field effect transistor (FET) or bipolar junction transistor (BJT) input stage (Ref. 5,13). The amplifier is analysed by referring all noise sources to its input, so that the amplifier noise $\langle i_A^2 \rangle$ can then be added to the detector noise. Assume the amplifier open-loop gain is large so that the input impedance is determined by the input capacitance C_A of the amplifier in parallel with the feedback resistance R_F and capacitance C_F. It can be shown that the amplifier noise current i_A has a mean-square value at frequency f of

$$\langle i_A^2 \rangle = \frac{4kT}{R_F}\, df + a_1 f^2\, df, \qquad (5)$$

with a FET input stage and

$$\langle i_A^2 \rangle = \left(\frac{4kT}{R_F} + a_2\right) df + a_3 f^2\, df \qquad (6)$$

with a BJT input stage. The first term in (5) and (6) is the thermal noise of resistor R_F, where k is Boltzmann's constant and T is the absolute temperature (taken as 298 K or 25°C in the calculations). The constants a_1, a_2 and a_3 depend on the amplifier characteristics, the detector capacitance C_D (Table 2) and the parasitic capacitance C_F (taken as 0.5 pF). The constants used in this paper are $a_1 = 2.5 \times 10^{-39} A^2 s^3$ for a low-noise JFET with g_m = 19 mmhos and C_A = 8.2 pF, and $a_2 = 2.5 \times 10^{-24} A^2 s$ and $a_3 = 3.9 \times 10^{-40} A^2 s^3$ for a silicon BJT with 300 μA d.c. emitter current, β = 38 and f_T = 1.6 GHz. This data is typical for transimpedance amplifier designs, although better performance can be achieved with a MESFET input stage for which typically g_m = 40-50 mmhos and C_A = 1.5 pF. In the performance calculations of Section 4, the amplifier design which gives the lowest $\langle i_A^2 \rangle$ integrated over the receiver bandwidth, is chosen separately for each evaluation of an allowable fibre loss. The receiver baseband bandwidth at which the amplifier noise from the FET input stage exceeds that from the BJT input stage is 60 MHz. To simplify notation in the SNR expressions the amplifier noise is denoted as

$$\langle i_A^2 \rangle = u_1\, df + u_2 f^2\, df, \qquad (7)$$

where the constants u_1 and u_2 are given by (5) or (6) for the respective input stage.

The total mean-square receiver noise in a frequency interval df, which is integrated to determine the SNR, is

$$\begin{aligned} \langle i_n^2 \rangle &= \langle i_Q^2 \rangle + \langle i_D^2 \rangle + \langle i_A^2 \rangle \\ &= (2qR_o G^{2+x} P_o + 2qI_D G^{2+x} + u_1)\, df \\ &\quad + u_2 f^2\, df \end{aligned} \qquad (8)$$

The bandwidth of the transimpedance amplifier in Fig. 3 is determined by the total input capacitance ($C_D + C_A + C_F$), the feedback resistance R_F and the amplifier gain A. The bandwidth can be increased by reducing R_F, but at the expense of increased thermal noise. In practice, the performance of the transimpedance amplifier can be improved by the use of an equalisation stage to achieve the same bandwidth requirement using a larger value for R_F (Ref. 14). In this report, a modest improvement in the thermal noise by a factor of five is assumed. The maximum feedback resistance R_F for an amplifier bandwidth B is given by

$$R_F = \frac{5A}{2\pi B(C_D + C_A + C_F)} \qquad (9)$$

For the example of B = 5 MHz, R_F is taken as 330 kΩ.

4. THE ALLOWABLE FIBRE LOSS AND TRANSMISSION BANDWIDTH

The previous section outlined a system model which is used in this section to compare the performance of the modulation techniques in terms of the allowable fibre loss for a given transmission bandwidth and SNR of the video signal. The allowable fibre loss can be converted to a transmission distance based on the fibre attenuation, provided the CTV system is not operating in a dispersion-limited environment.

The total allowable fibre loss L, excluding a system operating margin but including losses due to splices and connectors, is given by

$$L = \frac{P_T}{P_R}, \qquad (10)$$

where P_T is the peak power coupled into the fibre (from Table 1) and P_R is the minimum peak-power required at the receiver for a specified SNR. In the following subsections for each modulation technique (except PCM), an expression is given for the SNR in terms of P_R. The expression is derived in the general form of

$$SNR = \frac{(\alpha_1 R_o G P_R)^2}{[2q R_o G^{2+x}(\alpha_2 P_R) + 2q I_D G^{2+x} + u_1] B_1 + u_2 B_2^3}, \qquad (11)$$

where α_1, α_2, B_1 and B_2 depend on the modulation technique. The factor α_1 relates I_{BW} to P_R, α_2 determines the average detector shot noise (or the shot noise at the decision instant in the case of the pulse techniques - PIM, PPM, PFM) and B_1, B_2 are equivalent noise bandwidths. The optical power P_R is determined by solving the quadratic equation (11). For a PIN detector G = 1, while for an APD the expression for P_R is minimized with respect to G. For PCM, the minimum optical power at the receiver for an error rate of 10^{-6} is determined based on the work of Personick (Ref. 14), with the same parameters assumed for the optical receiver as for the other modulation techniques.

TABLE 3 - Comparison of Modulation Techniques for Transmission of a Single TV Channel

Modulation Technique	Allowable Loss (dB) for SNR ≥ 40 dB				Transmission Bandwidth B_T (MHz)
	High-Speed LED_2 (P_T =-16 dBm)		LD (P_T=2 dBm)		
	PIN Detector	APD	PIN Detector	APD	
AIM	20*	21*	38	39	5
PWM-B	18	18	38	39	5
PFM-B	18	18	36	37	5
FM	15	22	33	40	50
FM+Emphasis	22	30	39	47	50
PIM	21	28	38	46	50
PPM	20	28	38	45	50
PFM	17	24	35	42	50
PCM	21	30	39	48	50

*Using low-bandwidth LED_1(P_T=-11 dBm), allowable loss is 25 and 26 dB with PIN detector and APD, respectively.

TABLE 4 - Comparison of Modulation Techniques for Multichannel TV Transmission

Modulation Technique	Allowable Loss (dB) for SNR ≥ 40 dB				Transmission Bandwidth B_T (MHz)
	High-Speed LED_2 (P_T=-16 dBm)		LD (P_T=2 dBm)		
	PIN Detector	APD	PIN Detector	APD	
2 TV Channels					
AIM	11	14	28	31	27
PWM-B	-	-	30	31	13
PFM-B	7	8	28	29	13
FM	8	12	25	29	60
PCM	-	-	36	46	100
3 TV Channels					
AIM	7	10	25	28	41
PWM-B	-	-	26	27	20
PFM-B	-	-	24	24	20
FM	-	-	23	26	90
PCM	-	-	35	45	150
6 TV Channels					
AIM	-	-	19	20	83
PFM-B	-	-	15	15	41
FM	-	-	18	21	180

The results for each modulation technique are compared in Tables 3 and 4. The transmission bandwidth, denoted B_T, is the 3 dB electrical bandwidth of the optical transmission channel (including source, fibre and photodetector). Provided the response times of the source and detector are adequate, B_T can be interpreted as the minimum 3 dB electrical bandwidth of the optical fibre required. For the techniques of PIM, PPM, PFM and PCM the optical channel is modelled by a Gaussian pulse response that corresponds to the bandwidth B_T (Ref. 14). For the techniques of AIM, PWM-B, PFM-B and FM the bandwidth B_T of the optical channel can be related to a 'flatness' figure for the received TV channel frequency response. Alternatively, given a flatness specification for the TV channel (e.g. within ±1 dB of the low frequency value at 5 MHz) and again assuming a Gaussian frequency response the minimum B_T can be determined. For simplicity with AIM, PWM-B, PFM-B and FM, the bandwidth B_T in Tables 3 and 4 is taken as the highest frequency in the TV channel allocations.

4.1 Analogue Intensity Modulation

A linear optical source and detector are assumed in the calculations for AIM. Consider transmission of the composite video signal, shown in Fig. 2. The video signal modulates the optical power of the source and produces an average current I_o at the detector, corresponding to the average incident power P_o. If the average video signal is midway between the blanking and white levels, then

$$I_{BW} = \left(\frac{m}{0.9}\right) I_o, \qquad (12)$$

where m is the modulation index for AIM, defined as the ratio of the peak signal current about I_o to the average current I_o at the detector. The definition of m can be equivalently defined in terms of the LED current, but for a LD its threshold current must be taken into account. The peak current

$$I_p = I_o + 0.83\, I_{BW} = R_o G P_R \qquad (13)$$

corresponds to the peak-power P_R, so that

$$I_{BW} = \left(\frac{m}{0.9 + 0.83m}\right) R_o G P_R, \qquad (14)$$

$$P_o = \frac{P_R}{1 + 0.92m} \qquad (15)$$

Integrating the receiver noise (8) over a baseband bandwidth B gives, together with (1), (14) and (15),

$$SNR = \frac{\left(\frac{m R_o G P_R}{0.9 + 0.83m}\right)^2}{\left[\frac{2q R_o G^{2+x} P_R}{1 + 0.92m} + 2q I_D G^{2+x} + u_1\right] B + \frac{u_2 B^3}{3}} \qquad (16)$$

A conservative assumption for transmission of n TV channels, which is reasonable for small n, is that the dynamic range of each TV signal is reduced by the factor n compared with a single TV signal and the same transmit power. Based on (16), the SNR for a TV channel occupying the frequency range f_1 to f_2 in a FDM arrangement of n channels is

$$SNR = \frac{\left[\frac{m R_o G P_R}{n(0.9 + 0.83m)}\right]^2}{\left[\frac{2q R_o G^{2+x} P_R}{1 + 0.92m} + 2q I_D G^{2+x} + u_1\right](f_2 - f_1) + \frac{u_2}{3}(f_2^3 - f_1^3)} \qquad (17)$$

No interference from adjacent TV channels is assumed in (17). Because the amplifier noise term in (17) depends on the frequency cubed, the SNR is lowest in the highest-frequency TV channel. For the results of Table 4, the minimum power P_R and hence the allowable loss L are determined to satisfy the SNR specification in all TV channels.

The modulation parameters chosen for the performance calculations are shown in Table 5. The channel frequency allocations are chosen to avoid harmonic distortion. The results for a single TV channel are with baseband transmission, although a more suitable scheme may be VSB modulation at higher frequencies.

TABLE 5 - Modulation Parameters for AIM

n	Channel Frequencies (MHz)	m
1	0-5	0.75
2	14-20, 21-27	0.75
3	21-27, 28-34, 35-41	0.75
6	42-48, 49-55, ..., 77-83	0.50

4.2 Pulse Width Modulation with Baseband Recovery

An analysis of the SNR performance for PWM-B is given by Takasaki *et al* (Ref. 4). Using the same notation, T_s is defined as the sampling time, T_w as the standard pulse width and a_w as the modulation index. The modulation parameters a_w, T_w and T_s are chosen to avoid any significant interference of the baseband signal from the higher-frequency components of the sampling process. The peak-to-peak amplitude of the desired baseband signal, in terms of the received power, is $2a_w T_w P_R/T_s$ (note, that equation (5) in (Ref. 4) is incorrect by a factor of two). This peak-to-peak amplitude is equated with the dynamic range 1.73 I_{BW} of the video signal. The receiver for PWM-B is the same as that for AIM. The average incident power determining the shot noise is $P_o = T_w P_R/T_s$ if the average video signal corresponds to the centre of the signal's

dynamic range, which is approximately true in practice. Applying these results, for multi-channel TV transmission

$$SNR = \frac{\left[\frac{1.16\, a_w\, T_w R_o G P_R}{n T_s}\right]^2}{\left[\frac{2qT_w R_o G^{2+x} P_R}{T_s} + 2q\, I_D G^{2+x} + u_1\right](f_2 - f_1) + \frac{u_2}{3}(f_2^3 - f_1^3)} \qquad (18)$$

If the optical fibre provides some filtering of the undesired higher-frequency components, but not the baseband signal, then equation (18) is valid with a redefinition of the power P_R. However, the allowable fibre loss remains unchanged.

The modulation parameters chosen for the performance calculations, satisfying distortion at least 50 dB below the signal level, are shown in Table 6. The modulation index a_w is chosen with regard to the optical source response time (Tables 1 and 2).

TABLE 6 - Modulation Parameters for PWM-B

n	Channels (MHz)	T_w (ns)	T_s (ns)	a_w
1	0-5	20	40	0.5 (LED_2)
1	0-5	20	40	0.8 (LD)
2	0-5, 7-13	7.7	15.4	0.7 (LD)
3	0-5, 7-13, 14-20	6.3	12.5	0.6 (LD)

4.3 Pulse Frequency Modulation with Baseband Recovery

The expression given for the SNR is based on the analysis of PFM-B in Ref. 4. Using the same notation, f_c is the centre pulse frequency, Δf is the maximum deviation of the pulse frequency and T_p is the pulsewidth. The peak-to-peak amplitude of the baseband signal component is equated with the dynamic range of the video signal, as for PWM-B. The average shot noise at the receiver is determined from

$$P_o = f_c T_p P_R, \qquad (19)$$

or with polarity inversion of the transmitted signal (Ref. 4),

$$P_o = (1 - f_c T_p) P_R. \qquad (20)$$

For multichannel TV transmission, using (17) and (19),

$$SNR = \frac{\left[\frac{1.16\Delta f\, T_p R_o G P_R}{n}\right]^2}{(2qf_c T_p R_o G^{2+x} P_R + 2q I_D G^{2+x} + u_1)(f_2 - f_1) + \frac{u_2}{3}(f_2^3 - f_1^3)} \qquad (21)$$

In this paper, the minimum shot noise given by (19) or (20) is used in the SNR expression to determine the allowable fibre loss.

The spectrum of the PFM signal which is derived by Takasaki *et al* (Ref. 4), is evaluated to establish realistic modulation parameters considering the response time of the optical sources (Tables 1 and 2). The chosen parameters are shown in Table 7.

TABLE 7 - Modulation Parameters for PFM-B

n	Channels (MHz)	f_c (MHz)	Δf (MHz)	T_p (ns)
1	0-5	20	10	24 (LED_2)
1	0-5	20	10	30 (LD)
2	0-5, 7-13	46	19.5	6 (LED_2)
2	0-5, 7-13	46	19.5	12 (LD)
3	0-5, 7-13, 14-20	70	30	6.5 (LD)
6	0-5, 7-13, ..., 35-41	123	41	3 (LD)

4.4 Frequency Modulation

Horak (Ref. 5) derives an expression for the SNR with FM, assuming white receiver noise. With the typical receiver bandwidths required for FM transmission of video signals, this assumption is invalid. The noise spectral density for the amplifier includes a term with a frequency squared dependence. Appendix I derives an expression for the SNR with non-white receiver noise. For a single video signal of baseband bandwidth B,

$$SNR = \frac{\left[1.42\left(\frac{M}{1+M}\right)\left(\frac{\Delta f}{B}\right) R_o G P_R\right]^2}{\left[\frac{2qR_o G^{2+x} P_R}{1+M} + 2q I_D G^{2+x} + u_1\right] B + u_2\left(Bf_c^2 + \frac{3B^3}{5}\right)}, \qquad (22)$$

where M is the amplitude modulation index of the FM carrier, f_c is the carrier frequency and Δf is the maximum deviation of the frequency. The equation, as for the other modulation techniques, includes allowance for the dynamic range 1.73 I_{BW} of the video signal.

There are different definitions of the required FM bandwidth, denoted B_{FM} (Ref. 6). The definition used in this report is that bandwidth for which the amplitude of the sideband frequencies is greater than 1% of the carrier amplitude, and B_{FM} is bounded by

$$2B\left(1 + \frac{\Delta f}{B}\right) \leq B_{FM} \leq 2B\left(2 + \frac{\Delta f}{B}\right) \qquad (23)$$

The bandwidth B_{FM} is centred on the carrier frequency, so that the required (baseband) transmission bandwidth B_T for the optical link

is $(f_c + B_{FM}/2)$. Horak (Ref. 5) states that to avoid interference from the spectrum of the carrier's first harmonic overlapping the carrier's fundamental spectrum (due for example to optical source nonlinearity), the carrier frequency f_c must not be less than B_{FM}.

For multichannel TV transmission, it is easily shown that there is a substantial performance gain using a separate FM carrier for each channel, compared with the alternative of transmitting a FDM arrangement of the channels (as for AIM) on a single FM carrier. The reason is that the noise power after FM demodulation has two components with a B^3 and B^5 dependence, where B is the baseband bandwidth of the signal. Using a separate FM carrier for each TV channel, equation (22) gives the SNR with the inclusion of a $1/n^2$ factor in the numerator to account for the reduction in the carrier power with n channels.

FM is suited to exploit the wide transmission bandwidth of optical fibres, as the transmission bandwidth can be increased for increases in the SNR or the allowable fibre loss. As a basis of comparison with the pulse techniques of PIM, PPM and PFM, a transmission bandwidth of 50 MHz is used for the results of Table 3. The modulation parameters chosen, with M=1 and B=5 MHz in each case, are shown in Table 8. There is a 10 MHz guard band between the channel spectra. For a single TV channel the analysis of Horak (Ref. 5) assuming white receiver noise gives a 3 and 4 dB higher allowable fibre loss for the PIN and APD respectively.

TABLE 8 - Modulation Parameters for FM

n	f_c (MHz)	Δf (MHz)	B_T (MHz)
1	32	6	50
2	20, 50	2.5	60
3	20, 50, 80	2.5	90
6	20, 50, ..., 170	2.5	180

The strong frequency dependence of the noise spectral density after FM demodulation means that substantial performance improvements can be achieved using emphasis techniques (Ref. 6), at the expense of a small increase in receiver complexity. Consider de-emphasis of the signal plus noise after demodulation, with a filter response

$$|W(f)|^2 = \frac{1}{1 + (f/f_o)^2} \qquad (24)$$

An expression for the SNR is derived in Appendix I. An example of the potential performance improvement is given in Table 3, where a SNR improvement of 6 to 8 dB is obtained, with a de-emphasis filter breakpoint of f_o = 0.8 MHz. Emphasis techniques can be applied to other modulation techniques such as AIM, but the performance improvement with AIM is much smaller than with FM (Ref. 6).

4.5 Pulse Modulation Techniques - PIM, PPM, PFM

Expressions for the SNR are derived in Ref. 7 for PIM, in Ref. 8 for PPM and in Ref. 9 for PFM. The derivations assume cosine squared pulses at the decision point in the receiver and no intersymbol interference between pulses. Expressions for the SNR are derived in Appendix II, with the more realistic assumption of a Gaussian pulse shape at the receiver. In both Appendix II and Refs. 7-9, white receiver noise is assumed in the analysis, although the receiver noise spectral density (8) has a frequency squared component. As discussed in the previous section for FM, the assumption of white receiver noise gives optimistic estimates of the allowable fibre loss. This may not be a severe limitation if the shot noise component dominates performance, as for example with an APD. However, the results for PIM, PPM and PFM, which are based on the above assumption, should be considered with some reservation as to their accuracy. A recent analysis of the noise spectrum for PFM has been reported by Webb (Ref. 15).

The respective SNR expressions from Appendix II, based on the general format of (11) and with allowance for a video dynamic range of 1.73 I_{BW}, are given below. For PIM,

$$SNR = \frac{\left[\dfrac{2.65\, a_p B_T R_o G P_R}{\sqrt{B f_s\,(1-\sin(2\pi B/f_s)/(2\pi B/f_s))}}\right]^2}{\left[1.21\, q R_o G^{2+x} P_R + 2q I_D G^{2+x} + u_1\right] B_T + \frac{u_2}{3} B_T^3}, \qquad (25)$$

where $f_s = 1/T_s$ is the sampling frequency, $a_p = 2T_{p\text{-}p}/T_s$ is the modulation index and $T_{p\text{-}p}$ is the peak-to-peak deviation of the pulse interval. For PPM, assuming synchronization of the receiver clock used to detect the time position of pulses and that $T_{p\text{-}p}$ is now the peak-to-peak variation in pulse position,

$$SNR = \frac{\left[5.29\,(\frac{a_p B_T}{f_s})\, R_o G P_R\right]^2}{\left[1.21\, q R_o G^{2+x} P_R + 2q I_D G^{2+x} + u_1\right] B_T + \frac{u_2}{3} B_T^3} \qquad (26)$$

For PFM,

$$SNR = \frac{\left[\dfrac{1.46 B_T\, \Delta f\, R_o G P_R}{B f_c}\right]^2}{\left[1.21\, q R_o G^{2+x} P_R + 2q I_D G^{2+x} + u_1\right] B_T + \frac{u_2}{3} B_T^3}, \qquad (27)$$

where f_c is the centre pulse frequency and Δf is the maximum deviation of the pulse frequency. In the above expressions, B is the baseband bandwidth of the video signal (5 MHz) and the receiver noise bandwidth is assumed to be the same as the transmission bandwidth B_T.

The modulation parameters chosen for the results of Table 3 based on B_T = 50 MHz are,

PIM: $a_D = 0.4$, $f_S = 15$ MHz; PPM: $a_D = 0.4$, $f_S = 15$ MHz; PFM: $f_C = 20$ MHz, $\Delta f = 5$ MHz.

As with FM, increases in B_T can be traded off for increases in the allowable fibre loss. The disadvantage of these modulation techniques for multichannel TV transmission is that they require wide transmission bandwidths, of the order of ten times the baseband bandwidth of the combined video signals.

4.6 Pulse Code Modulation

The results for PCM in Tables 3 and 4 are calculated based on the work of Personick (Ref. 14). The SNR of the video signal is determined by the quantization noise, provided the bit error rate of the PCM system is sufficiently small. Using the definition of the SNR in (1) and taking account of the desired dynamic range for the video signal of 1.73 I_{BW}, it is readily shown that the signal-to-quantization noise ratio is

$$SNR = \frac{(2^N/1.73)^2}{1/12} \simeq 2^{2N+2}, \qquad (28)$$

or expressed in dB,

$$SNR\ (dB) \simeq 6N+6 \qquad (29)$$

for N-bit, linear quantization of the video signal. To satisfy the CTV objective for the SNR of 40 dB, 6-bit quantization is used with a bit error rate of 10^{-6}. With 6-bit PCM coding of the video signal the addition of high frequency noise may be necessary to mask the effect of quantizing noise, which is seen as contouring of the image for low-bit resolution. The 5 MHz bandwidth video signal is assumed to be sampled at a rate of 13.3 MHz, giving a bit rate of approximately 80 Mbit/s per video channel (160 and 240 Mbit/s for n = 2 and 3 in Table 4).

The PCM results are based on the same receiver noise model as the other modulation techniques. A binary code and 50% duty-cycle transmit pulses, with an extinction ratio of zero, are assumed. The received pulses have a Gaussian pulse shape and are equalized for full raised-cosine pulses at the decision point (Ref. 14). Finer quantization of the video signal with PCM transmission can be adopted to improve the TV picture quality and allow less stringent demands on the subscriber's TV set. Studio quality signals for PAL TV can be achieved with 9-bit linear quantization, giving a bit rate of 120 Mbit/s per TV channel. The allowable fibre loss is within about 1 dB of that for 6-bit quantization (Table 3), with an increased transmission bandwidth of 75 MHz.

5. DISCUSSION OF RESULTS

Results for optical transmission of a single TV channel are given in Table 3. The modulation techniques can be classified in two groups, depending on the transmission bandwidth requirement. The first group of AIM, PWM-B and PFM-B require a B_T of 5 MHz and achieve similar allowable losses with a LD.

The LD modulation bandwidth is sufficient to not be a determining factor in the allowable fibre loss, which is in contrast to the case for a LED and PWM-B or PFM-B. With a LED, AIM has an advantage in the loss of about 7 dB. Part of this advantage is because PWM-B and PFM-B require LEDs with a wider modulation bandwidth and consequent reduction in the LED power output. However, there are questions on the suitability of AIM because of its stringent requirements on optical source linearity. There is little practical advantage in using an APD compared with a PIN detector for these modulation techniques. A PIN detector does not require the large bias voltages of an APD and can tolerate a larger incident power without distortion (useful for transmission over short distances) (Ref. 1).

The second group of modulation techniques is compared with a transmission bandwidth of 50 MHz. The allowable fibre loss for each technique can be improved to varying degrees by increasing the transmission bandwidth. The maximum allowable fibre losses are with FM using de-emphasis of the noise and PCM. The increase in allowable fibre loss for FM depends on the choice of pre-emphasis/de-emphasis network. PIM and PFM also have the potential for performance improvements with de-emphasis of the receiver noise, as their noise spectra are similar to that for FM (Refs. 7,9). The APD operating at its optimum avalanche gain, for the modulation techniques with a transmission bandwidth of 50 MHz, gives a 6 to 9 dB increase in the allowable fibre loss compared with that for the PIN detector.

The variation in the allowable fibre loss with SNR, for a fixed transmission bandwidth, is readily determined from the respective SNR expressions. For AIM, PWM-B and PFM-B, each 10 dB reduction in the SNR corresponds to an approximately 6 dB increase in the allowable fibre loss. A similar relationship applies to FM and the pulse techniques with an APD.

Results for multichannel TV transmission are given in Table 4. A subject requiring further investigation is the suitability of LEDs to transmission with PWM-B and PFM-B. For instance, will operating a LED close to its modulation bandwidth limit (such as for these techniques with n = 2) cause distortion of the analogue TV signals? A general conclusion from Table 4 is that PCM offers the best allowable fibre loss for multichannel transmission, even allowing for an improvement in FM with emphasis techniques. Results for more than six TV channels are not tabulated, because of the severe requirements on LD linearity (i.e. for AIM, PFM-B and FM) and optical fibre bandwidth (i.e. for FM and PCM). For the above reasons, a centrally-switched, star-distribution, optical CTV system is favoured in overseas investigations (Ref. 1). A typical approach is that only two or three TV channels are transmitted simultaneously to each subscriber.

The choice between the modulation techniques depends not only on the distribution area and transmission distance to be covered, but on many other factors, including some listed below.

(i) The cost and complexity of the receiver at each subscriber's premises is a major factor. It is expected that for all modulation techniques the receive signal must be converted to a VSB-modulated format to interface with conventional broadcast TV receivers. The simplest optical receiver is that with AIM, PWM-B or PFM-B. The next level of receiver complexity is that with FM and the pulse techniques. PCM receiver complexity is the greatest as a digital-to-analogue converter and synchronization circuits are necessary. However, digital transmission techniques have good potential for cost reduction with LSI technology.

(ii) A possible approach to offset the cost of a CTV distribution network is the inclusion of other services, such as telephony, data, audio services, etc. Digital transmission is particularly suited to this application.

(iii) The network configuration is closely inter-related with the choice of a modulation technique. For example, the use of remote switching units can significantly reduce the transmission distances in subscriber links. Signal conversion costs are avoided if the modulation technique for subscriber distribution is compatible with that in the trunk area.

(iv) Analogue modulation techniques are more susceptible to degradation in transmission performance when cascading links with the use of repeaters than digital transmission.

(v) Given the comments of the previous section that a centrally-switched CTV system is the most likely network implementation, some modulation techniques present problems in realising a suitable wideband switch (Ref. 1).

6. CONCLUSIONS

This paper has reviewed a representative sample of eight modulation techniques for CTV transmission on optical fibres. The modulation techniques are compared on the basis of transmission performance, i.e. the maximum allowable fibre loss and transmission bandwidth required to satisfy a specified SNR of the video signal. With the exception of PCM, a standard form of expression for the SNR is given for the modulation techniques. In summary, there is no one obvious modulation technique for all CTV distribution applications. Based on the calculations suitable modulation techniques are:

(i) Analogue intensity modulation (AIM) with a low-bandwidth optical fibre, such as step-index or low quality graded-index optical fibre. Nonlinearity of optical sources probably limits practical transmission with AIM to two or three TV channels and unrepeatered links.

(ii) Pulse frequency modulation and baseband recovery of the video signal (PFM-B), with a low-bandwidth optical fibre. PFM-B avoids the source linearity problems of AIM and is particularly suited to use with a laser diode.

(iii) Frequency modulation (FM) with a moderate-bandwidth, graded-index optical fibre. Fibre bandwidth can be traded off for increases in the SNR and/or fibre loss. The achievable transmission distance is greater than with AIM or PFM-B, especially if emphasis techniques are used.

(iv) Pulse code modulation (PCM) with a high-bandwidth optical fibre. The achievable transmission distance is the greatest with PCM, the advantage in fibre loss being larger for two or three TV channels than for a single TV channel. The significant disadvantage of PCM is the receiver circuit complexity, although this disadvantage may diminish in the future with the increased move to digitalize telecommunications networks.

7. ACKNOWLEDGEMENT

The assistance of Mr R.W. Ayre, Dr K.F. Barrell and Mr T.D. Stephens during the course of the review is gratefully acknowledged. The results for PCM are based on a computer program written by Mr T.D. Stephens.

8. REFERENCES

1. Mogensen, G., "Review : Wide-band optical fibre local distribution systems", J. Optical and Quantum Electronics, Vol. 12, 1980, pp. 353-381.

2. Peterman, K. and Arnold, G., "Noise and Distortion Characteristics of Semiconductor Lasers in Optical Fiber Communication Systems", IEEE Trans, Vol. QE-18, No. 4, April 1982, pp. 543-555.

3. Nagano, K., Takahashi, Y., Takasaki, Y., Maeda, M. and Tanaka, M., "Optimizing Optical Transmitters and Receivers for Transmitting Multi-channel Video Signals Using Laser Diodes", IEEE Trans., Vol. COM-29, No. 1, January 1981, pp. 41-45.

4. Takasaki, Y., Nakagawa, J. and Koya, M., "New Fiber Optic Analog Baseband Transmission Plan for Color TV Signals", IEEE Trans., Vol. COM-26, No. 6, June 1978, pp. 902-907.

5. Horak, W., "Analog TV Signal Transmission over Multimode Optical Waveguides", Siemens Forsch.-u. Entwickl.-Ber., Vol. 5, No. 4, 1976, pp. 194-202.

6. Schwartz, M., Information Transmission, Modulation and Noise, McGraw-Hill, New York, 1970.

7. Ueno, Y. and Yasugi, T., "Optical Fiber Communication Systems Using Pulse-Interval Modulation", NEC Research and Development, No. 48, January 1978, pp. 45-51.

8. Hubbard, W.M., "Utilization of Optical-Frequency Carriers for Low- and Moderate-Bandwidth Channels", B.S.T.J., Vol. 52, No. 5, May-June 1973, pp. 731-765.

9. Timmermann, C.C., "Noise in Receivers for Pulse Frequency Modulated Optical Signals", AEU, Vol. 31, 1977, pp. 285-288.

10. Sato, M., Murata, M. and Namekawa, T., "Pulse Interval and Width Modulation for Video Transmission", IEEE Trans., Vol. CATV-3, No. 4, October 1978, pp. 165-173.

11. Engineering Instruction EI 0 8010, "Television Relay Facilities : Operational Performance Objectives", Telecom Australia, Issue 2, 1973.

12. Broadcast Procedure 23, "Technical Standards and Procedures for Cable Television (CATV) Systems", Dept. of Communications (Canada), Issue 1, July 1, 1971.

13. Moustakas, S. and Hullett, J.L., "Noise modelling for broadband amplifier design", IEE Proc., Vol. 128, Pt. G, April 1981, pp. 67-76.

14. Kressel, H., ed., Topics in Applied Physics : Semiconductor Devices for Optical Communication, Vol. 39, Springer-Verlag, Berlin, 1980.

15. Webb, R.P., "Output Noise Spectrum from Demodulation in an Optical PFM System", Electronics Letters, Vol. 18, 8 July 1982, pp. 634-636.

16. Gradshteyn, I.S. and Ryztik, I.M., Table of Integrals Series and Products, Academic Press, New York, 1965.

APPENDIX I

SNR ANALYSIS FOR FREQUENCY MODULATION

1. Non-White Receiver Noise

The receiver noise is

$$\langle i_n^2 \rangle = (2qR_o G^{2+x} P_o + 2qI_D G^{2+x} + u_1)\,df + u_2 f^2\,df \qquad (8)$$
$$= v_1\,df + v_2 f^2\,df,$$

where v_1 and v_2 are constants independent of frequency. The following derivation for the SNR is based on that for FM in Schwartz (Ref. 6), but modified for non-white receiver noise. The receiver noise, before FM demodulation, may be represented as a narrowband process about the carrier frequency f_c, so that

$$i_n(t) = y(t)\cos w_c t - z(t)\sin w_c t,$$

where $w = 2\pi f$. If $G_n(f)$ denotes the single-sided power spectral density of $i_n(t)$, i.e.

$$G_n(f) = v_1 + v_2 f^2,$$

then the power spectral density of y(t) or z(t) is (Ref. 6)

$$G_y(f) = G_z(f) = G_n(f + f_c) + G_n(f - f_c)$$
$$= 2[v_1 + v_2 (f^2 + f_c^2)]. \qquad (30)$$

Schwartz shows that the output noise power N_o from the FM demodulator of baseband bandwidth B (which acts as a differentiator with transfer function $|H(w)|=w/a_c$, where A_c is the peak carrier amplitude) is for large carrier-to-noise ratios

$$N_o = \int_o^B \frac{w^2}{A_c^2} G_y(f)\,df \;,\; A_c^2 >> N_o, \qquad (31)$$

Substituting for $G_y(f)$ and integrating,

$$N_o = \frac{2(2\pi)^2}{A_c^2} \left[(v_1 + v_2 f_c^2) \frac{B^3}{3} + v_2 \frac{B^5}{5} \right]$$

The signal power S_o after demodulation corresponding to the video current I_{BW}, taking account of the desired dynamic range 1.73 I_{BW}, is

$$S_o = \left[\frac{2(2\pi\Delta f)}{1.73} \right]^2 , \qquad (32)$$

so that

$$SNR = \frac{S_o}{N_o} = \frac{\left[1.42(\frac{\Delta f}{B})\, A_c \right]^2}{v_1 B + v_2 (B f_c^2 + \frac{3B^3}{5})}$$

If M is the amplitude modulation index of the FM carrier (defined as for m in AIM), $A_c = MR_o G P_R/(1+M)$, and on replacing v_1, v_2 by their respective expressions from (8), equation (22) follows.

2. De-emphasis at the Receiver

It is assumed that the video signal is pre-emphasized at the transmitter with a response which is the inverse of W(f) specified in (24). The overall effect of pre-emphasis and de-emphasis is that the output signal remains unchanged and its power is given by (32). The

noise power after demodulation and subsequent de-emphasis is from (31)

$$N_o = \int_o^B \frac{w^2}{A_c^2} G_y(f) |W(f)|^2 \, df.$$

Substituting from (30) for $G_y(f)$ and (24) for $W(f)$,

$$N_o = \frac{2(2\pi)^2}{A_c^2} \int_o^B \left[\frac{(v_1 + v_2 f_c^2) f^2 + v_2 f^4}{1 + (f/f_o)^2} \right] df$$

Evaluating the integral (Ref. 16), and replacing v_1 and v_2 from (8) gives, in place of (22),

$$SNR = \frac{[0.82 \frac{M}{1+M} \frac{\Delta f}{f_o} R_o G P_R]^2}{XY + u_2 Z}$$

where

$$X = \frac{2qR_o G^{2+x} P_R}{1+M} + 2qI_D G^{2+x} + u_1,$$

$$Y = B - f_o \arctan \frac{B}{f_o},$$

$$Z = f_c^2 (B - f_o \arctan \frac{B}{f_o}) + \frac{B^3}{3} - B f_o^2 + f_o^3 \arctan \frac{B}{f_o}$$

APPENDIX II

SNR ANALYSIS FOR PIM, PPM AND PFM

The SNR expressions derived for PIM, PPM and PFM, in Refs. 7-9 respectively, assume $\cos^2$ pulses at the receiver decision point. The SNR analysis developed in this Appendix is based on Refs. 7-9, but with the more realistic assumption of Gaussian pulses and inclusion of the dynamic range 1.73 I_{BW} for the video signal. Also the SNR expressions are arranged in the general form of (11).

Consider a Gaussian pulse response, where the signal current at the receiver decision point due to a transmitted pulse is

$$i_r(t) = \exp\left(\frac{-t^2}{2\sigma^2}\right) i_m, \qquad (33)$$

i_m is the peak current and 2σ is the r.m.s. pulsewidth. The corresponding frequency response is

$$I_r(f) = \beta \exp\left[\frac{-(2\pi\sigma f)^2}{2}\right], \qquad (34)$$

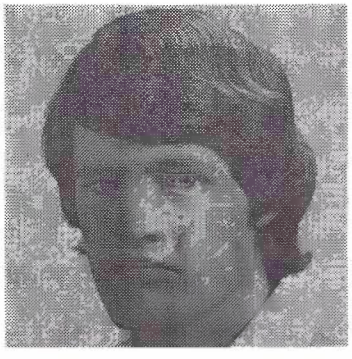
where β is a constant. For narrow transmit pulses, the optical channel frequency response can be equated with (34), so that $\sigma = \sqrt{\ln 2}/2\pi B_T$. Assuming negligible intersymbol interference (see later comment on distortion), the optimum decision instant is taken as the time t at which the slope of $i_r(t)$ is a maximum. The slope, denoted $i_r'(t)$, at the optimum decision instant is then

$$i_r'(t=-\sigma) = \frac{I_m}{\sigma\sqrt{e}} \simeq 4.577 \, i_m B_T, \qquad (35)$$

with the decision threshold at

$$i_r(t=-\sigma) = i_m/\sqrt{e} \simeq 0.6065 \, i_m \qquad (36)$$

Substituting the slope (35) in place of the slope at the decision instant for $\cos^2$ pulses in the analyses of Ref. 7-9, then after some manipulation the SNR expressions (25), (26) and (27) follow. The shot noise component at the decision instant in these expressions is changed because of the increased decision threshold level (36), compared with the threshold level of 0.50 for $\cos^2$ pulses. With each modulation technique there is a possibility of a false pulse being detected at the receiver due to the noise exceeding the threshold level. This effect is not included in the SNR expressions as it is negligible for most practical situations where the system has a moderate receiver bandwidth (Refs. 7-9). The modulation parameters in Section 4.5 are chosen to satisfy the criterion that the distortion or nonlinearity of the video signal is at least 50 dB below the signal level. This criterion is equivalent to restricting the intersymbol interference between Gaussian pulses at the decision instant.



BIOGRAPHY

GRANT NICHOLSON was educated at the University of Tasmania, Hobart, Tasmania, where he was awarded the degrees of Bachelor of Engineering with Honours (Electrical) in 1976 and Master of Engineering Science in 1979. He commenced work as an Engineer with Telecom Research Laboratories, Line and Data Systems Section in 1978, working in the area of PCM digital transmission systems. Since 1981 he has been in the Optical Systems Section where his main activities have been concerned with the characterisation of single mode fibre and the application of optical fibres in subscriber distribution networks.

Switched-Capacitor Equaliser Structures For A Digital Telephone

A. JENNINGS

Telecom Australia Research Laboratories

Equalisation of existing subscriber lines is an important problem in the application of time-compression multiplex digital transmission. We show that suitable switched-capacitor equaliser structures can be derived from simple analog prototypes. A two-stage optimisation procedure is described for the design of these equalisers. Coupled with existing adaptive switched-capacitor equalisers, this approach can form a complete equalisation scheme for TCM transmission.

1. INTRODUCTION

The possibility of digital transmission at rates of 64 kb/s to 144 kb/s over existing subscriber lines has attracted the interest of telecommunication administrations throughout the world (Refs. 1-6). Transmission schemes proposed include time-compression multiplexing, echo cancellation and frequency-division multiplexing. Ref. 6 provides a discussion of these techniques.

In the time-compression multiplexing (TCM) method the bidirectional digital information passing between the digital telephone and the exchange terminal is multiplexed into 'bursts' of bits and transmitted at a higher bit rate in alternating send and receive cycles. Fig. 1 illustrates this transmission scheme.

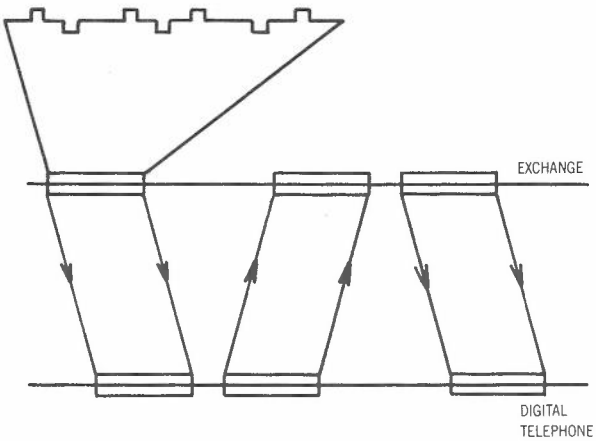
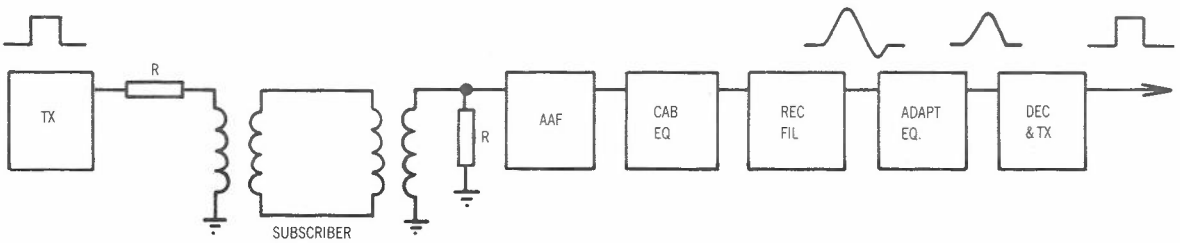


Fig. 1 - Time-compression multiplex transmission



TX – CODED OUTPUT TO LINE
AAF – ANTI-ALIASING FILTER
CAB EQ – CABLE EQUALISER & AGC
ADAPT EQ – ADAPTIVE EQUALISER
REC FIL – RECONSTRUCTION LOW-PASS FILTER
DEC & TX – DECISION CIRCUITRY & REGENERATION

Fig. 2 - Line receiver

Since the existing subscriber lines have highly dispersive transmission characteristics, equalisation becomes an important problem in the application of the TCM technique - typically up to 46 dB attenuation at a frequency corresponding to half the bit rate may be encountered for limit lines. In addition various imperfections are present such as gauge changes, bridged taps, etc.. Fig. 2 shows a block diagram of a line receiver for this application. A variable equaliser with nominal cable characteristics and AGC action is followed by an adaptive equaliser to deal with line imperfections.

It is highly desirable that an equaliser for TCM be capable of realisation in integrated form, preferably in MOS. This enables a single-chip realisation of the digital telephone since the digital

functions can be realised on the same IC chip. Some authors (Refs. 2,5) have claimed that it is difficult to realise equalisation in integrated form for the TCM technique, and use this as an argument in favour of either unequalised TCM or echo cancellation techniques. In this paper we present structures for switched-capacitor (SC) equalisers suitable for the TCM technique. We present an SC structure developed from a simple analog prototype to equalise nominal cable attenuation and show that it is readily coupled with known realisations of SC transversal equalisers to form a complete TCM equaliser.

There is an extensive literature devoted to the analysis of SC networks (Refs. 7-10) and to the design of SC filters (Refs. 11-20). However little work appears to have been devoted to the use of SC networks as equalisers. Martin and Sedra present an all-pole equaliser adjusted by zero-shifting (Ref. 21). Whilst this approach is not directly applicable to cable equalisation, it opens some of the possibilities explored in this paper. Suzuki and Shirasu (Ref. 22) mention a 5th order SC cable equaliser in their survey paper but give only a brief description.

2. SWITCHED CAPACITOR BUILDING-BLOCKS

In the realisation of switched-capacitor circuits it is desirable that the circuit be insensitive to stray capacitance to ground. If the circuit is sensitive to strays this places a lower limit on the size of capacitance that can be realised accurately. This increases the chip area the circuit occupies and we shall see that it also limits the speed of operation of the circuit. Fig. 3 shows a strays-insensitive switched-capacitor amplifier and Fig. 5 (Ref. 13) shows a strays-insensitive switched-capacitor integrator. The circuits require a two-phase clock and the notation is that a switch labelled ϕ_1 is closed when the clock ϕ_1 is in the high state and open otherwise, i.e. a switch labelled ϕ_1 is closed during time intervals $I_1, I_3, \ldots$ and open otherwise; the clock frequency $f_c = 1/T$.

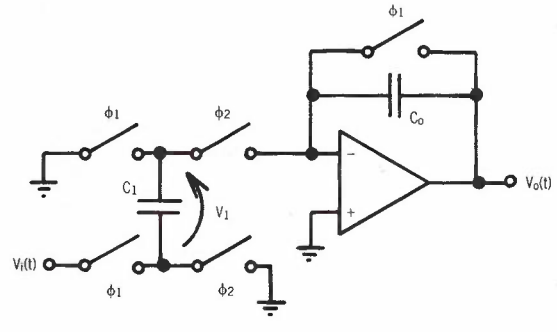


Fig. 3 - SC amplifier

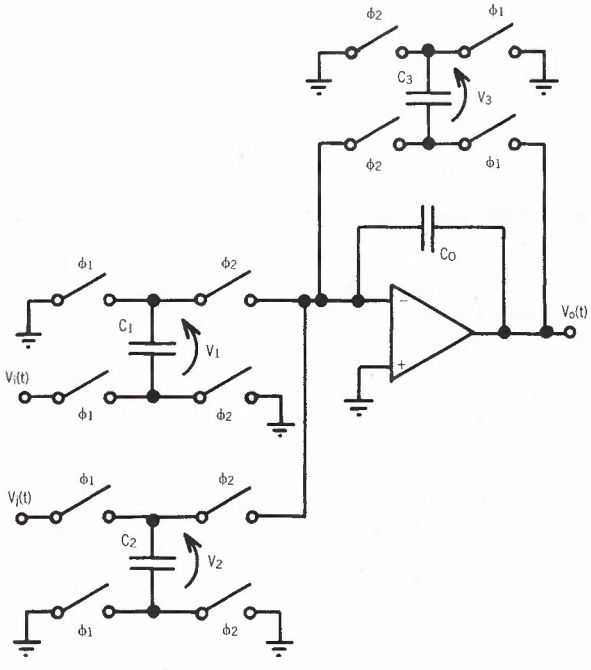


Fig. 5 - SC integrator

In Appendix A1 we show that the discrete-time amplifier of Fig. 3 has transfer function

$$\frac{V_o}{V_i} = \frac{C_1}{C_0} z^{-\frac{1}{2}} \qquad (1)$$

and the integrator of Fig. 5 has transfer function

$$V_o = \frac{C_1}{C_0} \frac{z^{-\frac{1}{2}}}{1-z^{-1}} V_i - \frac{C_2}{C_0} \frac{z^{-\frac{1}{2}}}{1-z^{-1}} V_j - \frac{C_3}{C_0} \frac{z^{-1}}{1-z^{-1}} V_o \qquad (2)$$

In any practical circuit finite switch resistance and op-amp gain-bandwidth will

cause deviations from these ideal transfer functions. Martin and Sedra (Ref. 15) have considered the effect of op-amp gain-bandwidth on integrator performance, and in Appendix A2 we give expressions for the performance of the SC amplifier in Fig. 3 with finite switch resistance and finite op-amp gain-bandwidth. Substitution of practical values in these expressions gives as a rule of thumb that op-amp gain-bandwidth of approximately $10f_c$ will give satisfactory performance.

3. ANALOG PROTOTYPES

To develop an SC equaliser using these building blocks we start first with an analog prototype and then transform this prototype to a suitable form. This approach has the advantage that we are working with familiar circuit relations, and some classical circuit theory results can be brought to bear on the problem.

For convenience the cable equaliser can be split into a nominal fixed equaliser at mid-range followed by a variable equaliser with AGC. Fig. 6 shows the desired equalisation characteristics for 0-4 km of 0.40 mm line with 120 Ω terminating resistances (i.e. R = 120 Ω in Fig. 2).

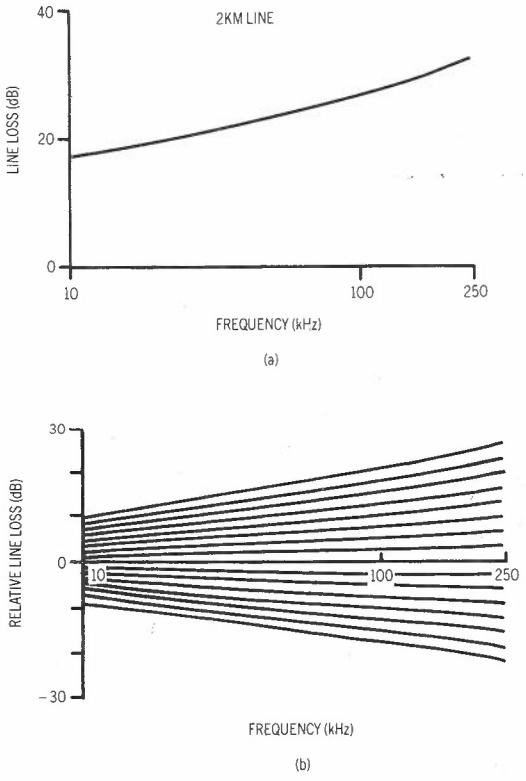


Fig. 6 - Equalisation characteristic

All of the prototypes of Fig. 7 are suitable for realising the equalizer characteristics of Fig. 6, but prototype C has a ladder structure and is preferred for development of the transformation. By selecting a ladder structure we can directly apply the signal flow graph method of Ref. 11.

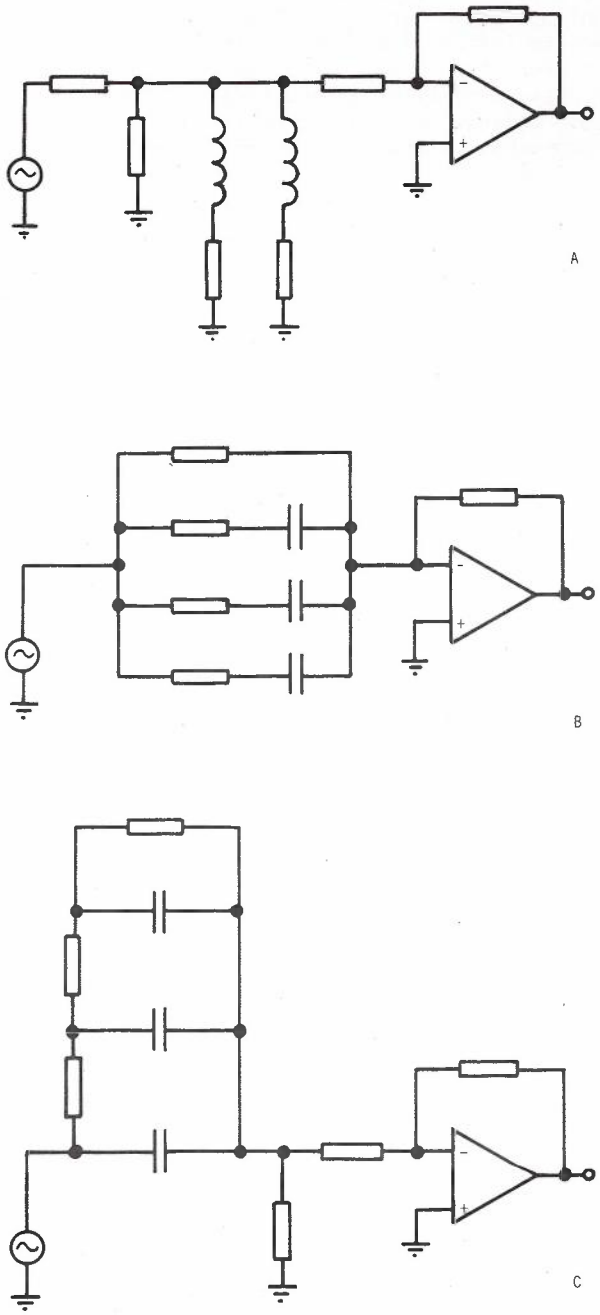


Fig. 7 - Analog prototypes

The transfer function will be of the form

$$G(s) = \frac{(s+z_1)(s+z_2)\ldots(s+z_n)}{(s+p_1)(s+p_2)\ldots(s+p_n)}$$

where

$$z_1 < p_1 < z_2 < p_2 \ldots\ldots < z_n < p_n \qquad (3)$$

and the ladder network can be synthesised using standard techniques.

Since the transfer function G(s) is minimum-phase (has no zeros in the right-half s-plane), to a first approximation we need only design the equaliser to satisfy a magnitude criterion, i.e. $|G(j\omega)|$ has the characteristics of Fig. 6 and the phase characteristics will be such as to equalise the phase distortion of the line. This result follows from the property that a minimum-phase transfer function is completely specified by its magnitude on the jω-axis (Ref. 23). Since we are not approximating the equalisation characteristics over the entire jω-axis but only over part of the axis, the phase equalisation will not be exact. It is interesting here to contrast the analog prototype design approach with the direct time-domain approach. For equalisation of the cable characteristic we can proceed more directly with the analog prototype, however for adaptive equalisation where no *a priori* characteristics are available the direct time-domain approach is more efficient.

4. TRANSFORMATION OF THE ANALOG PROTOTYPE

To obtain an SC network that simulates the prototype C we apply the lossless-discrete integrator (LDI) transform (Ref. 24):

$$S \rightarrow \frac{1}{T}\,\frac{1 - z^{-1}}{z^{-\frac{1}{2}}} = \frac{2}{T}\,\sinh\,\frac{ST}{2} \qquad (4)$$

to the network. First re-draw the passive part as in Fig. 8. This network can then be represented by the signal flow graph of Fig. 9 (applying the method of Ref. 11). Summing the currents at each node we obtain the equations

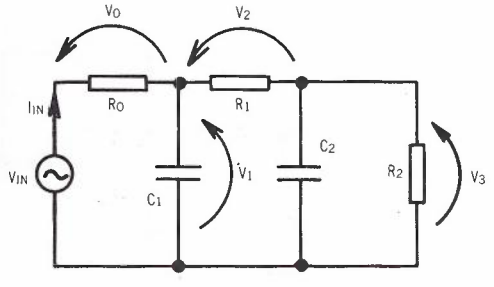


Fig. 8 - Analog prototype C

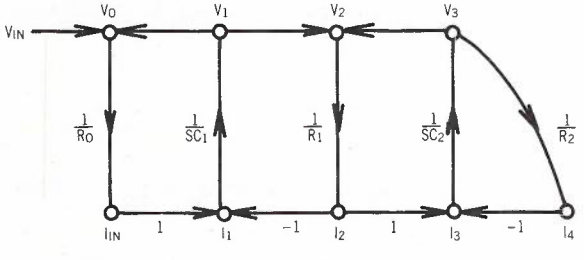


Fig. 9 - Signal flow graph representing Fig. 8

$$I_1 = \frac{(V_{IN} - V_1)}{R_0} - \frac{(V_1 - V_3)}{R_1}$$

$$I_3 = \frac{(V_1 - V_3)}{R_1} - \frac{V_3}{R_2} \qquad (5)$$

and

$$V_1 = \frac{1}{SC_1}\left[\frac{V_{IN}}{R_0} - \left(\frac{1}{R_0} + \frac{1}{R_1}\right) V_1 + \frac{V_3}{R_1}\right]$$

$$V_3 = \frac{1}{SC_2}\left[\frac{V_1}{R_1} - \left(\frac{1}{R_1} + \frac{1}{R_2}\right) V_3\right] \qquad (6)$$

Now re-arranging (6) to a form suitable for the application of the LDI transform,

$$V_1 = \frac{1}{ST}\left[\frac{T}{R_0C_1} V_{IN} - \left(\frac{1}{R_0} + \frac{1}{R_1}\right)\frac{T}{C_1} V_1 + \frac{T}{R_1C_1} V_3\right]$$

$$V_3 = \frac{1}{ST}\left[\frac{T}{R_1C_2} V_1 - \left(\frac{1}{R_1} + \frac{1}{R_2}\right)\frac{T}{C_2} V_3\right] \qquad (7)$$

and define

$$\alpha_1 = \frac{T}{R_0C_1},\ \alpha_2 = \left(\frac{1}{R_0} + \frac{1}{R_1}\right)\frac{T}{C_1},\ \alpha_3 = \frac{T}{R_1C_1}$$

$$\alpha_4 = \frac{T}{R_1C_2},\ \alpha_5 = \left(\frac{1}{R_1} + \frac{1}{R_2}\right)\frac{T}{C_2} \qquad (8)$$

Then applying the LDI transform (4) to (7) we obtain

$$\begin{bmatrix} V_1 \\ V_3 \end{bmatrix} = \frac{z^{-\frac{1}{2}}}{1 - z^{-1}} \begin{bmatrix} -\alpha_2 & \alpha_3 \\ \alpha_4 & -\alpha_5 \end{bmatrix} \begin{bmatrix} V_1 \\ V_3 \end{bmatrix} + \alpha_1 \frac{z^{-\frac{1}{2}}}{1 - z^{-1}} \begin{bmatrix} V_{IN} \\ 0 \end{bmatrix} \qquad (9)$$

which is representable by the signal flow graph (sfg) of Fig. 10. However this sfg is unrealisable (Ref. 25), and the sfg must be modified to the form of Fig. 11. The extra half-cycle delay has the effect of converting the resistors to frequency-dependent complex

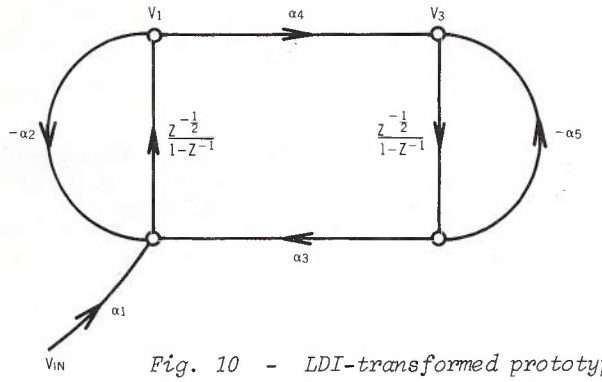


Fig. 10 - LDI-transformed prototype

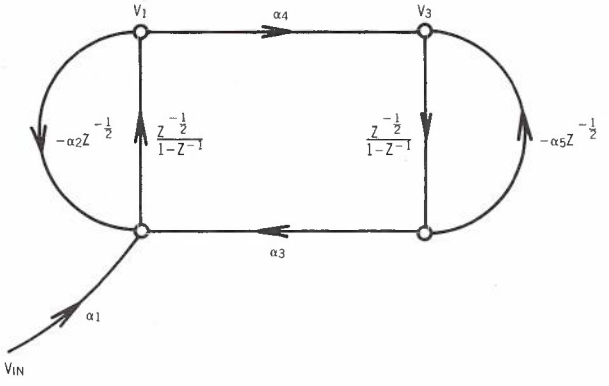


Fig. 11 - LDI-transformed prototype modified for realisability

impedances and this must be taken into account in the design process. Equation (9) then becomes

$$\begin{bmatrix} V_1 \\ V_3 \end{bmatrix} = \frac{z^{-\frac{1}{2}}}{1 - z^{-1}} \begin{bmatrix} -\alpha_2 z^{-\frac{1}{2}} & \alpha_3 \\ \alpha_4 & -\alpha_5 z^{-\frac{1}{2}} \end{bmatrix} \begin{bmatrix} V_1 \\ V_3 \end{bmatrix} + \alpha_1 \frac{z^{-\frac{1}{2}}}{1 - z^{-1}} \begin{bmatrix} V_{IN} \\ 0 \end{bmatrix} \quad (10)$$

The SC network of Fig. 12 realises the sfg using the building blocks of Section 1, and includes a gain stage to compensate for the cable flat loss. Equation (10) gives the transfer function of this network as

$$H(z) = \frac{k\, z^{-\frac{1}{2}} [a_0 + a_1 z^{-1} + a_2 z^{-2}]}{[1 + b_1 z^{-1} + b_2 z^{-2}]}$$

where

$$a_0 = 1 - \alpha_1$$

$$a_1 = [(\alpha_2 - 1) + (\alpha_5 - 1) - \alpha_3\alpha_4] - \alpha_1 (\alpha_5 - 1)$$

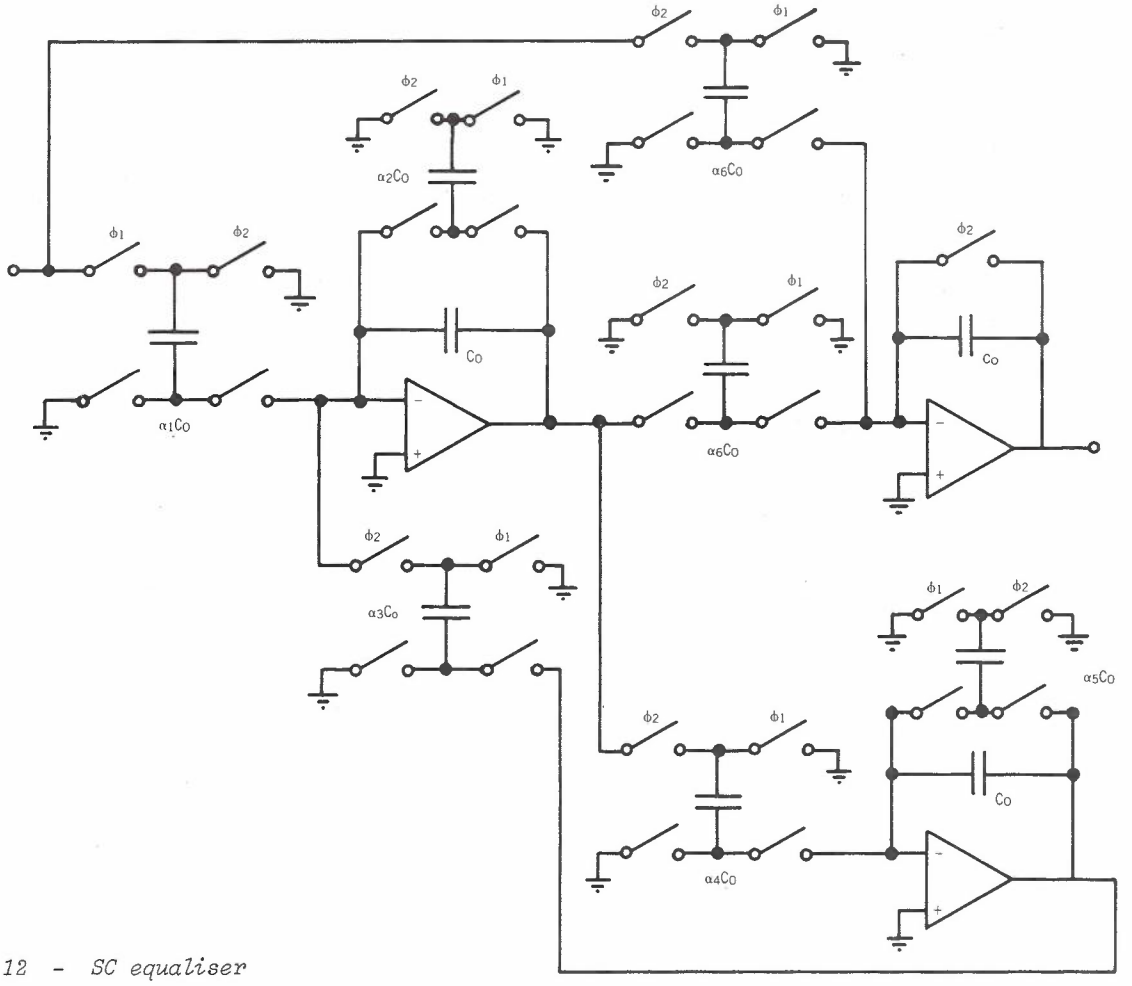


Fig. 12 - SC equaliser

$$a_2 = (\alpha_2 - 1)\ (\alpha_5 - 1)$$

$$b_1 = (\alpha_2 - 1) + (\alpha_5 - 1) - \alpha_3\alpha_4$$

$$b_2 = (\alpha_2 - 1)\ (\alpha_5 - 1) \qquad (11)$$

and these expressions can be used to locate poles and zeros of the transfer function. Recall that for stable operation only the poles must be within the unit circle, but to preserve the minimum-phase property of prototype C both the poles and zeros must be within the unit circle.

5. OPTIMISATION

If the clock frequency f_c is substantially greater than the highest frequency of interest then the LDI transform (4) reduces to

$$j\Omega \rightarrow \frac{2}{T} \sinh\left(\frac{j\omega T}{2}\right) = j\frac{2}{T} \sin\left(\frac{\omega T}{2}\right)$$

$$\simeq j\omega \qquad (12)$$

and the derived SC network closely approximates the analog prototype frequency response. However in our case the highest frequency of interest is about 200 kHz and this would require a clock frequency of approximately 8 MHz. This is extremely difficult to realise, and obviously it is more desirable to tolerate some warping of the frequency axis at the expense of a slightly more difficult design procedure. In theory the clock frequency can be reduced to the Nyquist limit of $2f_{max}$, but this places very stringent requirements on the anti-aliasing filter (Fig. 2). A compromise must be reached between these conflicting requirements, and $f_{max}/f_c \simeq 1/5$ gives good results. In our case this results in a clock frequency of 1 MHz. This is realisable with NMOS technology (Ref. 26).

To take into account the warping of the LDI transform, the analog prototype transfer function is predistorted as follows:

$$G(j\Omega) = G\left[j\frac{2}{T}\sin\left(\frac{\omega T}{2}\right)\right] \qquad (13)$$

and optimised to give the desired equalisation characteristic $D(j\omega)$, i.e.

$$\min \left|\left| G\left[j\frac{2}{T}\sin\left(\frac{\omega T}{2}\right), \underline{z}, \underline{p}\right] - D(j\omega) \right|\right|_r \qquad (14)$$

where

$$||.||_r = \sum_{f=f_1}^{fmax} (.)^r$$

$$\underline{z} = [z_1, z_2, z_3 \cdots z_n]$$

$$\underline{p} = [p_1, p_2, p_3 \cdots p_n]$$

using standard techniques (Ref. 27). The speed of convergence of this optimisation is greatly improved (Ref. 28) by defining auxiliary variables

$$x_{2i-1} = (z_i + p_i)/2$$

$$x_{2i} = p_i/z_i \qquad i = 1, \ldots n$$

and finding

$$\min \left|\left| G\left[j\frac{2}{T}\sin\left(\frac{\omega T}{2}\right), \underline{x}\right] - D(j\omega) \right|\right|_r$$

where

$$\underline{x} = [x_1, x_2 \ldots\ldots x_{2n}]. \qquad (15)$$

For G to be realisable in the form of prototype C, the RC - realisability constraint

$$z_1 < p_1 < z_2 < p_2 \ldots\ldots < p_n$$

must be satisfied.

The analog prototype is then transformed and the SC network of Fig. 12 obtained. This network has the transfer function H(z) of equation (11), and will diverge from the desired equalisation characteristic due to the $z^{-\frac{1}{2}}$ delays added to satisfy realisability. Evaluating the transfer function H(z) of the SC network using equation (11), the optimised design becomes

$$\min \left|\left| H(z,\underline{\alpha}) - D(j\omega) \right|\right|_r$$

where

$$\underline{a} = [\alpha_1, \ldots\ldots \alpha_5] \qquad (16)$$

subject to the constraint that the poles and zeros of $H(z,\underline{\alpha})$ lie within the unit circle.

To summarise, the design procedure is as follows:

1. Optimise to obtain a predistorted analog prototype $G(j\omega)$ - equation (15).

2. Transform this analog prototype to obtain the SC network of Fig. 12 - equation (8).

3. Optimise the SC network to account for extra delays - equation (16).

Fleischer and Laker's biquadratic building block (Ref. 34) could be used to realise (11) directly. However the LDI transformed SC network of Fig. 12 has the advantage of tending to retain the performance of the analog prototype even for high frequency applications where the clock frequency is only 4 or 5 times the highest frequency of interest (Ref. 12), and consequently sensitivity properties are comparable with RC ladder network sensitivity. It appears that the network of Fig. 12 offers a saving of at least one capacitor and four switches over the Fleischer and Laker biquad for this application. For second order equalisers the choice between the two circuits could only be based on a detailed comparison. To construct higher order equalisers we could cascade Fleischer and Laker biquadratic building-blocks, or the method of equations (4) - (11) could be used to directly realise the higher order equaliser.

6. GENERAL EQUALISER STRUCTURE

To this point we have only considered the design of a fixed equaliser such as the mid-range equaliser for Fig. 6(a). In order to equalise the nominal cable characteristic over the full 0-4 km range a variable equaliser with the characteristics of Fig. 6(b) is required. The preferred method (Ref. 13,33) for varying the SC network transfer function is the introduction of variable capacitors of the type in Fig. 13. Here a combination of binary-weighted capacitors

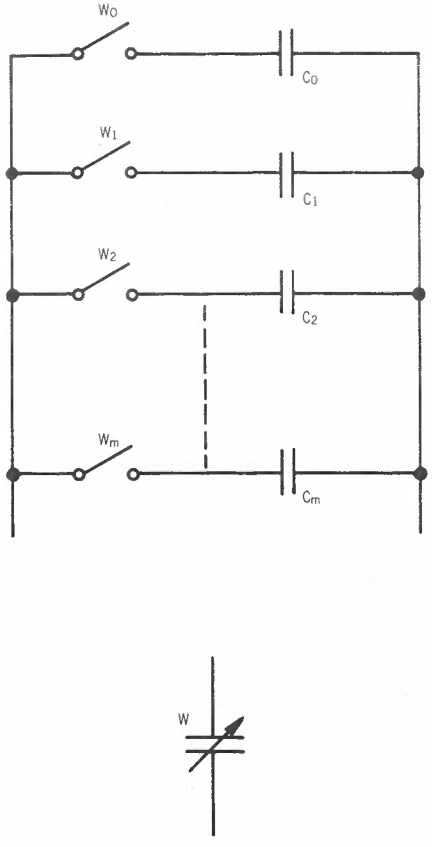


Fig. 13 - Binary-weighted variable capacitor

$$C_o = 2^m C_m , C_1 = 2^{m-1} C_m \ldots .$$

has its value controlled by a digital control word w where

$$w = w_0 w_1 w_2 w_3 \ldots w_m = 1011 \ldots 1$$

is composed of m bits.

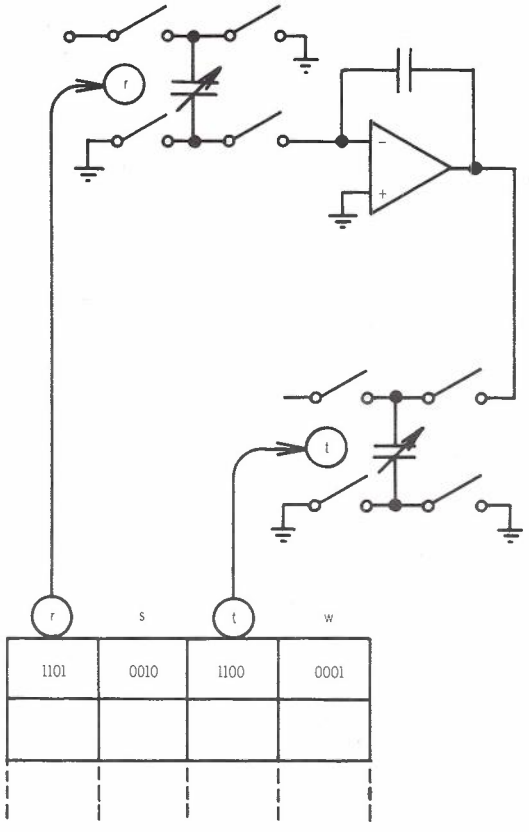


Fig. 14 - Control-word capacitor variation

We can then develop a general equaliser structure as in Fig. 14 where a number of control words r,s,t,w control a number of variable capacitors in the SC network. These control words may be derived from a digital store, and for example one row of the store could correspond to a particular equaliser gain setting - to adjust the equaliser setting we simply refer to another row in the table. This provides great flexibility in the adjustment of the equaliser since the control words in each row are limited only by the wordlength of the store. A general structure such as Fig. 14 is directly suggested by the developments presented in Ref. 13.

The structure of Fig. 14 includes transversal equalisers, recursive equalisers and decision-feedback equalisers as special cases since the coefficient store can be changed adaptively. It is interesting to contrast the

flexibility of Fig. 14 with classical analog equaliser structures (Ref. 29-32).

7. EXAMPLE

There are many different ways of producing a variable equaliser from the SC network of Fig. 12. Traditionally variable equalisers are designed with the minimum of variable components. However this criterion may not be appropriate in an integrated realisation such as Fig. 14. Since the operational amplifiers consume a large part of the chip area it may be desirable to minimise the number of op-amps.

The SC network of Fig. 15 can provide a full-range variable equaliser for .40 mm subscriber line with a maximum equalisation error < .3 dB in the frequency range 10 kHz - 160 kHz. The switch L is closed when the line length is greater than 2 km and open otherwise. The SC network is particularly economical in its use of op-amps to perform the variable equalisation.

Obviously the word length requirement of the control words is an important design criterion. If the control words vary rapidly over the equalisation range then a long wordlength will be required. Fig. 16 shows the variation in $(\alpha_1 \ldots \alpha_5)$ over the 0-4 km range, and indicates that an 8-bit wordlength is sufficient.

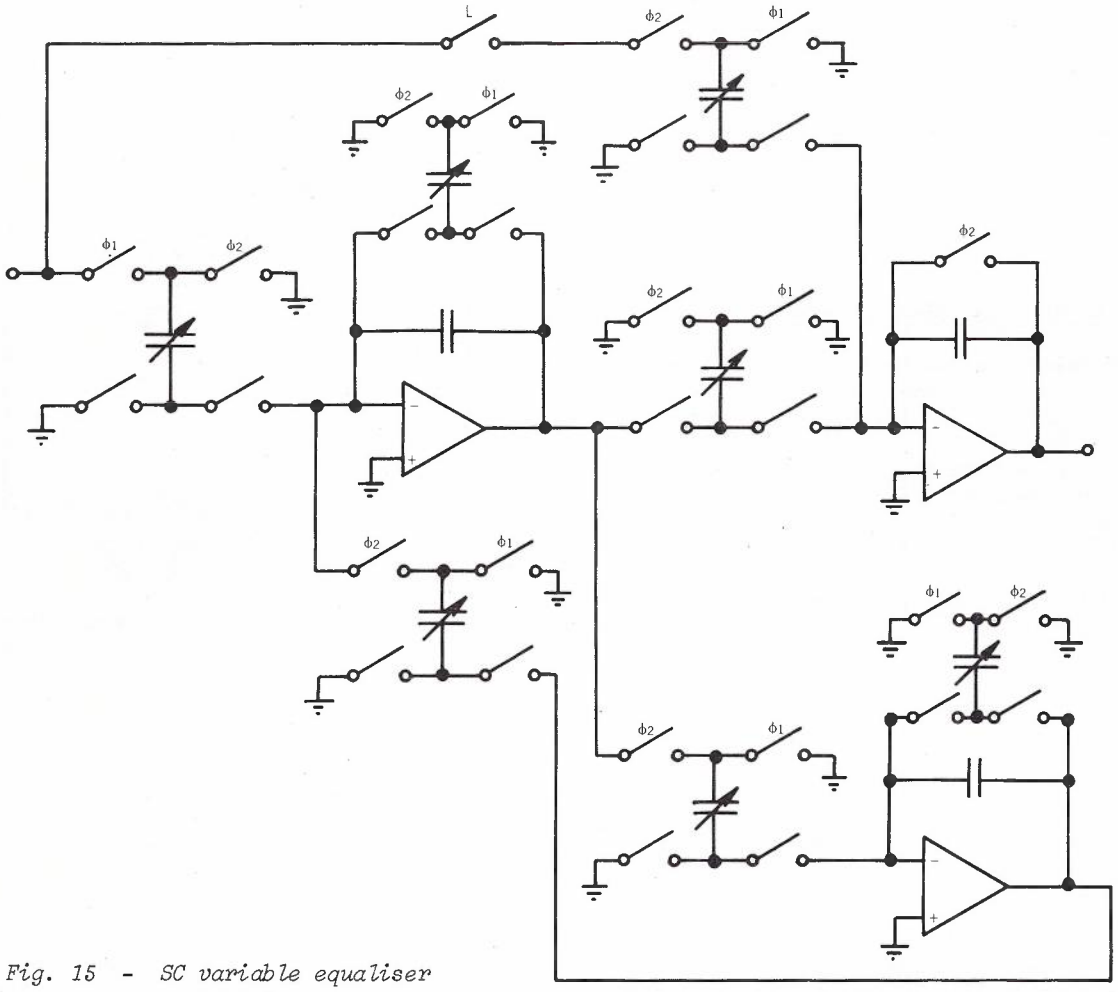


Fig. 15 - SC variable equaliser

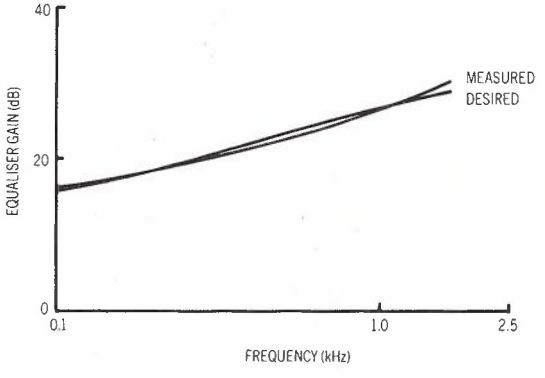


Fig. 16 - Comparison of experimental SC equaliser performance with desired equalisation characteristic

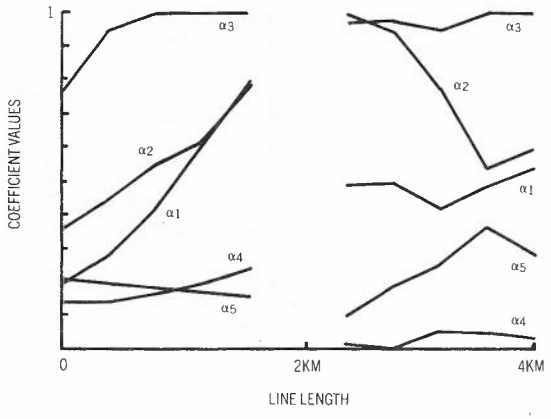


Fig. 17 - Control word variation 0-4 km equalisation

In order to verify the design procedure a discrete prototype of the fixed equaliser was constructed and tested. Fig. 17 shows the theoretical and actual equaliser characteristics with a frequency scaling factor of 1/100, i.e. f_c = 10 kHz. The experimental results have been corrected for sampling distortion (i.e. sin (ωT) / (ωT) distortion).

It would be preferable to construct a full clock rate discrete prototype, i.e. f_c = 1 MHz, but there are many practical obstacles. Referring to Appendix A2 we see that the settling time within a half clock cycle of the SC circuit is determined by the capacitance values, the switch resistance and the op-amp gain-bandwidth. In a discrete prototype the minimum practical capacitance value is much larger than for an integrated SC circuit and consequently the switch resistance must be reduced or the op-amp gain-bandwidth increased. Overriding these considerations is clock feedthrough via the switches. In an integrated version we can match devices and effectively eliminate clock feedthrough by the 'balanced switch' arrangement of Figs. 3 and 5. Again this is not possible in a discrete prototype. The most critical of the SC circuits is the high gain SC amplifier to compensate for cable flat loss. Particular care is needed to minimise clock feedthrough in this stage.

8. OTHER STRUCTURES

Since a typical subscriber line will not have nominal characteristics and may have gross imperfections, the cable equaliser is typically followed by some form of adaptive equalisation to improve performance. Fig. 18

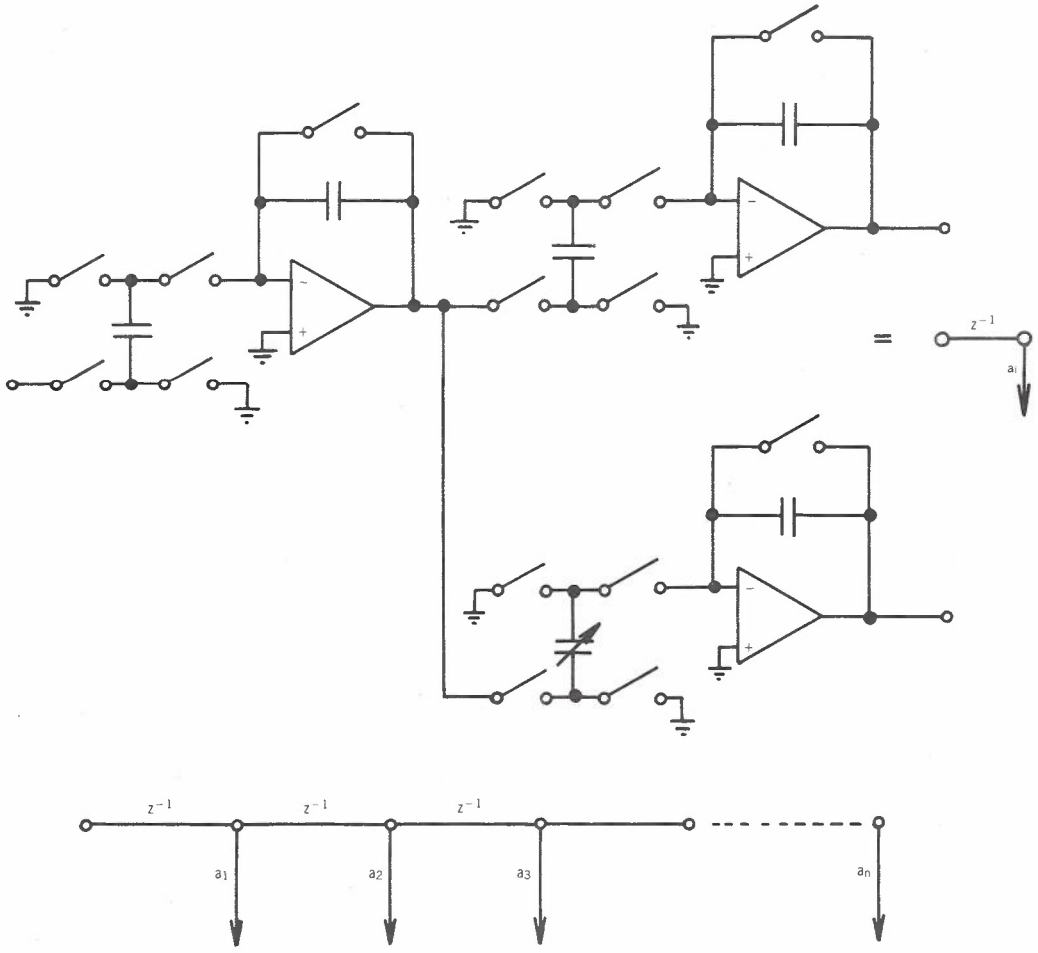


Fig. 18 - SC transversal equaliser

shows a transversal equaliser implemented using basic SC building blocks (Ref. 26). The transversal equaliser may be readily coupled with the nominal cable equaliser to give a complete TCM equalisation scheme. Obviously other adaptive equalisers can be constructed from the basic cell of Fig. 18.

9. CONCLUSIONS

Switched-capacitor equalisers suitable for cable equalisation in a TCM mode digital telephone can be derived by transformation of simple analog prototypes. This results in a variable equaliser that uses a minimum number of op-amps and consequently is attractive for integrated realisation. Coupled with existing SC adaptive equaliser structures this can form a complete equalisation scheme for TCM transmission.

The general structure of Fig. 14 admits many possible solutions to the equalisation problem and it is interesting to speculate that superior solutions to Fig. 15 may exist. The number of variable capacitors may be reduced, possibly at the expense of adding extra op-amps. Also the SC equalisers may possibly be combined with classical active RC equalisers if they can be realised easily on the same chip.

10. ACKNOWLEDGEMENTS

R. Swinton and R. Owers constructed and measured the prototype equaliser.

11. REFERENCES

1. Inoue, N., Komiya, R. and Inoue, Y., "Time-shared two-wire digital subscriber transmission system and its application to the digital telephone set", IEEE Trans. Communication, Vol. 29, No.11, November 1981, pp. 1565-1572.

2. Holte, N. and Stueflotten, "A new digital echo canceller for two-wire subscriber lines", IEEE Trans. Communication, Vol. 29, No.11, November 1981, pp. 1573-1581.

3. Falconer, D.D. and Mueller, K.H., "Adaptive echo cancellation AGC structures for two-wire, full-duplex data transmission", BSTJ, Vol. 58, No.7, September 1979, pp. 1593-1615.

4. Ahamed, S.V., "Simulation and design studies of digital subscriber lines", BSTJ, Vol. 61, No.6, July-August 1982, pp. 1003-1077.

5. Wurzburg, H. and Hillman, G., "CMOS chip set gives new life to twisted pairs for local networks", Electronics, September 22 1982.

6. Court, R.A. and Gale, N.J., "Completing the digital telecommunications network", Telecommunication Journal of Australia, Vol. 32, No.1, 1982, pp. 21-33.

7. Sun, Y., "Direct analysis of time-varying continuous and discrete difference equations with application to nonuniformly switched-capacitor circuits", IEEE Trans. Circuits and Systems, Vol. CAS-28, No.2, February 1981, pp. 93-100.

8. Lee, C.F. and Jenkins, W.K., "Computer-aided analysis of switched-capacitor filters", IEEE T. C&S, Vol. CAS-28, No.7, July 1981, pp. 681-691.

9. Tanaka, M and Mori, S., "Topological formulations for the coefficient matrices of state equations for switched-capacitor networks", IEEE T. C&S, Vol. CAS-29, No.2, February 1982, pp. 106-115.

10. Kurth, C.F. and Moschytz, G.S., "Nodal analysis of switched capacitor networks", IEEE T. C&S, Vol. CAS-26, No.2, February 1979, pp. 93-104.

11. Jacobs, G.M., Allstot, D.J., Broderson, R.W. and Gray, P.R. "Design techniques for MOS switched-capacitor ladder filters", IEEE T. C&S, Vol. CAS-25, No.12, December 1978, pp. 1014-1021.

12. Choi, T.C. and Broderson, R.W., "Considerations for high-frequency switched-capacitor ladder filters", IEEE T. C&S, Vol. CAS-27, No.6, June 1980, pp. 545-552.

13. Martin, K., "Improved circuits for the realisation of switched-capacitor filters", IEEE T. C&S, Vol. CAS-27, No.4, April 1980, pp. 237-244.

14. Haug, K., "Design, analysis and optimization of switched-capacitor filters derived from lumped analog models", AEU, Vol. 35, 1981, pp. 279-287.

15. Martin, K. and Sedra, A.S., "Effects of the op-amp finite gain and bandwidth on the performance of switched-capacitor filters", IEEE T. C&S, Vol. CAS-28, No.8, August 1981, pp. 822-829.

16. Luder, E., "Switched-capacitor filters insensitive to parasitics", AEU, Vol. 34, 1980, pp. 501-506.

17. Nossek, J.A. and Temes, G.C., "Switched-capacitor filter design using bilinear element modelling", IEEE T. C&S, Vol. CAS-27, No.6, June 1980, pp. 481-491.

18. Scanlan, S.O., "Analysis and synthesis of switched-capacitor state-variable filters", IEEE T. C&S, Vol. CAS-28, No.2, February 1981, pp. 85-93.

19. Hokonek, E. Brugger, U.W. and Moschytz, G.S., "New frequency transformation for the accurate design of SC ladder filters", Electronics Letters, Vol. 18, No.6, March 1982, pp. 276-278.

20. Herbst, D. et al., "VIS-SC-filters with reduced influences of parasitic capacitances", IEE Proc., Vol. 129, Pt.G, No.2, April 1982, pp. 29-39.

21. Martin, K. and Sedra, A.S., "Switched-capacitor building blocks for adaptive systems", IEEE T. C&S, Vol. CAS-28, No.6, June 1981, pp. 576-584.

22. Suzuki, T. and Shirasu, H., "Recent work on switched-capacitor circuits in Japan", IEEE Conference on Circuits and Systems, Conference Record, 1981.

23. Bode, H.W., "Network analysis and feedback amplifier design", Van Nostrand, N.Y., 1945.

24. Bruton, L.T., "Low-sensitivity digital ladder filters", IEEE T. C&S, Vol. CAS-22, No.3, March 1975, pp. 168-176.

25. Fettweis, A., "Realizability of digital filter networks", AEU Vol. 30, No.2, 1976, pp. 90-96.

26. Enomoto, T., Ishihara, T. and Tasumoto, M., "Integrated tapped delay-line using switched-capacitor technique", Electronics Letters, Vol. 18, No.5, pp. 193-194, 4 March 1982.

27. Chen, R.M-M., "A least-path optimisation algorithm without calculating derivatives", IEEE T. C&S, Vol. CAS-28, No.4, April 1981, pp. 331-336.

28. McGregor, I.M., private communication.

29. Bode, H.W., "Variable equalisers", BSTJ, Vol. 17, 1938, pp. 229-244.

30. Brglez, F, "Inductorless variable equalisers", IEEE T. C&S, Vol. CAS-22, No.5, May 1975, pp. 415-419.

31. Takasaki, Y., "Simple inductorless automatic line equaliser for PCM transmission using new variable transfer function", IEEE Trans. Communications, Vol. COM-26, No.5, May 1978, pp. 675-678.

32. Takasaki, Y., "Generalized theory of variable equalisers", Proceedings IEEE ISCAS 1979.

33. Allstot, D.J., Broderson, R.W. and Gray, P.R., "An electrically-programmable switched-capacitor filter", IEEE J. Solid-State Circuits, Vol. SC-14, December 1979, pp. 315-321.

34. Fleischer, P.E. and Laker, K.R., "A family of active switched capacitor biquad building blocks", Bell Syst. Tech. J., Vol. 58, No.10, December 1979, pp. 2235-2269.

APPENDIX A1

DERIVATION OF TRANSFER FUNCTIONS

The discrete-time transfer function of the amplifier of Fig. 3 can be obtained using Sun's exact method (Ref. 7), but here we present a less rigorous derivation. The reader is referred to Ref. 7 for further details.

If we assume that the amplifier is driven by a sinusoidal input signal

$$V_i = V_i \exp(j\omega t) \qquad (A1.1)$$

then during the time interval I_1 we can write the following state equation:

$$\begin{bmatrix} V_1(t) \\ V_o(t) \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} V_1(t_o^-) \\ V_o(t_o^-) \end{bmatrix} + \begin{bmatrix} -1 \\ 0 \end{bmatrix} V_i(t)$$

$$t \in I_1 \qquad (A1.2)$$

where t_o^- denotes the instant prior to t_o. It is assumed that the switches have zero on-resistance, and therefore the capacitors charge and discharge instantaneously. At $t = t_1$ we have

$$\begin{bmatrix} V_1(t_1^-) \\ V_o(t_1^-) \end{bmatrix} = \begin{bmatrix} -V_1(t_1^-) \\ 0 \end{bmatrix} \qquad (A1.3)$$

During I_2 we have

$$\begin{bmatrix} V_1(t) \\ V_o(t) \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ -C_1/C_0 & 0 \end{bmatrix} \begin{bmatrix} V_1(t_1^-) \\ V_o(t_1^-) \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \end{bmatrix} V_i(t)$$

$$t \in I_2 \qquad (A1.4)$$

$$\begin{bmatrix} V_1(t_3^-) \\ V_o(t_3^-) \end{bmatrix} = \begin{bmatrix} 0 \\ -C_1 V_1(t_1^-)/C_0 \end{bmatrix} . \qquad (A1.5)$$

where Fig. 4 shows the charge conservation principle applied to derive equation (A1.4). Substituting (A1.3) in (A1.5)

$$V_o(t_3^-) = (C_1/C_0)\, V_1(t_1^-) = (C_1/C_0)\, z^{-\frac{1}{2}}\, V_i(t_3^-) \qquad (A1.6)$$

since from (A1.1):

$$V_i(t_1^-) = V_i \exp(j\omega t_1^-)$$

$$= V_i \exp(j\omega t_3^-) \exp(-j\omega T/2)$$

and $z = \exp(j\omega T)$

Similarly for the integrator of Fig. 5,

$$\begin{bmatrix} V_1(t) \\ V_2(t) \\ V_3(t) \\ V_0(t) \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} V_1(t_o^-) \\ V_2(t_o^-) \\ V_3(t_o^-) \\ V_0(t_o^-) \end{bmatrix} + \begin{bmatrix} -1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} V_i(t) \\ V_j(t) \end{bmatrix}$$

$t \in I_1$ (A1.7)

and

$$\begin{bmatrix} V_1(t_1^-) \\ V_2(t_1^-) \\ V_3(t_1^-) \\ V_0(t_1^-) \end{bmatrix} = \begin{bmatrix} -V_i(t_1^-) \\ V_j(t_1^-) \\ -V_o(t_o^-) \\ V_o(t_o^-) \end{bmatrix} \quad \text{(A1.8)}$$

During I_2 we have

$$\begin{bmatrix} V_1(t) \\ V_2(t) \\ V_3(t) \\ V_0(t) \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -C_1/C_0 & -C_2/C_0 & C_3/C_0 & 1 \end{bmatrix} \begin{bmatrix} V_1(t_1^-) \\ V_2(t_1^-) \\ V_3(t_1^-) \\ V_0(t_1^-) \end{bmatrix}$$

$t \in I_2$ (A1.9)

$$V_0(t_3^-) = [-C_1V_1(t_1^-) - C_2V_2(t_1^-) + C_3V_3(t_1^-)]/C_0 + V_0(t_1^-) \quad \text{(A1.10)}$$

Substituting (A1.8) in (A1.10) we find

$$V_0 = \frac{C_1}{C_0} \frac{z^{-\frac{1}{2}}}{1-z^{-1}} V_i - \frac{C_2}{C_0} \frac{z^{-\frac{1}{2}}}{1-z^{-1}} V_j - \frac{C_3}{C_0} \frac{z^{-1}}{1-z^{-1}} V_0 \quad \text{(A1.11)}$$

APPENDIX A2

PRACTICAL LIMITATIONS OF SC AMPLIFIERS

Fig. A2.1 shows the equivalent circuit of the SC amplifier when $t \in I_2$. The state equations then become

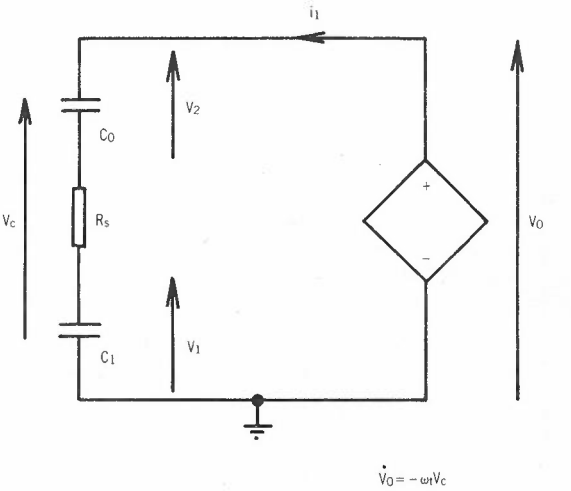


Fig. A2.1 - SC amplifier - equivalent circuit

$$C_1\dot{V}_1 = \frac{1}{R_s}(V_0 - V_2 - V_1)$$

$$C_0\dot{V}_2 = \frac{1}{R_s}(V_0 - V_2 - V_1)$$

$$\dot{V}_0 = -\omega_+(V_0 - V_2)$$

or

$$\begin{bmatrix} \dot{V}_1 \\ \dot{V}_2 \\ \dot{V}_0 \end{bmatrix} = \begin{bmatrix} \frac{-1}{R_sC_1} & \frac{-1}{R_sC_1} & \frac{1}{R_sC_1} \\ \frac{-1}{R_sC_0} & \frac{-1}{R_sC_0} & \frac{1}{R_sC_0} \\ 0 & \omega_+ & -\omega_+ \end{bmatrix} \begin{bmatrix} V_1 \\ V_2 \\ V_0 \end{bmatrix} \quad \text{(A2.1)}$$

with initial conditions $V_1 = -V_i$, $V_2 = V_0 = 0$. Taking the Laplace transform of (A2.1) we have

$$\begin{bmatrix} V_1 \\ V_2 \\ V_0 \end{bmatrix} = \begin{bmatrix} S+p_1 & p_1 & -p_1 \\ p_2 & S+p_2 & -p_2 \\ 0 & -\omega_+ & S+\omega_+ \end{bmatrix}^{-1} \begin{bmatrix} -V_i \\ 0 \\ 0 \end{bmatrix} \quad \text{(A2.2)}$$

where $p_1 = \frac{1}{R_s C_1}$ and $p_2 = \frac{1}{R_s C_0}$

Solving A2.2 we obtain

$$V_0 = \frac{p_2\omega_t}{S[S^2 + (p_1 + p_2 + \omega_t)S + p_1\omega_t]} V_i \qquad (A2.3)$$

and expanding

$$V_0 = \frac{1}{S}\left(\frac{p_2}{p_1}\right) + \frac{K_1}{S + a_1} + \frac{K_2}{S + a_2}$$

where

$$a_1, a_2 = \frac{-(p_1 + p_2 + \omega_t) \pm \sqrt{(p_1 + p_2 + \omega_t)^2 - 4p_1\omega_t}}{2}$$

$$K_1 = \frac{-p_2\omega_t}{a_1(a_2 - a_1)}$$

$$K_2 = \frac{-p_2\omega_t}{a_2(a_1 - a_2)}$$

The circuit response for $t \in I_2$ in response to an initial charge on capacitor C_1 then becomes

$$V_0 = V_i \left(\frac{C_1}{C_0} + K_1 e^{-a_1 t} + K_2 e^{-a_2 t}\right) \qquad (A2.5)$$

Here the first term is the desired steady-state value and the decaying exponential terms are due to the finite op-amp gain and finite switch resistance. The deviation in the transfer function can be obtained directly from equation A2.5 by the magnitude of the exponential terms at the end of the time interval I_2,

$$\frac{\Delta T_A}{T_A} = \frac{|K_1 e^{-a_1 t_3^-} + K_2 e^{-a_2 t_3^-}|}{C_1/C_0} \qquad (A2.6)$$

BIOGRAPHY

ANDREW J. JENNINGS was born in Melbourne in 1952. He received the B.Eng. Hons degree in 1975 from Monash University and the Ph.D degree in 1979. The Ph.D thesis concerned the analysis of acoustic waveguides and surface acoustic wave devices. In 1979 he joined Telecom Australia, initially in the Engineering Department where he worked on problems of primary level PCM transmission. In 1980 he joined the Circuit and System Theory Section, Transmission Branch, Research Department where he is concerned with the application of advanced signal processing techniques in telecommunication networks. Dr. Jennings is currently an acting Class 3 Engineer in the Transmission Branch, Telecom Australia Research Laboratories.

Optimal Capacity Assignment In Packet-Switching Networks

M.J.T. NG
D.B. HOANG

Department of Electronic and Communication Science
La Trobe University

In this paper we consider the capacity assignment problem in a packet-switching network. Each communication link in the network is realised by a collection of parallel transmission lines. To obtain the optimal capacity assignment, i.e. finding the optimal number of transmission lines for each link, network queueing models are first derived to describe the network in terms of the design parameters, namely: link flow, link capacity, and the average message delay in the network. The capacity assignment problem is then formulated as an optimization problem which minimizes the overall capacity cost, subject to an average message delay constraint. A continuous variable approximation method based on the Lagrange Multiplier technique and a dynamic programming method with efficient boundings are developed to solve the optimization problem. These techniques are illustrated with one of the planned topologies of the AUSTPAC network. The dynamic programming technique yields the optimal capacity assignment.

1. INTRODUCTION

An important design issue in packet-switching networks is the capacity assignment of the communication subnet. Indeed, as the communication subnet is used for transporting data and network control traffic - in the form of data packets - and in a store-and-forward fashion, cost-effective design of the subnet is a basic requirement in optimizing the network throughput and minimizing packet delay. Over the past decade many design techniques have been developed for solving the capacity assignment problems in various computer networks, and had different degrees of success in finding optimal solutions (Refs. 1-8).

The general design problem is to find the optimal channel capacities of the subnet which optimize a well-defined objective function for a given network topology and subnet traffic flow, subject to a set of design constraints. The ability to solve the design problem depends critically on how successfully one can express both objective functions and constraints in analytically manageable form as a function of the capacity variable. Network models are therefore an essential prerequisite to the design problem. For continuous capacity variable problems, the well known "square-root" assignment was developed in Kleinrock (Ref. 10) using linear costs and the Lagrange Multiplier technique. The design approach and the underlying network model was successfully applied in the design of the ARPA network. In (Ref. 6) the important concave cost function and the power law cost function for capacity assignment were discussed. Linear programming techniques were developed in (Ref. 8) for general computer networks and in (Ref. 7) for hybrid voice-data networks.

Channel capacities are, however, usually only available in a finite and discrete set (e.g. 48, 64, 96 kbps). Computationally efficient methods of choosing these capacities approximate the discrete capacity/cost functions by continuous functions (e.g. concave cost function) and solve the continuous variable problem. In a refinement stage, the optimal continuous variable solution would be discretized (Ref. 1). For centralized computer networks with tree structures, Frank *et al* (Ref. 2) developed an efficient dynamic-programming-like algorithm for searching the optimal discrete capacities. Integer programming (such as the branch and bound technique) can also be used for the discrete variable design; however, these techniques are usually rather mathematically intractable (Ref. 2) and computationally inefficient for networks of practical sizes. A dynamic programming approach applied to the capacity assignment was reported in Ref. 6.

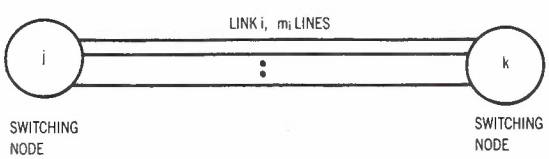


Fig. 1 - A communication link with m_i transmission lines

In this paper we consider the capacity assignment problem for a distributed packet-switching network in which each channel in the communication subnet is realized by a collection of identical parallel transmission lines, each having a fixed capacity C (an

example of such a network is AUSTPAC, (Ref. 15), see Figure 1. The channel implementation is based on a modular hardware and software design which provides the flexibility for network expansion and topology changes. The capacity assignment is therefore a discrete variable problem of finding the number of transmission lines per channel. Although techniques for solving discrete capacity problems are available, this paper contributes in the following aspects:

1) Some basic network queueing models for modelling the communication subnet based on the well-known "independence assumption" (Ref. 9) are developed. The queueing models developed not only reflect the operational features of the network, but are also incorporated in the capacity assignment problem to yield an analytically manageable objective function and constraints for optimization;

2) A detailed dynamic programming algorithm with efficient bounds for finding the optimal discrete capacities is developed. The dynamic programming algorithm can also be used with other suitable queueing models;

3) A capacity assignment technique based on a continuous variable approximation and the Lagrange Multiplier technique is also developed. These techniques are illustrated through their application to the study of the AUSTPAC network.

In Section 2 the network queueing model used for the capacity assignment is presented. A multiple queue/server structure is used to model each communication channel and find expressions for the average message delay in the network. In Section 3 the capacity assignment problem is formulated. In Section 4, we attempt to solve the capacity assignment problem by a continuous variable approximation method; closed-form analytical solutions are obtained. In Section 5, we proceed to first illustrate the dynamic programming approach by a simple network design example; then we present a detailed algorithm with bounding techniques for improving the computation efficiency. Applications of both techniques to a proposed AUSTPAC topology are presented in Section 6.

2. NETWORK QUEUEING MODEL

Perhaps one of the most useful and far-reaching assumptions employed in the design of packet-switching networks is the independence assumption, used by Kleinrock throughout his research with the ARPA network (Ref. 9). The independence assumption suggests that although a computer network has a complex collection of interdependent queues and serving disciplines for data packets at each switching node, the overall behaviour of the network can be analysed in much the same way as if each switching node had independent queues. The assumption, which has been successfully applied in the design of the ARPA network, allows the simplification of the otherwise complicated and usually analytically difficult queueing problem, to an analytically manageable form, suitable for many network optimization problems. In this section, by employing the same assumption together with a "random queue assignment" strategy, a multiple queue/server model is developed for finding the average message delay T in the network. The average message delay, an important performance index in many network design problems (Refs. 12-14), is used as a constraint function in the capacity assignment.

We consider a partially distributed network with N switching nodes and M full-duplex links (channels). It is more convenient to consider these M full-duplex links as 2M simplex links. Label these 2M simplex links by i = 1,2, ... 2M, such that the ith and M+i th simplex links represent one full-duplex link. Suppose m_i transmission lines are used to implement the ith simplex link, then the capacity of link i will be $m_i C$. If the traffic requirement of the network, i.e. the amount of traffic generated by each origin-destination node pair, and the traffic routing are known, the traffic flow λ_i (in packets per second) on each link i can be obtained.

At each switching node, data packets arriving from different incoming links and hosts are transferred to outgoing links or hosts according to traffic routing (bifurcated or non-bifurcated). This kind of traffic mixing provides much of the traffic decorrelation required in establishing the independence assumption introduced by Kleinrock - the traffic entering a link can be well approximated by a random Poisson process. To find out the average message delay T in a network having the parallel-transmission-lines structure (Figure 1), suppose a data packet entering a link is assigned randomly to one of the transmission lines in the link with equal probability, then the arrival process to each transmission line with also be a Poisson process (a decomposition theorem stated in Kuehn, Ref. 13). In much the same way that Kleinrock used a M/M/1 queue (i.e. Poisson input, unlimited buffers and a server with exponentially distributed service times) to model a link with one transmission line, we model each transmission line by a M/M/1 queue. A link implemented by m_i transmission lines is therefore modelled as a m_i-M/M/1 queue, see Figure 2. The average packet delay in link i is given by (Ref. 14)

$$T_i = \frac{m_i}{m_i \mu C - \lambda_i} \qquad (1)$$

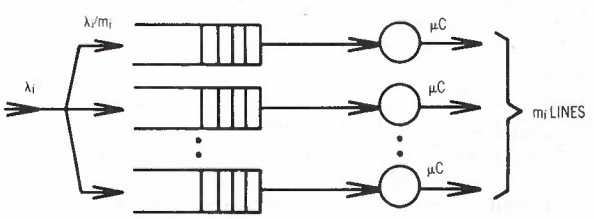


Fig. 2 - A m_i-M/M/1 queueing model

where μ^{-1} is the average packet size in bits.

Thus the average delay in a m_i-M/M/1 queueing network with a total traffic requirement γ (in packets per second) is given by

$$T = \frac{1}{\gamma} \sum_{i=1}^{2M} \frac{m_i \lambda_i}{m_i \mu C - \lambda_i} \qquad (2)$$

This model can be further enhanced by considering the propagation delay t_i for each link i and nodal processing delay K as follows:

$$T = (\bar{n}+1)K + \frac{1}{\gamma} \sum_{i=1}^{2M} \lambda_i \left[\frac{m_i}{m_i \mu C - \lambda_i} + t_i \right] \qquad (3)$$

where $\bar{n} = \frac{1}{\gamma} \sum_{i=1}^{2M} \lambda_i$ is the average number of hops per message in the network.

3. THE CAPACITY ASSIGNMENT PROBLEM STATEMENT

The capacity assignment problem for the packet-switching network studied in Section 2 can be stated as follows:

Given: Network topology: N nodes, M full-duplex links (2M simplex links);

Link traffic flow λ_i, i = 1,2, ... 2M.

Minimize: Total cost D of all link capacities over the discrete variable $m_i, m_i = 1,2, \ldots$; i = 1,2, ... 2M,

where $D = \sum_{i=1}^{2M} D_i(m_i)$

$D_i(m_i)$ = cost of implementing m_i transmission lines for the simplex link i.

Subject to: Average message delay T to be less than or equal to a specified parameter T_{MAX}.

Apart from the above formulation, a dual capacity assignment problem by interchanging the functions employed in the objective function (for minimization) and the constraint function can also be formulated (Ref. 14).

4. A CONTINUOUS VARIABLE APPROXIMATION METHOD FOR DISCRETE CAPACITY ASSIGNMENT

Although the problem formulation presented is clearly a discrete variable problem, we shall attempt to solve it analytically by approximating the discrete variables with continuous functions, saving a discrete variable optimization method for the next section.

Suppose each design variable m_i can take on any real positive value which can be decomposed into an integer n_i, $n_i = 0,1,2, \ldots$ and a positive residue ϕ_i, $0 < \phi_i \leq 1$, i.e.

$$m_i = n_i + \phi_i \qquad (4)$$

Then the link i can be implemented by n_i transmission lines each of capacity C, plus one more line with capacity $\phi_i C$. Furthermore, suppose data packets are also assigned randomly to these lines with probability proportional to the line capacity, then each of the n_i transmission lines with capacity C will receive a packet arrival rate λ_i/m_i while the line with capacity $\phi_i C$ will receive a rate $\phi_i \lambda_i/m_i$. The average packet delay on the ith link is given by (Ref. 14)

$$T_i = \frac{1}{\mu C} \left(\frac{n_i+1}{m_i} \right) \left(\frac{1}{1-\rho_i} \right) \qquad (5)$$

where $\rho_i = \lambda_i/\mu C m_i$.

Typical values of T_i are plotted in Figure 3 with normalized $\mu C = 1$. T_i is a discontinuous function of m_i, which can be approximated by the continuous function T_i' (shown by a dotted line).

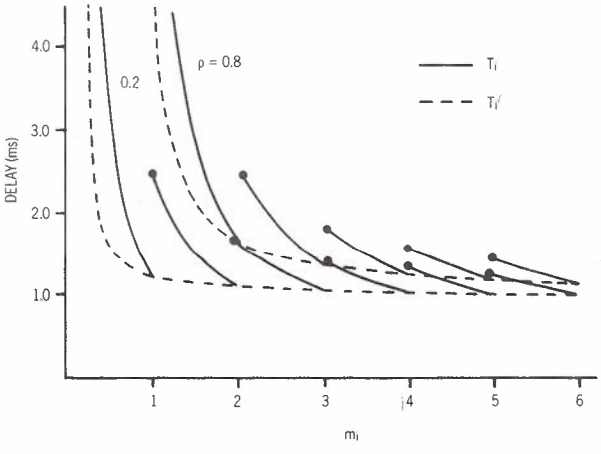


Fig. 3 - Delay characteristics T_i and T_i'

$$T_i' = \frac{1}{\mu C} \left(\frac{1}{1-\rho_i} \right) \qquad (6)$$

T_i' approximates T_i well for large values of m_i, and yields exactly the same value as T_i when m_i takes an integer value.

4.1 Linear Cost Function and the Lagrange Multiplier Method

Suppose the cost of each simplex link is directly proportional to the physical length

DX_i of the link and the number of transmission lines in the link, given by

$$D_i(m_i) = m_i\, e_i\, DX_i + D_{io}$$

$$= m_i\, E_i + D_{io} \qquad (7)$$

where $D_i(m_i)$ is the cost of link i with m_i transmission lines,
e_i is the cost per transmission line per unit distance for the simplex link i,
D_{io} is an initial set up cost,

and

$$E_i = e_i DX_i \text{ (cost per line).} \qquad (8)$$

The Lagrange Multiplier method (Ref. 11) can be employed here to yield the optimal values of m_i (in a continuous variable sense with respect to the delay expression T_i') as follows:

Forming the Lagrangian G,

$$G = \sum_{i=1}^{2M} D_i(m_i) + \beta T \qquad (9)$$

where β is the Lagrange Multiplier,

and T is the average message delay, given by

$$T = (\bar{n}+1)K + \frac{1}{\gamma} \sum_{i=1}^{2M} \lambda_i\, (T_i' + t_i) \qquad (10)$$

and setting the partial derivatives of G with respect to m_i equal to zero; i.e.

$$\frac{\partial G}{\partial m_i} = 0, \quad i = 1,2, \ldots 2M,$$

we obtain, after some algebraic manipulations, the optimal value of m_i given by

$$m_i = \frac{\lambda_i}{\mu C}\left[1 + \frac{1}{\gamma\sqrt{E_i}} \frac{\sum_{j=1}^{2M} \lambda_j \sqrt{E_j}}{(\mu C T'_{MAX} - \bar{n})}\right] \qquad (11)$$

where

$$T'_{MAX} = T_{MAX} - (\bar{n}+1)K - \frac{1}{\gamma} \sum_{i=1}^{2M} \lambda_i t_i \qquad (12)$$

The cost of assigning these m_i's to the network is a minimal cost given by

$$D = \sum_{i=1}^{2M} D_{io} + \frac{1}{\mu C} \sum_{i=1}^{2M} \lambda_i E_i + \frac{1}{\gamma \mu C} \frac{\left[\sum_{i=1}^{2M} \lambda_i \sqrt{E_i}\right]^2}{(\mu C T'_{MAX} - \bar{n})} \qquad (13)$$

The m_i in (11) can be seen proportional to the flow λ_i - this is in contrast to the well-known square-root assignment obtained by using a M/M/1 queueing network (Ref. 10). Equation (11) also shows an interesting property i.e. to ensure all m_i are positive, T'_{MAX} should be greater than $\bar{n}/\mu C$. In other words the average message delay in the network will be always greater than $\bar{n}/\mu C$ no matter how many lines (positive m_i) are to be assigned. This property is of course rather obvious since (i) the average number of hops $\bar{n} \geq 1$; and (ii) the average packet delay is always greater than or equal to the average service time $1/\mu C$.

Finally, notice that equation (11) yields different optimal numbers of transmission lines m_i and m_{M+i} for each full duplex link. The value $\bar{m}_i$ which is the smallest integer greater than or equal to the largest value of m_i and m_{M+i} can be adopted for discrete capacity assignment.

5. THE DYNAMIC PROGRAMMING METHOD

Although the Lagrange Multiplier technique yields rather quick analytical solutions for capacity assignment, the results obtained are sub-optimal because of the approximation steps taken in the method. In this section, we relax these approximations by developing an efficient dynamic programming technique to obtain the true optimal discrete capacity solutions. The dynamic programming technique developed is not only useful for the parallel-transmission-line network model, but in fact useful for any networks which satisfy the following three constraints:

1) Finite and Discrete Capacity Options such that

$$C_k(1) > C_k(2) > \ldots C_k(i_k) > \ldots C_k(L_k) \qquad (14)$$

where $C_k(i_k)$ is a capacity option for link k, $i_k = 1, \ldots, L_k$
L_k is the number of options for link k, k = 1, ..., M.

2) If $D_k(i_k)$ is the cost of the capacity $C_k(i_k)$, then

$$D_k(1) > D_k(2) > \ldots D_k(i_k) > \ldots D_k(L_k) \qquad (15)$$

In other words, higher capacity corresponds to higher cost; and

3) If the average delay for link k is $T_k(i_k)$ with capacity $C_k(i_k)$, then

$$T_k(1) < T_k(2) < \ldots < T_k(i_k) \ldots < T_k(L_k) \qquad (16)$$

5.1 A Multistage Decision Process Formulation for Capacity Assignment

Dynamic programming has been traditionally used for solving multistage decision processes by decomposing a problem with M decision variables into M stages of smaller sub-problems (Ref. 16). It is expected that these sub-problems can be solved with better efficiency and their solutions can be combined forming an overall solution to the original problem.

Most of the studies of the dynamic programming are based on Bellman's principle of optimality (Ref. 16) which is used below in formulating a multistage decision process for the discrete capacity assignment (notice that a full-duplex link is handled directly in the formulation below):

Minimize Terminal Cost D

$$\min D = \min_{I^M} J_M(I^M) = \min_{I^M} \sum_{k=1}^{M} D_k(i_k) \tag{17}$$

$$I^M = [i_1, i_2, \ldots i_M] \tag{18}$$

where $D_k(i_k)$ is the cost of a full-duplex link k with capacity option i_k for both directions of traffic flow, and I^M is a decision vector for capacity assignment. Since the cost function D is separable, by the Principle of Optimality, min D can be decomposed into

$$\min_{I^k} J_k(I^k) = \min_{I^k} \sum_{\ell=1}^{k} D_\ell(i_\ell)$$

$$= \min_{i_k} [J_{k-1}(I^{k-1}) + D_k(i_k)]; \tag{19}$$

$$J_0(I^0) = 0 \tag{20}$$

for $k = 1, 2, \ldots M$

where $J_k(I^k)$ is the partial cost function.

Subject to Terminal Delay T

$$S_M(I^M) = T = \frac{1}{\gamma} \sum_{k=1}^{M} [\lambda_k^1 T_k^1(i_k) + \lambda_k^2 T_k^2(i_k)]$$

$$\leq T'_{MAX} \tag{21}$$

$$S_k(I^k) = \frac{1}{\gamma} \sum_{\ell=1}^{k} [\lambda_\ell^1 T_\ell^1(i_\ell) + \lambda_\ell^2 T_\ell^2(i_\ell)]$$

$$= S_{k-1}(I^{k-1}) + T_k(i_k) \tag{22}$$

$$S_0(I^0) = 0 \tag{23}$$

where $S_k(I^k)$ is the partial delay;

λ_k^1 and λ_k^2 are used to denote the traffic flows for each direction of the full-duplex link k;

$T_k^1(i_k)$ and $T_k^2(i_k)$ denote the delay for each traffic direction if the capacity option i_k is used.

$T_k(i_k)$ is a delay function for each full-duplex link, given by

$$T_k(i_k) = \frac{1}{\gamma} [\lambda_k^1 T_k^1(i_k) + \lambda_k^2 T_k^2(i_k)]; \tag{24}$$

and

$$T'_{MAX} = T_{MAX} - (\bar{n}+1) K - \frac{1}{\gamma} \sum_{k=1}^{M} [\lambda_k^1 + \lambda_k^2] t_k \tag{25}$$

to include the effect of nodal processing delay and propagation delay for a specified average delay T_{MAX}.

5.2 The Dynamic Programming Technique - An Illustrative Example

The technique for solving (17)-(25) is first illustrated with a simple network example. A generalized algorithm is then described. Consider a 3-link network with the available cost $D_k(i_k)$ and delay function $T_k(i_k)$ shown in Table 1.

TABLE 1 - Cost and Delay Functions for the 3-Link Network

	i_1	1	2	3
Link 1	$T_1(i_1)$ ms	100	80	70
k = 1	$D_1(i_1)$ \$	10	20	35

	i_2	1	2	3
Link 2	$T_2(i_2)$ ms	70	65	60
k = 2	$D_2(i_2)$ \$	5	20	30

	i_3	1	2
Link 3	$T_3(i_3)$ ms	45	40
k = 3	$D_3(i_3)$ \$	5	25

The average packet delay T of the network is required to be $T \leq T_{MAX} = 200$ ms, with

5.1 A Multistage Decision Process Formulation for Capacity Assignment

Dynamic programming has been traditionally used for solving multistage decision processes by decomposing a problem with M decision variables into M stages of smaller sub-problems (Ref. 16). It is expected that these sub-problems can be solved with better efficiency and their solutions can be combined forming an overall solution to the original problem.

Most of the studies of the dynamic programming are based on Bellman's principle of optimality (Ref. 16) which is used below in formulating a multistage decision process for the discrete capacity assignment (notice that a full-duplex link is handled directly in the formulation below):

Minimize Terminal Cost D

$$\min D = \min_{I^M} J_M(I^M) = \min_{I^M} \sum_{k=1}^{M} D_k(i_k) \tag{17}$$

$$I^M = [i_1, i_2, \ldots i_M] \tag{18}$$

where $D_k(i_k)$ is the cost of a full-duplex link k with capacity option i_k for both directions of traffic flow, and I^M is a decision vector for capacity assignment. Since the cost function D is separable, by the Principle of Optimality, min D can be decomposed into

$$\min_{I^k} J_k(I^k) = \min_{I^k} \sum_{\ell=1}^{k} D_\ell(i_\ell)$$

$$= \min_{i_k} [J_{k-1}(I^{k-1}) + D_k(i_k)]; \tag{19}$$

$$J_0(I^0) = 0 \tag{20}$$

for $k = 1, 2, \ldots M$

where $J_k(I^k)$ is the partial cost function.

Subject to Terminal Delay T

$$S_M(I^M) = T = \frac{1}{\gamma} \sum_{k=1}^{M} [\lambda_k^1 T_k^1(i_k) + \lambda_k^2 T_k^2(i_k)]$$

$$\leq T'_{MAX} \tag{21}$$

$$S_k(I^k) = \frac{1}{\gamma} \sum_{\ell=1}^{k} [\lambda_\ell^1 T_\ell^1(i_\ell) + \lambda_\ell^2 T_\ell^2(i_\ell)]$$

$$= S_{k-1}(I^{k-1}) + T_k(i_k) \tag{22}$$

$$S_0(I^0) = 0 \tag{23}$$

where $S_k(I^k)$ is the partial delay;

λ_k^1 and λ_k^2 are used to denote the traffic flows for each direction of the full-duplex link k;

$T_k^1(i_k)$ and $T_k^2(i_k)$ denote the delay for each traffic direction if the capacity option i_k is used.

$T_k(i_k)$ is a delay function for each full-duplex link, given by

$$T_k(i_k) = \frac{1}{\gamma} [\lambda_k^1 T_k^1(i_k) + \lambda_k^2 T_k^2(i_k)]; \tag{24}$$

and

$$T'_{MAX} = T_{MAX} - (\bar{n}+1) K - \frac{1}{\gamma} \sum_{k=1}^{M} [\lambda_k^1 + \lambda_k^2] \dagger_k \tag{25}$$

to include the effect of nodal processing delay and propagation delay for a specified average delay T_{MAX}.

5.2 The Dynamic Programming Technique - An Illustrative Example

The technique for solving (17)-(25) is first illustrated with a simple network example. A generalized algorithm is then described. Consider a 3-link network with the available cost $D_k(i_k)$ and delay function $T_k(i_k)$ shown in Table 1.

TABLE 1 - Cost and Delay Functions for the 3-Link Network

	i_1	1	2	3
Link 1	$T_1(i_1)$ ms	100	80	70
k = 1	$D_1(i_1)$ \$	10	20	35

	i_2	1	2	3
Link 2	$T_2(i_2)$ ms	70	65	60
k = 2	$D_2(i_2)$ \$	5	20	30

	i_3	1	2
Link 3	$T_3(i_3)$ ms	45	40
k = 3	$D_3(i_3)$ \$	5	25

The average packet delay T of the network is required to be $T \leq T_{MAX} = 200$ ms, with

negligible nodal processing and propagation delay.

Using the dynamic programming formulation, this example is now considered as a 3-stage decision process. At each stage, the capacity for one particular link is determined, called a partial capacity assignment. Such a partial assignment sets the multistage decision process into states characterized by the partial delay $S_k(I^k)$ and the partial cost $J_k(I^k)$.

Stage 1 - Link 1

The partial assignment of link 1 can be made by choosing from $i_1 = 1,2,3$; see Table 2. Each choice of these values is in fact an optimal choice i_1^*. All the optimal values i_1^* and their corresponding optimal partial cost and delay are stored for stage 2 computation.

TABLE 2 - Partial Assignment - Stage 2

i_1^*	i_2	$S_2(I^2)$ (ms)	$J_2(I^2)$ ($)	Optimal State
3	3	130	65	✓
3	2	135	55	✓
2	3	140	50	
3	1	140	40	✓
2	2	145	40	
2	1	150	25	✓
1	3	160	40	
1	2	165	30	
1	1	170	15	✓

Stage 2 - Link 2

The partial assignment of link 2 together with the optimal partial assignment of link 1 are shown in Table 2. The partial delays $S_2(I^2)$ in the third column are arranged in a monotonic non-decreasing order. Column 1 and Column 2 indicate the value of i_1^* and i_2 respectively. The partial cost $J_2(I^2)$ is shown in Column 4. Column 5 indicates which partial assignment is an optimal state (with a ✓).

States Elimination

Table 2 shows that there are 9 states in stage 2. There are some non-optimal states, i.e. non-optimal partial assignment, among these states. From Bellman's principle of optimality, the non-optimal states can be eliminated from any further consideration.

Take the partial assignment in the 3rd row as an example. This assignment is non-optimal because one can find another partial assignment, in the 4th row, which yields the same partial delay yet at a smaller partial cost.

The partial assignment in the 7th row is another non-optimal assignment because the partial assignment in the 6th row yields a smaller delay at a smaller cost. Rows 5 and 8 are non-optimal for the same reason. The number of the optimal states at this stage is therefore 5.

In general, if the partial delay $S_k(I^k)$ and the partial cost $J_k(I^k)$ are also labelled as $S_k(\ell)$ and $J_k(\ell)$ respectively for $\ell = 1,2, \ldots \Pi_k$ where Π_k is the total number of such partial states in stage k; then the optimal states $[S_k^*(\ell), J_k^*(\ell)]$, $\ell = 1,2, \ldots \Pi_k^*$, satisfy the following criterion.

$$\text{If} \quad S_k^*(\ell) < S_k^*(\ell')$$

$$\text{Then} \quad J_k^*(\ell) > J_k^*(\ell') \qquad \ell \neq \ell'$$

$$\ell, \ell' = 1,2, \ldots \Pi_k^* \qquad (26)$$

Stage 3 - Link 3

Now apply the same procedure to the 3rd stage. The results obtained are shown in Table 3.

TABLE 3 - Partial Assignment - Stage 3

i_1^*	i_2^*	i_3	$S_3(I^3)$ (ms)	$J_3(I^3)$ ($)	Optimal State
3	3	2	170	90	✓
3	3	1	175	70	✓
3	2	2	175	80	
3	2	1	180	60	✓
3	1	2	180	65	
3	1	1	185	45	✓
2	1	2	190	50	
2	1	1	195	30	✓
1	1	2	210	40	
1	1	1	215	20	✓

Again, the partial delays $S_3(I^3)$ are arranged in a monotonic non-decreasing order. It can be seen that there are 6 optimal partial assignments in this final stage. The solution for the overall capacity assignment problem with $T \leq T_{MAX} = 200$ ms is given by the assignment in the 8th row.

The efficiency of the dynamic programming depends on the number of states that can be eliminated at each stage. By calculating all the possible combinations of the option i_k to find the optimal capacity assignment, the number of states obtained for this example will be 3x3x2 = 18, as compared to the 6 optimal states in the last stage of the dynamic

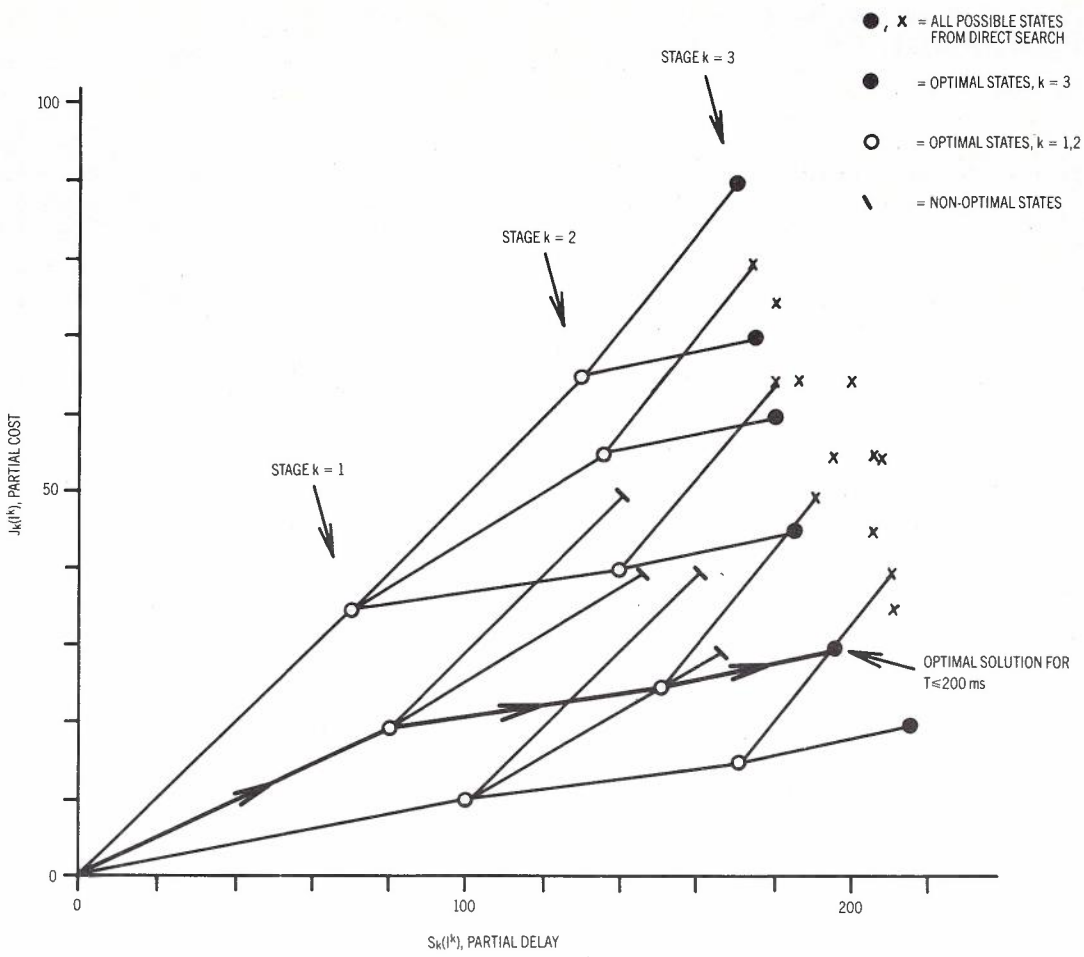


Fig. 4 - Cost-delay characteristic of the 3-link network

programming solution. In general, an even larger percentage of the states can be eliminated when there are more stages in the calculation. Figure 4 shows a plot of all the 18 possible states and the optimal states derived from the dynamic programming.

5.3 A Generalized Dynamic Programming (DP) Procedure

A generalized dynamic programming procedure is presented.

Step 1

First, denote by $[S^*_{k-1}(I^{*k-1}), J^*_{k-1}(I^{*k-1})]$ and I^{*k-1} the optimal state and decision sequence at stage k-1, i.e.

$$I^{*k-1} = [i^*_1, i^*_2 \ldots i^*_{k-1}] \qquad (27)$$

The admissible states $[S'_k(I^{*k-1}, i_k), J'_k(I^{*k-1}, i_k)]$ of stage k are obtained from the optimal states of stage k-1 as follows:

$$S'_k(I^{*k-1}, i_k) = S^*_{k-1}(I^{*k-1}) + T_k(i_k) \qquad (28)$$

$$J'_k(I^{*k-1}, i_k) = J^*_{k-1}(I^{*k-1}) + D_k(i_k) \qquad (29)$$

$$i_k = 1,2, \ldots L_k.$$

Step 2

Rearrange these admissible states and relabel them by $[S'_k(\ell), J'_k(\ell)]$, $\ell = 1,2, \ldots \Pi'_k$, such that

$$S'_k(1) \leq S'_k(2) \leq \ldots \leq S'_k(\Pi'_k) \qquad (30)$$

where Π'_k is the number of the admissible states.

Step 3

The optimal states $[S^*_k(\ell), J^*_k(\ell)]$, $\ell = 1,2, \ldots \Pi^*_k$, of the stage k, where

$\Pi^*_k \leq \Pi'_k$, are derived from the admissible states $[S'_k(\ell), J'_k(\ell)]$, $\ell = 1,2, \ldots \Pi'_k$ as follows:

Let

$$Jmin_k(0) = \infty;\ S'_k(0) = 0;\ U(0) = 0;\ V(0) = 0 \tag{31}$$

Step 4

Let

$$Jmin_k(\ell) = \min\,[Jmin_k(\ell-1),\ J'_k(\ell)] \tag{32}$$

Construct two index functions $U(\ell)$ and $V(\ell)$:

$$U(\ell) = \begin{cases} \ell, & \text{if } J'_k(\ell) < Jmin_k(\ell-1) \\ U(\ell-1), & \text{otherwise,} \end{cases} \tag{33}$$

and

$$V(\ell) = \begin{cases} V(\ell-1)+1, & \text{if } J'_k(\ell) < Jmin_k(\ell-1) \\ & \quad \text{and } S'_k(\ell) \neq S'_k(\ell-1) \\ V(\ell-1), & \text{otherwise,} \end{cases} \tag{34}$$

and let

$$S^*_k(V(\ell)) = S'_k(U(\ell)) \tag{35}$$

$$J^*_k(V(\ell)) = J'_k(U(\ell)) \tag{36}$$

Repeat step 4 for $\ell = 1,2, \ldots \Pi'_k$.

Repeat step 1 to step 4 for $k = 1,2, \ldots M$.

5.4 Bounds for Optimal State Elimination

The generalized dynamic programming procedure described above yields the entire set of the optimal capacity assignment. The end-product is a cost-delay plot like the one shown in Figure 4. If we are only interested in solving one particular optimal capacity assignment with terminal delay parameter T_{MAX}, then it is possible to introduce bounds to the dynamic programming for improving the efficiency computationally. Some of the optimal states at each stage, which are not required in such a dynamic programming solution, can be eliminated by the two bounding techniques described below.

1. Bounds on the partial delay $S^*_k(\ell)$ - minimum delay bound

At stage k, consider the partial delay for the remaining unassigned links, k+1, k+2, ... M. The minimum partial delay for these unassigned links at stage k+1, k+2, ... M is given by

$$T^k_{BMIN} = \sum_{j=k+1}^{M} T_j(1)\ ;\ T^M_{BMIN} = 0 \tag{37}$$

since $T_k(1)$ is the minimum delay function for each stage k (see (16), (24)). Now for any optimal state at stage k, if the sum of the partial delay $S^*_k(\ell_1)$ and the minimum partial delay T^k_{BMIN} for the remaining unassigned stages (links) exceeds the delay parameter T_{MAX}, the optimal state considered will not lead to any optimal state at the final stage of the DP solution that meets the design specification $T \leq T_{MAX}$. In other words, that optimal state at stage k can be eliminated.

Thus if

$$S^*_k(\ell_1) + T^k_{BMIN} > T_{MAX} \tag{38}$$

then the states $[S^*_k(\ell_1), J^*_k(\ell_1)]$, and from (26) all the states such that $S^*_k(\ell_2) > S^*_k(\ell_1)$, can be eliminated; see Figure 5(a).

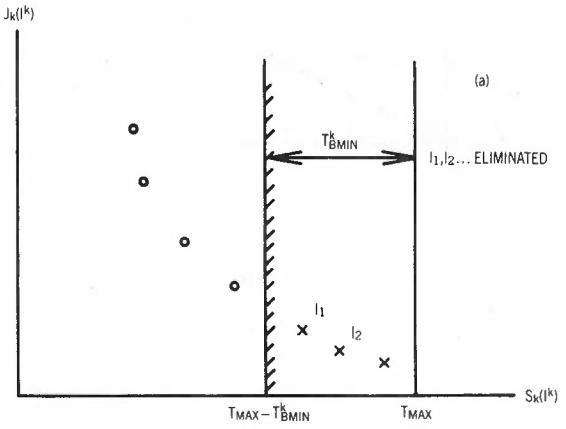


Fig. 5 - Bounds on the dynamic programming: (a) Partial delay bound

Notice that the bound $T_{MAX} - T^k_{BMIN}$ can be derived recursively and prior to the dynamic programming procedure as follows:

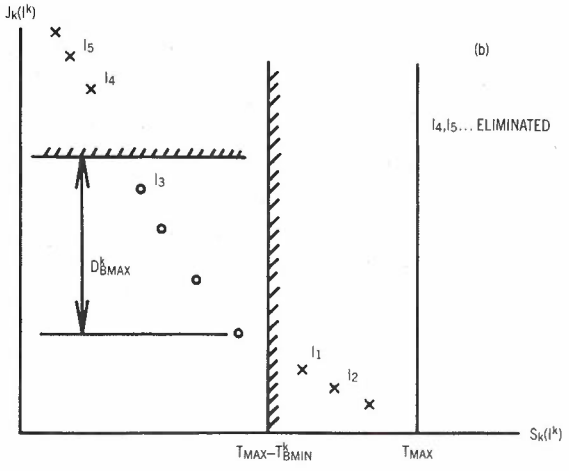
$$T^o_{BMIN} = \sum_{j=1}^{M} T_j(1) \tag{39}$$

$$T^k_{BMIN} = T^{k-1}_{BMIN} - T_k(1) \tag{40}$$

It is interesting to note that the bound $T_{MAX} - T^k_{BMIN}$ eliminates the optimal states in a way similar to the conventional dynamic programming approach to linear resource allocation problems in which the state variable is formulated as $(L_k - i_k)$.

2. Bounds on the partial cost $J^*_k(\ell)$ - maximum cost bound

The maximum cost bound assumes the knowledge of the minimum delay bound. Suppose the partial delay $S^*_k(\ell_3)$ of the state ℓ_3 is the largest partial delay that is smaller or equal to the minimum delay bound $T_{MAX} - T^k_{BMIN}$, see Figure 5(b). Then from (26), the partial cost $J^*_k(\ell_3)$ is also the smallest among those optimal states that stay inside the partial delay bound.



(b) Partial cost bound

The cost of assigning capacities with the minimum partial delay to the unassigned links at stage k+1, k+2, ... M will be a maximum partial cost (see (14), (15)) given by

$$D^k_{BMAX} = \sum_{j=k+1}^{M} D_j(1) \; ; \; D^M_{BMAX} = 0 \qquad (41)$$

The partial assignment of the state ℓ_3 at stage k together with the minimum partial delay assignment to the rest of the links represents a candidate for the overall solution to the capacity assignment of the network. The cost of such assignment is given by

$$D^k_{BMAX}(\ell_3) = D^k_{BMAX} + J^*_k(\ell_3) \qquad (42)$$

Thus, the cost $D^k_{BMAX}(\ell_3)$ represents an upper bound to the cost of the states at stage k. Suppose there is a state ℓ_4, at stage k, such that

$$J^*_k(\ell_4) \geq D^k_{BMAX}(\ell_3) \qquad (43)$$

Then clearly the state ℓ_3 is always a better candidate than ℓ_4 to meet the terminal delay specification T_{MAX} at a smaller cost. In other words, the state ℓ_4, and from (26) those states ℓ_5 such that

$$\ell_5 > \ell_4 \qquad (44)$$

can be eliminated, see Figure 5(b). Notice that at the last stage M, the state ℓ_3 is in fact the optimal solution for the overall network capacity assignment problem with $T \leq T_{MAX}$.

The overhead of implementing both types of bounds studied is the computation required to find out all those optimal states outside the bounds. This overhead is, however, relatively small because the partial delay and the partial cost are already arranged in sequential order during the procedure of eliminating the non-optimal states. A sequential search can therefore be employed to effectively locate those optimal states outside the bounds and eliminate them.

6. APPLICATIONS TO THE AUSTPAC NETWORK

Both capacity assignment techniques have been applied to study one of the planned topologies of the AUSTPAC network, shown in Figure 6. Each full-duplex link in the network is implemented by a number of parallel transmission lines. The normal link traffic flow and approximate physical length are shown in Table 4. Using a normalized cost figure e_k = \$1 for a simplex line (i.e. \$2 for a full-duplex line) per 100 km, D_{io} = 0, and

TABLE 4 - Link Flows and Distances

μ^{-1} = 512 bits	Flow λ (packets/sec)		Distance (km)
Link (full-duplex)	East bound	West bound	
1	74.0	74.1	680
2	28.5	27.3	420
3	48.5	38.3	1140
4	13.0	10.2	570
5	31.9	27.1	2470
6	24.1	21.0	230
7	45.1	34.1	570
8	47.9	38.7	1060
9	27.1	21.2	1900

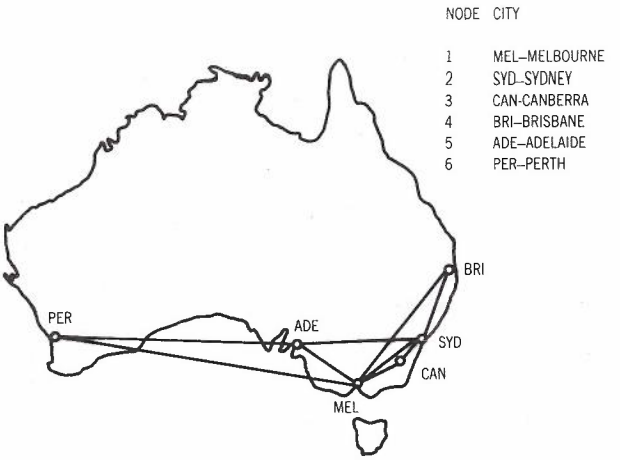


Fig. 6 - A proposed AUSTPAC network topology

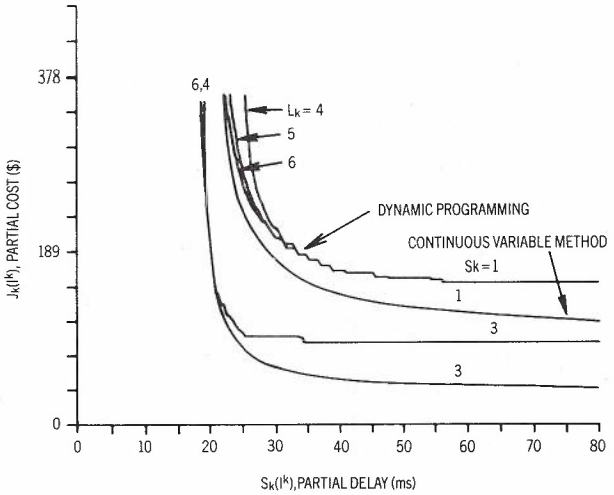


Fig. 7 - Cost-delay characteristics: the dynamic programming compared to the continuous variable approximation method

negligible nodal processing time, the capacity assignments using these techniques are obtained. Figure 7 shows the cost-delay characteristics obtained: the continuous curves are calculated by the Lagrange Multiplier method, the step-wise curves (joining the discrete optimal points) by the dynamic programming method, for different scaled level (Scale factor Sk) of the nominal traffic flow.

The continuous variable Lagrange Multiplier solution has a lower-cost delay characteristic. This is, however, expected as the continuous capacity variable can always be optimized to the desired value. The dynamic programming solution, on the other hand, depends on the number of options L_k available in the design.

With specified message delay constraint T_{MAX}, the value m_i for each link is obtained, see Table 5. Recall that in the Lagrange Multiplier method, $\bar{m}_i$ is the smallest integer equal to or greater than the largest value of m_i and m_{M+i}. Table 5 shows that the Lagrange Multiplier method yields the optimal assignment for traffic levels Sk=1 and Sk=3. When Sk=2, the Lagrange Multiplier method does not yield the optimal assignment. The dynamic programming yields optimal capacity assignment in all cases.

The effectiveness of the dynamic programming technique over the continuous variable approximation method is thus demonstrated. The continuous variable approximation method can yield optimal results in certain cases; these however depend critically on how one discretizes the continuous solutions. The dynamic programming on the other hand yields directly optimal results. In fact, the dynamic programming can be also applied for continuous capacity variable problems by discretizing the continuous capacity into many finite steps.

TABLE 5 - Capacity Assignment

	Sk=1; T_{MAX}=21.4 ms			Sk=2; T_{MAX}=29.3 ms			Sk=3; T_{MAX}=58.5 ms		
Link (full-duplex)	Lagrange (m_i/m_{M+i})	$\bar{m}_i$	DP m_i	Lagrange (m_i/m_{M+i})	$\bar{m}_i$	DP m_i	Lagrange (m_i/m_{M+i})	$\bar{m}_i$	DP m_i
1	(2.3/2.3)	3	3	(2.65/2.65)	3	4	(2.49/2.49)	3	3
2	(1.02/1.06)	2	2	(1.09/1.13)	2	2	(0.93/0.97)	1	1
3	(1.16/1.30)	2	2	(1.41/1.58)	2	2	(1.45/1.61)	2	2
4	(0.45/0.61)	1	1	(0.50/0.68)	1	1	(0.45/0.61)	1	1
5	(0.58/0.76)	1	1	(0.78/1.02)	2	1	(0.89/1.16)	2	2
6	(0.98/1.12)	2	2	(0.97/1.12)	2	1	(0.73/0.84)	1	1
7	(0.97/1.15)	2	2	(1.08/1.68)	2	2	(0.98/1.53)	2	2
8	(1.06/1.17)	2	2	(1.27/1.41)	2	2	(1.29/1.43)	2	2
9	(0.49/0.55)	1	1	(0.64/0.71)	1	1	(0.70/0.79)	1	1
Cost D($)		138.9	138.9		163.7	142.4		156.1	156.1

However, for such applications the Lagrange Multiplier method will offer a straightforward analytical result.

The bounding technique of the dynamic programming method is illustrated in Figure 8. Each line in the figure joins a pair of optimal states at consecutive stages. In Figure 8(b), many optimal states are eliminated by the bounds in calculating a solution for $T_{MAX} \leq 30$ ms.

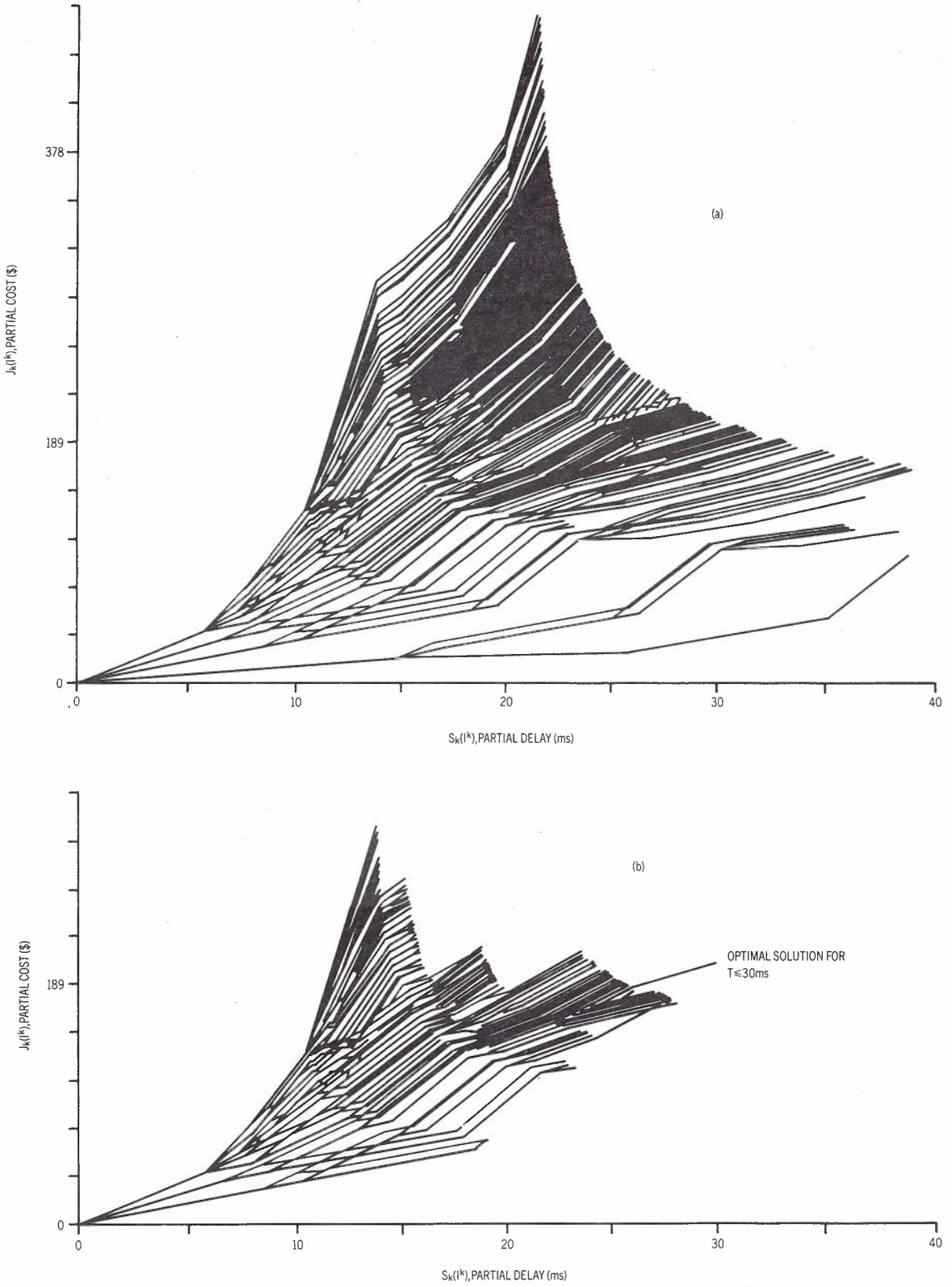


Fig. 8 - Dynamic programming applied to the AUSTPAC topology:
(a) bounds not applied
(b) bounds applied

7. CONCLUSION

The complexity and cost of packet switching networks demand that existing networks be efficiently used and that new networks be rationally designed. To meet this demand many factors have to be studied and analysed.

Some design aspects and techniques of the capacity assignment of a m-M/M/1 queueing network have been presented. The dynamic programming technique yields the optimal capacity assignment solution. The Lagrange Multiplier technique, on the other hand, yields closed-form analytical solutions which are close to optimal.

Although capacity assignment itself is a sub-problem in many network design problems when topology and traffic routing are also considered in the network optimization - a problem usually too complicated and too difficult to solve analytically, the capability to solve the capacity assignment problem optimally is definitely a step forward in solving the larger scale problems.

8. ACKNOWLEDGEMENT

Support of this work by the Radio Research Board is gratefully acknowledged.

9. REFERENCES

1. Kleinrock, L., "Analytic and Simulation Methods in Computer Network Design", Proceedings of the Spring Joint Computer Conference, 1970, pp. 568-579.

2. Frank, H., Frisch, I.T., Van Slyke, R., Chou, W.S., "Optimal Design of Centralized Computer Networks", Networks 1, pp. 43-57, 1971.

3. Zadeh, N., "On Building Minimum Cost Communication Networks Over Time", Networks 4, pp. 19-34, 1974.

4. Rubin, I., "Optimal Link Capacity Assignments in Teleprocessing and Centralized Computer Networks", Proc. Int. Telemetry Conf., Los Angeles, Calif., 1976.

5. Maruyama, K. and Tang, D.T., "Discrete Link Capacity Assignment in Communication Networks", in Proc. 3rd ICCC, pp. 92-97, Aug. 1976.

6. Gerla, M. and Kleinrock, L., "On Topological Design of Distributed Computer Networks", IEEE Trans. Commun., Vol. COM-25 pp. 48-60, Jan. 1977.

7. Avellaneda, O.A., Hayes, S.F. and Nassehi, M.M., "A Capacity Allocation Problem in Voice - Data Networks", IEEE Trans. Commun., Vol. COM-30, pp. 1767-1772, July 1982.

8. Frank, H. and Frisch, I.T., "Planning Computer Communication Networks", in Computer Communication Networks, N. Abramson and F.F. Kuo, Eds. Englewood Cliffs, NJ: Prentice-Hall, 1973.

9. Kleinrock, L., "Queueing Systems: Vol. 2, Computer Applications", Wiley - Interscience New York, 1976.

10. Kleinrock, L., "Communication Nets: Stochastic Message Flow and Delay", New York: McGraw-Hill, 1964.

11. Hildebrand, F.B., Methods of Applied Mathematics (2nd Ed.), Prentice-Hall (Englewood Cliffs, New Jersey), 1965.

12. Gallager, R.G., "A Minimum Delay Routing Algorithm using Distributed Computation", IEEE Trans. Commun., Vol. COM-25, pp. 113-126, Jan. 1979.

13. Kuehn, P.J., "Approximate Analysis of General Queueing Networks by Decomposition", IEEE Trans. Commun., Vol. COM-27, pp. 113-126, Jan. 1979.

14. Hoang, D.B. and Ng, M.J.T., "Delay Analysis, Routing and Capacity Assignment in Packet Switching Network", Report to Radio Research Board, Australia; Department of Electronic and Communication Science, La Trobe University, Australia, 1983.

15. Harrison, M.J., "DPS 25: Distributed Packet Switching Technology for the AUSTPAC Network", Telecommunication Journal of Australia, Vol. 32, No. 1, 1982.

16. Bellman, R.E. and Dreyfus, S.E., "Applied Dynamic Programming", Princeton, NJ: Princeton University Press, 1962.

BIOGRAPHIES

M.J.T. NG

TOMY NG received the B.Sc. degree in Computer and Communication Engineering from the University of Essex, England, in 1978. From 1978 to 1979, he joined the ITT Standard Telecommunication Laboratories, Ltd., Harlow, England, as a Graduate Research Engineer, working on microprocessor based modems and data transmission systems. He is currently with the Department of Electronic and Communication Science, La Trobe University, working for the Ph.D. degree in digital signal processing techniques, modelling and optimization applied to vehicular and computer network traffic engineering studies.

DOAN B. HOANG

DOAN B. HOANG received the Bachelor of Electronic Engineering degree with first class honours from the University of Western Australia, 1975; the M.E. and Ph.D. from the University of Newcastle, 1976 and 1978 respectively. He was a Radio Communication Engineer at Telecom Australia, 1978; a Computer Engineer at Sydney University 1978-1980. Since 1980 he has been a Lecturer in the Department of Electronic and Communication Science, La Trobe University, Australia. His current research interests are in the fields of computer communication networks, local area networks, analogue and digital signal processing techniques and their applications.

A Non-Destructive Method Of Monitoring Internal Parameter Drifts in CMOS Integrated Circuits

J. THOMPSON
T. ROGERS
R.A. GALEY
Telecom Australia Research Laboratories

It is shown that the set of characteristic curves given by plotting the supply current versus input voltage on CMOS integrated circuits may be readily analysed to reveal the electrical characteristics of individual transistors on the chip. Both P and N channel transistors may be characterised and thus direct measurement of drifts in transistor parameters which affect reliability is possible. Results are presented of the use of this technique to quantitatively determine the effect of radiation on MOS threshold voltage. These results are compared with accelerated bias-temperature stress ageing where the degradation of P channel transistor threshold voltage by the slow hole trapping mechanism is investigated. The method reported does not require special test device structures and may be applied to almost any standard commercial CMOS integrated circuit. It would therefore be of value to the end user of high reliability CMOS ICs in that it enables a determination of the susceptibility of the IC to long term threshold drift and also an easy and quick determination of the radiation sensitivity of the device.

1. INTRODUCTION

Many papers have been published on instability mechanisms in MOS devices but in almost all cases these investigations are carried out on specialised test structures which are not generally available to the component user. In order for an IC customer to make his own determination of the effect of instability mechanisms it is necessary to be able to isolate individual transistors such that their characteristics may be precisely measured. In the case of CMOS devices, which are used in many applications requiring high reliability, it is possible to isolate individual transistor characteristics in such a way as to be able to directly measure the effect of charge-induced instability on the device.

This paper describes the technique used to obtain data on instability induced by both radiation and bias-temperature stressing, and presents results which demonstrate the quantitative measurements of threshold instability mechanisms in such devices. This method is applicable to many types of CMOS integrated circuit and provides a simple procedure for the comparative assessment of devices from different manufacturers.

2. MEASUREMENT TECHNIQUE

The technique is based on the fact that the majority of CMOS digital integrated circuits have little static power dissipation, and except for sub-threshold conduction and leakage, only draw current from the supply during switching. Thus if the supply current for a fixed value of supply voltage is monitored as a function of input voltage on one or more selected inputs, then the characteristic obtained will depend on the currents flowing through all transistors affected by that input or inputs. As will be shown later, the current characteristic will be dominated by the input stage and subsequent stages will not affect the validity of the measurements.

As an initial example, one of the simplest of all CMOS circuits, the 4049UB type unbuffered hex inverter has been chosen. Fig. 1 shows the circuit layout of a single gate in this device. Each inverter stage consists of a single P channel transistor with source and substrate to V_{DD} and a single N channel transistor with source and substrate to V_{SS}. If the supply current vs input voltage is plotted a curve similar to that in Fig. 2 is obtained. The interpretation of this curve proceeds as follows.

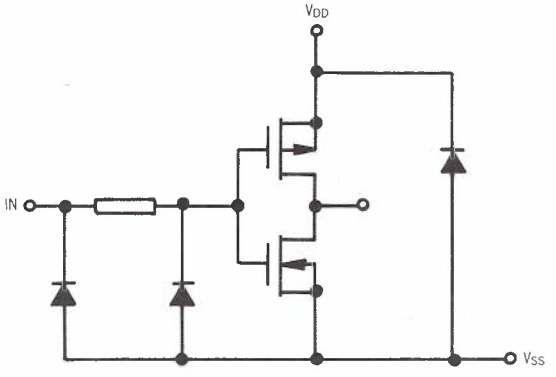


Fig. 1 - Single inverter of 4049 UB type

The first order characteristic equation for a MOS device in saturation with zero source-substrate bias, is given by:

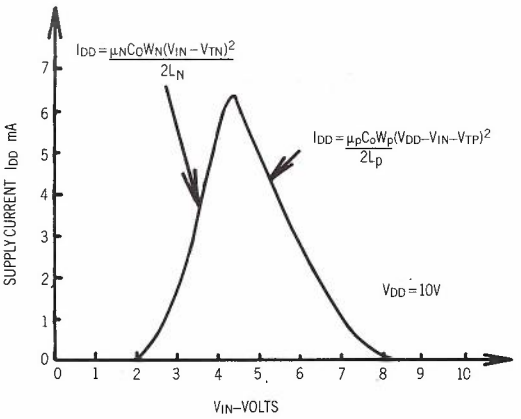


Fig. 2 - Supply current vs input voltage for a single inverter in a hex unbuffered inverter package (4049 UB)

$$|I_{DS}| = \frac{\mu C_o W}{2L} (V_{GS} - V_T)^2 \qquad (1)$$

Where I_{DS} = drain to source current, μ = channel carrier mobility, C_o = gate capacitance per unit area, W = gate width, L = channel length, V_{GS} = gate to source voltage and V_T = threshold voltage.

This equation is valid for $|V_{DS}| > |V_{GS} - V_T|$ (where V_{DS} = drain to source voltage) for long channel devices operating with gate voltages a few volts above threshold (Ref. 1).

Consequently in the circuit of Fig. 1 which consists of an N channel and a P channel transistor in series, the current drawn from the supply will be the lesser of the two values:

$$\frac{\beta_N}{2} (V_{IN} - V_{TN})^2 \text{ and } \frac{\beta_P}{2} (V_{DD} - V_{IN} - V_{TP})^2 \qquad (2)$$

where gain factors β_N and β_P are defined by:

$$\beta_N = \frac{\mu_N C_o W_N}{L_N} \text{ and } \beta_P = \frac{\mu_P C_o W_P}{L_P} \qquad (3)$$

Here N and P subscripts refer to the N and P channel transistors respectively and V_{IN} is the input voltage. This will be the case since that transistor which is driven least hard, and hence determines the supply current, will have at least half the supply voltage dropped across it and so will be in saturation. In this region, the drain current is independent of drain voltage and hence the resulting supply current versus input voltage characteristic will consist of two intersecting square law characteristics as shown, for a practical device, in Fig. 2.

As can be seen from Fig. 2, for a fixed supply voltage of 10 volts, the point of maximum supply current (and hence of output logic swing) does not occur at precisely half the supply voltage, but at a value determined by the equality of the two terms (expression 2). Hence the input voltage for maximum current is given by:

$$V_{IN}(I_{DD}\,MAX) = \frac{\left[\sqrt{\frac{\beta_P}{\beta_N}} (V_{DD} - V_{TP})\right] + V_{TN}}{\left(1 + \sqrt{\frac{\beta_P}{\beta_N}}\right)} \qquad (4)$$

This fact is of importance where transistor parameters and geometries may not be matched and also for more complicated gate structures where series and parallel combinations of transistors are used, the latter case resulting in an input logic threshold which is pattern sensitive.

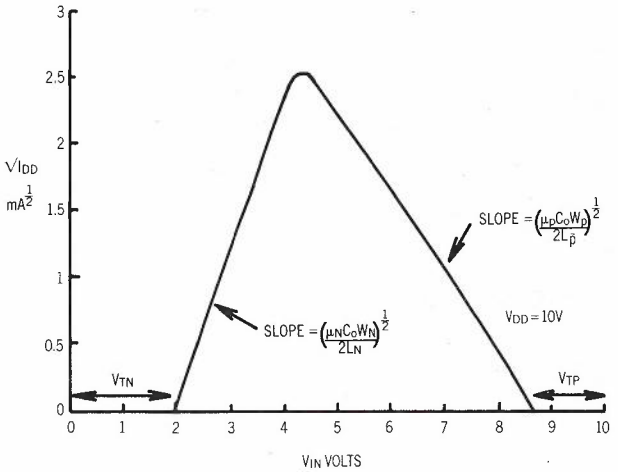


Fig. 3 - Square root of supply current vs input voltage for a single inverter in a hex unbuffered inverter package (4049 UB)

Fig. 3 shows a plot of $\sqrt{I_{DD}}$ vs V_{IN} for the device of Fig. 2, and indicates that equation (1) is obeyed for the whole range of operation of the N channel device and for much of the range of operation of the P channel device. The change of slope in the P channel characteristic may be due to the reduction in carrier mobility which occurs at high gate fields (Ref. 2). It can be seen from Fig. 3 however that the threshold voltages of the P and N channel transistors may be readily obtained from the intercepts with the X axis, and their gain factors obtained from the respective slopes. Hence V_{TN}, V_{TP}, β_N and β_P may be derived directly from the supply current versus input voltage characteristic.

These four figures provide a first order characterisation of the CMOS device and provide a means of monitoring the effect of changes in these parameters, as a result of ageing or radiation, on the operation of the device.

As an example of a slightly more complicated device the 4011B Quad 2 input buffered NAND gate was used in this study. The package contains 32

transistors, there being four of each type per gate as shown in Fig. 4.

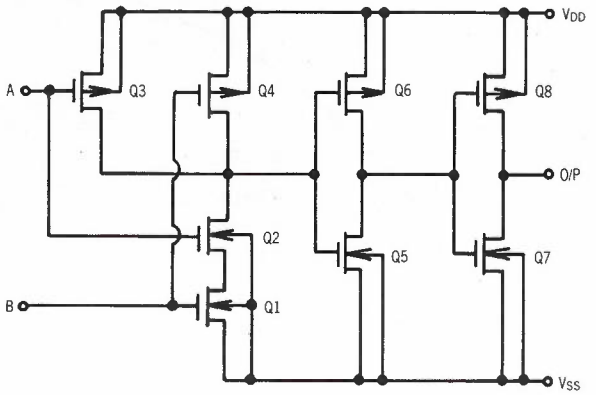


Fig. 4 - Circuit schematic of 4011 B quad dual input nand gate (single gate only shown)

Each gate may be operated in three ways, each of which will produce substantially different supply current versus input voltage characteristics. Referring to Fig. 4, if input A is held at V_{DD} and input B driven by the ramp, then transistor Q3 will be off and transistor Q2 will be in a low resistance unsaturated state, such that the observed characteristic will result from the series combination of Q1 and Q4, with the characteristics of Q5,Q6,Q7 and Q8 superimposed on the mid portion of the characteristic. The resulting characteristic is shown in curve A of Fig. 5. Also in Fig. 5, curve B shows the corresponding case where input A is driven by the ramp voltage and input B is held to V_{DD}, and curve C shows the case where

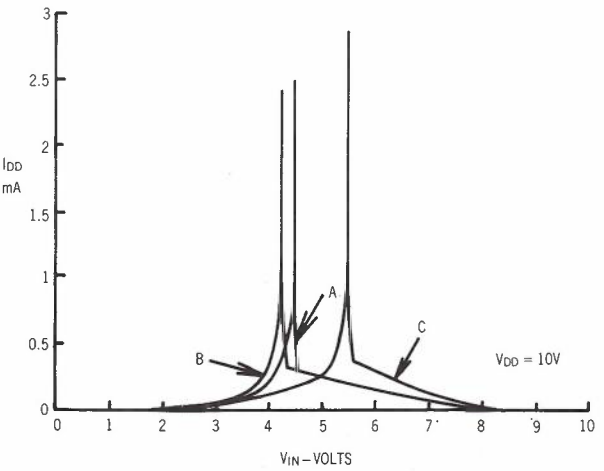


Fig. 5 - Supply current vs input voltage for a single 2 input nand gate (¼ 4011 B)

both inputs A and B are driven by the ramp. The differences between the curves is explained by the fact that a series or parallel combination of MOS transistors is equivalent to a single device of altered geometry and hence gain factor. An extra complication in the case of curve C is that the threshold voltage of Q2 is increased due to the backgating effect of the voltage drop across Q1, and so the composite device departs from the square law characteristic.

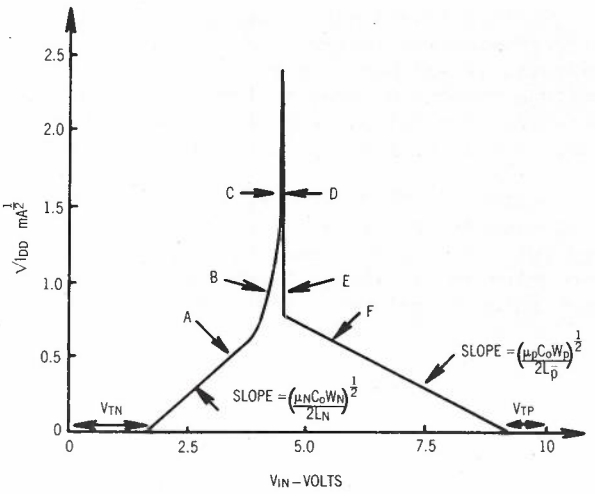


Fig. 6 - $\sqrt{I_{DD}}$ vs input voltage for a single input of a 2 input nand gate with other input to V_{DD}. V_{DD}=10 V

Fig. 6 is a square root plot of the data from curve A of Fig. 6 and clearly shows the six different areas of operation of the device, labelled A-F. Regions A and F represent conduction of the input transistors only, regions B and E are dominated by the intermediate transistors Q5 and Q6, and regions C and D result from switching of the output transistors Q7 and Q8.

The above discussion relates specifically to devices based on the most common arrangement for a 2 input NAND gate, as shown in Fig. 4. However in studies of this kind it is important to decapsulate and inspect the chip as other configurations are often used. Fig. 7 shows the actual arrangement of the components on the chip for the 4011B devices subsequently referred to as batch 3.

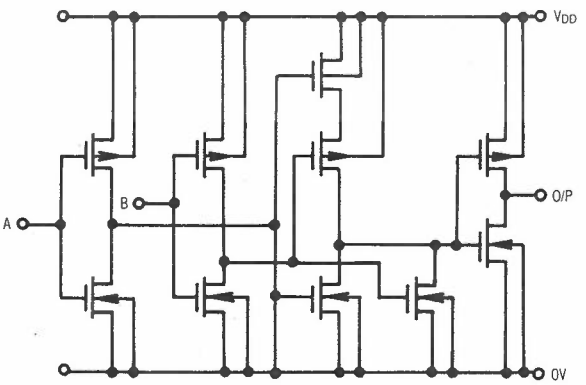


Fig. 7 - Alternative layout for 2 input nand gate used in devices from batch 2

3. THRESHOLD VOLTAGE DRIFT UNDER BIAS-TEMPERATURE STRESS

It is well known that the threshold voltage of MOS devices is affected by a number of instability mechanisms. A voluminous literature exists on the subject, a useful summary of which is given in (Ref. 3).

In this investigation using simple devices free from excess leakage and sub-threshold effects, it was possible to resolve threshold voltage changes of greater than 5 mV by the parallel translation along the X axis of one or other halves of the supply current characteristic.

A group of 20 devices were placed under bias-temperature stress at 200°C. Each chip had two pairs of inputs connected to zero volts and two pairs to +15 volts. Test devices were from four batches, batches 1 and 2 being different diffusion batches of 4011B type devices from the same manufacturer, batch 3 consisting of 4011B devices from a different manufacturer, and batch 4 being 4011A (unbuffered) devices from a third manufacturer. All devices were mounted in cerdip packages. By measuring the current response of each input pin separately, 16 individual transistors may be resolved in each package making a total of 320 transistors under test. Under the applied bias conditions each package has two pairs of inputs at zero volts and two pairs at +15 volts. This results in four N channel transistors with zero volts gate bias and four with +15 volts gate bias. Similarly four P channel transistors per package have zero volts gate bias and four have -15 volts effective gate bias. Thus all conditions of practical importance are covered in this simple arrangement. The resulting mean threshold drifts obtained for batches 1-3 are given in Figs. 8-10 and will be discussed below.

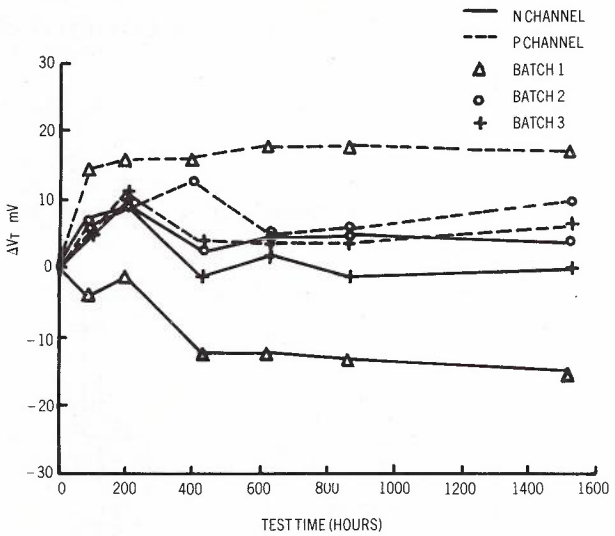


Fig. 8 - V_T shifts for N channel (unbroken lines) and P channel (dashed lines) transistors versus stress time at 200°C for zero volts effective gate bias.

Fig. 8 shows mean drifts for both N and P channel devices under zero bias, and as would be expected they were negligible. Maximum drifts recorded after 1520 hours were -15 mV, and +17 mV, for N and P channel devices respectively, both from batch 1.

Fig. 9 provides a surprising result with a large positive threshold drift under +15 V bias for devices from batch 3, with a mean drift of

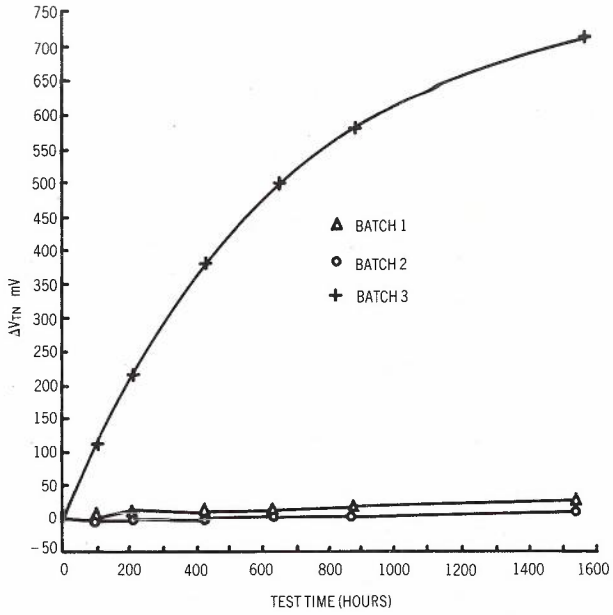


Fig. 9 - V_T shifts for N channel transistors at 15 V gate bias at 200°C

714 mV after 1520 hours at 200°C. Devices from batches 1 and 2 show a small initial fall, probably due to polarisation of the phospho-silicate glass gate passivation (Ref. 4), followed by a very slow increase in threshold which is still negligible however after 1520 hours. The reason for the large positive increase in threshold for devices from batch 3 is difficult to explain by any of the standard theories for MOS instability and will be the subject of further investigation. Electron injection is unlikely in this case since the transistors are turned on and hence have near zero drain voltage.

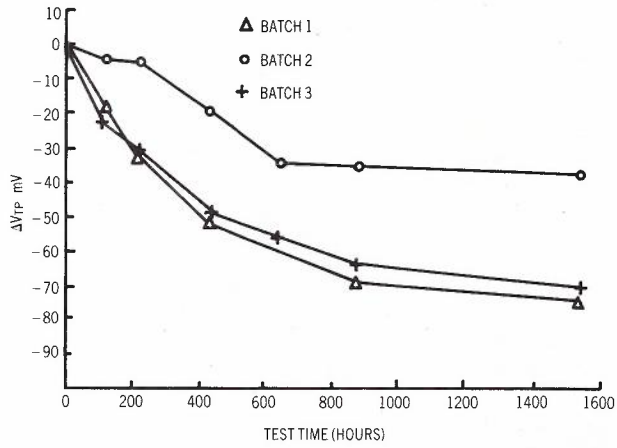


Fig. 10 - V_T shifts for P channel transistors at -15 V effective gate bias at 200°C

Fig. 10 shows the result of -15 volts gate bias on P channel transistors and displays the expected negative threshold shift under negative gate bias, which is a result of the slow hole trapping phenomena (Ref. 3). While batches 1 and 3 show the normally observed saturating characteristic drift, batch 2 shows a delay in the onset of negative drift due to

a competing mechanism occurring over the first 200 hours. It can be seen that for batches 1-3 total drift due to this mechanism at -15 V bias is unlikely to be much in excess of -100 mV.

The results obtained from batch 4 are shown in Fig. 11, and demonstrate a much greater degree of hole trapping, with a mean drift of -2.34 volts after 1715 hours at 200°C for a bias of -15 volts. Due to a peculiarity in the design of this chip where one of the N channel transistors has an independent non-grounded P well, it was not possible to obtain stable results for N channel transistors under zero bias. The effect will not reduce the validity of the technique since only unbiassed transistors are affected.

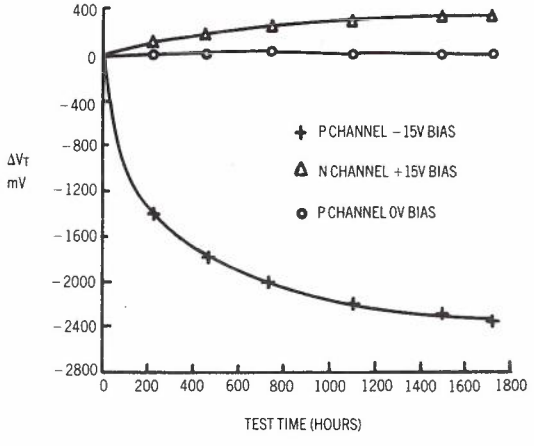


Fig. 11 - V_T shifts for transistors from batch 4 under bias at 200°C

4. EFFECT OF RADIATION

High energy radiation is known to have the effect of increasing both fast and slow interface state densities at the Si-SiO_2 interface of MOS devices (Ref. 5). These effects can lead to shifts in threshold voltage and reduction of channel mobility, resulting in increased propagation delay and reduced noise immunity. The mechanism for interface state buildup is reported to be radiation-induced generation of electron-hole pairs in the gate dielectric, and subsequent trapping of holes at neutral centres close to the Si-SiO_2 interface. (Refs. 6-10). The degree of hole trapping has been found to be very dependent on both the growth conditions and post-oxidation annealing treatment of the gate dielectric (Refs. 11,12), and also on oxide thickness (Refs. 13,14). It has been reported that the buildup of interface states after irradiation is independent of the energy of the radiation, provided it is able to penetrate into the gate oxide region and ionise electron-hole pairs (Refs. 5,15). In this investigation the devices were irradiated using 20 kVp X-rays generated by an HP Model 43805N Faxitron system, after careful removal of the cerdip package lid. The 4011B devices were maintained under bias during the irradiation with both inputs of each gate at the same bias. Bias voltages of 0, 5 V, 10 V and 15 V were applied to the inputs of the four gates in each package and the power supply input was connected to 15 V. Consequently in each package there were pairs of N channel transistors with a gate bias of 0, 5, 10 or 15 V, and similarly pairs of P channel transistors with an effective gate bias of 0, -5, -10 or -15 V.

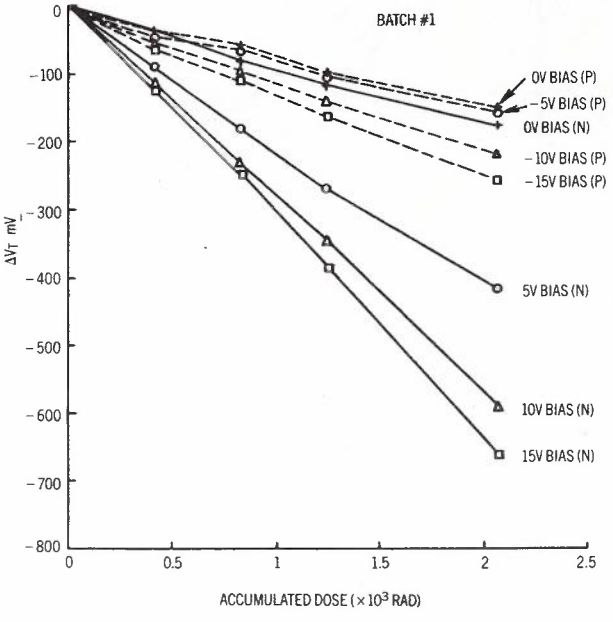


Fig. 12 - V_T shifts vs 20 kVp X Ray dose for N channel (unbroken lines) and P channel devices (dashed lines) from batch 1 with gate bias as parameter.

Measurement of threshold voltage of both N channel and P channel transistors was made both initially and after accumulated X-ray dosages. Results for batch 1 are plotted in Fig. 12 which shows the mean shift in threshold voltage against dose for a sample of eight packages. For N channel devices the usual pattern with increasing negative shift for positive gate bias is followed. It can be seen that the effect of bias tends to saturate at about 10 V and that threshold shift is about three times greater than in the absence of bias. This bias dependence results from the effect of a positive gate field in attracting the very low mobility holes towards the Si-SiO_2 interface where they are trapped. It is also seen from Fig. 12 that P channel transistors at low bias, having effectively negative gate field under normal operating conditions, show a drift very similar to that of N channel transistors at zero bias. An increase in trapping for P channel transistors with larger negative bias is however contrary to what would be expected by the simple theory (Ref. 9) and requires further experimentation.

Fig. 13 contains mean results obtained from a sample of four devices from batch 3, and shows drifts very similar to those obtained for batch 1.

With regard to batches 1 and 3, X-ray exposure was terminated when leakage currents resulting from junction degradation interfered with the measurement of threshold voltage shift.

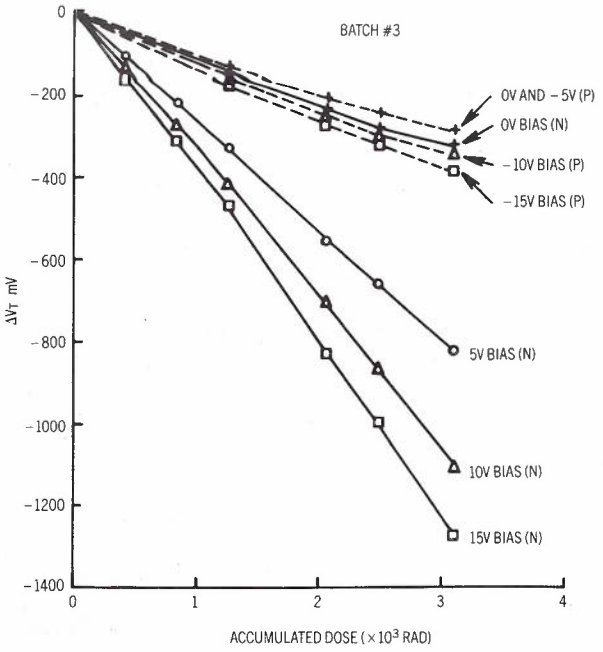


Fig. 13 - V_T shifts vs 20 kVp X Ray dose for N channel (unbroken lines) and P channel devices (dashed lines) from batch 3 with gate bias as parameter.

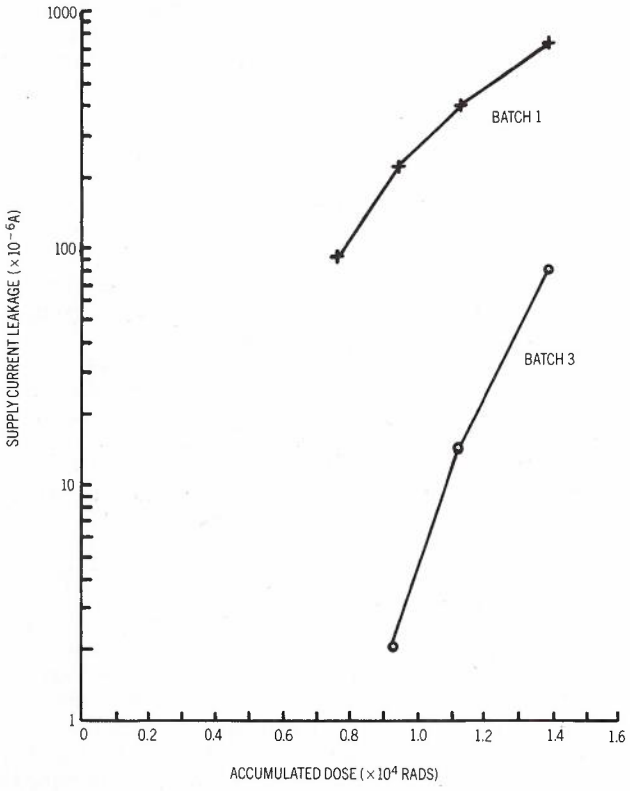


Fig. 14 - Excess power supply current vs accumulated dose

Fig. 14 compares the increase in excess power supply current due to leakage currents between a sample of 8 devices from batch 1 and 4 devices from batch 3. Measurements were made with all inputs to zero volts, the actual leakage current having been observed to be highly pattern sensitive. No significant reduction in carrier mobility for either N or P channel transistors was observed.

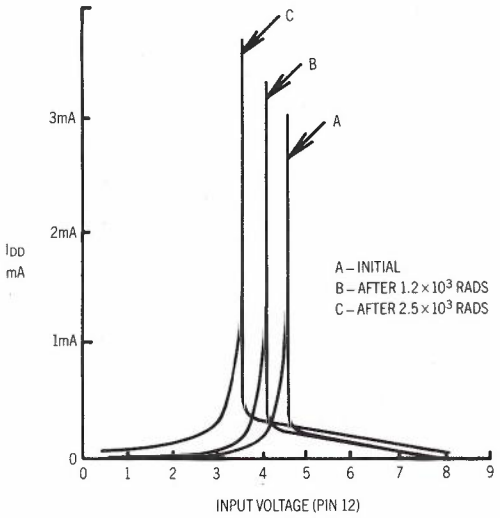


Fig. 15 - Effect of radiation on supply current characteristic for device from batch 1. V_{IN}= +15 V during irradiation

Fig. 15 shows the effect of radiation on the supply current response for Pin 12 of a device from batch 1. After 1.2 x 10^3 Rads at +15 V input bias, curve B shows a shift of approximately 500 mV to the left, indicating a reduction in noise immunity of this magnitude. At the same time the worst case propagation delay (when both inputs are rising simultaneously) has fallen by approximately 10%, indicating a small improvement in operating speed. The principle feature of curve C, taken after an exposure of 2.5 x 10^3 Rads, is the further degradation of 500 mV in noise immunity and the appearance of junction leakage, producing a vertical displacement of the curve. Worst case propagation delay is now on the falling edge of one input with the other input of the gate at V_{DD}. This has increased by 26% over its initial value.

5. DISCUSSION OF RESULTS

The results of this investigation are summarised in Tables 1 and 2 and it can be seen

TABLE 1 - Bias-Temperature Stress Results

Device Type	Effective Gate Bias	Mean ΔV_T after 1520* hours at 200°C			
		Batch 1	Batch 2	Batch 3	Batch 4
N Channel	0 volts	-15mV	4mV	0mV	/
	+15 volts	21mV	11mV	714mV	349mV
P Channel	0 volts	17mV	10mV	7mV	24mV
	-15 volts	-74mV	-37mV	-70mV	-2271mV

*1500 hours for batch 4

TABLE 2 - X-Ray Stress Results

Device Type	Effective Gate Bias	Mean ΔV_T after 2070 rads X-Ray dose (20kVp)	
		Batch 1	Batch 3
N Channel	0 volts	-177mV	-228mV
	5 volts	-418mV	-549mV
	10 volts	-590mV	-703mV
	15 volts	-665mV	-824mV
P Channel	0 volts	-153mV	-206mV
	-5 volts	-157mV	-206mV
	-10 volts	-221mV	-242mV
	-15 volts	-259mV	-263mV

from Table 1 that negative bias instability of P channel transistors (assumed to result from slow hole trapping) has been found to be a significant degradation mechanism in one out of four batches of the CMOS integrated circuits investigated. The generally accepted activation energy for this process is 1.0 eV (Ref. 3), resulting in an acceleration factor of 88 for devices operating at 50°C. This effectively means that for batches 1-3 slow hole trapping will not be a significant failure mode whereas for devices from batch 4 rapid degradation by this mechanism is predicted. In the case of batch 3, where a positive bias instability of N channel transistors has been detected, a serious source of degradation appears to be present but further work is required to determine the activation energy and voltage dependence of this phenomenon.

The most interesting feature of the X-ray studies, from a reliability point of view, is that the principle failure mechanism is radiation-induced junction leakage rather than excessive threshold voltage drift. Significant differences exist between manufacturers in the sensitivity of their product to radiation of this type.

6. CONCLUSION

This investigation has shown that it is feasible to use the supply current response of commercial CMOS circuits to internally characterise the devices on the chip, and this gives the circuit user the capability to quantitatively assess the relative reliabilities of devices on a batch by batch basis, without recourse to specially processed test patterns. A number of new questions have been raised which will require further investigation, the most significant being the nature of the positive bias instability detected in one batch of devices.

7. REFERENCES

1. S.R. Hofstein and G. Warfield, "Carrier Mobility and Current Saturation in the MOS Transistor", IEEE Trans. ED., March 1965, pp. 129-138.

2. R.F. Pierret and C.T. Sah, "An MOS-Orientated Investigation of Effective Mobility Theory", Solid State Electronics, Vol. 11, pp. 279-290, 1968.

3. J.R. Gray, "Instabilities in MOS Devices", Gordon and Breach 1981.

4. J.M. Eldridge, R.B. Laibowitz, and P. Balk, "Polarization of Thin Phosphosilicate Glass Films in MGOS Structures", Journal of Applied Physics, Vol. 40, No. 4, pp. 1922-1930, March 1969.

5. P.S. Winokur, J.M. McGarrity and H.E. Boesch Jr., "Dependence of Interface-State Buildup on Hole Generation and Transport in Irradiated MOS Capacitors", IEEE Trans. Nucl. Sci. Vol. NS-23, No. 6, Dec. 1976, pp. 1580-1585.

6. H.E. Boesch, F.B. McLean, J.M. McGarrity and G.A. Ausman Jr., "Hole Transport and Charge Relaxation in Irradiated SiO_2 MOS Capacitors", IEEE Trans. Nucl. Sci. NS-22, p. 2163, 1975.

7. R.C. Hughes, E.P. Eer Nisse and H.J. Stein, "Hole Transport in MOS Oxides", IEEE Trans. Nucl. Sci. NS-22, p. 2227, 1975.

8. H.H. Sander and B.L. Gregory, "Unified Model of Damage Annealing in CMOS from Freeze-In to Transient Annealing", IEEE Trans. Nucl. Sci. NS-22, p. 2157, 1975.

9. F.B. McLean, H.E. Boesch Jr., J.M. McGarrity, "Hole Transport and Recovery Characteristics of SiO_2 Gate Insulators", IEEE Trans. Nucl. Sci. NS-23, No. 6, Dec. 1976, pp. 1506-1512.

10. C.T. Sah., "Origin of Interface States and Oxide Charges Generated by Ionizing Radiation", IEEE Trans. Nucl. Sci. Vol. NS-23 No. 6, Dec. 1976, pp. 1563-1568.

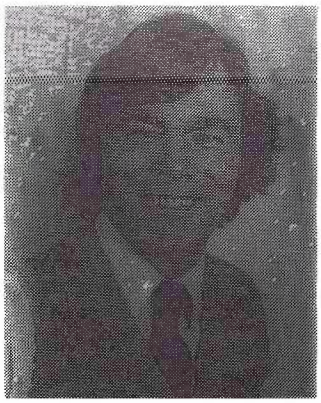
11. K.M. Schlesier and C.W. Benyon., "Processing Effects on Steam Oxide Hardness", IEEE Trans. Nucl. Sci. Vol. NS-23, No. 6, Dec. 1976, pp. 1599-1603.

12. E.P. Eer Nisse and G.F. Derbenwick., "Viscous Shear Flow Model for MOS Device Radiation Sensitivity", IEEE Trans. Nucl. Sci. Vol. NS-23, No. 6, Dec. 1976, pp. 1534-1539.

13. C.R. Viswanathan and J. Maserjian., "Model for Thickness Dependence of Radiation Charging in MOS Structures", IEEE Trans. Nucl. Sci. Vol. NS-23, No. 6, Dec. 1976, pp. 1540-1545.

14. G.W. Hughes and R.J. Powell., "MOS Hardness Characterisation and its Dependence upon some Process and Measurement Variables", IEEE Trans. Nucl. Sci. Vol. NS-23, No. 6, Dec. 1976, pp. 1569-1572.

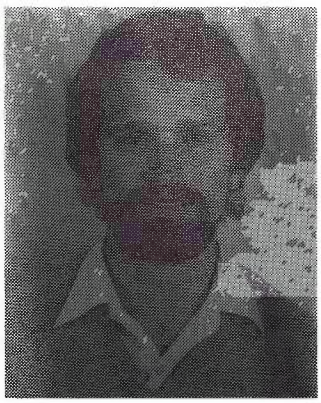
15. P.S. Winokur and M.M. Sokolosky., "Comparison of Interface-State Buildup in MOS Capacitors Subjected to Penetrating and Non-Penetrating Radiation", Appl. Phys. Lett., Vol. 28, p. 627 (1976).



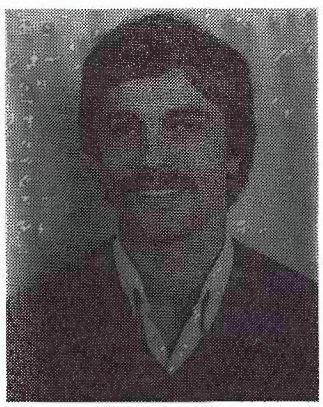
BIOGRAPHIES

JAMES THOMPSON graduated in 1968 from Lancaster University, England, with a B.A. (Hons.) in Physics. He spent ten years in the Microtechnology Section of the British Post Office Research Labs at Dollis Hill and Martlesham involved in the development of M.O.S. processing, and two years with L.M. Ericsson in Broadmeadows before joining the Telecom Research Labs in 1981.

He is currently engaged in studies of semiconductor reliability and the development of failure analysis techniques.



TIM ROGERS graduated with a B.App.Sc. (Physics) from Caulfield Institute of Technology in 1976 and is currently working towards an M.App.Sc. degree. He was employed by Melbourne University, School of Physics, until joining the Telecom Research Laboratories in January 1983 to work on semiconductor reliability.



ROY GALEY joined the Research Laboratories in 1974 as a Trainee Technical Officer. He has been involved in the design and construction of electronic equipment and investigating the reliability aspects of electronic components. He is currently a Senior Technical Officer in a team investigating semiconductor device reliability.

Radio Frequency Interference From An Incandescent Lamp — A Curious Case

P. MURRELL

Telecom Australia Research Laboratories

An incandescent household-type lamp was found to cause radio frequency interference. An explanation for its strange behaviour is proposed.

1. INTRODUCTION

Incandescent lamps, as opposed to arc/vapour lamps, are generally supposed to be electromagnetically quiet. However, a 240 volt, 60 watt incandescent lamp which produces R.F. energy at 60 to 80 MHz and harmonics thereof, has been identified as a source of television interference.

2. OBSERVATIONS

The lamp, which was manufactured in England in 1968, consists of an inverted "U" coiled filament in an evacuated glass envelope (Fig. 1). More modern incandescent lamps are generally coiled-coil filaments in an inert gas.

When placed in the vicinity of a TV antenna, a single horizontal interference bar is visible on the screen (Fig. 2). This bar, being in synchronism with the mains supply rather than the video frame rate, is seen to move vertically on the screen.

As the manufacturer confirmed that the lamp appeared to be unique in its behaviour*, a closer examination of the lamp was made to determine the most plausible mechanism of oscillation, and the following data was obtained.

1. The lamp radiates only at the peak of one half cycle when connected to 240 V A.C. mains (Fig. 3).

2. It oscillates with only one orientation of the lamp's electrodes with respect to the active and neutral lines of the power mains.

3. The frequency of oscillation increases with the applied instantaneous voltage.

4. A significant amount of harmonic R.F. energy is produced (Fig. 4) causing interference at higher bands.

5. R.F. bypassing at the lamp's socket has negligible effect.

6. A weak magnetic field placed near the filament will stop the oscillations.

7. A visible layer of vacuum-deposited metal exists on the glass inner surface.

8. The radiation intensity increases as the lamp warms up.

9. Radiation appears to be emitted from the lower filament area.

*It was later pointed out to the author by Mr I. Macfarlane that this effect, without explanation of its cause, was mentioned in the British Standard Code of Practice, "General Aspects of Radio Interference Suppression". CP 1006 : 1955

3. LIKELY CAUSE

The above observations strongly suggest the existence of "Barkhausen-Kurz" oscillations (Ref. 1) which were first reported around 1920. Here, thermally emitted electrons accelerate across the lamp from one filament arm towards the other in an electric field formed by the applied potential difference between them. Because the filament is coiled, some electrons may pass through the coil of the higher potential filament arm and enter a reverse field formed by a low electric potential on the metal deposited glass inner surface. These electrons are reflected and may oscillate several times about the high potential filament arm before collection. Their frequency of oscillation is dependent only on the lamp geometry and electrode potential. The collective, coherent, non-sinusoidal movement of these electrons produces harmonic-rich electromagnetic radiation.

The failure of the lamp to oscillate on alternative half cycles of the applied A.C. mains is due to the slightly asymmetric nature of the two filament arms within the glass enclosure (Fig. 1). Whilst the conditions for self oscillation appear to be only just met on one half cycle of the A.C. mains (evidenced by a weak magnetic field arresting the oscillations), the different electrical geometry when the electric field between the filament arms is reversed is such that oscillations do not occur. Similarly, changing the orientation of the lamp's electrodes with respect to the active and neutral lines of the A.C. power mains alters the electric field geometry and again prevents oscillation.

4. ACKNOWLEDGEMENT

The author wishes to acknowledge the Australian Department of Communications for making the lamp available for investigations, and for permission to reproduce the photograph of Figure 2.

5. REFERENCE

1. H.E. Hollman, "Electron Oscillations in a Triode", Proc IRE, Vol. 17, No.2, pp 228-251, February 1929.

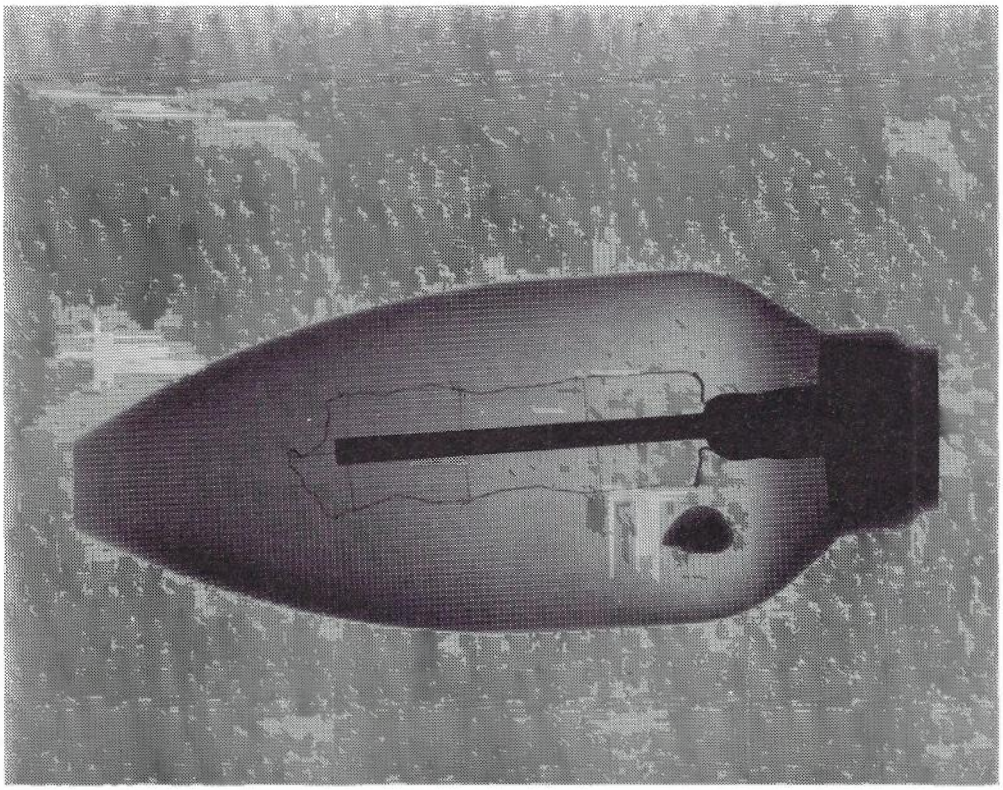


Fig. 1 - X-Ray View of Lamp

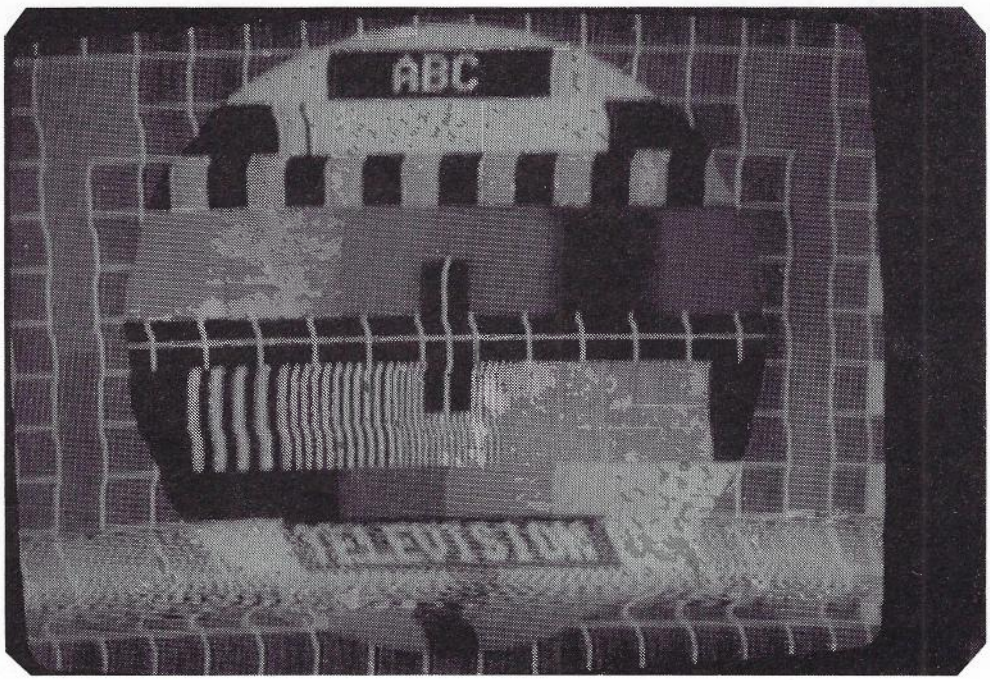


Fig. 2 - Interference Bar Due to Lamp

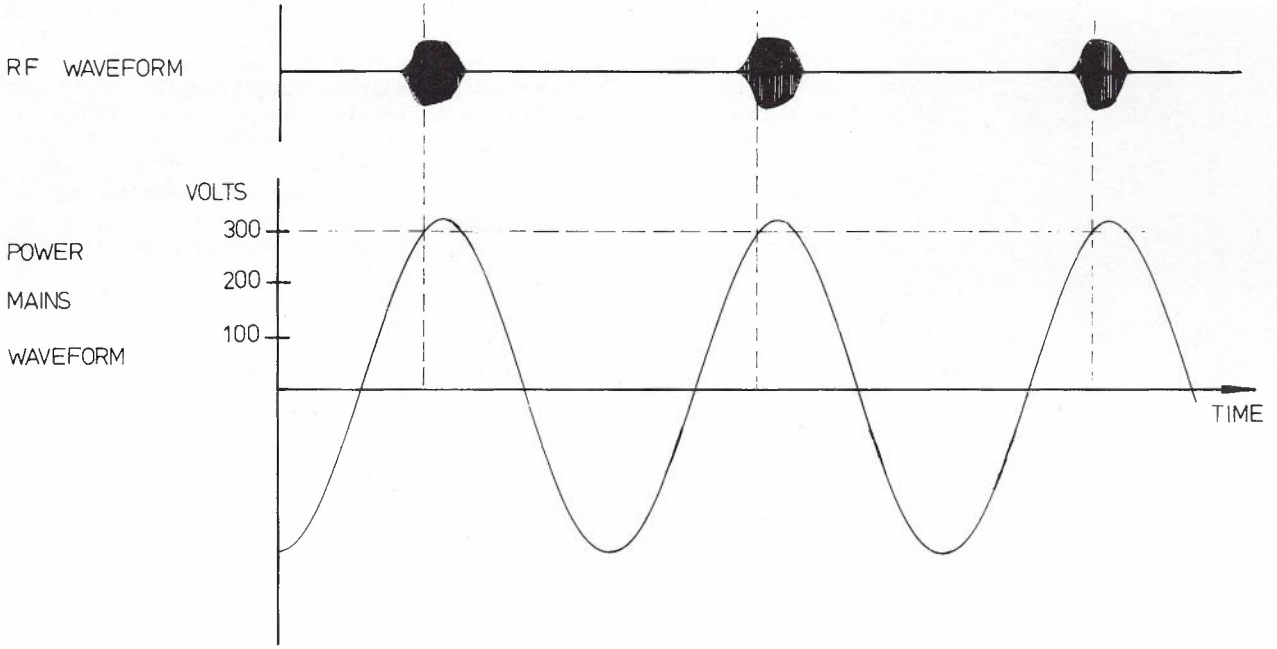


Fig. 3 - Waveform of mains supply and RF bursts (sampled with E-field probe) showing phase relationship and commencement of oscillations around 300 Volts.

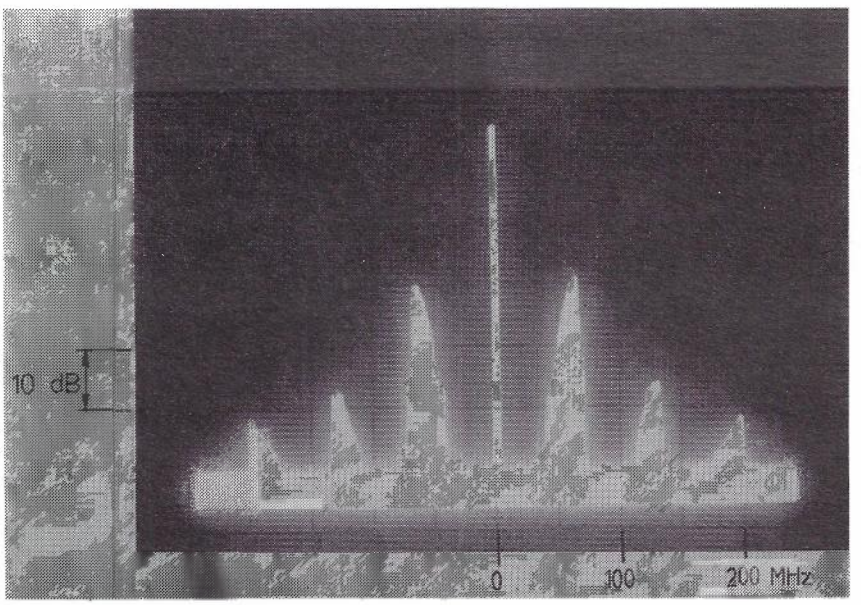


Fig. 4 - Spectrum of Lamp Emission (the centre line is a reference mark generated by the Spectrum Analyser)

BIOGRAPHY

PETER R. MURRELL graduated in 1976 from the University of Melbourne with a Bachelor of Engineering (Electrical). Mr Murrell joined the Telecom Australia Research Laboratories in 1977 where he worked in the field of microelectronics as an Electrical Engineer for a period of two years. In 1979 he was seconded to the Department of Science, Antarctic Division and spent 15 months at Davis Base, Antarctica, where he worked as an Electrical Engineer on upper atmosphere and magnetic measurement projects.

Mr Murrell rejoined the Telecom Research Laboratories in 1981 and is currently working in the field of microwaves in the Satellite and Electromagnetic Environment Section.

Information for Authors

ATR invites the submission of technical manuscripts on topics relating to research into telecommunications in Australia. Original work and tutorials of lasting reference value are welcome.

Manuscripts should be written clearly in English. They must be typed using double spacing with each page numbered sequentially in the top right hand corner. The title, not exceeding two lines, should be typed in capital letters at the top of page 1. Name(s) of author(s), in capitals, with affiliation(s) in lower case underneath, should be inserted on the left side of the page below the title. An abstract, not exceeding 150 words and indicating the aim, scope and conclusions of the paper, should follow below the affiliation(s).

ATR permits three orders of headings to be used in the manuscript. First-order headings should be typed in capitals and underlined. Each first-order heading should be prefixed by a number which indicates its sequence in the text, followed by a full stop. Second-order headings should be underlined and typed in lower case letters except for the first letter of each word in the heading, which should be typed as a capital. They should be prefixed by numbers separated by a full stop to indicate their hierarchical dependence on the first-order heading. Third-order headings are typed as for second-order headings, underlined and followed by a full stop. The text should continue on the same line as the heading. Numbering of third-order headings is optional, but when used, should indicate its hierarchical dependence on the second-order heading (i.e. two full stops should be used as separators).

Tables may be included in the manuscript and sequentially numbered in the order in which they are called up in the text. The table heading should appear above the table. Figures may be supplied as clear unambiguous freehand sketches. Figures should be sequentially numbered in the order in which they are called in the text using the form: Fig. 1. A separate list of figure captions is to be provided with the manuscript. Equations are to be numbered consecutively with Arabic numerals in parenthesis, placed at the right hand margin.

A list of references should be given at the end of the manuscript, typed in close spacing with a line between each reference cited. References must be sequentially numbered in the order in which they are called in the text. They should appear in the text as (Ref. 1). The format for references is shown in the following examples.

> Wilkinson, R.I., "Theories for Toll Traffic Engineering in the U.S.A.", BSTJ, Vol. 35, No. 2, March 1956, pp.421-514.
>
> Abramowitz, M. and Stegun, I.A., (Eds), Handbook of Mathematical Functions, Dover, New York, 1965.

Three copies of the paper, together with a biography and clear photograph of each of the authors should be submitted for consideration to the secretary (see inside front cover). All submissions are reviewed by referees who will recommend acceptance, modification or rejection of the material for publication. After acceptance and publication of a manuscript, authors of each paper will receive 50 free reprints of the paper and a complimentary copy of the journal.

Benefits of Authorship for ATR.

- ATR is a rigorously reviewed journal with international distribution and abstracting.
- Contact with workers in your field in the major telecommunication laboratories in Australia can improve interaction.
- Contributions relevant to the Australian context can contribute to the viability and vitality of future Australian research and industry.
- Publication is facilitated because ATR arranges drafting of figures at no cost to the author, and with no need for the author to spend time on learning journal format standards.

ATR AUSTRALIAN TELECOMMUNICATION RESEARCH
ISSN 0001-2777

VOLUME 17, NUMBER 2,
1983

Titles (Abbreviated)