Founded 1925

Incorporated by Royal Charter 1961

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

VOLUME 42 No. 12 DECEMBER 1972

THE RADIO AND ELECTRONIC ENGINEER

The Journal of the Institution of Electronic and Radio Engineers

Integrating Communications

IN 1946, in his Presidential Address to the Institution, the Earl Mountbatten of Burma made a plea for the three fighting services of the British Commonwealth to implement their discussions on a single unified and integrated network of world wide communications. Fifteen years later, in an Address to the Radio and Electronics Industries Councils, he found it necessary to speak in almost the same terms: it is only in the fast few years—a quarter of a century after Lord Mountbatten first publicly raised the subject—that such an integrated network has approached implementation. It is probably fair to say that the great potentialities offered by the use of artificial earth satellite repeaters for communications links of all kinds—voice, data and vision—have been a deciding factor in achieving the present extent of interservice co-operation.

What, then, may be regarded as the vital technical factors which will bring to reality an even greater project for integrating civil communications in Great Britain which is now coming to the fore and is popularly known as the 'electronics grid'? The National Electronics Council, under the chairmanship of Lord Mountbatten, is giving considerable encouragement to the British Post Office in overcoming what seems to many to be the disappointingly slow progress of the proposals set out in a Report prepared by the Post Office and published in the *NEC Review* in October 1967.*

In its most ambitious form the 'electronics grid' is envisaged as providing the conventional telephone, teleprinter and data networks as parts of an integrated system which also includes entertainment and educational and business video services, industrial control functions, and such applications as remote reading of domestic meters for public utilities; two-way operation of many of these services is a further extension and indeed new possible uses for such a system are being put forward continuously. Some of these applications are already being individually employed to limited extents but, as far as the linking of other than conventional telecommunications services to the private subscriber is concerned, Post Office involvement is strictly at the 'pilot project' stage. Systems are being installed at several new towns in different parts of the United Kingdom, for in such developing areas the 100% penetration desirable for the 'wired city' can be most easily achieved. At Milton Keynes in Buckinghamshire a coaxial cable network capable of handling a frequency band up to 270 MHz carrying several high quality television channels (on both the present UK line standards) as well as sound radio channels, is being installed, and a brief description of the scheme appears in a recent issue of the Post Office *Telecommunications Journal.*†

An assessment of these developments of the telecommunications networks was made last year by Professor H. M. Barlow,‡ in which he pointed out that neither cables nor free space could be expected to cope with the various facilities: microwave waveguides now and optical-fibre guides in, say, thirty years' time are two technological developments which will go far to making practical a 'wired city' and the 'electronics grid'.

But the complexities of the schemes call for the most advanced electronic circuitry: in fact, as was the case in achieving practical pulse code modulation, the answer to the integration of communication networks is—the integrated circuit. Some of the background to the technological problems involved will be dealt with in the January/February 1973 issue of *The Radio and Electronic Engineer*, which will mark the 25th anniversary of the invention of the transistor—and the first quarter-century of the semi-conductor revolution.

* 'Telecommunications services of the United Kingdom', *NEC Review*, **3**, No. 2, pp. 37–41, October 1967.
† Granger, S. H., 'The city of tomorrow', *P.O. Telecommun. J.*, **24**, No. 3, pp. 2–4, Autumn 1972.
‡ 'Telecommunication services in the United Kingdom—future and overall policy', *Nat. Electronics Rev.*, **7**, No. 2, pp. 30–31, March 1971.

Contributors to this issue

Professor D. A. Bell (Fellow 1967) received his first degree in physics from Oxford University in 1933 and the degree of B.Sc. for research in the Electrical Engineering Department in 1934. He was employed in the electronics and communications industry for the next 14 years and in 1949 he joined the Electrical Engineering Department of the University of Birmingham, where he became Reader in Electromagnetism and and in 1953 received the degree of Ph.D. In 1959 he initiated a postgraduate course in information engineering in that Department. From 1961 to 1964 he was Director of AMF British Research Laboratory. He spent three months in 1965 as visiting professor of telecommunications at McGill University and subsequently joined the University of Hull to set up a new Department of Electronic Engineering. Professor Bell was a member of the Post Office Economic Development Committee from 1966 to 1970. He is the author of several books and of many papers and articles on communications.

Mr. Ulrich Haller studied at the Technical Universities of Stuttgart and Darmstadt and graduated in 1961 as Diplomingenieur in communication engineering. Since 1961 he has been a member of the scientific staff of the Research Institute of AEG-Telefunken, Ulm. For the past four years he has led the research group for protected data transmission.

Mr. Hans-Jurgen Matt graduated from the Technical University Darmstadt in 1968 as Diplomingenieur in communication engineering and joined AEG-Telefunken in 1969 as a member of the scientific staff of the Research Institute, Ulm. He is engaged in protected data transmission and is especially concerned with coding theory and computer simulations of systems.

Mr. Max Prögler graduated as Diplomingenieur in communication engineering from the Technical University of Munich in 1968. He has been with AEG-Telefunken as a member of the scientific staff of the Research Institute, Ulm, since 1968. His work in the field of protected data transmission has been especially concerned with coding theory and error-correction procedures.

Mr. Y. Sunthankar received his bachelor's degree in physics in 1949 and M.Sc. in electronics in 1951 from Bombay University, the M.Sc. in crystallography from London University in 1963 and the M.Phil. from Reading University in 1971. He worked as a development engineer with Automatic Telephone and Electric Co. Ltd. at Liverpool and in associated Companies from 1953 to 1962 on the development of v.h.f. transmitters and receivers. He joined Oxford Polytechnic as a lecturer in electronics in 1967 and since October 1972 he has been a Senior Lecturer in electronics at Sunderland Polytechnic. His fields of research have included ultrasonics, measurement of magnetostriction and transformer noise and the development of a matric computer in collaboration with Washington University, St. Louis, Missouri, U.S.A.

Mr. R. Longbottom (Member 1971) joined the Government's computer engineering consultative service, the Technical Support Unit, in 1960, after initial technical training as a Post Office Engineer in the Bradford Telephone Area. Throughout his career in TSU he has been involved in reliability and maintainability assessment of computer systems; at first this was in conjunction with duties concerning environmental and accommodation factors, then with computer system acceptance testing and later with design appraisal studies. He was promoted to Executive Engineer in 1963 and Senior Executive Engineer in 1969. In 1972, with transfer of TSU into the Central Computer Agency, he joined the Civil Service as a Professional and Technology Officer Grade 1 and is currently responsible for appraisal and testing of large scientific computer systems.

Dr. G. B. Lockhart graduated with honours (B.Sc.(Eng.)) from Aberdeen University in 1965 and was awarded M.Sc. (Aberdeen 1966). He joined the Communication Section of the Electrical Engineering Department at Imperial College in 1966 to work on s.s.b. techniques for broadcasting with the support of a BBC research scholarship, and he was awarded a Ph.D. degree in 1970. From 1969 to 1971 Dr. Lockhart was employed as a research fellow at the University of Technology, Loughborough, sponsored by the Joint Speech Research Unit (Eastcote, London), working on digital encoding of speech, particularly delta modulation techniques. He has been employed as a Lecturer at Leeds University Department of Electrical and Electronic Engineering since October 1971, and his special interests are in the areas of communication and digital signal processing.

Dr. C. F. Ho (Member 1965, Graduate 1959) is a lecturer in the Electrical Engineering Department at the University of Hong Kong. (See also *Journal* for June 1971.)

U.D.C. 621.384.182

A Forward Error Correction System for Heavily Disturbed Data Transmission Channels

U. HALLER, Dipl.-Ing.*

H. J. MATT, Dipl.-Ing.*

and

M. PRÖGLER, Dipl.-Ing.*

Based on a paper presented at the IERE Conference on Digital Processing of Signals in Communications held in Loughborough from 11th to 13th April 1972

SUMMARY

This paper describes a new forward error correction system suitable for protected data transmission on heavily disturbed transmission channels, such as radio channels or dialled telephone lines. The system consists of two independent coding stages, one designed to combat random errors due to a binomial distribution and the other to correct the remaining bursts. This configuration adapts the system to channels, which contain random errors as well as bursts. In addition, a signal quality detector and a device designed for channel status measurements control the two coding stages in a way to achieve optimum adaption of the decoders to the actual channel disturbances. The channel status measurement makes a better use of the redundancy of the burst error correction stage.

Furthermore, the system's error detection characteristics avoid the delivery of data to the sink, which cannot be corrected with sufficient residual error probability with high security. Assuming a required residual error rate of $p_{rest} \sim 10^{-8}$, random errors with error rates up to $p_{ist} \sim 10^{-3}$ and bursts with lengths up to about 2900 bits are correctable.

* *AEG-Telefunken Research Institute, D-7900 Ulm (Donau), West Germany.*

1. Introduction

Recently, a considerable increase in remote data processing has taken place. For this purpose the data transmission is accomplished via channels contained commonly in existing transmission networks, i.e. telephone or radio channels using various kinds of transmission speeds. Most of these channels are disturbed in such a way that data protection is necessary.

With redundant coding[1] and signal quality detectors[2] the error rates of transmission channels can be reduced to nearly any value. In most protection systems the correction of erroneous received data is performed by their retransmission. Those *feedback systems* require a feedback channel for the transmission of receipt signals produced by aid of error detecting means. Such systems are available already, and they can be adapted to most common transmission channels to give a high performance of transmission rate and security.[3]

If there is no feedback channel available, the necessary data protection can be obtained by *forward error correction systems*. Here the correction is accomplished by use of the transmitted code redundancy. For this purpose coding theory offers various kinds of correction procedures and algorithms, suited to combat most different disturbance and error structures.[4] Here too, the efficiency of the decoding procedures can be improved considerably using the analogue redundancy of the channel with signal quality detectors. This so-called soft decision enables the decoding process to be controlled, increasing the correction capacity of the code and shortening the decoding time.[5–7]

The problem for realization of forward error correction systems consists of finding an adequate block length: to reduce the amount of hardware the block length should be as small as possible without a substantial loss of security. Moreover, the information rate should be as high as possible, a requirement which is incompatible with the demand for small block length, too. Thus in developing forward error correcting systems, one has to find a compromise between the amount of hardware, the attainable security and channel utilization.

This paper describes a new forward error correction system suitable for protected data transmission on heavily disturbed transmission channels, such as radio channels or dialled telephone lines. The system consists of two independent coding stages, one designed to combat random errors due to a binomial distribution and the other to correct the remaining bursts. This configuration allows the system to be adapted to channels, which contain random errors as well as bursts. Additionally, a signal quality detector and a device designed for channel status measurements control the two coding stages in a way to achieve optimum adaption of the decoders to the actual channel disturbances. The channel status measurement makes a better use of the redundancy of the burst error correction stage as it nearly doubles the correctable burst length. Furthermore, the system's error detection characteristics avoid the delivery of data to the sink, which cannot be corrected with sufficient residual error probability with high security. Assuming a required residual error rate of

$p_{rest} \simeq 10^{-8}$, random errors with error rates up to $p_{ist} \simeq 10^{-3}$ and bursts with lengths up to about 2900 bits are correctable.

The following section contains a short survey of the error structures which have been considered for this system and of suitable protection methods. A description of the system's basic functioning and its components follows. The last section deals with the estimation of the system's performance and the devices to measure it.

2. Transmission Channels

As shown in the literature, radio channels may be even more disturbed than bad telephone channels. K. Brayer and O. Cardinale made extensive studies of h.f. radio connexions using frequencies from 2 to 30 MHz at data rates of 1200 bit/s and 2400 bit/s.[8, 9] They found error rates between 10^{-5} during good and 10^{-2} during bad transmission conditions. The variable error rates substantially result from the extreme time dependence of the propagation conditions. During transmission times with error rates of $\geqslant 10^{-3}$ they found an increase of error bursts with densities of $\geqslant 0{\cdot}1$. Due to fading or atmospherics these bursts last from a few seconds to longer channel interruptions. During transmission times with error rates of $\leqslant 10^{-4}$ the channels are mostly disturbed by random errors. Similar results have been obtained by Cohn *et al.*[10]

The same authors investigated the problems of suitable protection procedures by aid of computer simulations. They compared normal and concatenated block coding with interleaving in one or both stages to convolutional codes with threshold decoding methods.[8, 10] To reduce error rates from $10^{-2} \ldots 10^{-3}$ to $10^{-5} \ldots 10^{-6}$ these methods need codes with 50% redundancy and constraint lengths between 5000 and 63 000 bits according to the actual error structures. Both authors mention the advantages of concatenated or tandem coding because each stage can be designed according to various error structures.

The correction system to be presented is able to cope with mixed error structures, i.e. to correct simultaneously appearing bursts and random errors. Compared to the systems mentioned in literature it uses smaller block lengths and should attain better residual error rates for the data delivered to the sink by using a channel status measurement which among other things indicates uncorrectable data blocks.

3. Outlines of the System

Figure 1 demonstrates the block diagram of the forward error correction system containing sending and receiving data protection device, each of which consists of two coding stages. In addition the receiver contains a signal quality detector, a device to measure the channel status, and synchronization circuits. The transmission starts with bit and block synchronization followed by an indication of the start of information. At first, source data are delivered to coder 2 for burst error protection and then to coder 1 to combat random errors. The two-stage encoded data are transmitted through a disturbed digital channel consisting of modulator, analogue channel and demodulator. At the receiver decoder 1 corrects random errors first. This circuit and the signal quality detector indicate disturbed, erroneous or uncorrectable blocks to the channel status measuring device. This circuit locates error bursts, measures their lengths and controls the correction of burst errors in decoder 2. Furthermore, the circuit may control decoder 1 to work either in correction or detection mode to guarantee that the correction capacity of code 1 will not be exceeded, even during high error density regions of the channel. It is also possible to cause decoder 1 to correct random errors or short bursts. The corrected information of decoders 1 and 2 is delivered via an output circuit to the sink, where the channel status measuring device marks uncorrectable information sequences.

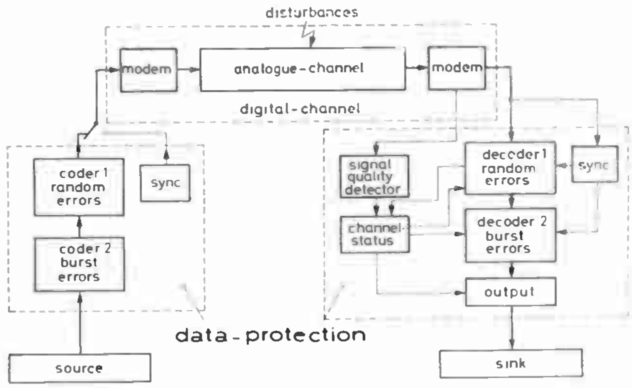


Fig. 1. Forward error correction system.

Figure 2 shows the system's information and redundancy flow. Coder 2 receives data bits from the source, calculates redundancy, forms data blocks protected against burst errors and delivers these blocks to coder 1. Coder 1 separates these blocks into shorter sequences and evaluates their respective redundancies, thus protecting the data against random errors. The constraint length consists of the source data and the accompanying redundancy of both encoders.

The receiver first checks the incoming blocks, i.e. the signal quality detector indicates error bursts with high security and simultaneously decoder 1 marks all small blocks with detected errors by the code, especially all weak disturbed blocks. The channel status measuring device combines both indications and defines the error

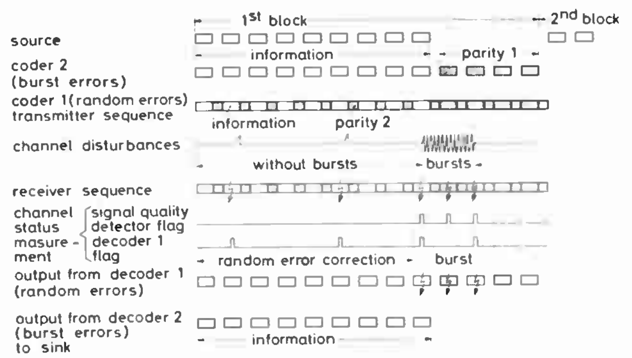


Fig. 2. Two-stage forward error correction system (System flow).

burst regions in location and length. The system's special advantage is the combined application of decoder 1 with a signal quality detector because during burst error regions the error detection capability of the quality detector increases whereas the capability of code 1 decreases.

Simultaneously with the processes of the channel status measuring device, decoder 1 corrects random errors and feeds the corrected data to decoder 2. If decoder 1 cannot correct a block, the channel status measuring device receives a signal again and adds it to the measured burst region. Decoder 2 now corrects all indicated correctable bursts and passes the information to the sink.

4. The Components of the System

In the following section the most important components and parameters of the system are described.

4.1. *Synchronization*

A basic requirement for every digital transmission system is the synchronization of emitter and receiver. In the receiver the synchronization has to guarantee:

(i) the correct bit clock,
(ii) the correct block phase for the decoders,
(iii) the exact starting-point of the transmission of information:

An additional difficulty in forward error correction systems is the fact that the transmitter has no information about whether the receiver succeeded or failed in synchronizing. Therefore, the receiver must synchronize correctly at the start of transmission, even during heavy disturbances. This synchronization must be maintained during the whole transmission even during heavy disturbed periods.

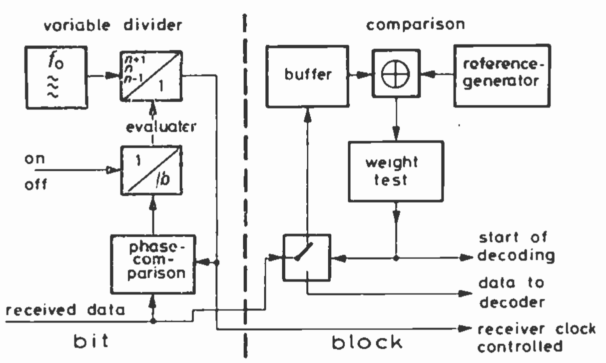


Fig. 3. Synchronization.

The synchronization of the bit clock uses a digital approach on the basis of a variable frequency divider, usually dividing the frequency f_0 of the timing generator by n, thus giving the bit clock rate f_B (Fig. 3).[11] This clock is compared in phase with the received data and its deviation adjusts the division ratio. A leading phase causes the circuit to divide by $(n+1)$ and a lagging phase by $(n-1)$; i.e. the actual edge of the bit clock is delayed or accelerated. An evaluator suppresses most events of disturbances.[12] The evaluator changes the division ratio only if b consecutive phase deviations occur with the same direction. If the direction of phase deviations changes after less than b events the evaluator is reset. Thus the probability for wrong adjustments of the clock phase due to disturbances decreases proportionally to about 2^{-b}. At the beginning of transmission the evaluator is switched off to allow a quick initial synchronization. During heavy disturbances the signal quality detector interrupts the whole readjustment of the phase. This calls for a good stability of the timing generators in both terminals.

The block synchronization and the marking of the exact start of information transmission is achieved by vsing a maximal-length code (pseudo-noise sequence) with length $n = 2^i - 1$. Two identical code-vectors are emitted consecutively; the beginning of the first activates the receiver and its bit synchronization. After this the receiver compares sequences of n received bits with the known synchronization sequence which is produced by an internal reference generator (Fig. 3). Every incoming bit shifts the compared sequence one bit ahead. The Hamming distance between every code word of length n —resulting from those cyclic shifts—and the known synchronization sequence is exactly $d = (n+1)/2$ for the undisturbed case.[4] The distance is zero only if both sequences are identical, i.e. if the block synchronization has been found. In case of transmission errors the weight of the syndrome (syndrome = difference between received and synchronization word) usually will be $w \gtrless d$ and only in case of exact block synchronization is $w > 0$. Therefore, a reliable detection needs a threshold S with the bounds $0 < S < d$. The optimum value of S will be in the range of approximately $d/2 < S_{opt} < d$ because errors in the non-synchronized state may increase (advantageous case) or decrease (disadvantageous case) the syndrome weight whereas in the synchronized state errors increase the weight only. Considering a block length of $n = 511$ and a threshold value of $S \simeq 200$, the probability for false block synchronization and the probability for undetected block synchronization will be beyond 10^{-7} up to an average bit error rate of about 0·3.

Transmission of data blocks begins immediately after the block synchronization. Resynchronization must be provided if a loss of block synchronization is unavoidable, e.g. due to long lasting heavy disturbances during which the phase readjustment will be interrupted. The correction of small deviations of the block phase may be accomplished with specific synchronization patterns, which are coded with the information; some methods are described by Stiffler.[13] Another possibility is to perform a complete block synchronization—as described above, for instance—in periodical distances which enables a protected information transmission, even in case of unsuccessful or wrong initial synchronization.

4.2. *Correction Stage for Random Errors*

The inner stage optionally uses two modified BCH-Codes, a (23, 11)-code and a (63, 44)-code, with information rates of $(n-k)/n = 0{\cdot}5$ and $0{\cdot}7$. The Hamming distance of both codes is $d = 8$, therefore correcting up to $e = 3$ errors and detecting 4 errors simultaneously.[4, 14]

Error correction is achieved by cyclic permutation combined with error word superposition.[15, 16] At first errors are located followed by the actual correction procedure. In polynomial description a received word is

$$c_f(x) = c(x) + f(x) \qquad \ldots\ldots(1)$$

with the emitted code word $c(x)$ and the so-called error word $f(x)$. At first, the decoder calculates the syndrome by dividing $c_f(x)$ by the generator polynomial $g(x)$ of the code:

$$s_0(x) = R\left\{\frac{c_f(x)}{g(x)}\right\} = R\left\{\frac{f(x)}{g(x)}\right\} = R\left\{\frac{f_i(x)}{g(x)}\right\} + f_r(x) \quad \ldots(2)$$

$\left(R\left\{\frac{f(x)}{g(x)}\right\}\right.$ denotes the remainder of the division.$\left.\right)$

The syndrome weight is $w(s_0) > e$ if $f_i(x) \neq 0$. The error word is $f(x) = f_i(x) + f_r(x)$ with its information part $f_i(x)$ and its redundancy part $f_r(x)$. To accomplish the correction it is sufficient to determine $f_i(x)$ and to subtract it from the received word, which means for the syndrome

$$s(x) = f_r(x) \qquad \ldots\ldots(3)$$

with

$$w(s) = w(f_r) \leqslant e - w(f_i)$$

Because $f_i(x)$ is unknown, different possible $f_j(x)$ must be tested:

$$s_{0j}(x) = R\left\{\frac{f(x)}{g(x)}\right\} - R\left\{\frac{f_j(x)}{g(x)}\right\} \quad j = 1, 2, \ldots j_{\max} \qquad (4)$$

The correct $f_j(x) = f_i(x)$ is found if the syndrome weight satisfies $w(s_{0j}) \leqslant e - w(f_j)$.

It is possible to reduce the number $j_{\max}$ of tested error patterns substantially by comparing a certain $f_j(x)$ with $f(x)$ and all its cyclic shifts $x^a.f(x)$ $(a = 1, 2, \ldots (n-1))$. The aim of this process is to move the maximum possible number of errors into the redundancy thus having a $f_j(x)$ with minimum weight. We form

$$s_{aj}(x) = R\left\{\frac{x^a.f(x)}{g(x)}\right\} - R\left\{\frac{f_j(x)}{g(x)}\right\}$$

$$= R\left\{\frac{x^a s_{0j}(x)}{g(x)}\right\} - R\left\{\frac{f_j(x)}{g(x)}\right\} \qquad \ldots\ldots(5)$$

Equation (5) implies the algorithm for finding out the errors in the received word: Calculate for $a = 1, 2, \ldots (n-1)$ and $j = 1, 2, \ldots j_{\max}$,

$$s_{aj} = R\left\{\frac{x^a s_{0j}(x)}{g(x)}\right\} - R\left\{\frac{f_j(x)}{g(x)}\right\} \qquad \ldots\ldots(6)$$

until

$$w(s_{aj}) \leqslant e - w(f_j)$$

If for a certain a the equations (6) are valid the actual error pattern has been found,

$$f(x) = (f_j(x) + s_{aj}(x))x^{-a} \qquad \ldots\ldots(7)$$

Thus, the received information will be corrected by superposition of mod 2 of the actual error pattern on to the received block. To initiate the correction the blocks are fed from the buffer storage to the syndrome and the information register, as shown in Fig. 4. The

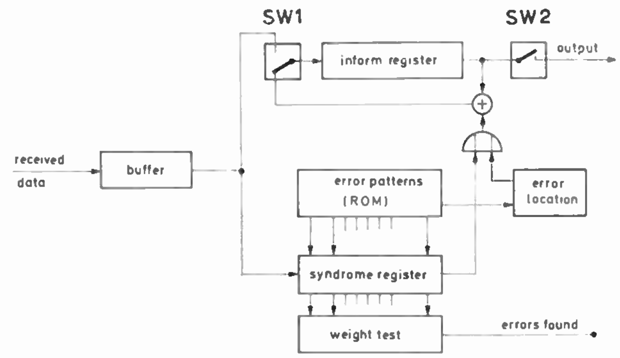


Fig. 4. Random error correction.

syndrome calculating register is connected with a weight evaluating circuit and furthermore with a read-only memory (r.o.m.) where the different vectors corresponding to $R\{f_j(x)/g(x)\}$ and to $f_j(x)$ are stored.

The (63, 44)-code needs 13 error patterns and the (23, 11)-code 1 pattern, respectively, which are to be stored into the r.o.m. If the error pattern is found and its position stored in the error location circuit the vectors $s_{aj}(x)$ and $f_j(x)$ will be added into the corresponding places of the information sequence, which will circulate in its register for this purpose. If the weight of the syndrome exceeds the given threshold $w(s_{aj})$ for all $(n-1)$ cyclic shifts, uncorrectable errors have been detected, which will be indicated to the next decoding stage. The information is read out synchronously with the transfer of the next block from the buffer to decoder 1.

An advantage is that the described procedure needs a relatively low amount of hardware. In contrast with more algebraic methods—as treated by Gallager, for instance[4] —this procedure needs no arithmetic elements for computation. A simple read-only memory to store a few patterns is sufficient. Furthermore, the speed is no problem, because only 14×63 and 2×23 respectively comparisons are needed.

4.3. *Correction Stage for Burst Errors*

The decoder 2 realizes two functions; first, it corrects all bursts with lengths up to a certain limit, and secondly, it assigns a fidelity signal to all blocks which are delivered to the sink. The corresponding information will not be delivered without being noted as uncorrectable if the channel disturbances exceed the correction capacity of the code. This component of the whole system is designed to operate with two modes, i.e. it is programmable to work with the code only or to use the signals of the channel status measuring device. The decoder 2 uses a BCH-code with blocklength $n' = 15$ and a known burst correction capacity b'. This code has been interleaved cyclically to a block length n of about 2600 bits, thus guaranteeing a burst error capability of $b = sb'$ due to interleaving factor s.[17] This coding stage also uses two codes with the information rates of 0·5 or 0·7 respectively. These codes have been chosen for the following reasons: (1) The burst error capacity is known. (2) These codes can be implemented with relatively complex but cheap integrated m.o.s.-registers due to the

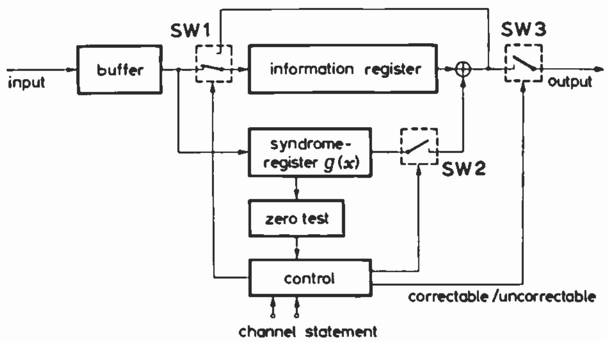


Fig. 5. Burst error correction.

small number of feedback connexions necessary for the generator polynomials of the interleaved codes.

Two cases must be distinguished for correcting burst errors. If the burst location is known it is sufficient to move the whole burst into the redundancy of the cyclic code, to read out its pattern from the syndrome and to correct the burst. This mode allows burst lengths to be corrected up to the length of the redundancy. If the burst is unknown, the syndrome must be checked by a zero test logic to find the burst location and pattern, as described by Gallager,[4] for instance. In this case the correctable burst length is reduced to half of the redundancy length, at least.[21]

Figure 5 shows the block diagram of decoder 2 of the outer stage. Data blocks from the channel are fed simultaneously (via a buffer for matching the speed) into the information and syndrome register. After the complete inputting of an information sequence the information register stops and the remaining redundancy bits enter the syndrome register. Now the decoding procedure starts. During the loading process a counter checks the feedback line of the syndrome register for a sequence of $n-k-b = m-b$ consecutive zeros. Every intermediately occurring 'one' clears the counter. After the entrance of a whole block the counter is interrogated for the first time. Indicating a zero chain with adequate length it signifies the discovery of the unknown burst. Then switch SW2 is closed and the content of the syndrome register will be added to the information sequence.

Usually, the burst will not be detected at once. In this case switch SW1 is commutated to provide a circulation of the information sequence and both the information and syndrome register are shifted by the same clock to maintain the correct phase relationship between them. After each clock pulse the zero test logic is interrogated

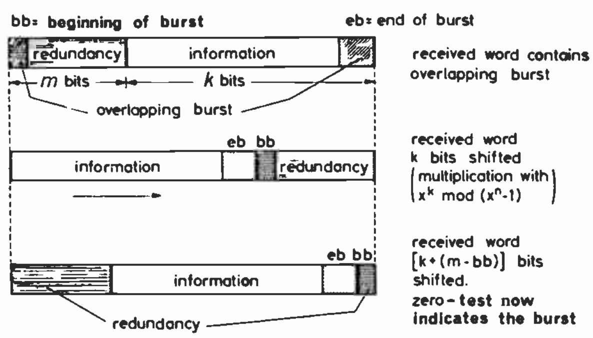


Fig. 6. Correction of overlapping burst.

again until the burst is detected because each correctable burst appears once in the syndrome register exactly.[18] If, during n consecutive pulses, no adequate zero chain has been found an uncorrectable error pattern is detected. In this case the fidelity signal marks the information to be uncorrectable. The organization oi decoder 2 implies always a circulation of the information in its register. This allows correction of so-called overlapping bursts, too, i.e. errors have occurred at the beginning and the end of a data block as shown in Fig. 6. Such overlapping bursts will be indicated only after the information has circulated in its register, for the syndrome register is now searching for bursts beginning in the redundancy where all overlapping bursts begin. During this examination the information register is stopped. After having found the start bb of an overlapping burst, switch SW2 is closed and the beginning of the syndrome pattern leaves its register until m clock pulses of exclusive syndrome register working are finished. Then switch SW3 is closed and the rest of the syndrome pattern eb is added to the information thus correcting the information delivered to the output of decoder 2. The control circuit of Fig. 5 supervises this algorithm and can be programmed to control the switches for the superposition of the syndrome pattern on the information either by instructions from the zero test logic—as described—or from the channel status measuring device, which will be discussed later.

4.4. *The Signal Quality Detector*

In addition to redundant coding signal quality detectors are well suited to recognize channel disturbances and their locations. Various principles of signal quality estimation and its performance have been investigated.[2] It has been found that signal quality detectors using so-called separate areas of correct reception are very efficient. Our system makes use of such a device to gain information about the location of bursts.

The following section deals with signal quality detectors for frequency and differential phase-shift keying. In both cases the efficiency in disturbance detection results from thresholds which limit substantially the decision areas for the analogue signals compared to the reception areas of the demodulator, i.e. the analogue signals are allowed to have small tolerances from the nominal value only. Every exceeding of the tolerances leads to a disturbance indication. For the adjustment of the thresholds (tolerances) a compromise must be found to achieve a minimum number of false alarm and a maximum number of correct detected disturbances. Furthermore, the thresholds must be self-adapting to the received signals. The signal quality detector for f.s.k. monitors the distortions of the frequencies (Fig. 7). For this purpose an upper and lower threshold U_o and U_u limit the output voltage u_D of the discriminator corresponding to both nominal frequencies. If the actual discriminator voltage leaves one of the tolerance areas, disturbance indications follow. The number of thresholds and the amount of hardware is reduced to nearly one half by making both tolerance areas symmetrical

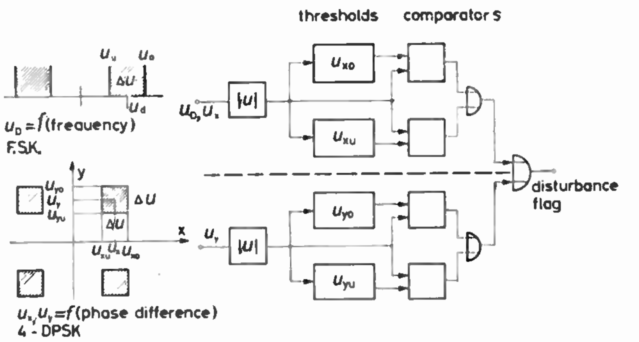


Fig. 7. Signal quality detector.

with respect to the average value of the discriminator output u_D. In this case only the deviations of the absolute value of u_D must be supervised by the thresholds.

The signal quality detector for phase modulated signals (d.p.s.k.) uses the same principles. The voltages u_x and u_y depending on the phase difference of two consecutive signal elements must form a vector which lies in one of the four receiving areas of Fig. 7. If the resulting vector exceeds the tolerance areas, a disturbance alarm will be given. Using the same idea of making the tolerance areas symmetrical the signal quality detector for d.p.s.k. signals consists of two detectors in parallel (for u_x and u_y) as described for the supervision of f.s.k. signals.

Our system uses the signal quality detector in a special application. It should indicate with high security all those code words of stage 1 containing more than $d/2$ errors, for in these cases the code may be not able to correct or even detect these errors. This is also very important for the channel status measuring device, which otherwise would not be able to detect short bursts, for instance. Measurements with a signal quality detector for f.s.k. have shown[19] that the probability for undetected blocks of length 63 containing more than three errors is smaller than 10^{-3}.

4.5. *The Channel Status Measuring Device (C.S.M.D.)*

The correction ability of the burst error correcting code may be increased considerably if the location of the burst is known.[20] Therefore, it is the task of the c.s.m.d. to determine location and length of any occurring burst with maximum reliability. The accuracy of this measurement influences the residual error probability and the actual information rate of the system, because any wrong burst location may lead to a wrong correction.

The information about burst location and length will be gained from the following indications:

(i) by the signal of a signal quality detector (s.q.d.)
(ii) by the indications of decoder 1 about
  (a) erroneous received words
  (b) uncorrectable received words.

The c.s.m.d. exploits the properties of the two-stage coding scheme by using indications from the signal quality detector as well as signals from the decoder 1, which reflect the actual error distribution in the channel. Long error-free regions lead to a sequence of received words with syndrome $S = 0$. Random errors cause a randomly distributed sequence of words with $S \neq 0$, too. Regions of disturbance bursts yield sequences of wrong received words with $S \neq 0$, exceeding the correction capacity of decoder 1. During these regions the decoder will also find uncorrectable blocks more frequently. These blocks obviously belong to the wanted burst pattern.

The c.s.m.d. has to calculate the actual error density of the channel from the frequency of wrong received words indicated by decoder 1. Though it seems to be possible to distinguish between different disturbance states of the channel, accurate measurements, however, require sufficiently long intervals of the disturbance regions considered. For practical reasons our c.s.m.d. has been designed to detect one region of high error density only. Such a zone is assumed if S or more erroneous blocks have occurred in a sequence of L consecutive blocks.

The signal quality detector's indications per bit and block are counted and compared to a value i. If the number of indications exceeds the value i, it is assumed that the corresponding block contains too much errors and therefore belongs obviously to the burst, too.

Figure 8 shows the block diagram of the c.s.m.d. The channel data are fed to the signal quality detector and to decoder 1. The different flags then are processed —as described—and combined in an OR-gate, whose output leads to a zero-sequence counter. The counter searches for the longest undisturbed region within a block of the burst error correcting code thus determines the rest of the block to be the burst. The circuit processes two programmable parameters ε and a to increase the reliability of the indications. Parameter ε denotes the number of blocks of decoder 1 which will be added on both sides of the burst and parameter a assures that no bursts with length greater than $(m-a)$ will be corrected (m = redundancy). This allows the last a bits of the syndrome to be examined for zeros, thus getting some information of the reliability of the correction. The signals of the c.s.m.d. concerning burst location and length directly control the correction procedure of decoder 2. This procedure allows the correction of normal and overlapping bursts.

5. Estimation of the System's Performance

Finally, some remarks will be made about the performance expected for this forward error correction

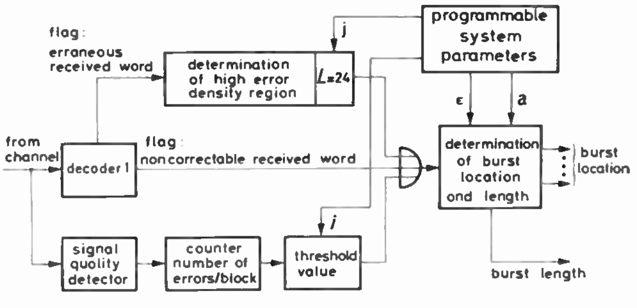


Fig. 8. Channel status measurement.

Table 1

$n = 23$

L	24			30			60		
s	1	2	3	1	3	4	2	6	8
P_N	4×10^{-3}	3×10^{-2}	1×10^{-1}	1×10^{-3}	4×10^{-2}	1×10^{-1}	2×10^{-5}	9×10^{-3}	5×10^{-2}
P_F	4×10^{-1}	1×10^{-1}	2×10^{-2}	5×10^{-1}	3×10^{-2}	5×10^{-3}	4×10^{-1}	2×10^{-3}	6×10^{-5}

$n = 63$

L	24			30			60		
s	1	6	10	2	7	12	4	6	8
P_N	3×10^{-7}	3×10^{-3}	2×10^{-1}	2×10^{-7}	2×10^{-3}	2×10^{-1}	8×10^{-13}	1×10^{-10}	6×10^{-9}
P_F	8×10^{-1}	8×10^{-3}	6×10^{-7}	6×10^{-1}	2×10^{-3}	8×10^{-8}	5×10^{-1}	2×10^{-1}	3×10^{-2}

system. Since theoretical estimations are most difficult and can be obtained for simple models only, the performance of the decoders will be measured, too.

5.1. *Theoretical Estimation of the System's Security*

The residual error rate of decoder 1 using its information rate 0·7 is smaller than 10^{-8} assuming a binomial error distribution of $p \simeq 10^{-3}$ in the channel. In this case uncorrectable blocks are expected to be indicated with a probability of 7×10^{-7}. Using the rate of 0·5 the same values are obtained for an average error probability of about 3×10^{-3}. The burst correction capability of decoder 2 depends from the rates of both decoders. The lengths vary from about 700 bits up to 2900 bits requiring guard spaces of about 1900 up to 2500 bits. These values represent an upper bound which can be reached by the c.s.m.d. only.

Now some possibilities of the c.s.m.d. will be discussed with the associated probabilities. It is one task of the c.s.m.d. to determine regions of high actual error probability of the channel. The circuit should recognize regions with error probabilities of 10^{-2} or greater, whereas error rates of 10^{-3} should not cause an indication because decoder 1 can handle it with high efficiency.

The error rate P_s determines the block error rate P_B which depends on the block length n. For simplification it is assumed in the following that all erroneous blocks will be detected by code 1 (i.e. the probability for indication of an erroneous block P_A is $P_A \sim P_B$) and that the different disturbance regions of the channel have binomial error distributions. Thus

$$P_A \sim P_B = 1-(1-P_s)^n \qquad \ldots\ldots(8)$$

The c.s.m.d. has to distinguish between the bit error probability $P_{s,h}$ (e.g. 10^{-2}) if there are s or more erroneous received blocks during a sequence of L blocks, and the probability $P_{s,l}$ (e.g. 10^{-3}). The probability for a correct indication of the heavily disturbed region is

$$P_R = \sum_{i=s}^{L} \binom{L}{i} P_A^i (1-P_A)^{L-i} \quad \text{for } P_A = f(P_{s,h}), \ \ldots(9)$$

and the probability that the high error region will not be detected:

$$P_N = \sum_{i=0}^{s-1} \binom{L}{i} P_A^i (1-P_A)^{L-i}$$

$$= 1-P_R \quad \text{for } P_A = f(P_{s,h}) \qquad \ldots\ldots(10)$$

Furthermore, the error rate $P_{s,l}$ may cause an erroneous indication of a high error region with the probability

$$P_F = \sum_{i=s}^{L} \binom{L}{i} P_A^i (1-P_A)^{L-i} \text{ for } P_A = f(P_{s,l}) \qquad \ldots\ldots(11)$$

The c.s.m.d. is required to indicate

(1) the regions with high security, i.e. P_R should be as large, P_N and P_F as small as possible;

(2) even short regions having a $P_{s,h}$ error rate, i.e. L should be as short as possible.

Equations (9) to (11) show that these requirements are contradictory; P_N and P_F will decrease with increasing L only, as well as a decreasing P_N results in an increasing P_F. Table 1 illustrates these dependencies for two block lengths $n = 23$ and $n = 63$. It is obvious that P_N and P_F decrease with increasing block lengths. Furthermore, the efficiency of the c.s.m.d. will grow with increasing difference between $P_{s,h}$ and $P_{s,l}$.

A good compromise seems to be $P_F \sim P_N$. If the probabilities P_N and P_F do not exceed a certain limit, a minimum value for the parameter L results, thus determining the minimum length of high error regions. It seems to be most difficult to calculate the influence of the c.s.m.d. on the residual error rate of the whole system, because little is known about the probability distribution of different error regions. Therefore, the next section deals with the possibilities of analysing the system by measurements.

5.2. *Measurements for Analysing the System's Properties*

To determine the system's quality concerning the most important parameters for reliable data transmission (i.e. the effective information rate V_{eff} and the residual error probability) a circuit for measuring the various parameters has been designed.

For this purpose the transmitter uses a maximal-length sequence generator as information source which enables the receiver to compare the received words with the true information, generated by an own reference source. We define the reduction factor R_{bit} of the bit error probability as the ratio

$$R_{bit} = \frac{P_r}{P} \qquad \text{......(12)}$$

where P_r is the output bit error and P the input bit error rate of the decoder. Figure 9 shows the block diagram of the system's evaluation circuitry, containing the reference generator and the counters for the most

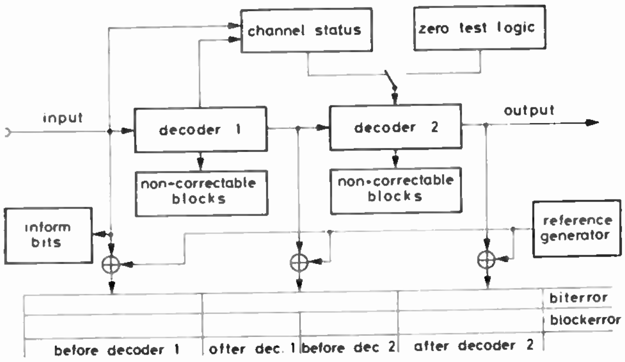


Fig. 9. Block diagram of evaluation circuit.

important events. This circuit measures the number of information bits transmitted as well as the number of bit and block errors at the input and output of every decoding stage. Thus, we find

$$P = \frac{\text{number of bit errors before decoding}}{\text{number of information bits}} \qquad \text{...(13)}$$

and the corresponding reduction factors

$$R_{bit} = \frac{\text{number of bit errors after decoding}}{\text{number of bit errors before decoding}} \qquad \text{...(14)}$$

$$R_{block} = \frac{\text{number of block errors after decoding}}{\text{number of block errors before decoding}} \qquad (15)$$

These values represent an objective measure for the security gain of the two stages. To calculate the actual transmission rate of the system all blocks which are detected uncorrectable are counted. Thus the information rate of the system is given by

$$V_{eff} = \left(1 - \frac{\text{number of uncorrectable blocks}}{\text{number of transmitted blocks}}\right) \cdot I_c \qquad (16)$$

I_c denoting the code rate.

The reduction factors and the actual information rate will be measured for each stage as well as for the whole system. By repeating a certain measurement, the influence of the various system parameters can be studied. This allows the performance of the different parts of the c.s.m.d. and its influence on the whole system to be examined. The measurements are expected to control our estimations, to yield results where calculations cannot be carried out and to help optimizing the system for the use on real channels.

6. Acknowledgment

The authors are grateful to Dr. H. Ohnsorge for his substantial encouragement and critical review of this paper.

7. References

1. Peterson, W. W., 'Error Correcting Codes' (Wiley, New York, 1961).
2. Ohnsorge, H., 'Zusammenhänge zwischen Redundanz und Fehlerwahrscheinlichkeit bei gesicherter Datenübertragung', Thesis, TH-Aachen, 1967.
3. Haller, U. and Matejka, D., 'A loop-time controlled data protection system for optimum data transmission in the telephone network', *NachrTech. Z.*, **23**, No. 3, pp. 134–42, 1970.
4. Gallager, R. G., 'Information Theory and Reliable Communication', (Wiley, New York, 1968).
5. Chase, D., 'Decoding with channel measurement information', IEEE Symposium on Information Theory, 1970, Noordwijk, Netherlands.
6. Ohnsorge, H., 'Redundant coding and its economical aspects', Digest of Eurocon, B 6-3, Lausanne, October 1971.
7. Dorsch, W. B., 'Maximum likelihood decoding of binary group codes for the Gaussian channel', IEEE Symposium on Information Theory, 1970, Noordwijk, Netherlands.
8. Brayer, K., 'Error-correction code performance on hf, troposcatter and satellite channels', IEEE Trans. on Communication Technology, **COM-19,** No. 5, Part II, pp. 781–9, October 1971.
9. Brayer, K. and Cardinale, O., 'Evaluation of error correction block encoding for high-speed hf data,' *IEEE Trans.*, **COM-15,** No. 3, pp. 371–82, 1967.
10. Cohn, D. L., *et al.*, 'Performance of selected block and convolutional codes on a fading hf channel', *IEEE Trans. on Information Theory*, **IT-14,** No. 5, pp. 627–40, September 1968.
11. Oetinger, R., 'Digitale Synchronisierung bei Datenübertragung über gestörte Übertragungskanäle', Festschrift zum 100-jährigen Bestehen der Ingenieurschule, Esslingen, Württemberg.
12. Swoboda, J., 'Ein Vorschlag zur Taktsynchronisation bei Datenübertragung', *Arch. Elekt. Übertrag.*, **22,** No. 11, pp. 509–13, 1968.
13. Stiffler, J. J., 'Theory of Synchronous Communications' (Prentice-Hall, Englewood Cliffs, N.J., 1971).
14. Stenbit, J. P., 'Table of generators for Bose-Chaudhuri codes', *IEEE Trans.*, **IT-10,** pp. 390–1, 1964.
15. Kasami, T., 'A decoding procedure for multiple-error-correcting cyclic codes', *IEEE Trans.*, **IT-10,** pp. 134–8, April 1964.
16. Matt, H. J. and Prögler, M., 'A generalization of permutation decoding and its combination with error pattern superposition', IEEE Symposium on Information Theory, 1970, Noordwijk, Netherlands.
17. Kasami, T., 'Cyclic codes for burst-error-correction', *J. Inst. Elec. Commun. Engrs*, **45,** pp. 9–15, 1962.
18. Ohnsorge, H., 'Erkennbarkeit von Fehlern bei Fehlerkorrektur durch zyklische Codes nach Peterson', *Arch. Elekt. Übertrag.*, **20,** No. 12, pp. 692–3, 1966, and **24,** No. 7/8, p. 380, 1970.
19. 'Signal Quality Detector with Self-Adjusting Thresholds', CCITT-Doc. COM Sp. A-71, September 1970.
20. Ohnsorge, H., 'Ein Verfahren zur Korrektur von Fehlerbursts größtmöglicher Länge, Bull. SEV (Schweiz), 61, pp. 720–4, 1970.
21. Matt, H. J., 'Verfahren zur Korrektur von Bündelfehlern', AEG-Telefunken Patentanmeldung Nr. P 2 037 877.2.

Manuscript received by the Institution on 21st January 1972. (Paper No. 1484/Com 59.)

U.D.C. 534.232:534.321.9

A Vibrating-blade, End-fire Ultrasonic Radiator

Y. SUNTHANKAR,
M.Phil., M.Inst.P.†

SUMMARY

The technique of producing highly directional ultrasonic radiation in the end-fire direction consists of exciting a beam (blade) clamped at one end by applying an a.c. signal to piezo-electric strain gauges placed in the best position to produce flexural vibration in the beam. Due to the acoustic impedance mismatch between the material of the beam and the medium to which the beam is coupled standing waves develop on the beam. A general theoretical expression for the acoustic radiation pressure, based on wavelength coincidence principle and the piston type theory has been derived. A design philosophy for the radiator has also been developed.

† *Department of Electrical and Electronic Engineering, Oxford Polytechnic; now at Sunderland Polytechnic, Sunderland, Co. Durham.*

1. Introduction

An ultrasonic radiator has been developed on a rather different principle to other acoustic transducers with fluid coupling. In other transducers or radiators usually operating in air an impedance match is used to maximize the energy transfer to the medium and so the formation of standing waves is avoided. The device developed, on the other hand, depends for its operation on the formation of standing waves. This can be achieved provided there is only a relatively weak coupling between the radiator and the medium to be excited. In other words, there must be a mismatch of impedances. The impedance relationships must be just like those in an organ pipe. Further, to obtain the maximum radiation, the wavelength in the transducer must be the same as the wavelength in the medium (wavelength coincidence condition).

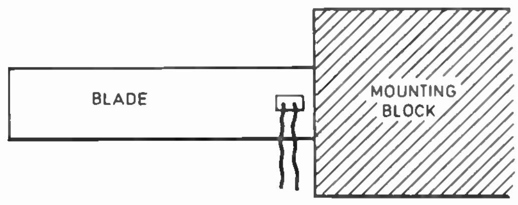


Fig. 1. The vibrating-blade end-fire ultrasonic radiator

The ultrasonic radiator which has highly directional properties produces standing waves in a thin rectangular blade clamped at one end. The blade is excited by applying a sinusoidal signal to piezoelectric strain gauges attached on either side of the blade at the clamped end (Fig. 1).

To obtain maximum radiation in an axial direction dimensions are chosen so that the wavelength in the blade is the same as the wavelength in air thus satisfying the wavelength coincidence condition.

2. Basic Theory

The basic principle governing the emission pattern from the radiator can be illustrated by considering a combination of the radiation emitted by simple sources. In Fig. 2 let $S_1, S_2, \ldots$ be a series of such sources in a straight line separated from each other by a half-wavelength. Provided the radiation from each source has the same amplitude and all are in phase, then in the direction $\theta = 0$ or $180°$ the response is a minimum and in a direction $\theta = 90°$ or $270°$ a maximum. If on the other hand adjacent sources are vibrating in opposite phase the response in the direction $\theta = 0°$ or $180°$ becomes a maximum. The standing waves in a transversely vibrating blade are in anti-phase at successive

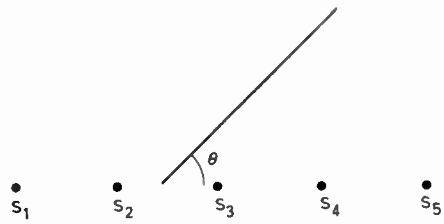


Fig. 2. Radiation from simple sources.

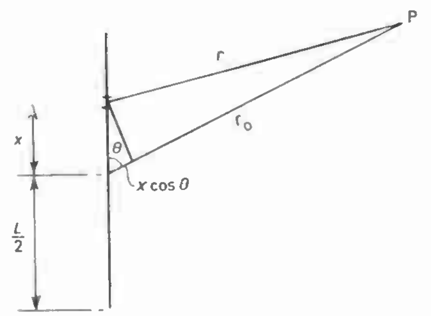


Fig. 3. Radiation from a thin blade.

half-wavelength intervals and the transmitted radiation is a maximum along the axis corresponding to the above second condition.

Figure 3 represents a thin blade of length L, and unit breadth. The pressure variation dp at a point P distance r from a vibrating element dx is[1]

$$dp = -\frac{j\rho c\beta v_x}{2\pi r}\exp\{j(\beta r - \omega t)\,dx,$$

where ρ = density of the material of the blade,
c = velocity of sound in air,
$\beta = 2\pi/\lambda_a$ where λ_a = wavelength in air,
v_x = velocity of element dx when distance x from the blade centre.

From Fig. 3, $r = r_0 - x\cos\theta$

$$dp = -\frac{j\rho c\beta v_x \exp\{j[\beta(r_0 - x\cos\theta) - \omega t]\}}{2\pi(r_0 - x\cos\theta)}\,dx.$$

When $r_0 \gg x\cos\theta$,

$$dp = -\frac{j\rho c\beta v_x}{2\pi r_0}\exp[j(\beta r_0 - \omega t)].\exp(-j\beta x\cos\theta)\,dx.$$

Put $v_x = v_0 \sin\beta_b x$ where $\beta_b = 2\pi/\lambda_b$, λ_b being the wavelength in the blade, and

$$dp = \frac{-j\rho c\beta \exp[j(\beta r_0 - \omega t)]}{2\pi r_0} v_0 \sin\beta_b x \times$$
$$\times \exp(-j\beta x\cos\theta)\,dx$$
$$= K\sin\beta_b x\exp(-j\beta x\cos\theta)\,dx$$
$$p = K\int_{-L/2}^{L/2} \sin\beta_b x\exp(-j\beta x\cos\theta)\,dx.$$

There are two conditions to be applied, first that the length of the blade is an integral number n of wavelengths, thus

$$L = n\lambda,\ \beta = \frac{2\pi}{\lambda} = \frac{2\pi n}{L} \text{ and } \beta\frac{L}{2} = n\pi$$

and second that $\beta = \beta_b$ thus satisfying the wavelength coincidence condition that $\lambda_a = \lambda_b$.

Integrating and applying these conditions gives

$$p = \frac{K2j}{\beta\sin^2\theta}\sin\left(\frac{n\pi}{2}\cos\theta.\right)$$

Substituting

$$K = \frac{-j\rho c\beta\exp[j(\beta r_0 - \omega t)]}{2\pi r_0} v_0$$

gives

$$p = \frac{v_0\rho c\beta\exp[j(\beta r_0 - \omega t)]}{r_0}\,\frac{\sin(n\pi/2\cos\theta)}{\sin^2\theta}$$

which is a maximum when $\theta = 0, 180°$ (see Appendix 1).

3. Design Philosophy for the Blade

A frequency of 40 kHz was chosen for the radiator and the dimensions and mode of vibration of the blade were calculated to comply with this figure and also with the condition $\lambda_a = \lambda_b$. The equation of motion of beam flexure is[2]

$$\frac{\partial^4 Y}{\partial x^4} = -\frac{A\rho}{EI}\frac{\partial^2 Y}{\partial t^2}.$$

where Y = deflexion
A = cross-sectional area
ρ = density of the material
E = Young's modulus of the material
I = second moment of area.

A standing wave solution of this equation is of the form
$Y = (a\cos\alpha x + b\sin\alpha x + c\cosh\alpha x + d\sinh\alpha x)\exp(j\omega t)$
where

$$\alpha^4 = \frac{\omega^2 A\rho}{EI}$$

and

$$\alpha^2 = \omega\sqrt{\frac{A\rho}{EI}}.$$

For a rectangular blade breadth b, thickness d, $A = bd$ and $I = bd^3/12$ thus

$$\alpha^2 = \frac{\omega}{d}\sqrt{\frac{12\rho}{E}}$$
$$= \frac{6{\cdot}92\pi f}{d}\sqrt{\frac{\rho}{E}}.$$

If the mode number n is greater than 5 (Ref. 2),

$$\alpha_n L \simeq (2n-1)\pi/2$$

giving

$$(\alpha_n L)^2 = (2n-1)^2\frac{\pi^2}{4} = \frac{6{\cdot}92\pi f L^2}{d}\sqrt{\frac{\rho}{E}}$$

whence

$$d = \frac{8{\cdot}81 f L^2}{(2n-1)^2}\sqrt{\frac{\rho}{E}}.$$

By applying the wavelength coincidence condition L can be expressed in terms of the frequency f since

$$\frac{2L}{n} = \frac{c}{f} \text{ and } L = cn/2f.$$

Substituting for L and assuming $2n \gg 1$,

$$d = \frac{8{\cdot}81 f\,c^2 n^2}{4n^2\ 4f^2}\sqrt{\frac{\rho}{E}}$$
$$= 0{\cdot}55\lambda_a\frac{c}{c_m}$$

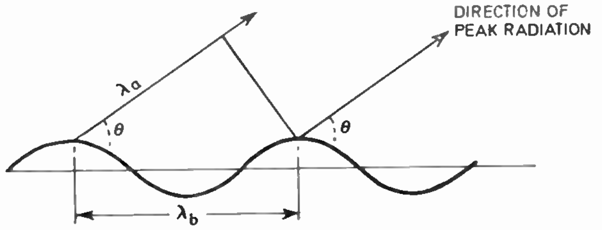


Fig. 4. Radiation from blade

where $c_m = \sqrt{\frac{E}{\rho}}$ = velocity of sound in blade material.

Too high a value of c_m results in an impractically thin blade. Aluminium (c_m = 6400 ms^{-1}), steel (c_m = 6000 ms^{-1}) and brass (c_m = 3500 ms^{-1}) were considered and brass finally chosen. The constants used for the radiator were

$$c_m = 3500 \text{ ms}^{-1},\ n = 24,\ f = 40 \text{ kHz},\ \lambda_a = 8{\cdot}25 \text{ mm}$$

giving $d = 0{\cdot}43$ mm and $L = \frac{cn}{2f} = 99$ mm.

The breadth b was chosen to be 19 mm which is greater than the thickness and the wavelength λ_b and minimizes any tendency to develop cross-vibrations.

4. Variation of Beam Angle with the Wavelength in the Blade

Let us consider the radiation from two effective sources separated by a wavelength on the blade, λ_b (Fig. 4). At a distant point at an angle θ, the radiation from these sources will reinforce and this will have a peak value if their path difference is a full wavelength in air.

Thus wavelength in air = $\lambda_a = \lambda_b \cos\theta$.

In order to determine the variation of θ for maximum radiation with λ_b, differentiate the above equation with respect to λ_b. Rearranging as $\theta = \cos^{-1} \lambda_a/\lambda_b$,

$$\left(\frac{d\theta}{d\lambda_b}\right)_{\lambda_a \text{ const}} = \frac{1}{\sqrt{\left(\frac{1}{\lambda_a}\right)^2 - \left(\frac{1}{\lambda_b}\right)^2}} \cdot \frac{1}{\lambda_b^2}.$$

A computed plot of $\left(\frac{d\theta}{d\lambda_b}\right)_{\lambda_a \text{ const}}$ with λ_b is shown in Fig. 5.

It will be seen that if $\lambda_a = \lambda_b$, $d\theta/d\lambda_b$ becomes large. Thus small changes in λ_b will produce large changes in θ. Therefore the ratio of wavelengths in air and in the blade is very critical in adjusting the directional properties of the radiator.

5. Excitation of the Beam

The blade was excited sinusoidally by the application of an alternating voltage of 25 V r.m.s. from a variable frequency oscillator to piezoelectric strain gauges (type, Brush-Clevite BSG/1-3) placed in the best position on either side of the blade, to produce flexural vibrations at the fixed end. This position was determined by

X	FN
1.0025975	13.829255
1.0050950	9.8435655
1.0075925	8.0386828
1.0100900	6.9516291
1.0125875	6.2046932
1.0150850	5.6503877
1.0175825	5.2176436
1.0200800	4.8674290
1.0225775	4.5762870
1.0250750	4.3291570
1.0275725	4.1158663
1.0300700	3.9292625
1.0325675	3.7641494
1.0350650	3.6166486
1.0375625	3.4837984
1.0400600	3.3632927
1.0425575	3.2533050
1.0450550	3.1523676
1.0475525	3.0592855
1.0500500	2.9730742
1.0525475	2.8929149
1.0550450	2.8181195
1.0575425	2.7481055
1.0600400	2.6823755
1.0625375	2.6205023
1.0650350	2.5621163
1.0675325	2.5068959
1.0700300	2.4545601
1.0725275	2.4048614
1.0750250	2.3575818
1.0775225	2.3125275
1.0800200	2.2695261
1.0825175	2.2284234
1.0850150	2.1890813
1.0875125	2.1513752
1.0900100	2.1151928
1.0925075	2.0804323
1.0950050	2.0470014

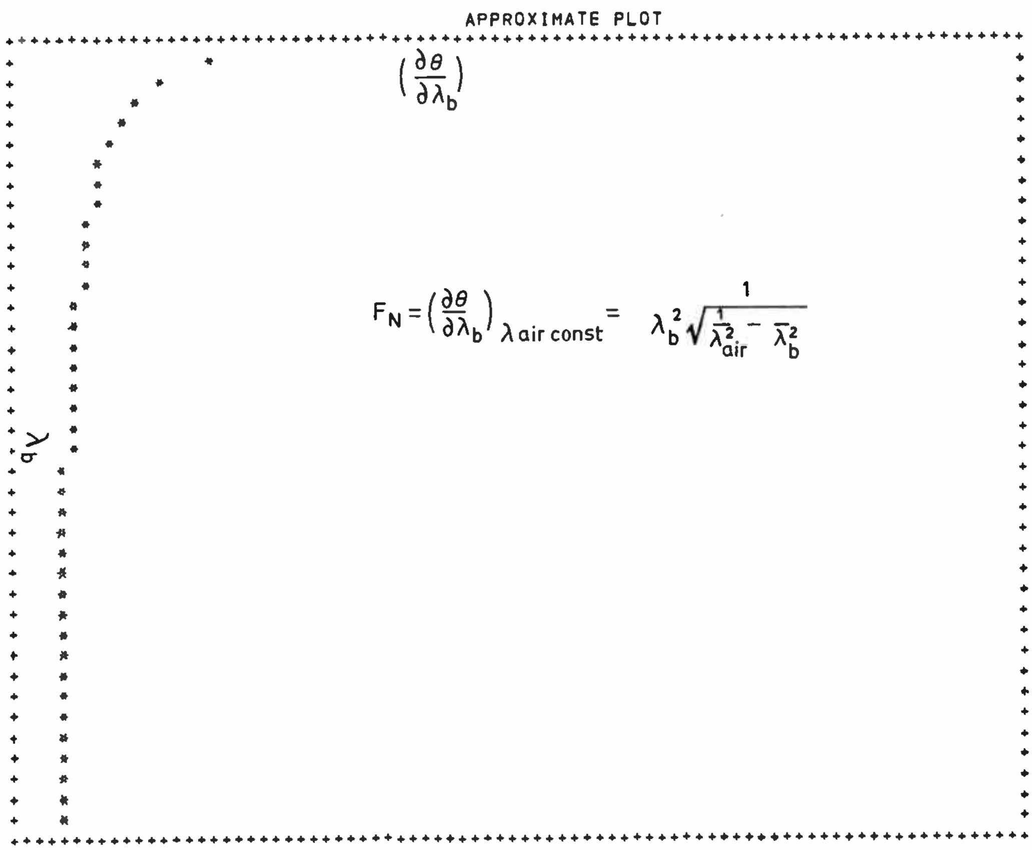


Fig. 5. Computed plot of variation of maximum direction of radiation with wavelength.

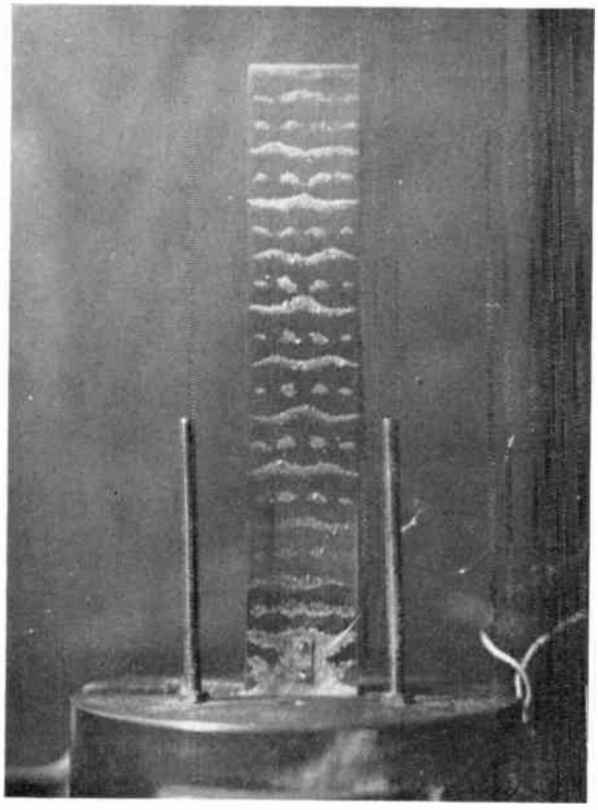


Fig. 6. Lycopodium powder showing pattern on blade.

observing the pattern of nodes with lycopodium powder. The nodes should be well-defined straight lines perpendicular to the centre line of the blade (Fig. 6). Due to the acoustic impedance mismatch between the blade and medium, standing waves develop on the blade as desired.

The radiating blade transducer was mounted on a rotary table so that it could be moved through 360°. The beam pattern (Fig. 7) was monitored by a half-inch Bruel and Kjaer condenser microphone with a frequency analyser.

6. Improvements in the Directionality

The physical configuration is antisymmetrical about the plane of the blade and so two acoustic beams are emitted, one from each side of the blade, in antiphase with one another. (See Appendix 1.) The second beam may be unwanted in practice, and two methods of eliminating it have been investigated.

First, if one side of the blade is shielded with an acoustic absorber one may expect an improvement in the in-line radiation emitted and a suppression of the radiation pattern on the shielded side of the transducer, so producing an asymmetric radiation pattern but with better directionality.

Secondly, it was possible that a standing wave of opposite phase would be produced as a result of reflexion from the holder of the blade. To cancel out this spurious reflected wave a reflector was introduced at a point shown in (Fig. 8a). With this reflector it was possible to tune the main beam for maximum acoustic pressure. Also the introduction of the reflector perpendicular to the blade generates a virtual image of the blade, so increasing the effective length of the transducer (Fig. 8b).

7. Q-factor Measurement

The mechanical Q-factor of the vibrating beam was obtained by using the usual expression

$$Q = \frac{\omega_0}{\omega_2 - \omega_1}$$

where ω_1 and ω_2 represent, respectively, the angular frequencies one on each side of ω_0 where a signal is $1/\sqrt{2}$ of the maximum. The Q-factor for the transducer was about 400. The Q-factor was rather high and indicates poor acoustic coupling of the transducer (See Appendix 2).

8. Electrical Impedance Measurement

The electrical input impedance of the transducer was measured by series resonance (impedance measured 250 Ω) as well as parallel. In the second case a 250 Ω resistance was connected across the parallel combinations to protect strain gauges and also for bandwidth considerations.

Because piezoelectric strain gauges are used, the transducer is capacitive with a value of 1370 pF. To balance out this capacitance, a coil of about 15 mH was connected in parallel. This reduced the input impedance to 250 Ω.

9. Measurements Relating to the Use of the Transducer in Detection Systems

The compact, inexpensive nature of this transducer would seem to recommend its possible use as a directional detection or sensing device; to this end measurements were made on the short and long range response, and an attempt made to normalize the measurements in such a way as to make them comparable with those for other types of transducers.

In these measurements the detecting microphone was of the Bruel and Kjaer half-inch condenser type with a

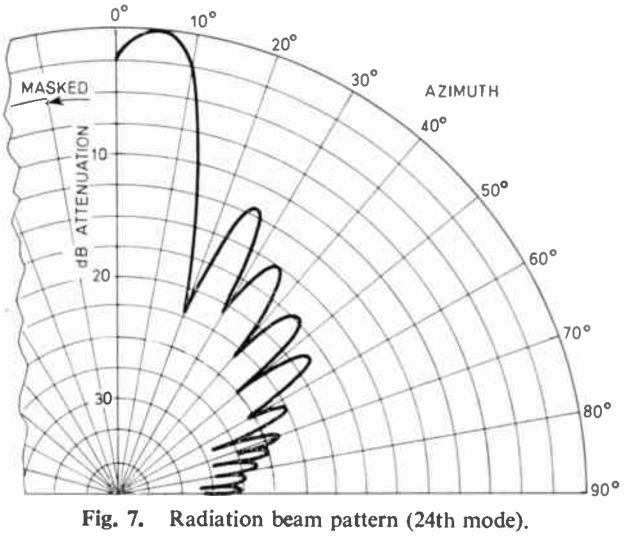


Fig. 7. Radiation beam pattern (24th mode).

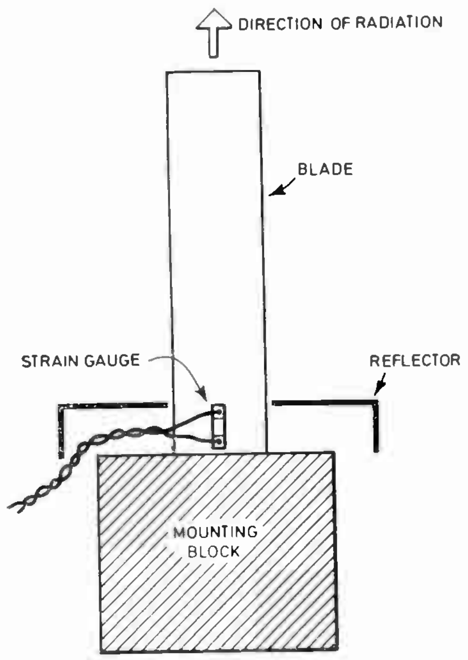
matched amplifier. The output voltage was calculated[3] approximately from the sound level figures obtained.

9.1 *Short Range Measurements*

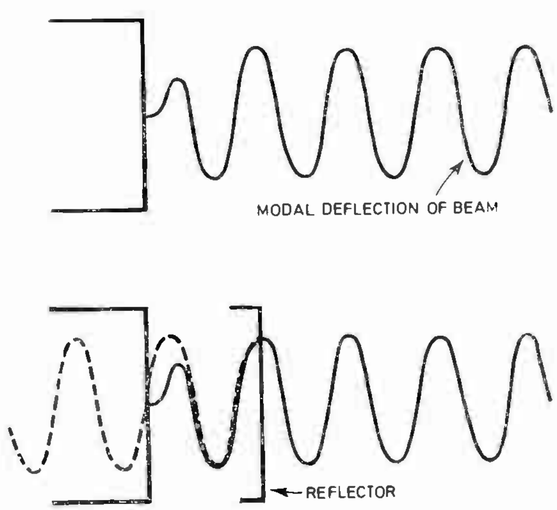
The transducer was driven at 25 V r.m.s. (10 mA) and produced a radiation pressure at the receiver placed 1 foot (0·305 metres) from the radiator as indicated below:

Sound level for 250 mW at 1 foot = 105 dB.

Reference pressure = 2×10^{-4} µbar.



(a) Arrangement of transducer with reflectors.



(b) Effect of reflector in generating virtual image of blade

Fig. 8.

Now, as 1 µbar corresponds to 74 dB, this corresponds to +31 dB relative 1 µbar. But the receiving sensitivity of the B & K microphone (from data sheet) is −60·4 dB relative to 1 volt per µbar.

Output for 1 µbar is given by

$$20 \log V = -60{\cdot}4 \text{ dB}$$

whence $V = 0{\cdot}955$ mV.

Pressure corresponding to 105 dB is given by

$$20 \log \frac{p}{2 \times 10^{-4}} = 105$$

i.e. $p = 35{\cdot}6$ µbar.

Microphone output for 105 dB is $35{\cdot}6 \times 0{\cdot}955$ or 33·95 mV.

9.2 *Long Range Measurements*

Measurements under long range conditions were made at night when the conditions were calm (still air). As the microphone was moved away from the blade transducer it was found that the signal strength dropped by 7 dB each time the distance between the transmitter and the microphone was doubled (see Table 1). This was partly due to the absorption of acoustic energy in the medium at ultrasonic frequencies and partly due to inverse square law effect. With 64 feet range, which is a factor of 2^6 greater than 1 foot, and 25 V r.m.s. drive at the transmitter, it may be possible to receive a voltage of about 63 dB ($= 20 \log p/2 \times 10^{-4}$). $p = 0{\cdot}282$ µbar. Hence $V = 0{\cdot}269$ µV.

Table 1

Distance (feet)	Pressure (dB)
1	105
2	98
4	91
8	84
16	77
32	70
64	63

Table 2

	End-fire standing wave transducer	Brush Clevite broadside coherent disk†
Operating frequency	39·461 kHz	40 kHz
Bandwidth	100 Hz	500 Hz
Beamwidth at 3 dB points	12°	60°
Mechanical *Q*	400	150
Electrical impedance	250 ohms (tuned)	650 ohms ± 350 ohms (+1800 pF) ± 400 pF
Input power	250 mW	1 W
Transmitting acoustic radiation	+31 dB relative to 1 microbar at 1 foot (0·305 m) with 250 mW input	+33 dB relative to 1 microbar at 1 foot (0·305 m) with 1 W input

† These figures were obtained from Brush Clevite.

9.3 Comparison with Conventional Transducer

Table 2 presents a comparison between the transducer, and a conventional commercial transducer with directional properties.

10. Conclusions

While the model is more directional than the one available for comparison operating at the same frequency, it still has a radiation emission pattern with undesirable side-lobes.

The radiator produces a low-power ultrasonic beam and cannot be used in its present form for high-power applications such as ultrasonic cleaning or mixing, but can be used as a sensing device or in burglar alarm systems.

It would seem that this transducer represents a useful, robust and potentially inexpensive solution to the problem of producing a highly directional ultrasonic radiation.

11. Acknowledgments

The author is indebted to Mr. H. V. Piling, Head of the Department of Science at Oxford Polytechnic, who helped with constructive criticisms and suggestions.

The permission of the Director of Oxford Polytechnic to publish this paper is gratefully acknowledged.

12. References

1. Morse, P. M., 'Vibration and Sound' (McGraw-Hill, New York, 1948).
2. Bishop, R. E. and Johnson, D. C. 'The Mechanics of Vibration' (Cambridge University Press, London, 1960).
3. Beranek, L., 'Acoustics' (McGraw-Hill, New York, 1954)..
4. Timoshenko, S. P. and Young, D. H., 'Elements of Strength of Materials' (D. Van Nostrand, New York, 1962).

13. Appendix 1: Directional properties of the radiator

From the expression for p

$$p = K \int_{-L/2}^{L/2} \sin \beta_b x \exp(-j\beta x \cos \theta)\, dx.$$

Substituting for

$$\sin \beta_b x = \frac{\exp(j\beta_b x) - \exp(-j\beta_b x)}{2j}$$

$$p = \frac{K}{2j} \int_{-L/2}^{L/2} [\exp(j\beta_b x) - \exp(-j\beta_b x)] \exp(-j\beta x \cos \theta)\, dx$$

$$= \frac{K}{2j} \left[\int_{-L/2}^{L/2} \exp[j(\beta_b - \beta \cos \theta)]\, dx - \right.$$

$$\left. - \int_{-L/2}^{L/2} \exp[-j(\beta_b + \beta \cos \theta)]\, dx \right]$$

$$= \frac{K}{j} \left[\frac{\sin(\beta \cos \theta + \beta_b)}{\beta \cos \theta + \beta_b} \cdot \frac{L}{2} - \frac{\sin \beta \cos \theta - \beta_b}{(\beta \cos \theta - \beta_b)} \cdot \frac{L}{2} \right]$$

Applying the wavelength coincidence condition

$$\beta_b = \beta \text{ and } \theta \text{ is very small.}$$

Then the right-hand term is large compared with the left-hand term.

As $\theta \to 0$ right-hand term $\to L/2j$

as $\theta \to 180°$ left-hand term $\to L/2j$.

Therefore two main beams are at $\theta = 0$ and $\theta = 180°$.

If $\theta = 90°$, $\cos \theta = 0$

$$p \propto \frac{2 (\sin \beta_b)\, L/2}{\beta_b}.$$

This is equal to zero if

$$\frac{\beta_b L}{2} = n\pi; \frac{\pi L}{\lambda} = n\lambda \text{ or } L = n\lambda.$$

14. Appendix 2: Choice of materials

If two beams of different materials are to radiate identically their surface motion must be the same at the same frequency. The strain energy stored is given by[4]

$$U_{max} = \frac{EI}{2} \int_0^l \left(\frac{\partial^2 V}{\partial x^2} \right)^2_{max} dx.$$

If the beams differ in material and thickness then

$$U_{max} = Ed^3 \times \text{constant}.$$

But the previous design criterion for the given frequency is that

$$d = \frac{\text{constant}}{c_m} = \text{constant} \times \sqrt{\frac{\rho}{E}}.$$

Thus

$$U_{max} = \rho^{\frac{3}{2}} E^{-\frac{1}{2}} \times \text{constant}.$$

Thus to reduce the Q value for a given configuration a material should be chosen which will provide as low a value as possible for U_{max}. Table 3 evaluates the physical properties of some common materials which might be considered for the beams.

Table 3
Physical properties of some beam materials

Material	Young's modulus $N/m^2 \times 10^{-10}$	Density g/cm^3	U_{max} arbitrary units	Design thickness for 40 kHz cm
Spruce (parallel to grain)	1·31	0·5	0·71	0·0306
Magnesium	4·15	1·7	1·09	0·0317
Glass	6·9	2·5	1·51	0·0299
Aluminium alloy	7·05	2·79	1·76	0·0312
Polystyrene	0·28	1·05	2·03	0·0960
Hard rubber	0·35	1·2	2·22	0·0918
Nylon	0·276	1·15	2·35	0·1012
Epoxy resin	0·314	1·22	2·41	0·0978
Steel	20·9	7·80	4·76	0·0303
Tungsten carbide	62·1	14·9	7·30	0·0243
Brass	10·06	8·53	7·86	0·0457

Manuscript first received by the Institution on 27th September 1971 and in revised form on 20th September 1972. (Paper No. 1485/IC 68).

U.D.C. 681.32:621.37/9.004

Analysis of Computer System Reliability and Maintainability

R. LONGBOTTOM,
M.B.C.S., C.Eng., M.I.E.R.E.*

SUMMARY

Equipment failures on computer systems give rise to more interruptions to service and excessive periods of repair time than a user would expect. This paper examines theoretical and actual reliability and repair statistics, showing how the shortcomings are influenced by test techniques, maintenance engineers, application of the system and spares availability. Various reliability and maintainability parameters are derived from established computer systems, further developed for modern equipment and then considered in relation to the latest computers.

* *Central Computer Agency, Civil Service Department, 207 Old Street, London EC1V 9PS.*

1. Introduction

Reliability specifications in terms of mean time between failures (m.t.b.f.) and mean time to repair (m.t.t.r.) are sometimes available for particular computer configurations prior to delivery. When the systems are installed and fully utilized it is often found that the m.t.b.f. and m.t.t.r., as seen by the computer user, bear no resemblence to the initial specifications. The purpose of this paper is to show the differences between true and apparent reliability and to give formulae to enable these to be predicted.

The Central Computer Agency's Technical Support Unit (TSU) provides, amongst its activities, a maintenance service responsible for advising computer users in Government departments on reliability and maintenance problems. Fault returns issued by the users' operators are submitted for every fault with repair details added by the engineers after fault clearance.

In order to evaluate the statistics, fault returns were selected from five identical computer systems, each consisting of a central processing unit, a line printer and card reader, and giving a total of more than 20 000 hours of system utilization. Each system had on-site maintenance engineers and was used on a single shift basis 5 days per week. One further useful point was that systems designated 2A and 3A were at one site and 4B and 5B at another site.

The predicted figures used in the exercise were produced via one of the TSU computer terminals using specially written programs.

2. Fault Distribution Theory

It is usually assumed that complex equipment, such as a computer system, will suffer from a constant failure rate $\lambda = (1/\text{m.t.b.f.})$. Failures are assumed to be random and, as such, follow an exponential distribution:[1-3]

$$f(t) = \lambda\, e^{-\lambda t}$$

where t is the duration of the measurement period. The probability of success (P_s) or failure free operation for this period is:

$$P_s = e^{-\lambda t}.$$

One of the requirements of the above distribution is that usage of equipment is constant over the period. At a complex computer installation the equipment utilization tends to vary from hour to hour but is fairly constant over long periods. In the limit, an unreliable piece of equipment may have only one period of high activity per month, distorting any resemblance to an exponential distribution. (The exponential distribution of one of the simple systems is discussed in Section 4.)

For general-purpose computer configurations, in most cases, a number of failures of the various devices may be expected every month; in this case the Poisson distribution is more applicable,[3] where the probability of 0 failures is again e^{-a} (where $a = \lambda t$) and the probability of 1, 2, 3,... etc. failures is:

$$a\, e^{-a}, \frac{a^2}{2!} e^{-a}, \frac{a^3}{3!} e^{-a} \ldots, \text{etc.}$$

The distribution has a mean or expected number of failures of a and standard deviation of $\sqrt{a}$.

3. Measured Fault Distribution of the Five Identical Systems

The fault reports, details of which are given in Table 1, were analysed for each item and for the complete configuration to show the number of months during which 0, 1, 2, 3 etc. fault reports were issued.

Table 1

Number of fault reports and periods of operation

System	Periods of use		Number of fault reports		
	Months	Serviceable hours	Processor	Printer	Card reader
1	39	5002	18	29	90
2A	39	5185	23	20	52
3A	27	3611	12	31	51
4B	34	4142	46	52	46
5B	33	3928	50	59	33

Predicted figures were calculated for each item according the the Poisson distribution. An example of the calculations is given in Table 2.

The calculations for the individual items were similar to those for the complete systems, as shown in Table 2, indicating that in some cases the Poisson predictions gave reasonable representations of the measured figures and others did not. It was also noted that, in every case, the standard deviations of the measured figures were greater than the predicted ones, indicating that the former has a larger dispersion than the Poisson distribution.

Examination of details of many of the fault reports showed that the fault was investigated and no fault was found or, alternatively, an adjustment was made or part changed, yet the fault re-occurred. Further investigation revealed a high proportion of these intermittent faults affected one user program only and not others using the same device, nor the engineering diagnostic tests run by the engineers in attempting to repair the faults.

4. True Faults

The fault reports were analysed to separate faults as first observed from repeated occurrences. To distinguish between the events, the initial observation of component or part failure is defined as a fault (also used in m.t.b.f.) and the total fault reports are defined as incidents. Thus:

incidents = faults (prime occurrence) + repeated occurrences when intermittent.

The distribution of months with 0, 1, 2, etc., faults was established and compared with Poissonian predictions, based on the new m.t.b.f. figures. (The number

Table 2

All equipment fault reports showing the number of months for which 0, 1, 2, 3—15 fault reports were issued

System	Incidents	M or P	Mean	Standard deviation	Number of fault incidents 0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	Fit
1	137	M	3·513	2·56	2	6	10	6	4	3	1	3	2	1	1						
	,,	P1	,,	1·87	1·2	4·1	7·2	8·4	7·4	5·2	3·0	1·5	0·7	0·3							LIKELY
	,,	P2	,,	2·53	3·2	5·7	6·7	6·4	5·3	4·0	2·8	1·9	1·2	0·7	0·4	0·3	0·1	0·1			LIKELY
2A	95	M	2·436	1·86	7	6	10	5	4	4	3										
	,,	P1	,,	1·56	3·4	8·3	10·1	8·2	5·0	2·4	1·0	0·4									LIKELY
	,,	P2	,,	2·30	8·4	8·1	7·0	5·3	3·7	2·5	1·6	1·0	0·6	0·4	0·2	0·1	0·1				LIKELY
3A	94	M	3·481	2·36	1	7	3	4	1	6	2	2	–	1							
	,,	P1	,,	1·87	0·8	2·9	5·0	5·8	5·1	3·5	2·1	1·0	0·4	0·2	0·1						POSSIBLY
	,,	P2	,,	2·63	2·6	4·1	4·6	4·2	3·5	2·7	1·9	1·3	0·8	0·5	0·3	0·2	0·1	0·1			LIKELY
4B	144	M	4·325	3·39	4	2	6	5	4	2	4	3	1	1	–	–	1	–	–	1	
	,,	P1	,,	2·06	0·5	2·1	4·4	6·2	6·6	5·6	3·9	2·4	1·3	0·6	0·4						UNLIKELY
	,,	P2	,,	2·89	2·0	3·8	4·9	5·1	4·7	3·9	3·1	2·3	1·6	1·1	0·7	0·4	0·3	0·2	0·1	0·1	LIKELY
5B	142	M	4·303	2·95	3	3	5	4	2	6	2	2	2	1	1	1					
	,,	P1	,,	2·07	0·4	1·9	4·1	5·9	6·4	5·5	3·9	2·4	1·3	0·6	0·4						UNLIKELY
	,,	P2	,,	2·73	1·4	3·3	4·6	5·1	4·9	4·1	3·2	2·3	1·6	1·0	0·6	0·4	0·2	0·1	0·1		LIKELY

M = measured. The predicted values (P1) are derived from the Poisson distribution. The predicted values (P2) are derived from the distribution given in Section 6. The last column gives an indication of 'goodness of fit' based on standard statistical methods.

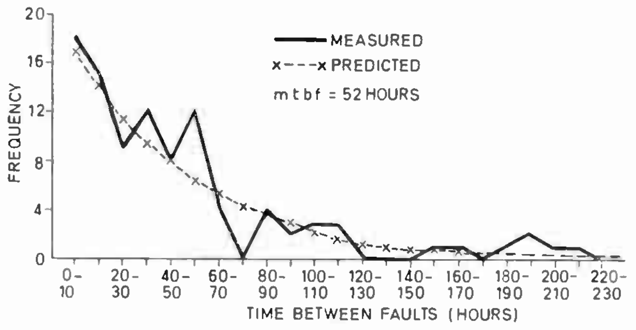


Fig. 1. Distribution of time between faults.

of faults is shown in Table 4). For 14 of the 15 items and for all 5 complete systems the predictions proved to be reasonable estimates.

For these simple systems each device was in use all of the time and the work load was fairly constant on a daily basis, so the distribution of time between faults was examined in relation to the exponential distribution.

Figure 1 shows the number of periods between faults, at 10-hour intervals, for the 97 faults at installation 1 over the 5000-hour period. Also shown is the exponential distribution based on the same fault rate. It seems reasonable to assume that the measured figures are exponential. However, it has been pointed out that this method may give rise to false assumptions as trends are not taken into account.[5]

The distribution of time between the 137 fault incidents at the same site was also examined and showed that there was a greater number of shorter periods between incidents than the predicted exponential figures. The effect of intermittent faults on the distribution of time between failures is explored in a comprehensive paper.[4]

5. Incidents per Fault

If a large number of fault incidents are examined, a proportion of them will constitute faults, as defined above, and the probability p that any randomly selected incident will be a fault is the ratio of total faults/total incidents. Furthermore, if all the faults are repaired p could also represent the probability that the incident will constitute a repair.

For any subsequent incidents, if it could be considered that p is constant, then the events could be described by the binomial distribution:

$$(q+p)^n$$

where the probability that y ($y = 0, 1, 2, \ldots, n$) of the incidents will constitute repairs is:

$$\binom{n}{y} q^{n-y} p^y.$$

For only one incident $(q+p)^1$ applies and if the repair of one fault is considered then:

p, pq and pq^2 are the probabilities that it will be repaired at the first, second or third attempts respectively, and q, q^2 and q^3 are the probabilities that it will *not* be repaired at the first, second or third attempts respectively.

Thus, the probabilities that the fault will give y ($y = 1, 2, 3, 4, \ldots$) incidents $P(y)$ are the respective terms of:

$$P(y) = p + pq + pq^2 + pq^3 + \ldots.$$

The mean of the distribution is $1/p$ (mean incidents/fault) and standard deviation $\sqrt{q}/p$.

The relationship between the actual and incidents/fault predicted by the above distribution were compared and although some of the samples were rather small, overall the predictions gave a reasonable representation. The worst case predictions were for installation 1, as shown in Table 3; in other cases the standard deviations of the predictions varied more uniformly about the measured figures.

At first sight, the suggestion that the probability that a fault will be repaired is constant could be refuted on

Table 3

Number of faults with 1, 2, 3—etc. incidents

Item	Total incidents	Total faults	M or P	p or 1/mean	Standard deviation	Number of incidents 1	2	3	4	5	6	7	Fit
1	18	12	M	0·667	1·17	9	2			1			
C.P.U.	,,	,,	P	,,	0·86	8·0	2·7	0·9	0·3	0·1			LIKELY
1 Card	90	67	M	0·744	0·88	56	3	6		2			
reader	,,	,,	P	,,	0·68	49·9	12·8	3·3	0·8	0·2	0·1		POSSIBLY
1 Printer	29	18	M	0·621	1·46	13	3	1				1	
	,,	,,	P	,,	0·99	11·2	4·2	1·6	0·6	0·2	0·1		LIKELY

M = measured values. The predicted figures (P) are based on a constant probability of repair p (See Section 5). The last column gives an indication of 'goodness of fit'.

the basis that as the engineer investigates the fault, the chance of repairing it must improve. However, as the faults tended not to be reproduced by the engineering test programs, attempts at repairing them tended to be by inspired guesswork. Hence, a major contribution to the probability of repairing a fault is the probability of reproducing it. The latter tends to minimize the difference in improving the chance of repair at successive attempts.

6. Fault Incident Distribution

In order to predict the number of incidents occurring in a period, it is desirable to know the distribution of time between repeated incidents after the first occurrence. Some of the times to repeat incidents appeared to be quite random; a large number, however, appeared to be repeated on an hourly, daily, or weekly basis, for example the next time a particular program was run.

The repeated incidents at site 1 were examined, giving a mean time to repeat of about 14 hours but did not appear to follow an exponential distribution. For all installations the majority of faults were repaired within a week of first being seen; in fact, only 5 of the 421 faults appeared to be repaired in a different month. Therefore, in considering the prediction of the number of incidents per month, the time between repeated incidents was ignored. In this case, the probability distribution of incidents can be shown to be:

No. of Incidents y	Probability $P(y)$
0	e^{-a}
1	$pa\,e^{-a}$
2	$pqa\,e^{-a}+\binom{1}{1}p^2\frac{a^2}{2!}e^{-a}$
3	$pq^2a\,e^{-a}+\binom{2}{1}p^2q\frac{a^2}{2!}e^{-a}+\binom{2}{2}p^3\frac{a^3}{3!}e^{-a}$
4	$pq^3a\,e^{-a}+\binom{3}{1}p^2q^2\frac{a^2}{2!}e^{-a}+\binom{3}{2}p^3q\frac{a^3}{3!}e^{-a}+\binom{3}{3}p^4\frac{a^4}{4!}e^{-a}$
5	$pq^4a\,e^{-a}+\binom{4}{1}p^2q^3\frac{a^2}{2!}e^{-a}+\binom{4}{2}p^3q^2\frac{a^3}{3!}e^{-a}+\binom{4}{3}p^4q\frac{a^4}{4!}e^{-a}+\binom{4}{4}p^5\frac{a^5}{5!}e^{-a}$
.	
.	
.	
etc.	

There may be a simplification or approximation to this expression but this was not explored in detail as the expression is readily adaptable for a computer program. The mean of the distribution is a/p, or the expected number of incidents (b), and standard deviation

$$\frac{\sqrt{a(1+q)}}{p} \quad \text{or} \quad \sqrt{\frac{b(1+q)}{p}}.$$

The probabilities of incidents per month was again calculated, using the above distribution and, in almost every case, gave a more accurate representation than the Poisson distribution. An example of this is shown in Table 2.

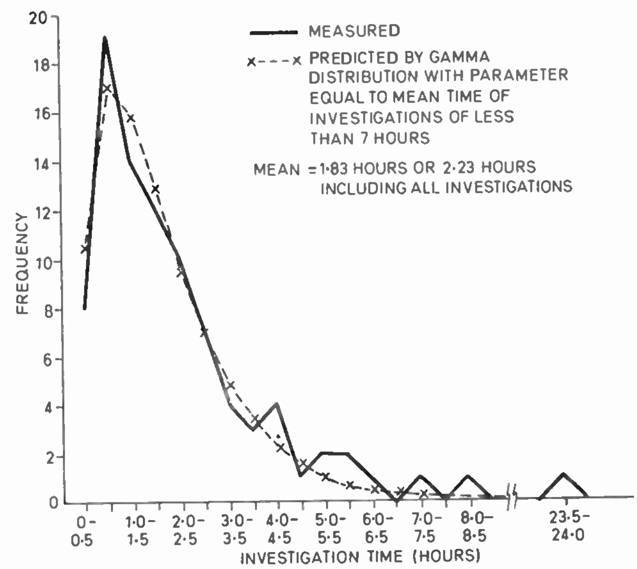


Fig. 2. Distribution of investigation time.

7. Repair-time Distribution

By definition, repair time consists of waiting time for engineering services, fault diagnosis, fault repair and proving time after repair.[7] The fault reports for the five installations gave repair times which fall within this definition, the waiting time being small as on-site engineers were available. For these simple systems, the maintenance engineers were encouraged to investigate all reported incidents but, usually, minor interruptions on the printer and card reader were not reported, unless they occurred frequently.

Figure 2 shows the distribution of investigation times (faults not necessarily repaired) at half-hourly intervals, for the 90 fault incidents on the card reader at installation 1, giving a mean investigation time of 2·23 hours.

In calculating repair times, various distributions, such as normal, exponential or gamma, have been suggested.[2,3] Using the mean and standard deviations for the actual results, none of these distributions appeared to fit. However, excluding long duration investigations, a gamma distribution with a parameter equal to the mean investigation of the remaining incidents would appear to apply (see Fig. 2). It is reasonable to exclude the long duration incidents as they were caused by exceptional events, such as no on-site spares, difficult faults which needed specialist assistance, or intermittent faults with serious consequences to user operations which required continuous investigation. Long duration investigations are considered later.

The gamma distribution has a probability density $f(t)$:

$$f(t) = \frac{e^{-t}\, t^{l-1}}{\Gamma(l)}$$

where the gamma function

$$\Gamma(l) = \int_0^\infty e^{-t}\, t^{l-1}\, dt$$

with a mean of l and standard deviation of $\sqrt{l}$.

Investigation times for all items were grouped similarly to Fig. 2 and compared with gamma distributions. To separate long duration investigations from others, a figure of 7 hours was taken as the lower limit of extended investigations for all items. In every case the gamma distribution was found to be a reasonable prediction.

Investigation times for all faults which were repaired at the first attempt were examined and also appeared to follow the same distribution.

The samples of long duration investigations were too small to enable distributions to be produced for individual items; however, examination of the figures showed that, besides a large number of investigations of between 7 and 8 hours, a peak was indicated in the 5–6 hour range. For any installation the distribution of exceptional faults could vary considerably depending on location of the site, hours worked, location of spares and engineering manpower resources etc.

The distribution of repair time per fault is dependent on the number of incidents per fault and can be shown to have a probability density of:

$$f(t) = p\,\frac{e^{-t}\, t^{l-1}}{\Gamma(l)} + pq\,\frac{e^{-t}\, t^{2l-1}}{\Gamma(2l)} + pq^2\,\frac{e^{-t}\, t^{3l-1}}{\Gamma(3l)} + pq^3\,\frac{e^{-t}\, t^{4l-1}}{\Gamma(4l)} + \ldots$$

with a mean of l/p and standard deviation of

$$\frac{l}{p}\sqrt{q + \frac{p}{l}}.$$

Predicted figures based on the above were compared with the actual results (excluding long duration investigations) and in every case gave a reasonable estimate.

8. Comparison of Results Obtained at Different Sites

8.1 *Observed Faults*

The expected fault rates were compared with the observed faults, over the complete periods, by means of confidence limits of the Poisson distribution.[1] The results are given in Table 4, which shows that there are significant differences between observed and expected

Table 4

Comparison of observed faults

Item and estimated m.t.b.f.	Installation	Expected number of faults	Expected 95% confidence limits	Observed faults	Remarks on difference in results
All C.P.U.s 277 hours	1	18·1	11–27	12	
	2A	18·7	12–28	8	Significant
	3A	13·0	7–21	9	
	4B	15·0	9–23	23	Borderline
	5B	14·2	8–22	27	Significant
All printers 155·1 hours	1	32·3	22–44	18	Significant
	2A	33·5	24–46	18	Significant
	3A	23·3	15–33	19	
	4B	26·7	18–38	41	Significant
	5B	25·3	17–35	45	Significant
All card readers 108·8 hours	1	45·9	34–60	67	Significant
	2A	47·6	36–62	34	Significant
	3A	33·1	23–45	35	
	4B	38·0	27–51	33	
	5B	36·0	26–48	32	

confidence limits and, therefore, the m.t.b.f. varies from installation to installation. The individual items at site A (systems 2 and 3) and at site B (systems 4 and 5), where the workload was shared equally between the two systems, were similarly compared; this showed that each of the items carrying out the same work could be considered to have equal fault rates.

The conclusion from the above is that reliability of computer equipment varies according to the application. This is due to the differing utilization and activity, which for the electro-mechanical input/output device is understandable; for the equipment being considered, the utilization was constant (all devices switched on and connected all the time) but the activity varied according to the amount of data handled by the programs at the different sites. For processors, application dependency is less understandable, but it can be dependent on heating effects and other factors; these are discussed further in Section 9.

8.2 *Normal Investigation Times*

Table 5 shows the 95% confidence limits of the mean investigation times (based on the t distribution), indicating that, as could be expected, different times were spent on investigation of faults on different devices. Processor investigation times appear fairly constant and overall confidence limits of approximately 1½ to 2 hours can be assumed. Differences on the other devices are, no doubt, due to the different site engineers. At site A, it was found that the two engineers tended to maintain and repair one particular system each and the difference can be seen in the Table.

As could be expected there were differences between the investigation times of the various engineers, but their contribution to overall performance was not as significant as that caused by the effect of the application on the m.t.b.f.

As shown later, at least for processors, the mean investigation times were rather high for all engineers. This was due to the inability of the test facilities to reproduce the faults.

8.3 *Long Duration Investigations*

When long duration investigations were necessary, in nearly every case each one applied to a different fault. A comparison of the results, similar to that shown in Table 4, gave an expected number of long duration investigations of 1–7 for processors and card readers and 0–4 for printers. Each of the recorded figures was within these limits.

The percentage of faults requiring these exceptional investigations was calculated as:

processors	8·9–26·5%	mean 16·5%
printers	3·5–10·5%	mean 6·5%
card readers	3·5– 9·5%	mean 6·0%

8.4 *Repair Times*

For processor faults, the actual time to carry out the repairs for the most frequently occurring faults, i.e. to change a circuit board, is only a matter of a few minutes and the time to run processor diagnostics is also of the same order. So often computer manufacturers claim an m.t.t.r. of the order of 20 minutes. For the systems being considered, taking into account the various normal and exceptional investigations, the mean time to repair faults was 4·8 hours for all systems, or about fifteen times what one may have expected.

8.5 *Probability that a Fault will be Repaired when Investigated* (p)

Table 6 shows the 95% confidence limits of p, derived from the binomial distribution $(q+p)^Y$, divided by Y, the total incidents.

Except for one of each of the input/output devices there is no significant difference between the installations, and the overall probability of repair could be considered to be between 0·45–0·67 for the processors and 0·68–0·80 for peripheral equipment.

The individual engineers did not appear to influence the results but on a collective basis they may all have changed the wrong components, maladjusted the wrong parts etc. However, when repairs were carried out, it

Table 5

95% confidence limits of mean investigation time

Installation	Processors		Printers		Card reader		All items
	Mean	95% confidence of mean	Mean	95% confidence of mean	Mean	95% confidence of mean	Mean
1	2·06	1·35–2·77	1·46	1·08–1·84	1·83	1·52–2·14	1·78
2A	2·0	1·39–2·61	1·38	0·60–2·16	1·80	1·39–2·21	1·76
3A	2·0	0·7–3·3	1·89	1·47–2·31	2·64	2·17–3·11	2·32
4B	1·89	1·49–2·29	1·0	0·74–1·26	1·50	1·04–1·96	1·43
5B	1·50	1·06–1·94	0·95	0·69–1·21	1·45	0·97–1·93	1·26
All installations	1·80	1·56–2·04	1·21	1·05–1·37	1·88	1·69–2·07	1·65

Table 6

95% confidence limits of p—the probability of repairing a fault

Installation	Processors		Printers		Card reader		All items
	Mean p	95% confidence limits	Mean p	95% confidence limits	Mean p	95% confidence limits	Mean p
1	0·667	0·44–0·89	0·621	0·45–0·79	0·744	0·64–0·83	0·708
2A	0·348	0·02–0·57	0·900	0·75–1·0	0·654	0·50–0·77	0·632
3A	0·750	0·50–1·0	0·613	0·45–0·78	0·686	0·57–0·80	0·670
4B	0·500	0·35–0·65	0·788	0·67–0·90	0·717	0·59–0·85	0·674
5B	0·540	0·40–0·68	0·763	0·64–0·86	0·970	0·91–1·0	0·732
All installations	0·53	0·45–0·67	0·738	0·68–0·80	0·739	0·69–0·80	0·688

was usual to re-run diagnostic programs to prove the repair. Hence, even for engineers' mistakes, the blame for the repeated incidents could be placed on the test features.

9. Modern Systems

The systems considered were chosen for analysis because of their simplicity, in order to show the various factors affecting engineering performance. Unfortunately, they were installed a few years ago. Of the modern systems it is very difficult to find two absolutely identical configurations. On the varying configurations it is quite usual for more equipment to be provided so that all devices are not in use all the time. Overall system m.t.b.f. figures are rather meaningless, since they not only prevent the comparison of complete systems but they also complicate the predictions of the number of faults occurring in a short period on one system. The varying utilization can make a known unreliable device appear to have an m.t.b.f. of thousands of hours, unless the actual utilization periods and activities are known.

However, the basic formulae given in the earlier sections have been checked against figures obtained from systems recently delivered and still been found to apply. The basic parameters for the formulae are discussed in the following sections.

9.1 *Mean Time between Faults*

The m.t.b.f. of equipment varies significantly from manufacturer to manufacturer and with the speed of operation, but the most frequently obtained figures for peripheral equipment are usually in the following ranges (with reasonable utilization):

Card readers, paper tape readers and printers	200–1000 h
Card and paper tape punches	100– 500 h
Magnetic tape units	500–2000 h
Exchangeable disks	1000–5000 h

Using modern technology, small processors having far greater throughput capacity than the systems mentioned in the earlier sections, achieve m.t.b.f.s in the range 2000 to 5000 hours or more. At the other end of the range, because of the extra speed, larger storage capacities and complex architectures, the most powerful computers currently being delivered give m.t.b.f.s of 200 hours or less.

For these modern systems, identical processors still appear to give m.t.b.f.s in the range 3 : 1, again influenced considerably by use-dependent faults. Use-dependency is not only noticeable in heat-sensitive areas, such as core storage, but also in other sections of processors, especially when associated with design deficiencies. This phenomenon of use-dependency has also been noted as arising when the equipment either is used continuously or is switched off overnight. The extra failures are here produced either by temperature cycling or by the effect of switching transients on the power supplies.

9.2 *Design Deficiencies*

All systems are subject to failures due to design margins being exceeded. Many of the problems are not resolved until thousands of hours of operation have been recorded on many systems; this is due to the time required to check all combinations of data patterns through all areas, in all sequences, in varying environments and, at the same time, to produce a statistically meaningful failure pattern. Thus, design changes to improve reliability are incorporated throughout a system's lifetime and sometimes there is a marked increase in m.t.b.f. after a single session of engineering change activity. Other engineering changes, giving similar effects, are found necessary due to component or manufacturing differences from the original system.

9.3 *Incidents per Fault*

The number of incidents per fault for modern equipment tends to be higher than the systems considered earlier and a factor of 10 is not unusual. However, due to the relatively easy restart facilities, many incidents are not investigated at the time, the users preferring to continue processing and engineers waiting until scheduled maintenance periods before investigating. For non-investigated incidents a certain amount of down time may be recorded for recovery/restart procedures and for the re-running of any programs, which were terminated by the incident; overall, the average time lost is usually in the range of 5–30 minutes. At some installations users insist that each incident is investigated at the time of occurrence, considerably reducing the incident rate but increasing investigation times. Thus, the computer users can have a significant effect on apparent reliability and a 3 : 1 difference in the number of recorded incidents per fault is not unusual.

To a user the time to failure may be expressed as mean-time-between-incidents (m.t.b.i.); so the net effect of the users', engineers' and test facilities' influence on the number of incidents per fault, and the effect of the application on the m.t.b.f., is to produce m.t.b.i. figures on identical systems over a wide range. An indication of this is given in Table 7.

Table 7

True and apparent reliability showing reliability figures for two identical modern processors, estimated over at least 5000 hours

System	Mean time between faults	Incidents per fault	Mean time between incidents	Mean down time per incident	Mean time to repair faults
	hours		hours	hours	hours
C	58	7·6	7·6	0·5	3·8
D	130	3·4	38·2	0·94	3·2

The effect of the wider range of m.t.b.i. on the number of incidents occurring in a month is shown in Table 8 (for same systems as shown in Table 7). The predicted figures were derived from the formula given earlier, based on values of p' of 0·13 and 0·29 respectively, where $p' = p''.p$ and p'' is the probability that an incident will be investigated or

$$p'' = \frac{\text{investigated incidents}}{\text{total incidents}}.$$

Also, as before,

$$p = \frac{\text{total faults}}{\text{investigated incidents}}.$$

As can be seen from Table 8, the predicted figures still give a reasonable estimation of the incident per month distribution.

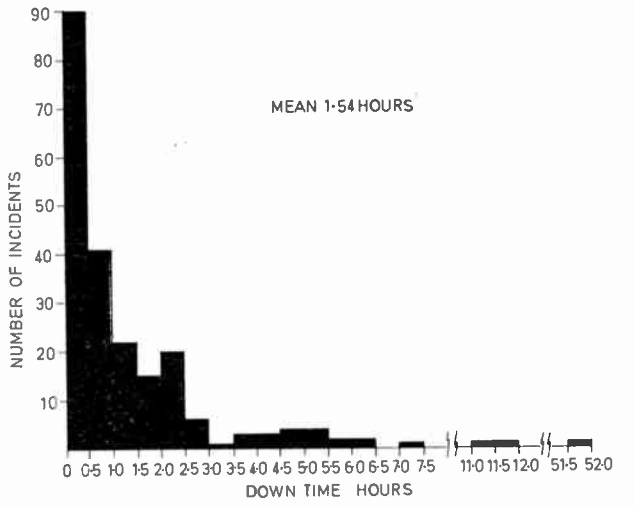


Fig. 3. Distribution of down-time on a modern processor.

Analysis of other fault returns from modern systems indicates that the ratio of investigated incidents per fault recorded during normal operational hours is usually in the range 1·5 to 3·0 for peripheral equipment and 2·0 to 5·0 for processors.

9.4 *Fault Investigation Times*

The mean fault investigation time for modern systems is similar to that considered earlier, being almost always between 1 and 2 hours, excluding waiting time where on-call maintenance services are provided. On some systems improvements have been made in maintenance aids, designed to reduce repair times, but this has been counteracted by increased complexity and lack of experience due to higher reliability.

Exceptional investigations are still necessary for 10–30% of processor faults and 5–20% of peripheral equipment faults, where, due to the lack of spares, equipment has been known to be out of service for a week or more.

Due to the introduction of shorter duration non-investigated incidents, the distribution of down time is different to that considered earlier; an example of this is shown in Fig. 3, which was derived from 217 incidents

Table 8

The number of months for which 0–4, 5–9 etc. incidents were recorded; predicted figures are calculated according to the formula given in Section 6

System	Measured or predicted	Mean	Number of incidents										
			0–4	5–9	10–14	15–19	20–24	25–29	30–34	35–39	40–44	45–49	50+
C	M	32·6	2	1	0	1	1	0	1	2	1	2	2
	P	32·6	0·61	0·89	1·15	1·31	1·35	1·30	1·19	1·05	0·90	0·74	2·51
D	M	4·2	19	4	3	1							
	P	4·2	17·34	5·97	2·39	0·86	0·30	0·09	0·03				

recorded on a modern processor. The recorded incidents have been checked against a distribution derived from gamma distributions for 126 non-investigated incidents at a mean of 0·5 hours, 80 investigated incidents at 2·0 hours and 10 exceptional investigations at 6 hours (excluding the one at 52 hours) and this assumption seemed to be valid. Excluding all exceptional investigations a gamma distribution with a parameter of 1·0 (same as exponential distribution with a repair rate of 1·0) also seemed to apply for the remaining 206 incidents.

10. Maintainability of Modern Systems

There is no doubt that higher reliabilities, in terms of m.t.b.f., are achieved on systems using modern technology. However, the problems associated with intermittent faults and exceptional long duration investigations still apply.

Regarding processors, analysis of the results of 100 acceptance trials, carried out by the Technical Support Unit,[8] shows that approximately 50% of faults were repaired at the first investigation (corresponding to $p = 0{\cdot}5$). With these systems installed in the field an average of 5 incidents per fault has been calculated from fault returns; so for 100 faults 500 incidents could be expected, or 450 failures would be the result of intermittent faults.

It has been claimed elsewhere that intermittents have comprised almost 90% of field failures on several known computer systems; also, the prevalence of intermittent failures of digital equipment is due to the relatively low efficiency of test techniques.[6] Test techniques, maintainability and availability features of systems currently being purchased are considered in the following sections.

10.1 *Test Programs*

Very few test programs can be classified as truly diagnostic, that is with the ability to diagnose faults down to component level, as the facilities to pinpoint failures are not usually provided in the basic system design. The worst form of test program, provided on many systems, will stop without giving details and leave the engineer to step through the program or interpret the state of the machine by means of lamps and switches. Somewhat better programs will indicate the area of failure and, with reference to logic diagrams, the fault can be diagnosed to a number of replaceable circuit boards; the fault is then rectified by board substitution.

In most cases test programs, as mentioned above, are adequate for diagnosing solid faults but, because of the interpretation required, test program re-running time and walking time to fetch spares, the mean time to repair normal solid faults is still quite high.

Regarding intermittent faults, the efficiency of test programs is questionable. An example of this is, again, available from experiences during acceptance trials, where engineering tests and customer programs are usually run alternately; only about 20% of intermittents observed during the running of customer programs were indicated in the engineering tests.[8] Some of these failures may have been reproduced if the test programs had been run for a longer time, for it was also calculated that, under trials conditions, the mean time to repeat intermittents was 20–30 minutes of running the program which reproduced the fault.

In an effort to overcome some of the difficulties associated with intermittent faults certain systems have pseudo-random programs available, which cover random instructions and random data in various sequences and run for several hours without repeating themselves. Others have complete system interaction tests to exercise the system with worst case interaction from all the various devices. Other facilities may be in the form of dynamic margins, that is, the ability to vary the supply voltages dynamically whilst running worst case programs.

10.2 *Availability Features*

Facilities are being introduced on many systems to increase availability or to enable the system to operate with either solid or intermittent faults present and to enable faults to be investigated without interrupting normal usage. However, in most cases, the facilities give rise to some degradation of throughput or a certain amount of unnoticeable down-time.

Perhaps the most common is the facility to reconfigure the system, either automatically or by operator intervention, then run the system without the faulty section of store or functional unit etc.; an 'off-line' test box may be available for investigation purposes. Next, 'on-line' test programs are often provided to assist in diagnosing intermittent faults on processors concurrently with user operations; these can also be used to ensure that any new solid faults are noticed. Others may be available for all scheduled and unscheduled maintenance purposes for peripheral equipment.

On earlier systems many processor intermittents caused the complete system to halt, but on modern systems features are often incorporated whereby only the programs directly affected are terminated, all other concurrent operations continuing unhindered. In some cases the failing program may be automatically restarted and/or the complete state of the system, at the time of the failure, may be recorded automatically for later analysis. Finally, facilities such as automatic 1-bit error correction on main storage are currently being provided. This enables programs to run through times of failure without being terminated.

10.3 *Results of Improved Maintainability and Availability Features*

No single system met so far has all the innovations mentioned above. However, all have been observed in operation on different systems during recent acceptance trials. The net results of all the features could be:

(i) A reduction of at least 50% of investigations during normal operations, the remainder being deferred until the scheduled maintenance periods, or repaired 'off-line'.

(ii) Mean investigation time for normal faults could be reduced to between ½ and 1 hour.

(iii) The number of non-investigated incidents may increase considerably due to automatic corrections or restarts and an average of at least 1·5 incidents per fault may still require operator intervention. However, the overall down-time should be reduced.

(iv) The number of exceptional long-duration investigations, when no spares are available or specialist assistance is required, may be reduced; however, although certain features may give improvements in this area, the added complications of automatic corrections or restarts and the possibility of further faults developing before the first is rectified, could lead to more severe problems.

11. Software

So far, only hardware failures have been considered but on many modern systems software problems give rise to more concern. For example, on a large configuration with a complex controlling system, besides restrictions in facilities and known deficiencies, the software may suffer from apparent random failures causing the whole system to stop. For these failures m.t.b.f. and m.t.t.r. figures are sometimes quoted by users, the repair time being the time to take dumps, software reloading and restart times, also program and file recovery times.

On new releases of the software, m.t.b.f.s as low as 2 hours may be observed and on some systems the best that has been achieved is in the range 10–20 hours. Software m.t.t.r. figures are usually in the range of 10–30 minutes but occasionally when say, the directory of disk-based files is lost, a few hours may be required to reconstitute the files from magnetic tapes. Hence, on many systems, software deficiencies give rise to more failures than the hardware and also contribute significantly to total down-time.

12. Conclusions

Reliability of computer systems, as seen by computer users, is extremely difficult to predict. However, models can be developed to give meaningful predictions, similar to those shown in this paper, if all the appropriate reliability, maintainability and utilization parameters are taken into account.

Regarding failure rates, factors other than the mean time between component failures must be used to represent the intermittent failure mode, including the effect of engineers, users, system concepts and, most of all maintainability features.

Because of problems with intermittent faults, times to repair must be based on the number of investigations per fault and user recovery times; exceptional fault times, due to availability of engineering expertise and spare parts, is also a significant factor.

On complex computer configurations, total system performance predictions may not be meaningful unless software reliability is taken into account.

13. Acknowledgment

Acknowledgment is made to the Department of Trade and Industry for permission to publish this paper.

14. References

1. Dummer, G. W. A. and Griffin, N., 'Electronic Equipment Reliability' (Pitman, London, 1960).
2. Barnett, V. D. and Ross, H. F., 'Statistical properties of computer acceptance trials', *J.R. Statist. Soc.*, A, **128**, pp. 361–94, 1965.
3. Bazovsky, I., 'Reliability Theory and Practice' (Prentice-Hall, London, 1961).
4. Lewis, P. A. W., 'A branching Poisson process model for the analysis of computer failure patterns', *J.R. Statist. Soc.*, B, 26, pp. 398–456, 1960.
5. Ascher, H. E., 'Evaluation of repairable system reliability using "bad-as-old" concept', *I.E.E.E. Trans. on Reliability*, **R-17**, pp. 103–10, 1968.
6. Ball, M. and Hardie, F., 'Effects and detection of intermittent failures on digital systems', A.F.I.P.S. Conf. Proc. No. 35, pp. 329–35, 1969.
7. 'Glossary of Terms Used in Automatic Data Processing', BS 3527: 1962. (British Standards Institution, London).
8. Longbottom, R. and Stoate, K. W., 'Acceptance trials of digital computer systems', *Post Office Elect. Engrs. J.*, **65**, pt. 2, pp. 91–7, 1972.

Manuscript first submitted to the Institution on 14th October 1970 in support of an application for exemption from the Graduateship Examination; received in revised form on 20th June 1972. (Paper No. 1486/Comp. 142).

U.L.C. 621.372.54: 621.374:621.376.9

Digital Encoding and Filtering Using Delta Modulation

GORDON B. LOCKHART, Ph.D.*

Based on a paper presented at the IERE Conference on Digital Processing of Signals in Communications held in Loughborough from 11th to 13th April 1972.

SUMMARY

Digital filters are described which are based on delta modulation rather than pulse code modulation. It is shown that delta modulation encoding may be combined with digital filtering to realize either non-recursive or recursive types. Conventional digital filter design techniques are applicable in both cases. Simulation of such filters on a small computer is discussed. Quantization noise is considered and comparisons made with the noise performance of the conventional digital filter.

* *Department of Electrical and Electronic Engineering University of Leeds, Leeds LS2 9JT; formerly at the Department of Electronic Engineering, University of Technology, Loughborough.*

1. Introduction

Conventionally, digital filters have been based on pulse code modulation (p.c.m.) techniques. Signal samples are encoded as groups of binary digits and linear difference equations implemented by digital arithmetic to effect filtering operations. If an analogue output is required the digital output from such a filter must be passed to a further stage of digital to analogue conversion. The hardware involved in these operations can be costly particularly in situations where both analogue to digital and digital to analogue conversion is necessary and 'time-sharing' techniques are inappropriate.

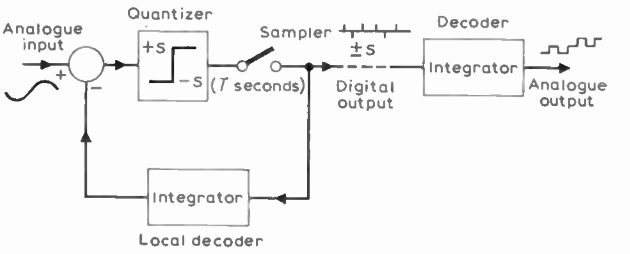


Fig. 1. Delta modulator and decoder.

Delta modulation[1] is an inexpensive means of analogue to digital conversion producing a binary sequence from which an approximate form of the analogue input can be recovered by simple integration (Fig. 1). It will be shown that the operations of delta modulation encoding and digital filtering may be efficiently combined to realise a simple digital filter which will be termed a 'delta modulation filter'.

Non-recursive and recursive types of delta modulation filter are realizable and conventional design techniques are applicable in both cases. The filters discussed here do not employ digital arithmetic for coefficient multiplication and therefore the primary source of noise is introduced by analogue to digital conversion. It will be shown that the noise performance of the delta modulation filter compares with that of the conventional filter.

2. Non-recursive Filtering

A non-recursive filter is formed by feeding the output of a delta modulator to a binary transversal filter (Fig. 2).

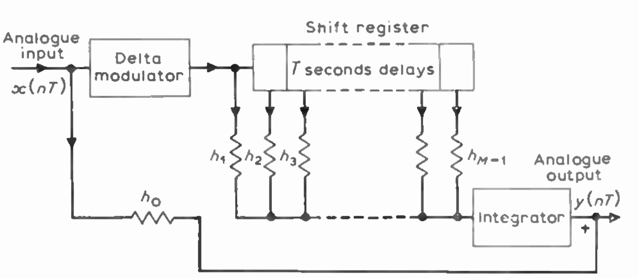


Fig. 2. Non-recursive delta modulation filter.

The latter consists of a shift register with tap outputs which are weighted according to the impulse response required. If the nth delta modulator input sample is

$x(nT)$, $(n = 0, \pm 1, \ldots)$ then, neglecting quantization noise and the possibility of overload, integrated output samples from the rth shift register tap are given by,

$$h_r x[(n-r)T]$$

where the constant h_r is the rth weight and T seconds is the delta modulator clock period. If M is the total number of taps then filter output samples are given by,

$$y(nT) = \sum_{r=0}^{M-1} h_r x[(n-r)T]$$

The pulse transfer function (p.t.f.) of the filter is of the form characteristic of non-recursive filters,[2] namely,

$$H(Z) = \sum_{r=0}^{M-1} h_r Z^{-r}$$

The frequency response of the filter is given by evaluating the p.t.f. on the Z-plane unit circle. Hence,

$$H(\exp(j\omega T)) = \sum_{r=0}^{M-1} h_r \exp(-rj\omega T) \quad \ldots\ldots(1)$$

Now M complex samples on the frequency response, $(H_k, k = 0, 1, \ldots M-1)$, are given at intervals of $(1/MT)$ Hz by taking the discrete Fourier transform (d.f.t.) of the set of weighting coefficients, i.e.

$$H_k = \sum_{r=0}^{M-1} h_r \exp(-j2\pi rk/M)$$

Conversely, if frequency samples are independently specified at intervals of $(1/MT)$ Hz on the frequency response, the corresponding set of weighting coefficients is given by the inverse transform,[2]

$$h_r = \frac{1}{M}\sum_{k=0}^{M-1} H_k \exp(j2\pi rk/M) \quad \ldots\ldots(2)$$

(Points on the frequency response between those specified will be trigonometrically interpolated according to eqn. (1).)

The d.f.t. relationships embodied in eqns. (1) and (2) are fundamental in the design of non-recursive filters from frequency domain specifications. If, for example, $M = 5$ as illustrated in Fig. 3 then the filter resolution is such that points on the frequency response (with exception of the point at $f = 0$ which is always real-valued), may be independently specified in magnitude and phase at the frequencies indicated extending from zero to the folding (Nyquist) frequency of $(1/2T)$ Hz.

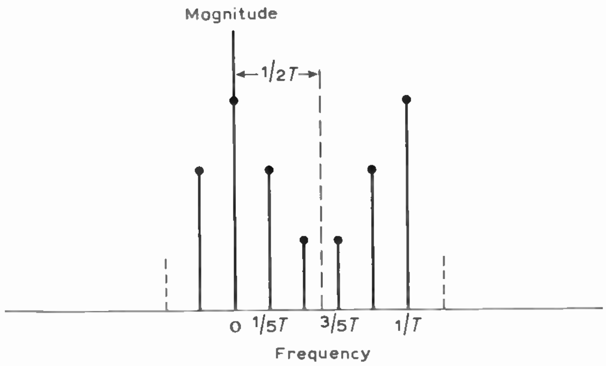


Fig. 3. Frequency samples, $M = 5$.

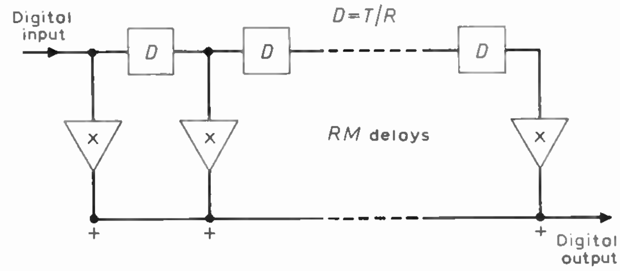


Fig. 4. 'Corresponding' conventional non-recursive filter.

In order to minimize slope overload distortion in the delta modulation encoding process[1] the bandwidth of the delta modulator input will normally be limited to some fraction R of this range. Frequency samples at frequencies greater than $(R/2T)$ Hz can therefore have no significance for input filtering purposes and may be usefully set to zero to effect a low-pass response with cut-off $(R/2T)$ Hz which will reduce delta modulation quantization noise above this frequency. The choice of the remaining sample values below $(R/2T)$ Hz is then made according to the frequency response required in the useable bandwidth.

A filter designed by this method may be compared with a conventional non-recursive filter having unit delays of (T/R) seconds and a total of RM delays (Fig. 4). Although the conventional filter will have a fold-over frequency of $(R/2T)$ Hz, its resolution and useable bandwidth will correspond since points on its frequency response can also be independently specified at intervals of $(1/MT)$ Hz up to $(R/2T)$ Hz. It follows that any conventional non-recursive filter can always be realized as a delta modulation type with a response passing through the same frequency sampling points in the useable bandwidth. Conventional design methods can therefore be applied directly to the delta modulation non-recursive filter or, if 'inter-sample' values deviate unacceptably, in conjunction with appropriate optimization techniques.[3]

The ratio, R, determines the level of quantization noise present in the output of the delta modulation filter and plays a role in design similar to the number of bits chosen to represent signal samples in p.c.m. designs. This is further discussed in Section 4.

3. Recursive Filtering

For recursive operation a feedback loop is introduced as illustrated in Fig. 5 which results in pulse transfer functions which incorporate poles in addition to zeros. Since the delta modulation step height is s volts the integrated output from a shift register tap will change by $\pm s$ every T seconds. If such a signal is weighted and fed back directly to the delta modulator input (or in combination with other weighted and integrated tap outputs) the delta modulator will tend to overload since there will be a high probability that the difference between T second samples of the combined delta modulator input (point Y in Fig. 5) will exceed s in magnitude. It follows that low-pass filtering of the feedback signal is desirable otherwise only small values of weighting coefficients can be used thereby restricting the variety of realizable filter characteristics.

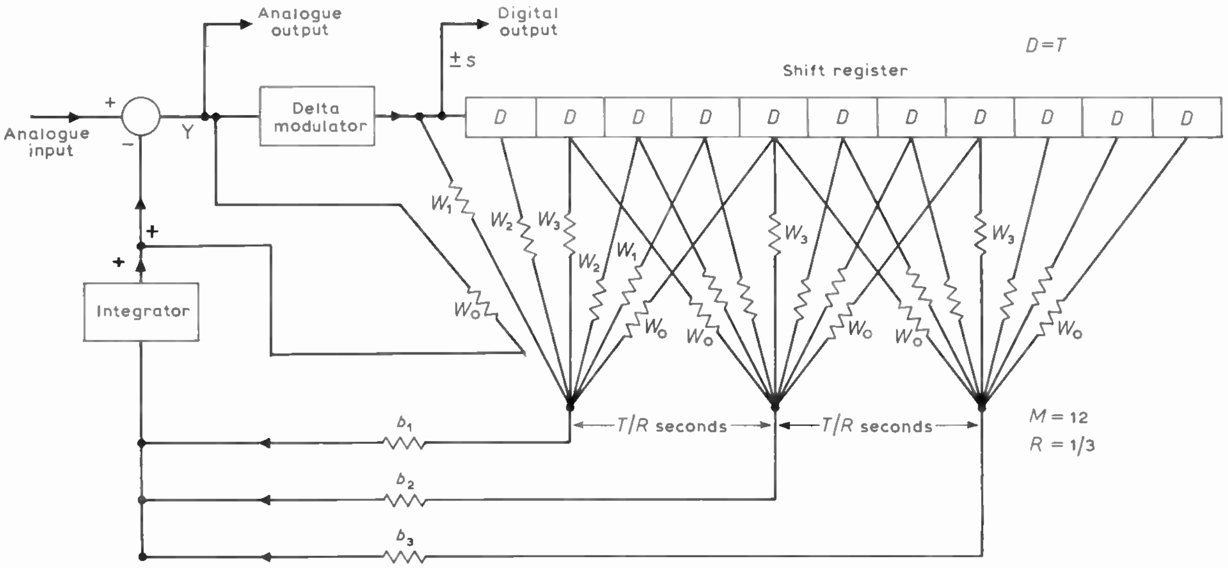


Fig. 5. Recursive delta modulation filter.

Ideally, a low-pass filter having unity gain up to the highest input frequency of $(R/2T)$ Hz and zero gain above this frequency should be inserted in each feedback path Also, no phase shift should be imposed at frequencies less than $(R/2T)$ Hz, otherwise the overall response of the recursive filter will be modified. A frequency response with zero phase shift implies a symmetrical 'non-causal' impulse response. (i.e. the response to an impulse commences before the impulse is applied). This can be accomplished if the low-pass filters are realized as $(RM-1)$ parallel non-recursive sections by combining sets of $(2/R+1)$ taps as illustrated in Fig. 5 for $M = 12$, $R = 1/3$. The low-pass filter weights $(W_0, W_1, \ldots, W_1, W_0)$ are chosen to give unity gain with zero phase shift at frequency samples below $(R/2T)$ Hz. Assuming that the low-pass filters are effective in reducing high-frequency quantization noise the design of the recursive filter can then proceed as if for a conventional recursive filter with respect to the $(RM-1)$ low-pass filtered tap outputs spaced at intervals of T/R seconds (Fig. 5). The weights $(b_1, b_2, \ldots,)$ are specified by the conventional design procedure, and if necessary can be combined with the low-pass filter weights to form a single set of M.

A recursive delta modulation filter designed by the above procedure will behave as a conventional recursive filter for input frequencies in the range below $(R/2T)$ Hz. If the frequency of the delta modulator input exceeds $(R/2T)$ Hz the feedback path will cut off and the behaviour of the circuit will revert to that of the simple delta modulator.

4. Noise Performance

Four sources of error are associated with digital filters.[2] These are:

(i) quantization in analogue to digital conversion,
(ii) overload distortion,
(iii) quantization of the weighting coefficients, and
(iv) quantization of the results of multiplication and summation.

The type of delta modulation filter discussed here performs multiplication and summation by analogue means and therefore error is introduced only by sources (i) and (ii) above. The level of the input signal, however, will normally be adjusted to minimize slope overload in the delta modulator and therefore, only analogue to digital conversion noise will be considered.

The essential operation of the non-recursive filter in Fig. 2 will be unaffected by removing the integrator from the filter output to the delta modulator output and passing the integrated delta modulator output to an ideal analogue delay line which replaces the shift register possessing identical delays and taps. Moreover, since the integrated output from the delta modulator may be regarded as the sum of the input signal and a noise component resulting from quantization[4] the delta modulator and output integrator may be replaced by a simple summation of input signal and noise leading to the equivalent circuit of Fig. 6. Noise at the output of the delta modulation filter is therefore equivalent to delta modulator noise filtered by the binary transversal filter.

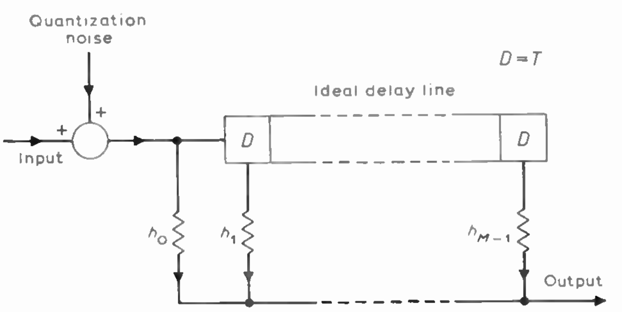


Fig. 6. Equivalent circuit of non-recursive delta modulation filter.

Alternatively, since the filter frequency response can be conceived as the response of the corresponding conventional filter imposed on a low-pass response with cut-off $(R/2T)$ Hz (Sect. 2), the equivalent noise output

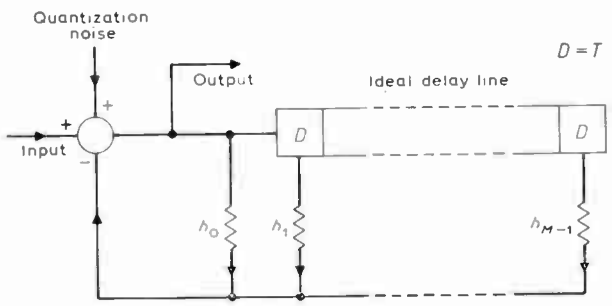


Fig. 7. Equivalent circuit of recursive delta modulation filter.

can also be obtained by passing delta modulator noise, low-pass filtered to $(R/2T)$ Hz, through the corresponding conventional filter. This implies that the anlogue-to-digital conversion noise at the outputs of the delta modulation filter and the corresponding conventional filter are proportional to the noise outputs of the p.c.m. and delta modulation encoders respectively where the latter output is low-pass filtered to the input signal bandwidth. The ratio of the noise outputs from the two filter types is therefore independent of the particular weighting coefficients employed.

In a conventional filter design the specification of an acceptable level of quantization noise for a given input signal bandwidth will determine the number of quantization levels required and hence the number of bits per sample value and the bit rate. For the delta modulation filter the same specification will determine the delta modulator clock rate (which is also the bit rate) and hence the value of the ratio R.

Noise performance in the case of the recursive delta modulation filter may be referred to the equivalent circuit of Fig. 7 which is derived by replacing the shift register in Fig. 5 by an ideal delay line and injecting delta modulator quantization noise at the input as discussed above for the non-recursive case. (Again it is assumed that the input signal level is adjusted so that overload is negligible.) Alternatively, because the input signal is band-limited to $(R/2T)$ Hz the output noise may be derived by injecting band-limited delta modulator noise at the input of the corresponding conventional filter. Since this applies equally in the non-recursive case the same conclusions follow regarding analogue to digital conversion noise and therefore, in general, the noise performance of any delta modulation filter can always be referred to the analogue-to-digital conversion noise model for the corresponding conventional filter.

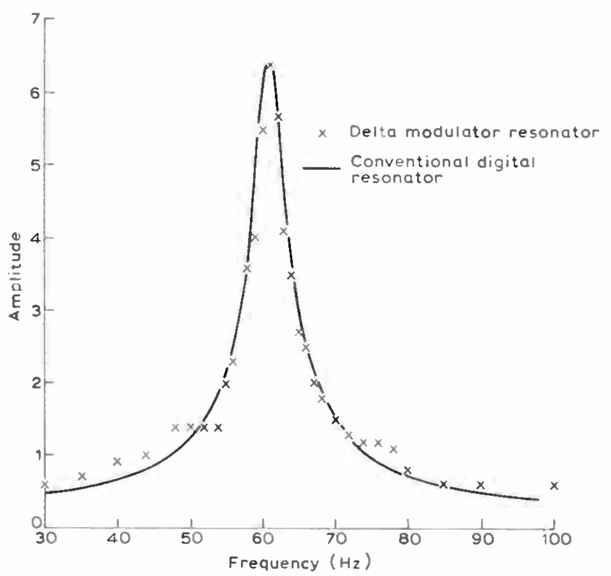


Fig. 8. Frequency response curves for a second-order delta modulator resonator and for a corresponding conventional digital resonator.

5. Simulation

It has been found by simulating a variety of delta modulation filters on a small computer that the approach to design outlined above is generally satisfactory. In particular, the simulation of a number of second-order resonators with $R = 1/16$ and $R = 1/8$ revealed no essential difference from the performance of 'corresponding' conventional filters. For example, the frequency response of a second-order delta modulator resonator is plotted in Fig. 8 with the theoretical response of the corresponding conventional filter. The structure of the conventional filter and associated pole positions are also illustrated in Fig. 9. Points on the

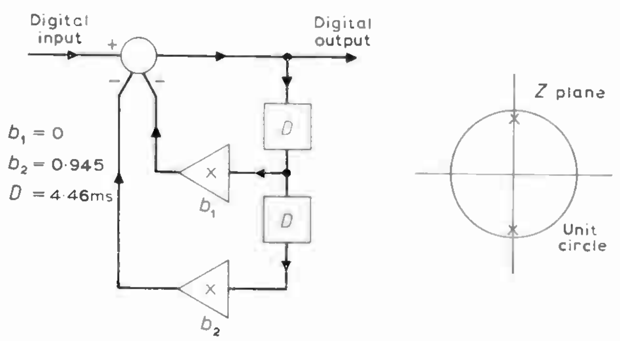


Fig. 9. Structure and pole positions of a digital resonator.

frequency response of the delta modulation filter ($R = 1/16$) were experimentally determined by feeding a sinusoidal input to the simulated filter and estimating output amplitude at a number of input frequencies. The two responses correspond satisfactorily at amplitudes above 1·5 but below this level there is some discrepancy due to quantization effects.

High-order pulse transfer functions are usually realized by placing elementary sections in parallel or cascade rather than utilizing the direct form as shown in Figs. 2 and 5 which is prone to high coefficient sensitivity.[2] Nevertheless, the direct form is suited to the type of delta modulation filter described since cascaded or parallel arrangements require additional delta modulators which increase the level of analogue to digital conversion noise. A variety of low-pass responses, mainly of the Butterworth type, have been successfully simulated in direct form. It has been found that the unwanted effects of tolerance errors on the weights can always be reduced by increasing the clock rate to permit a larger number of weights for realization of the same transfer function (i.e. smaller R).

6. Conclusion

In general delta modulation filters will comprise both non-recursive and recursive sections so that pulse transfer functions with both poles and zeros may be realized. If the delta modulator bit stream itself is to be transmitted over a digital channel then the poles can be realized by a recursive section at the transmitter and zeros by a non-recursive section at the receiver. In many applications, however, all-pole[5] or all-zero transfer functions will suffice. For example, the use of a recursive delta modulation filter in the remote collection of data (e.g. seismological data) appears to be a simple method of achieving a degree of frequency selectivity before transmission over a digital channel.

If the weighting coefficients of the delta modulation filter are quantized, then realizations become possible which eliminate resistor weighting in favour of digital multiplication.[6] Such filters have been implemented using m.o.s. recirculating shift registers operating in a serial processing mode.[7]

The delta modulation filters which have been described are particularly suited to applications where analogue-digital interfacing is involved and there is some advantage o be gained by further digital implementation. For example, delta modulation is already established as an inexpensive means of digitally encoding speech. Delta modulation filters can replace analogue types in such situations to permit spectrum shaping by digital techniques without recourse to p.c.m. encoding or digital arithmetic.

7. Acknowledgment

The writer would like to thank Professor J. W. R. Griffiths for his interest in the work described and for the use of computing facilities in the Department of Electronic and Electrical Engineering, University of Technology, Loughborough.

8. References

1. De Jager, F., 'A new method of p.c.m. transmission using the 1 unit code', *Philips Res. Repts.*, **7**, pp. 442–66, December 1952.
2. Bogner, R. E. and Constantinides, A., 'Introduction to Digital Filtering', (J. Wiley, New York. To be published.)
3. Rabiner, L. R., 'Techniques for designing finite-duration impulse response digital filters', *I.E.E.E. Trans. on Communication Technology*, COM-**19**, No. 2, pp. 188–95.
4. van de Weg, H., 'Quantising noise of a single integration delta modulation system with an N-digit code', *Philips Res. Repts.*, **8**, No. 5, pp. 367–85, October 1953.
5. Holmes, J. N., 'The advantages and disadvantages of the "matched" Z transform for digital filtering', Symposium on Digital Filtering, Imperial College of Science and Technology, London, July 1970.
6. Lockhart, G. B., 'Binary transversal filters with quantised coefficients', *Electronics Letters*, **7**, No. 11, pp. 305–7, 3rd June 1971.
7. Lockhart, G. B. and Babary, S. P. 'Binary transversal filters using recirculating shift registers', *The Radio and Electronic Engineer* **43**, 1973. To be published.

Manuscript first received by the Institution on 22nd December 1971 and in final form on 21st September 1972 (Paper No. 1487/CC153).

STANDARD FREQUENCY TRANSMISSIONS—November 1972

(Communication from the National Physical Laboratory)

November 1972	Deviation from nominal frequency in parts in 10^{10} (24-hour mean centred on 0300 UT)			Relative phase readings in microseconds N.P.L.—Station (Readings at 1500 UT)	
	GBR 16 kHz	MSF 60 kHz	Droitwich 200 kHz	GBR 16 kHz	†MSF 60 kHz
1	+0·1	−0·1	0	636	616·9
2	−0·1	0	0	637	616·6
3	−0·1	−0·1	0	638	617·2
4	+0·2	0	0	636	617·1
5	−0·1	0	0	637	617·3
6	0	0	0	637	617·1
7	+0·3	0	0	634	616·9
8	−0·3	0	0	637	616·7
9	0	−0·1	0	637	617·6
10	0	−0·1	−0·1	637	618·1
11	−0·2	−0·1	−0·1	639	619·4
12	−0·1	−0·1	−0·1	640	620·7
13	−0·3	−0·2	−0·2	674	622·4
14	+0·1	−0·1	−0·1	673	624·1
15	−0.2	−0·2	−0·1	675	625·0
16	+0·1	−0·1	−0·1	674	626·0
17	−0·3	−0·1	−0.1	677	626·6
18	−0·1	−0·1	−0·1	678	628·0
19	−0·2	−0·2	−0·2	680	630·1
20	0	−0·1	−0·3	680	631·3
21	−0·3	−0·1	−0·2	683	632·2
22	−0·1	−0·1	+0·1	684	632·2
23	+0·1	0	+0·1	683	632·4
24	−0·2	0	+0·2	685	632·4
25	0	−0·1	+0·2	685	633·4
26	−0·1	−0·1	+0·2	686	634·1
27	−0·1	−0·1	+0·2	687	635·2
28	0	−0·1	+0·2	687	636·1
29	−0·2	−0·1	+0·2	689	637·3
30	−0·1	−0·2	+0·1	690	639·4

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

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

Conference on Computers — Systems and Technology

Middlesex Hospital Medical School, London
24th to 27th October 1972

The Conference was organized by the IERE with the association of the Electronics Division of the Institution of Electrical Engineers and the British Computer Society.

This article is based on reports prepared by Dr. K. J. Dean, chairman of the Organizing Committee and other members of the Committee, namely Mr. K. D. F. Chisholm, Professor R. L. Grimsdale, Mr. D. M. Maclean, Mr. E. A. Newman, Dr. J. M. Pinkerton, Mr. A. L. Rowles and Mr. B. S. Walker. The full list of papers read at the Conference was published in the September 1972 issue of *The Radio and Electronic Engineer* (page S.141); the papers are available as IERE Conference Proceedings No. 25, which may be purchased from the Institution, price £8·00 post free.

In opening the Conference Sir Robert Cockburn, Chairman of the National Computing Centre, said that one third of the UK gross national product could be accounted for in the cost of data manipulation of one sort or another. Sir Robert, drawing on his own experience of ionizing radiations, spoke of the 'half life' of a growing industry—not of an exponential decay but of an exponential growth—not of the time taken for some activity to fall to one half, but to double, a 'twice life' rather than a 'half life'. The 'twice life' of computing, of world computing capacity, was two years, he said. For so rapid a rate of change we must be very alert, for there is so little time to put right what goes wrong! (Sir Robert's address appears in full at the end of this report.) The Conference, at which there were about 350 delegates, showed that they at least were alive to the problem. At the end of two very lean years for the computer industry there were more people at this conference than at either of the last two computer conferences in this series. A straw in the wind perhaps, but an encouraging one for all that.

The first session dealt with hardware, and in the first paper of the conference Mr. B. S. Walker of Reading University drew the attention of delegates to the very long way we had come down a particular path of hardware design: until now technology was such that minimization of gates is a waste of time and circuit design is for most people very largely a thing of the past. Whether his node, which he was describing, was the right path to take time would tell, but at least the challenge to think fundamentally was appreciated by his audience.

He was followed by Miss H. J. Kahn (Manchester University) who charmingly described the hardware simulator for the *MU5* of which we heard so much in the final conference session. She also defended herself from the question 'Which design takes longer, that of the simulator program or the logic which is being simulated?' Not only in this session, but throughout the conference the viewpoints of software and hardware supporters were shown more in contrast than in conflict; exponents of each seeking to explain and learn from the other.

Later that day hardware came into its own again with the papers from I. S. Williams and G. Lee of GEC Computers on circuit board design and testing. These two papers really opened up the questioning, both formally in the hall and over coffee and lunch; a man-to-man interaction which really gave value to the conference.

There was a good review of new hardware too, such as the bubble stores described by Dr. R. M. Pickard of Manchester University. In contrast there was still plenty of scope, it seemed, for improvement in established products, illustrated by Mr. J. A. Purdie of Watford College of Technology, describing keyboard layout and key switch techniques and various coding systems.

The peripherals session contained things both old and new, with Dr. H. C. A. Hankins' paper on cathodochromic imaging in visual displays and the paper by Mr. F. C. Evans of St. Andrews University on low-cost optical character recognition.

The chairman of the session on architecture, Mr. A. Rowles, had the temerity to define 'architecture' at the start of the session, when the papers had not only been written but were about to be given! However his definition of structure rather than performance or implementation met the common objectives of the speakers, for he was not gainsayed. There were lighter moments too, such as when Dr. A. L. Freedman of Plessey Radar cited an example in which a diagnostic program correctly diagnosed which of a dual processor system was faulty, inadvertently switching out the good one. This gave rise to a comment, believed to be from a Post Office delegate, that we might do better to make exchange equipment less reliable to give maintenance staff more experience at locating and removing faults.

It was in this paper too that the subject of new technologies came up once more. Dr. Freedman put in a plea for taking greater advantage of current technology, where we have experience, before plunging headlong down shining new avenues where we have none. Against this the fact that 1973 sees the 25th anniversary of the transistor should act as a reminder that not all those who plunge down new avenues are Gadarene swine. Discernment is needed to distinguish the best of the old, the most promising avenues of research, and of course the ability to grasp that engineering nettle, to make it NOW, and not to hang on for new technology and new developments that might be so much better a little later on. This applies to computer architecture, complete systems or whatever.

The second session on architecture dealt with the *System 250* developed by the Plessey Company. It was a credit to the authors that they fitted in so well what were really five presentations into the time allowed for three. It was, in fact, a feature of the architecture sessions that to a great extent a common thread could be detected in a number of the proposals for and descriptions of systems, tending to point a fairly clear path to the way that computer design is moving.

In the session on man-machine interaction we heard from both users and manufacturers: a conference like this is always a good forum for that sort of debate. But it was Dr. Freedman who made the point that specifications should not restrict the designer's choice of implementation. The more advanced the technology the greater the waste of resources otherwise. There were some present though who felt that not everyone could be expected to affect this type of decision. He also asked why man-machine interaction broke down in the wider sense, pointing to the failure so common, to follow a train of thought leading eventually to product innovation, in a society which is grounded in science- and engineering-based research.

In the operations management session 'The measurement of information processing power' presented by G. G. Scarrott (ICL), developed ideas which lead to a system of units for measuring processing power, derived from considerations of logical processes of human beings. Conclusions reached in the paper demonstrate the plausibility of the ideas on which the theory is based. The paper produced a lively discussion which, if at times lighthearted, certainly showed the general interest in the subject.

A paper presented by J. Garrett described measurement techniques, both hardware and software, developed by the Post Office Research Dept. for monitoring the performance of an experimental stored program control telephone system. Although some of the techniques were applicable only to the field of telephony, most could be applied to the wider field of process control.

With each new generation, computer installations grow in size, complexity and price. Development of modular simulation models such as the one described by M. Berjak, of Sussex University, is essential to provide an optimum system. A simulation model had been designed to investigate the operation of a computer network consisting of a number of processors. The model catered for a variety of hardware configurations, and operating modes, and in particular allowed for processing jobs originating at one processor being carried out at any other processor.

Professor Grimsdale, who chaired the session on applications, observed that no conference could be complete without reference to applications. Surely all those present were fascinated by what must have seemed to everyone a new computing application—that of cinema control, described by Mr. L. Forth of Essex University. If one had not read his paper, his first illustration would have led one to believe that he was about to talk on 'point of sale' systems with computer-controlled box offices, but in the event it was the house lights, the sound, the vision, in fact everything about the presentation except, as one questioner pointed out, the artistic merit of the art of cinema, which perhaps needed it most, which was under control.

Mr. K. H. C. Phillips of the Post Office outlined the use of computers in mechanized letter sorting and the routing of mail, together with pictures of the hardware. This session and the one following on software continued the dialogue of hardware with software. The organizing committee has consciously tried to throw software and hardware proponents together in the interests of better understanding. This was emphasized by the comment that the letter sorting equipment was not yet in service: the hardware was working but the equipment which should have been in operation last July still needed the software to be debugged.

The final session included five papers, all on the Manchester University *MU5* research machine. Introducing it Professor D. B. G. Edwards said it was denoted *MU5* reckoning *Atlas* as the fourth machine to have been designed and built at the University under the general guidance of Professor Kilburn. The aim was a machine with 20 times the power of *Atlas* at substantially the same cost. This aim was expected to be achievable (Kinninment & Edwards) by using a technology about eight times faster and an even more efficient internal organization and system architecture. Content addressable buffer memories between main store and registers were part of the answer as was the pipe line arithmetic unit organization (Ibbett, Phillips & Edwards). Considerable overlaps were possible in processing instructions as a result.

Whereas the *Atlas* was the first machine to introduce over ten years ago the virtual store concept, the *MU5* greatly extended the principle to offer users substantially independent virtual machines. These interacted only in a controlled manner by messages in a special format (Morris, Frank & Sweeney) so that misdemeanour of one program could have little or no deleterious effect on another. The secondary operand unit (Standeven, Lanyado & Edwards) was designed to help further in this respect by allowing data in store to be accessed indirectly in several different ways depending on its nature without recourse to software routines for generating the addresses. The so-called descriptors allowed the virtual addressing system of the *MU5* hardware to be fully exploited. Confidential *code* was thus expected to be substantially more efficient than in *Atlas* and all compilers were intended to 'target' onto a common intermediate level language (Capon & Wilson). Lively discussions on all the papers underscored the interest provoked amongst those present.

Throughout the conference the papers were of an exceptionally high standard: in fact a substantial conference could have been put together from papers which had to be left out or came in too late. As it was, the conference had to be extended by half a day in the early planning stage. The fact that the three institutions, IERE, IEE and BCS, were involved did a great deal to set the seal of success on this conference by bringing together delegates of a wide variety of backgrounds. The range of discussion at the conference can well be illustrated by reference to comments by two of the early speakers: Sir Robert Cockburn and Mr. R. W. Mitchell of International Computers Ltd. Mr. Mitchell reminded the delegates that computer systems are for people and not people for computers. One member commented that human values were of prime importance and that people should not be left with moronic tasks from which all challenge had been removed—whether they were employed by the largest of manufacturers or the humblest of users. If they were, then as Sir Robert said, 'The real danger is not that computers will become more like us, but that we will become more like computers.' The eye and brain that can appreciate all that is best in God's world have a value beyond computing systems and a skill beyond that which our technology can match.

Conference Opening Address

SIR ROBERT COCKBURN, K.B.E., C.B., Ph.D., C.Eng., F.R.Ae.S., F.I.E.E.*

Although I have been somewhat involved in the computer field for the last few years, I am certainly not a professional; and before such an audience it would be foolish to attempt to open this Conference on Computer Systems and Technology with a 'keynote' address. The range and authority of the papers must be intimidating even to you experts. Instead I will comment on some of the broader issues, particularly those which arise on the interface between the computer and the community, as seen by a 'lay' observer.

One of the most striking characteristics of the computer age is the speed at which it has engulfed us. In evidence to Sub-

* Chairman, National Computing Centre, Manchester.

Committee D of the Select Committee on Science and Technology of the House of Commons, the National Computing Centre estimated that for Europe as a whole the annual growth rate in computer investment was of the order of 20%. For this country it estimated the growth rate in the number of installations to be about 40% per annum and in number of users 30% per annum. It forecast that within a decade, data processing by computer could become the world's third largest industry, surpassed only by the petroleum industry and the motor-car industry.

Without in any way underestimating the outstanding developments in hardware and software technology which sustain this growth, the driving force behind it is the need for greater efficiency in handling the mass of information essential to modern society. All stable communities have depended on reliable recording and distribution of data; and the larger and more sophisticated they become the more effort is absorbed in sheer housekeeping. The Assyrians used clay tablets, the Egyptians papyrus, the Incas the quipu; and our technology had not advanced much further until the computer arrived.

In this country we spend about £10,600M per annum, that is nearly a third of our gross national product, in generating, transmitting and manipulating data of various kinds. So there has existed a large and readily accessible market for the more efficient processes which the computer can provide. Moreover, penetration of this market has proceeded simultaneously at many levels, in administration, in commerce and industry, and in the High Street; in design, production, teaching and training and scientific research. As a consequence growth in the field of computers far exceeds anything we have previously experienced; and it could give rise to problems of adjustment which if not anticipated could create intolerable strains in our economy, and indeed for society as a whole.

I have had some experience in other areas of technological growth—in radio communications, radar, atomic energy, and in various aspects of aviation; and provided appropriate normalizing factors are applied useful comparisons can be drawn between them. The history of aviation extends over about 60 years and throughout there has been a succession of major advances in design, from the biplane to the monoplane, from piston engine to jet, and from the conventional aerofoil to the slender wing. Technical progress can be quantified by various parameters, and two of the most important in aviation are speed and all-up-weight. Until quite recently these have double consistently every ten years. Demand in terms of passenger miles or ton-miles has doubled every five years.

The electronic computer as a practicable device has a history of about 30 years and during this time there have been major advances in technique such as from valves to transistors to integrated circuits; and from cards to tape, to drums and to exchangeable disks. Speed and storage capacity, although used in a quite different context, are as significant for computers as they are for aircraft, and both have been increasing ten-fold every 5 years, that is doubling every 18 months. World demand in terms of computers installed is doubling every two years.

But exponential growth can never continue indefinitely. Restraints imposed by the environment always lead to a slowing-down and to eventual stabilization; and the more rapid the growth, that is the shorter the half-life in the initial stages, the more abrupt in real-time is the levelling-off.

Those of you responsible for planning new developments and new investment for the future must be constantly concerned with this question: Is some recession such as that recently experienced in the computer field an incidental fluctuation, or is it the first sign that we are moving from a sellers' market in which growth is maintained by developments, to a buyers' market in which growth is limited by the willingness or ability of the customer to absorb further changes.

The aviation industry reached this stage during the last decade. Stabilization of the civil and military markets led to a slowing-down in technical innovation. Whereas in the post-war years the design of a new aircraft took about 5 years and might become obsolescent in 10 years, now as in the case of the *Concorde* it can take up to 20 years to satisfy the more rigorous demands of the operators, and once adopted the design is unlikely to be replaced for 25 years.

I am not suggesting that growth in the sales of computers and computer packages will not continue in those areas where need has been clearly established. I doubt whether an economy as large and complex as that of the European Community, for instance, could be viable without the fullest use of computerized data processing. But I do suggest that it is becoming more difficult to identify new applications, and taking longer to satisfy them. This is partly because technology, and particularly software technology, is becoming more sophisticated but mainly because of a reluctance to replace the huge investment in traditional techniques and patterns of behaviour.

In the engineering industry for example, the computer has improved management and housekeeping as it has in other sections of the economy; and in the more advanced fields it is revolutionizing design. Conventional lofting of an aircraft wing for instance, which at one time took 10 men 3½ months, can now be done in 3 weeks using a computer and a draughting machine.

Graphic techniques have been developed at Cambridge and elsewhere for composing complex shapes, for operating on them interactively, and for extracting the information needed for their construction. These are making an increasing contribution to three-dimensional design such as arises in aircraft, car bodies, ship building, and in highway construction.

But in mechanical engineering as a whole the computer has had only a limited impact on the design and production of the thousands of relatively simple piece-parts on which the bulk of the industry depends. The freedom from sectional drawings which the computer offers is not matched by the production methods available on the shop floor where for a hundred years machinery has relied on orthogonal or axially symmetric cutting. We may have to wait for entirely new methods such as chemical milling, explosive forming, or the weaving of fibres before the potentialities of computer-aided design can be fully realized.

It seems to me that the main problem of the future is essentially one of communication between man and the machine, and between the computer community and the rest of society. How do we reconcile the rigorously logical, arithmetical language of the computer, with the subtle conceptual language of the user?

The gap has been narrowed from both sides. On the one hand higher level languages like *Fortran* and *Cobol* allow instructions to be given to the computer without any knowledge, let alone understanding, of what goes on inside its black boxes. Despite their limited vocabulary and even more limited grammar, these languages, which resemble the technical shorthand which scientists and engineers use rather than a basic English, nevertheless allow quite sophisticated problems to be handled. On the other hand the computer insists on

Continued on page 561

U.D.C. 621.394.147:621.394.542.3

Multi-Level Codes

Professor D. A. Bell,
Ph.D., F.Inst.P., C.Eng., F.I.E.R.E., F.I.E.E.*

Based on a paper presented at the IERE Conference on Digital Processing of Signals in Communications held in Loughborough from 11th to 13th April 1972

SUMMARY

The use of multi-level p.s.k. arouses interest in non-binary error-correcting codes. This paper points out that the algorithm for constructing BCH codes can be applied when the number of levels is a prime or a power of a prime. Quaternary codes are examined. Possible extensions to higher powers of two and to decimal codes are indicated.

* *Department of Electronic Engineering, University of Hull, Hull HU6 7RX.*

1. Introduction

Shannon's mathematical theory of communication and in particular his geometric derivation of the channel-capacity theorem, can be paraphrased as follows. Let all the signals which it is physically possible to transmit through a given channel be regarded as a set. Then there is a maximum limit to the size of a sub-set of which all the members can be mutually distinguished in the presence of any specified signal/noise ratio. The transformation of the message to be communicated into members of a sub-set of all possible signals is called coding, and Shannon stated in 1948 that by sufficiently complicated encoding it is possible to achieve the maximum rate of communication for a channel of given bandwidth and signal/noise ratio, namely

$$C = W \log (1 + P/N)$$

Since 1948 there has been a continuing search for encoding procedures which approach the ideal and are not too difficult to implement both in encoding and in decoding. (The acceptable limit on complexity in encoding and decoding tends to be progressively relaxed with the growing use of computers and computer-like techniques.)

The majority of the work on error-correcting codes has hitherto been concerned with binary codes. There is now increasing use of multi-level signalling in the form of phase-shift keying (p.s.k.) with four or eight phases, for the transmission of digital data over telephone lines, and the use of 4-phase p.s.k. has been proposed for communication via satellites. It is at present usual to encode the signal in binary, allocate two binary digits to each phase of 4-phase p.s.k. (or 3 for 8-phase p.s.k.) and possibly use a binary error-correcting code. But one should ask whether it is better to use a specifically designed error-correcting multi-level code rather than a binary code translated into multi-level signals. A general type of code which is widely adaptable in its binary form is the set of Bose–Chaudhuri–Hocquenghem (BCH) codes. It is adaptable because BCH codes can be constructed of many lengths and various error-correcting capabilities up to about a quarter of the total number of digits. (Peterson[1] tabulates 114 binary BCH codes with lengths between 7 and 255 and graphs the performance of 4885 codes with lengths between 31 and 65535.) It is also adaptable because the BCH algorithm is not limited to binary codes but may be used whenever the number of levels is a prime or a power of a prime: hence it can be used to construct codes for p.s.k. with 2, 4, 8 . . . phases. It could also be used for coding signals with 3, 5, 7, 11, 13, 17 . . . levels and there is a possibility that codes based on 11 levels might be used for coding decimal numbers, using the eleven levels for 0 to 9 and the decimal point.

2. Quaternary BCH Codes

2.1. *Mathematical Background*

BCH codes are cyclic codes and the requirement is to find a non-repeating pattern so that successive shifts of the pattern and their modulo-two sums show enough differences from each other. The mathematical basis of most cyclic codes is that the position of an element in the code is identified with the power of the variable in a polynomial, and the value of the element is identified with the coefficient of that power in the variable. For binary codes the coefficients can only be 1 or 0, represented by the presence or absence of that power. For example, $x^{10} + x^8 + x^5 + x^4 + x^2 + 1$ represents 10100110101. It is necessary for the BCH algorithm that the coefficients be the elements of a Galois field, and the number of elements in the field must be a power of a prime. The elements must also be a closed set, including the values 0 and 1 and inverses of any other elements with respect to both addition and multiplication. When the number of elements is a prime, the elements are natural numbers and the operations of addition and multiplication obey the rules of arithmetic modulo the prime in question: for example the elements of GF(2) are 0 and 1 with addition and multiplication modulo 2 and the elements of GF(11) are the numbers 0 to 10 with arithmetic modulo 11. But this does not apply to a power of a prime and the elements of GF(4) may be denoted by 0, 1, a, b† with their arithmetic properties defined by two tables (Table 1).

Table 1. Addition and multiplication tables for GF(4)

+	0	1	a	b
0	0	1	a	b
1	1	0	b	a
a	a	b	0	1
b	b	a	1	0

×	0	1	a	b
0	0	0	0	0
1	0	1	a	b
a	0	a	b	1
b	0	b	1	a

For the construction of quaternary BCH codes it is necessary to know the irreducible (non-factorizable) polynomials with coefficients drawn from GF(4) and of degree n such that $4^{n-1}-1$ is equal to the length of the code. This means that for values of n of 2 and 3, with lengths 15 and 63 quaternary digits, one must use irreducible polynomials of degrees 2 and 3 respectively. Irreducible polynomials over GF(2) have been tabulated by Marsh[2] up to degree 19 and Peterson[1] reproduces these up to degree 16 and gives a limited selection up to degree 34. Church[3] has tabulated irreducible polynomials with coefficients modulo 2, 3, 5 and 7 to degrees 11, 7, 5 and 3 respectively. As no similar table was known for GF(4), the second and third degree polynomials irreducible over GF(4) were found by listing all monic‡ polynomials of the degree in question, evaluating all products of factors amounting to the like degree, and deleting the latter from the list. (Only monic polynomials were considered, because in a Galois field it is always possible to divide through by the leading coefficient, so that a polynomial with any other coefficient of the leading term is a constant multiple of some polynomial.) The number of irreducible polynomials to be expected can be checked by the theorem that if $Z = q^m-1$, where q is the number of elements in the field, then x^Z-1 is divisible at least by $x-1$ and all irreducible polynomials of degree m. In the present case, with $q = 4$ and $m = 2$, $x^{15}-1$ can have at most 7 factors of degree 2. But over GF(4) one quadratic factor of $x^{15}-1$ is equal to the product $(x+a)(x+b)$ so there can be at most 6 second-degree polynomials which are irreducible over GF(4), and these are given in Table 2.

† The elements 1, a, b may be identified with the three cube roots of unity and both a and b are primitive elements.

‡ Coefficient of highest power is unity.

Table 2. Second-degree polynomials irreducible over GF(4)

$x^2 + x + a$	$x^2 + ax + a$
$x^2 + x + b$	$x^2 + bx + 1$
$x^2 + ax + 1$	$x^2 + bx + b$

Applying Peterson's method (ref. 1, pp. 162–5) to the factorization of $x^{15}-1$ over GF(4) one can find the properties of all second-degree polynomials which are irreducible over GF(4). Let

$$x^{15}-1 = \psi_1(x).\psi_3(x).\psi_5(x).\psi_{15}(x)$$

$$\psi_1(x) = x-1$$

$$\psi_3(x) = \frac{x^3-1}{x-1}$$

$$\psi_5(x) = \frac{x^5-1}{x-1}$$

$$\psi_{15}(x) = \frac{x^{15}-1}{\psi_1(x).\psi_3(x).\psi_5(x)}$$

By inspection and trial division by irreducible polynomials the factors of Table 3 can be found.

Table 3. Factors of $x^{15}-1$

$\psi_3(x) = (x + a)(x + b)$	order of roots = 3
$\psi_5(x) = (x^2 + ax + 1)(x^2 + bx + 1)$	,, ,, ,, = 5
$\psi_{15}(x) = (x^2 + x + a)(x^2 + x + b)(x^2 + ax + a)\times$ $\times(x^2 + bx + b)$	roots are primitive

This accounts for all six irreducible polynomials of second degree. Table 3 differs from factorization over GF(2) in the degree of the factors and of course in the structure of the factors. For example, $\psi_3(x)$ in GF(2) is the irreducible quadratic factor x^2+x+1 but in GF(4) this is the product of two linear factors, $(x+a)(x+b)$.

The work was extended to degree 3 as follows. There are 12 monic quadratics over GF(4). Each of these was multiplied in turn by $x, x+1$, $x+a$ and $x+b$ so as to produce 48 polynomials of degree 3 which have factors. These and the single term x^3 were deleted from the total list of 64 monic polynomials of degree 3, leaving the 20 polynomials shown in Table 4.

Using the same method of factorization as was used with irreducible quadratics, let

$$x^{63}-1 = \psi_1(x).\psi_3(x).\psi_7(x).\psi_9(x).\psi_{21}(x).\psi_{63}(x)$$

Table 4. Third-degree polynomials irreducible over GF(4)

$x^3 + a$	$x^3 + x^2 + ax + b$
$x^3 + b$	$x^3 + x^2 + bx + a$
$x^3 + x + 1$	$x^3 + ax^2 + x + b$
$x^3 + ax + 1$	$x^3 + ax^2 + bx + a$
$x^3 + bx + 1$	$x^3 + ax^2 + ax + a$
$x^3 + x^2 + 1$	$x^3 + ax^2 + bx + b$
$x^3 + ax^2 + 1$	$x^3 + bx^2 + x + a$
$x^3 + bx^2 + 1$	$x^3 + bx^2 + ax + a$
$x^3 + x^2 + x + a$	$x^3 + bx^2 + ax + b$
$x^3 + x^2 + x + b$	$x^3 + bx^2 + bx + b$

One can immediately write down the degree of each $\psi(x)$, the order of its roots and the degree of the minimum functions of its roots as in Table 5.

There are 3 linear factors (since $\psi_3 = (x+a)(x+b)$) and 20 cubic factors which are the irreducible polynomials of degree 3 which are listed in Table 4. The latter can be sorted into groups as follows:

$$\psi_7(x) = (x^3+x+1)(x^3+x^2+1);$$
$$\psi_9(x) = (x^3+a)(x^3+b);$$

the product of the other four irreducible polynomials containing only three non-zero terms is

$$\psi_{21}(x) = (x^3+ax+1)(x^3+bx+1)\times \times(x^3+ax^2+1)(x^3+bx^2+1)$$

and the remaining 12 irreducible polynomials of degree 3 are factors of $\psi_{63}(x)$ and are primitive.

2.2. *The Construction of BCH Codes in GF(4)*

The algorithm for the construction of BCH codes in GF(q) requires that one takes one of the roots α of a primitive polynomial of degree m such that q^m-1 is the length of the code, that for the correction of t errors one takes $2t$ consecutive powers of α and that one finds the minimum function of each of the chosen powers of α. The product of these minimum functions gives a sequence which with the appropriate number of cyclic shifts forms the generating matrix for the code. The number of check digits is equal to the total degree of the product of minimum functions, and the number of information digits is the difference between length of code and number of check digits.

With the aid of Table 2 one can construct codes of length $4^2-1 = 15$, but first it is necessary to find the minimum functions of powers of a root of a chosen primitive polynomial. If α is a root of the primitive polynomial x^2+x+a, the minimum functions of $\alpha^0, \alpha^1, \alpha^2, \alpha^3 \ldots \alpha^{14}$ are given in Table 6.

Table 6. Minimum functions of powers of a root of x^2+x+a

Power of α	Minimum function
0	$x + 1$
1	$x^2 + x + a$
2	$x^2 + x + b$
3	$x^2 + bx + 1$
4	as 1
5	$x + a$
6	$x^2 + ax + 1$
7	$x^2 + ax + a$
8	as 2
9	as 6
10	$x + b$
11	$x^2 + bx + b$
12	as 3
13	as 2
14	as 11

In GF(q), the minimum function of α^k is also the minimum function of α^{qk} and need not be repeated when both powers of α are involved in the construction of a code. In GF(2) this means that all even powers except α^0 are covered by previous odd powers; and it is therefore usual to start with α. With GF(4) only every fourth term from 4 onwards is regularly covered by an earlier term, though Table 6 shows that with the chosen quadratic function powers of α of 6 and 9, 2 and 13 and 11 and 14 also coincide. To construct a 2-error-

Table 5. Factorization of $x^{63}-1$

$\psi(x)$	Degree of $\psi(x)$	Order of roots	Degree of minimum function
$\psi_1(x) = x - 1$	1	1	1
$\psi_3(x) = \dfrac{x^3 - 1}{\psi_1(x)}$	2	3	1
$\psi_7(x) = \dfrac{x^7 - 1}{\psi_1(x)}$	6	7	3
$\psi_9(x) = \dfrac{x^9 - 1}{\psi_1(x).\psi_3(x)}$	6	9	3
$\psi_{21}(x) = \dfrac{x^{21} - 1}{\psi_1(x).\psi_3(x).\psi_7(x)}$	12	21	3
$\psi_{63}(x) = \dfrac{x^{63} - 1}{\psi_1(x).\psi_3(x).\psi_7(x).\psi_9(x).\psi_{21}(x)}$	36	63	3

correcting code, requiring minimum functions of α, α^2, α^3, and α^4, reference to Table 6 indicates the combination

$$(x^2+x+a)(x^2+x+b)(x^2+bx+1)$$
$$= x^6+bx^5+x^4+x^3+ax^2+ax+1$$

The code generating matrix is therefore 1 b 1 1 a a 1 0 0 0 0 0 0 0 0 plus its 8 shifts.

Table 7. Generating matrix for (15, 9, 2) quaternary code

1	b	1	1	a	a	1	0	0	0	0	0	0	0	0
0	1	b	1	1	a	a	1	0	0	0	0	0	0	0
0	0	1	b	1	1	a	a	1	0	0	0	0	0	0
0	0	0	1	b	1	1	a	a	1	0	0	0	0	0
0	0	0	0	1	b	1	1	a	a	1	0	0	0	0
0	0	0	0	0	1	b	1	1	a	a	1	0	0	0
0	0	0	0	0	0	1	b	1	1	a	a	1	0	0
0	0	0	0	0	0	0	1	b	1	1	a	a	1	0
0	0	0	0	0	0	0	0	1	b	1	1	a	a	1

This code must have 6 check digits and therefore 9 information digits, and its members are the rows and linear combinations of the rows of the generator matrix. Many of these differ in 7 places, but adjacent rows differ in only 6 so that the code could not correct 3 errors. A 3-error-correcting code can be constructed in exactly the same way using the minimum functions of α, α^2, α^3, α^5, α^6, with total degree 9, so that it would be a (15, 6, 3) code. Increase to 4-error-correcting calls for the addition of α^7 and α^8, but the latter has the same minimum function as α^2, leading to a total degree of 11 and a (15, 4, 4) code. Reference to Table 6 shows that by the use instead of α^2, α^3, α^4, α^5, α^6, α^7, $[a^8 \equiv \alpha^2]$ $[\alpha^9 \equiv \alpha^6]$, one would also have 11 check digits. Either sequence can be extended to make (15, 3, 5) and (15, 1, 6) codes. The latter compares with the (15, 1, 7) which could be achieved by majority logic.

Codes of length 63 can be similarly constructed, but this will require the determination of the minimum functions of a root of a chosen primitive polynomial of degree three, e.g. of x^3+x^2+x+a, and this has not yet been done. A matter of practical importance is the point at which the linear factors $x+a$ and $x+b$ come in (cf. α^5 and α^{10} in Table 6) since the addition of one of these minimum functions will add only one check digit whereas the addition of a cubic adds three. It would therefore be advantageous to choose a primitive polynomial such that the linear factors come early in the sequence of minimum functions of powers of one of its roots. This consideration does not arise in GF(2) where the only linear factor is $x+1$ which is always the minimum function of α^0.

2.3. *Performance of Quaternary BCH Codes*

Only two kinds of errors are possible with binary digits, and these are associated with the binary symmetric channel and the binary erasure channel. In a quaternary channel one must consider (a) deletion of a digit, (b) change by ± 1 level, (c) change specifically by ± 2 levels (phase reversal in p.s.k.) or (d) change by an arbitrary number of levels. In cases (b) and (c) it is known that the transmitted signal must have been one of two, but if the change is arbitrary no information is conveyed so that (a) and (d) are equivalent. Binary codes can be converted to quaternary by allocating 2 binary digits to each of the 4 levels of a quaternary digit in such a way that a 1-level error in the quaternary digit invalidates only one binary digit. A single-error-correcting binary code is then still single-error-correcting after conversion to quarternary if one defines 'single error' as case (b) above. However, an error in the quaternary system of type (a) or (d) would invalidate two binary digits and to give protection in these cases one would need to use a binary code capable of correcting either single errors or bursts of two errors. In this paper comparison with binary codes is considered from the point of view of arbitrary errors. Since an undetected error of one quaternary digit may mean the loss of two bits of information, failure of the quaternary code leads to more rapid degradation of the signal than would occur in a binary system. But if one is considering the performance of a code within its error-correcting capacity, then a quaternary code needs only half as many information digits as a binary code to convey the same information and in the absence of modification for error-correcting codes the quaternary channel will use half the digit rate of the binary. Therefore, it is reasonable to compare codes with the same number of bits, i.e. twice as many digits for a binary code as for a quaternary code, on the basis of their information rates k/n.

Quaternary codes must conform to the general principle that code members can be identified in the presence of e errors if they are separated from each other by a 'distance' $2e+1$. In binary codes the distance between two members is the number of digit places in which they differ. For multi-level codes one sometimes uses the Lee distance which is the sum of the amounts of difference in each digit place, taking the less of the direct arithmetic difference or its complement (e.g. the distance between 1 and 4 is 3, but between 1 and 9 is 2).

Systematic correction of errors requires that the magnitude as well as location of digit errors be determinable from the check sums, and in principle this is provided for by BCH codes in GF(q), though the decoding procedure is not simple. With the present state of development of digital techniques and integrated circuits, it may be more practicable to decode short codes by an exhaustive search for the best match whenever an error is indicated.

There appears to be an anomaly in that the code shown in Table 7 has members which differ in a sufficient number of places to allow the nearest code member to be chosen in the presence of two errors without reference to the values of digits. This suggests that there might be better codes than BCH: a suggestion which is supported by (a) the fact that the BCH algorithm predicts a (23, 12, 2) code whereas there exists Golay's (23, 12, 3) code, and (b) by the approach of long binary BCH codes to the Varsharmov–Gilbert lower bound rather than the Hamming upper bound.

The method of constructing BCH codes of length q^m-1 leads to a number of check digits not greater than $2t(1-1/q)m$ for the correction of t errors. This is mt for $q = 2$ and $(3/2)mt$ for $q = 4$. However, the length of code q^m-1 will be approximately twice as great for $q = 4$ so that as a proportion of the code length the number of check digits in the quaternary code will be only 3/4 of that for a binary code. This refers in both cases to the upper limit on the number of check digits, allowing for the fact that every qth minimum function will be the same as an earlier one. In fact the number of check digits may be a little less owing to the occurrence of other coincidences between minimum functions and the occurrence of one or two functions of less degree than m. Table 8 gives a comparison between quaternary BCH codes and the binary BCH codes which are nearest in terms of number of erroneous digits which can be corrected and information rate k/n. The quaternary codes of length 15 have been specified in section 2 and those of length 63 are assumed to have the maximum number of check digits. The actual binary BCH codes are taken from the table on p. 166 of Peterson's book and the worst-case from the formula of mt check digits. In the absence of the construction of specific quaternary BCH codes of length 63 it is not known how far coincidences amongst the minimum functions will lead to improvement over the worst-case predictions.

Ulrich[4] states that a binary single-error-correcting code is more efficient than a non-binary code capable of correcting any single error of unrestricted magnitude. This would appear to be correct if one added to the binary column in Table 8 the code (31, 26, 1) with $k/n = 0{\cdot}84$ and in Table 9 (127, 120, 1) with $k/n = 0{\cdot}945$. However, the binary codes with only *single* error-correction but twice the length can deal with only half the per-digit frequency of occurrence of errors.

It is possible to construct an upper bound for quaternary codes in the same way as the Hamming upper bound for binary codes considering short or at least finite lengths of code. The check digits have to carry enough information to select from all error patterns which the code sets out to cover plus the one case of no error. In a code with q levels there may be any of $q-1$ different errors in each position, making a total for z digit errors of $(q-1)^z\binom{n}{z}$ different error patterns. For any number of errors up to t the general expression for the number of error patterns is

$$P = \sum_{z=0}^{t} (q-1)^z \binom{n}{z}$$

(In the binary case $q-1 = 1$ and this reduces to the familiar Hamming formula $1 + n + \binom{n}{2} + \ldots \binom{n}{t}$. An equivalent result was given by Verhoeff[5].)
The theoretical number of check digits is then

$$r_{th} = [\log_q P]$$

where the square bracket on the right-hand side means 'the least integer equal to or greater than' the content of the bracket. If one were able to use a fractional number of check digits, one would arrive at a number r_{fr} smaller than r_{th}. The smallness of the difference between r_{th} and r_{fr} is a measure of the density of packing

Table 8. Comparison of codes
$n = 4^2$ or 2^5 (Actual codes)

Quaternary			Binary		
(n, k, t)	$\frac{k}{n}$	$\frac{t}{n}$	(n, k, t)	$\frac{k}{n}$	$\frac{t}{n}$
15, 12, 1	0·8	0·067	31, 21, 2	0·68	0·065
15, 9, 2	0·6	0·13	31, 11, 5	0·36	0·16
15, 6, 3	0·4	0·2	31, 6, 7	0·19	0·23
15, 4, 4	0·27	0·27	—		
15, 3, 5	0·2	0·33	—		
15, 1, 6	0·067	0·4	—		

Table 9. Comparison of codes
$n = 4^3$ or 2^7

Quaternary worst case			Binary worst case			Binary actual		
n, k, t	$\frac{k}{n}$	$\frac{t}{n}$	n, k, t	$\frac{k}{n}$	$\frac{t}{n}$	n, k, t	$\frac{k}{n}$	$\frac{t}{n}$
63, 57, 1	0·906	0·016	127, 113, 2	0·89	0·016			
54, 2	0·86	0·032	99, 4	0·78	0·032	127, 99, 4	0·78	0·032
48, 3	0·76	0·048	85, 6	0·67	0·047	85, 6	0·67	0·047
45, 4	0·71	0·064						
42, 5	0·67	0·079	57, 10	0·45	0·079	64, 10	0·50	0·079
36, 6	0·57	0·095						
33, 7	0·52	0·11	29, 14	0·23	0·11	43, 14	0·34	0·11
18, 10	0·29	0·16				29, 21	0·23	0·17
15, 11	0·29	0·17				22, 23	0·17	0·18
6, 13	0·095	0·21				15, 27	0·12	0·21

of the code and the performance in this respect of quaternary codes of length 4^2-1 is shown in Table 10. Any difference between columns (2) and (3) indicates further departure from perfect packing by the BCH code. It does not follow, however, that better packing is possible with any *systematic* code.

Table 10. Packing density of quaternary codes

t	r_a	r_{th}	r_{fr}
1	3	3	2·76
2	6	5	4·98
3	9	7	6·85
4	11	9	8·46
5	12	10	9·85
6	14	12	11·05

t = number of random errors which can be corrected in quaternary code of length 15.

r_a = number of check digits achieved in quaternary BCH code.

r_{th} = theoretical least-integer number of check digits.

r_{fr} = fractional number of digits equivalent to information necessary for the correction of t errors.

3. Quaternary Analogues of Reed–Muller

Reed–Muller binary codes are based on successive binary partition of the length of code between ones and noughts, e.g. a code of length 16 is based on the sequences shown in Table 11 and certain combinations of them.

Table 11. Zero- and first-order elements of a Reed–Muller code

1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1
0	0	0	0	1	1	1	1	0	0	0	0	1	1	1	1
0	0	1	1	0	0	1	1	0	0	1	1	0	0	1	1
0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1

For an rth order code the basis vectors include products of the vectors in Table 11 up to r at a time.

A standard de-coding procedure depends on dividing the digits into groups which for all but one of the basis rows contain even numbers of each symbol; and since in GF(4) the sum of any pair of like symbols is zero the method of even numbers of symbols will be valid for every pair of symbols drawn from GF(4). In binary the linear combinations of the basis vectors are sums of rows each multiplied by 0 or 1, but in quaternary the individual rows may be multiplied by 0, 1, a or b. Peterson[1] gives a general formula for the number of information digits for a given length of code which is independent of the number of symbols for each digit:

$$k = 1+\binom{m}{1}+\ldots\binom{m}{r}$$

where the length $n = q^m$ and the code is of rth order, i.e. the basis vectors include products of the initial vectors taken up to r at a time. The minimum distance is $d = q^{m-r}$. Since d is even, the Reed–Muller codes have the capacity to detect one more error than they correct. But the comparison between binary and quaternary codes in Table 12 is based on their error-correcting capability, requiring only distance $d-1$.

Table 12. Properties of binary and quaternary Reed–Muller codes according to Peterson's generalized formulae for any number of levels

(a) $q = 2$

m	n	r	k	d	t
5	32	1	6	16	7
5	32	2	16	8	3
5	32	3	26	4	1
5	32	4	31	2	*
5	32	5	32	1	0
6	64	1	7	32	15
6	64	2	22	16	7
6	64	3	42	8	3
6	64	4	57	4	1
6	64	5	63	2	*
6	64	6	64	1	0
7	128	1	8	64	31
7	128	2	29	32	15
7	128	3	64	16	7
7	128	4	9	8	3
7	128	5	120	4	1
7	128	6	127	2	*
7	128	7	128	1	0

(b) $q = 4$

m	n	r	k	d	t
3	64	1	4	16	7
3	64	2	7	4	1
3	64	3	8	1	0
4	256	1	5	64	31
4	256	2	11	16	7
4	256	3	15	4	1
4	256	4	16	1	0

* Single-error-detecting.

It appears from Table 12 that quaternary R–M codes are much worse than binary: for example we have a (64, 7, 1) quaternary compared with (128, 99, 3) binary, and (64, 4, 7) quaternary against (128, 29, 15) binary. A feature of the generalized formulae given by Peterson is that the length of code $n = q^m$ is dependent on q, with given m, but the number of information digits k depends only on m and not on q. This naturally means that increasing q decreases the ratio k/n. The method suggested above, using the structure of a binary code but multiplying basis vectors by all coefficients drawn from GF(4), should at least give the same ratio k/n for quaternary and binary codes. Such a method is not general, since it relies on the sum of an even number of like elements being zero, and if q is a prime other than 2 only the sum of q like terms is zero. (Hence the greater length of codes based on a general q.) But codes of 2, 4, 8, 16 . . . levels are so potentially important for multi-level p.s.k. that it may be worth pursuing Reed–Muller codes relying on even sums being zero. Table 13 shows addition and multiplication tables for GF(8). It can be seen that pairs of like terms sum to zero.

4. Decimal Codes

There has been a good deal of work on error-*detecting* decimal codes[5] but no practical scheme for error-

Table 13. Addition and multiplication tables for GF(8)

+	0	1	*A*	*B*	*C*	*D*	*E*	*F*
0	0	1	*A*	*B*	*C*	*D*	*E*	*F*
1	1	0	*B*	*C*	*D*	*E*	*F*	*A*
A	*A*	*B*	0	*D*	*E*	*F*	1	*C*
B	*B*	*C*	*D*	0	*F*	1	*A*	*E*
C	*C*	*D*	*E*	*F*	0	*A*	*B*	1
D	*D*	*E*	*F*	1	*A*	0	*C*	*B*
E	*E*	*F*	1	*A*	*B*	*C*	0	*D*
F	*F*	*A*	*C*	*E*	1	*B*	*D*	0

×	0	1	*A*	*B*	*C*	*D*	*E*	*F*
0	0	0	0	0	0	0	0	0
1	0	1	*A*	*B*	*C*	*D*	*E*	*F*
A	0	*A*	*B*	*C*	*D*	*E*	*F*	1
B	0	*B*	*C*	*D*	*E*	*F*	1	*A*
C	0	*C*	*D*	*E*	*F*	1	*A*	*B*
D	0	*D*	*E*	*F*	1	*A*	*B*	*C*
E	0	*E*	*F*	1	*A*	*B*	*C*	*D*
F	0	*F*	1	*A*	*B*	*C*	*D*	*E*

correction has been reported. It would in principle be possible to make an 11-level BCH code and then discard all code members in which the 11th level appeared in any digit, which would mean discarding 1/11th of all members of the set. Moreover, for codes of reasonable length one would not go beyond polynomials of degree 2 for the basis, and this seems rather restrictive.

5. Conclusion

There is scope for the development of specific error-correcting codes based on GF(2^s) for use with multi-level p.s.k. There may possibly be a case for the development of codes based on GF(11) for use with decimal numbers; but the complexity of BCH codes based on GF(11) would be formidable, and Reed–Muller codes would give a low information rate. Other types of code have yet to be examined from this point of view.

6. References

1. Peterson, W. W., and Weldon, E. J., Jr., 'Error-Correcting Codes', 2nd edn. (M.I.T. Press, Cambridge, Mass., 1972).
2. Marsh, R. W., 'Table of Irreducible Polynomials over GF(2) Through Degree 19'. U.S. Dept. of Commerce, Washington, D.C., 1957.
3. Church, R., 'Tables of irreducible polynomials for the first four prime moduli', *Ann. Math.*, **36**, pp. 198–209, January 1935.
4. Ulrich, W., 'Non-binary error correction codes', *Bell Syst. Tech. J.*, 36, pp. 1341–88, 1957.
5. Verhoeff, J., 'Error detecting decimal codes', Mathematical centre tracts, no. 29, Mathematisch Centrum, Amsterdam, 1969.
6. Bell, D. A., 'Walsh functions and Hadamard matrixes', *Electronics Letters*, **2**, pp. 340–1, September 1966.

Manuscript first received by the Institution on 16th November 1971, and in final form on 15th September 1972. (Paper No. 1488/Com 60.)

Conference Report (*cont. from p.* 554)

great precision in the presentation of these problems. It is a common and salutary experience when commissioning a computerized system to discover how inadequate is the documentation on which it has to operate; and one valuable by-product of the computer is the more disciplined approach in the handling of data which it has imposed upon us.

But there is a real danger that this will be carried too far. The imprecise clerical systems of the past worked because they relied heavily on human understanding and intervention. Decisions may have been based on dubious data but they worked because they relied on years of practical experience. Value judgements and human understanding are still necessary in all social matters and these are impossible, or impossibly clumsy to computerize.

It is important to recognize the limitations as well as the potentialities of the computer. Compared with the weighting of data which we perform intuitively every day of our lives, the most advanced adaptive program is on the level of the simple governor on an early steam engine. Even more fundamental is the ability of the human mind to comprehend pattern. Until we know far more about a process we take so much for granted, we are in no position to call the computer to our aid. A young child can look casually at the night sky and immediately point out the constellation of Orion or the Plough. To search the whole hemisphere of stars with sufficient resolution and perform the millions of correlations is beyond the capacity of any computer we can imagine.

From time to time science fiction frightens us with the spectre of HAL 1, and the possibility of designing a computer whose cogitative processes will out-think its creator. The real danger we face is not that the computer will become more like a human being, but that we may become more like computers.

You have an authority and a responsibility which extends beyond your particular expertise; and the great majority of people who are outside the computer field must rely on you to see that we do not drift inadvertently into an obsessively numerate society which, to quote Oscar Wilde, 'Knows the price of everything and the value of nothing'.

U.D.C. 621-52

An Extension to Fukuma and Matsubara Jump Resonance Criterion by the Use of Describing Functions

C. F. HO,
M.Sc.(Eng.), Ph.D., C.Eng., M.I.E.E., M.I.E.R.E.*

SUMMARY

The jump resonance criterion presented by Fukuma and Matsubara normalizes the non-linearity characteristic and restricts its derivative (with respect to the input) to lie within the sector $0 \leqslant f'(e) \leqslant 1$. The region in which jump resonance can occur is localized in the domain near the origin in the inverse polar plane. As compared with the circle criterion which may also be applied in the study of jump resonance, this criterion is conservative. Based on the sector-bounded-non-linearity concept afforded by the circle criterion, the work of Fukuma and Matsubara has been extended and compared favourably with the circle criterion.

* *Department of Electrical Engineering, University of Hong Kong, Hong Kong.*

1. Introduction

Jump resonance is one of the common phenomena exhibited in non-linear control systems. It is observed in some non-linear systems under a periodic excitation that when the amplitude or period of the excitation is smoothly changed, the output undergoes a discontinuous change in amplitude. Therefore, in the study of stability of non-linear systems it is in most cases insufficient to be concerned only with stability in the bounded sense. Stability in the continuous sense must also be considered since it yields information concerning the existence of jump resonance.

2. The Circle Criterion

The circle criterion is a recently developed stability criterion through the use of functional analysis. It consists of two theorems. The first theorem concerns the sufficient conditions for stability of non-linear systems in the bounded sense and the second theorem concerns the sufficient conditions for stability in the continuous sense. Consider the feedback system shown in Fig. 1 with r, e, u and y real-valued measurable functions of t for $t \geqslant 0$.

The block labelled N represents a memoryless non-linearity, either time-invariant or time-varying, which satisfies the conditions

$$u(t) = N[e(t), t]$$
$$N[0, t] = 0 \qquad \text{for} \qquad t \geqslant 0 \qquad \ldots\ldots(1)$$

and there exists a positive constant β and a real constant α such that

$$\alpha \leqslant \frac{N[e(t), t]}{e(t)} \leqslant \beta \qquad \text{when} \qquad t \geqslant 0 \quad \ldots\ldots(2)$$
$$\text{for all real} \qquad e \neq 0.$$

The block labelled G is a linear time-invariant element with the constraint

$$y(t) = \int_0^t g(t-\tau)u(\tau)\,\mathrm{d}\tau - g_0(t) \qquad t \geqslant 0 \qquad \ldots\ldots(3)$$

where g and g_0 are real-valued functions such that

$$\int_0^\infty |g(t)|\,\mathrm{d}t < \infty$$

and

$$\int_0^\infty |g_0(t)|^2\,\mathrm{d}t < \infty. \qquad \ldots\ldots(4)$$

The function g_0 takes into account the initial conditions at $t = 0$.

The circle criterion thus defines sufficient conditions for stability as follows:

THEOREM 1. A basic feedback system of Fig. 1 is stable in the bounded sense if it has a non-linearity of the form stated and satisfies the conditions illustrated in Fig. 2.

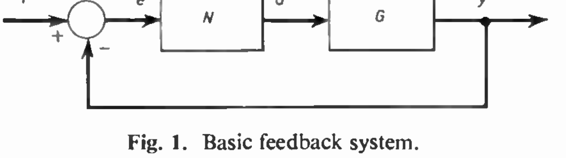


Fig. 1. Basic feedback system.

Expressing this mathematically,

(i) $\alpha > 0$, the Nyquist plot of $G(j\omega)$ lies outside the circle C_1 of radius $\frac{1}{2}[1/\alpha - 1/\beta]$ centered on the real axis of the complex plane at $[-\frac{1}{2}(1/\alpha + 1/\beta), 0]$ and does not encircle C_1.

(ii) $\alpha = 0$, the Nyquist plot of $G(j\omega)$ lies to the right of the line given by $\text{Re}\,[G(j\omega)] = -1/\beta$.

(iii) $\alpha < 0$, the Nyquist plot of $G(j\omega)$ is contained within the circle C_2 of radius $\frac{1}{2}[1/\alpha - 1/\beta]$ centered on the real axis of the complex plane at $[-\frac{1}{2}(1/\alpha - 1/\beta), 0]$.

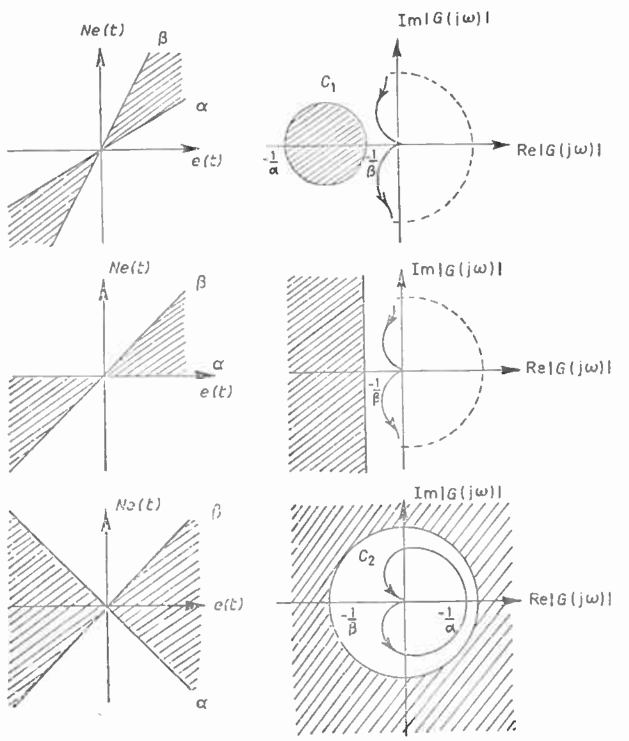


Fig. 2. Sectors of non-linearities and their corresponding critical regions in the $G(j\omega)$ plane.

THEOREM 2. A basic feedback system of Fig. 1 is stable in the continuous sense if

(i) $r(t)$ and $y(t)$ are square integrable functions of time,

(ii) the non-linearity introduces the constraints

$$u(t) = N[e(t), t]$$

$$N[0, t] = 0 \qquad \text{for} \qquad t \geqslant 0$$

$$\alpha \leqslant \frac{N[e_1(t), t] - N[e_2(t), t]}{e_1(t) - e_2(t)} \leqslant \beta \qquad \ldots\ldots(5)$$

for all real $e \neq 0$, $e_1 \neq e_2$ and $t \geqslant 0$ and

(iii) one of the three conditions given in Fig. 2 is satisfied.

The basic difference between the two theorems is the manner in which the non-linearity is bounded in a sector. The first requires that the instantaneous gain of the non-linearity

$$\frac{u(t)}{e(t)} \qquad \ldots\ldots(6)$$

be bounded within a sector by lines of slope α and β, whereas the second requires that the incremental gain of the non-linearity

$$\frac{u_1(t) - u_2(t)}{e_1(t) - e_2(t)} \qquad \ldots\ldots(7)$$

be bounded within the sector $[\alpha, \beta]$.

A system exhibits jump resonance if small changes in the input cause large changes in the output. Evidently, a system is stable in the continuous sense that by definition cannot possess jump resonance phenomenon. Thus, Theorem 2 of the circle criterion assures the non-existence of jump resonance.

3. Jump Resonance Criteria

West and others[1] obtained a jump resonance criterion of a control system with a non-linear element written by $y = e^3$. Their development revealed the non-existence of jump resonance if the locus of the linear part $G(j\omega)$ did not pass the region bounded by two half lines starting from the origin in the direction of $180^\circ \pm 30^\circ$ in the polar plane. Later Hatanka[2] obtained results for non-linearities with real describing function which showed that the critical region for jump resonance was bounded by the envelope of a family of circles centered on the real axis. Recently, Fukuma and Matsubara[3] obtained a more precise boundary curve for the family of circles on the basis of Hatanaka's work by normalizing the non-linearity characteristic and using the inverse polar plane of $G(j\omega)$. They restricted the derivative of the non-linearity with respect to its input to lie in the sector

$$0 \leqslant f'(e) \leqslant 1. \qquad \ldots\ldots(8)$$

The region in which jump resonance could occur was localized in the domain near the origin.

4. Comparison of the Circle Criterion and Fukuma and Matsubara's Jump Resonance Criterion

Because of the similarities of the Theorem 2 of the circle criterion and Fukuma and Matsubara's work, namely, both criteria impose a critical contour in the polar plane of $G(j\omega)$, it is worthwhile studying carefully and making a comparison. At first sight, one may expect that the ellipse given by Fukuma and Matsubara's jump resonance criterion lies inside the critical circle dictated by Theorem 2 of the circle criterion. This is not always true, however, since the circle criterion allows the incremental gain of the non-linearity to be bounded in a sector by a minimum of α, while Fukuma and Matsubara's criterion always takes the minimum bounding slope as zero. This is illustrated by considering the non-linearity characteristic of Fig. 3. By transforming the critical region of the circle criterion into the inverse polar plane (see Appendix), the contours of the circle

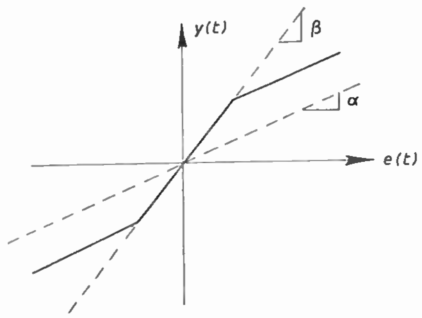


Fig. 3. Non-linearity characteristic.

criterion and the jump resonance criterion are as shown in Fig. 4.

It is also conjectured that in applying both criteria to systems, the circle criterion is more conservative due to more stringent restrictions on the systems. However, as seen from Fig. 4 the Fukuma and Matsubara's criterion is even more conservative in the estimate of the stability. This paper proposes to extend the Fukuma and Matsubara's jump resonance criterion to this type of non-linearity and to compare it with the circle criterion.

5. Extension of Fukuma and Matsubara's Criterion

Consider the feedback system of Fig. 1 with the following assumptions:

When the input $r(t) = R \sin \omega t$ is applied,

(i) the system remains stable,

(ii) no sub-harmonic oscillation appears, and

(iii) the higher harmonic oscillations are sufficiently attenuated by $G(s)$.

These conditions are essentially the same as those applied to the analysis using describing function.

In the steady state, $e(t)$ can be written as

$$e(t) = E \sin(\omega t + \phi). \qquad \ldots\ldots(9)$$

It follows that

$$E\,e^{j\phi} = \frac{R}{1+N(E)G(j\omega)} \qquad \ldots\ldots(10)$$

where $N(E)$ is the describing function of the non-linearity and is given by

$$N(E) = n_1(E) + jn_2(E). \qquad \ldots\ldots(11)$$

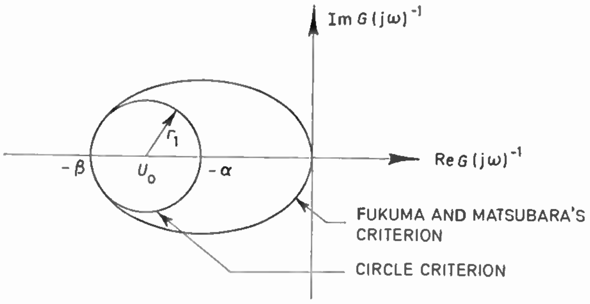


Fig. 4. Critical contours for circle criterion and Fukuma and Matsubara's criterion for jump resonance.

Introducing the notation

$$\frac{1}{G(j\omega)} = H(j\omega)$$

$$= h_1(j\omega) + jh_2(j\omega) \qquad \ldots\ldots(12)$$

the magnitude of eqn. (10) becomes

$$(h_1^2+h_2^2)R^2 = [(h_1+n_1)^2+(h_2+n_2)^2]E^2. \qquad \ldots\ldots(13)$$

Taking the partial derivative with respect to E, one has

$$(h_1^2+h_2^2)R\,\frac{\partial R}{\partial E} = \left[\left(h_1+\frac{n_1+\gamma_1}{2}\right)^2+\left(h_2+\frac{n_2+\gamma_2}{2}\right)^2 - \frac{(n_1-\gamma_1)^2+(n_2-\gamma_2)^2}{4}\right]E \qquad \ldots(14)$$

where

$$\gamma_1 = n_1 + E\frac{dn_1}{dE}$$

$$\gamma_2 = n_2 + E\frac{dn_2}{dE}.$$

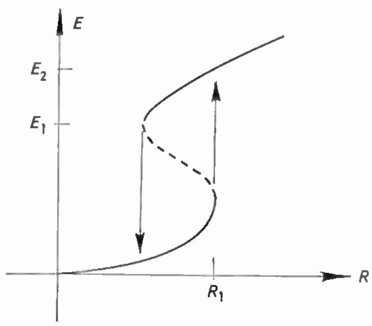


Fig. 5. Graph of E against R.

Since the system is stable, $1+N(E)G(j\omega) \neq 0$ and since $E=0$ for $R=0$, $E\to\infty$ for $R\to\infty$, the plot of E against R for some fixed ω must be similar to Fig. 5 and there must exist an ω, R and E which satisfy

$$\frac{\partial R}{\partial E} < 0 \qquad \ldots\ldots(15)$$

for jump resonance to exist.

This condition is simply a statement of the fact that if the system is stable in the bounded sense, then jump resonance occurs only if for some value of ω, the output of the system changes simultaneously for a small change in the input. The discontinuous jump is evident from Fig. 5 by allowing R to increase monotonically. As R passes through the value R_1, E jumps from the value E_1 to value E_2.

Equations (14) and (15) indicate the inequality

$$(h_1-C_1)^2+(h_2-C_2)^2 < \rho^2 \qquad \ldots\ldots(16)$$

where

$$C_1 = -\tfrac{1}{2}(n_1+\gamma_1)$$

$$C_2 = -\tfrac{1}{2}(n_2+\gamma_2)$$

$$\rho = \tfrac{1}{2}\sqrt{(n_1-\gamma_1)^2+(n_2-\gamma_2)^2}$$

which represents the interior of a circle in $H(j\omega)$ plane. Since C_1, C_2 and ρ are all functions of E, the envelope of circles for various values of E is the boundary curve of the region in which jump resonance can occur. Let the non-linearity be single-valued, continuous and odd $f(e)$. Following the derivation afforded by Fukuma and Matsubara, it can be shown that

$$C_1 = -\frac{1}{\pi}\int_{-\pi/2}^{\pi/2} f'(E\sin\theta)\,d\theta \qquad \text{......(17)}$$

$$C_2 = 0 \qquad \text{......(18)}$$

$$\rho = \frac{1}{\pi}\left|\int_{-\pi/2}^{\pi/2} f'(E\sin\theta)\cos 2\theta\,d\theta\right|. \qquad \text{......(19)}$$

The boundary curve of the circles is found by maximizing ρ for a given C_1. Now, consider the derivative of the non-linearity with respect to its input to lie in the sector

$$\alpha \leqslant f'(e) \leqslant \beta. \qquad \text{......(20)}$$

Thus, for a constant E, let

$$f'(E\sin\theta) = g(\theta). \qquad \text{......(21)}$$

Equations (17), (19) and (20) become

$$\alpha \leqslant g(\theta) \leqslant \beta \qquad \text{......(22)}$$

$$C_1 = -\frac{2}{\pi}\int_0^{\pi/2} g(\theta)\,d\theta \qquad \text{......(23)}$$

$$\rho = \frac{2}{\pi}\left|\int_0^{\pi/2} g(\theta)\cos 2\theta\,d\theta\right|. \qquad \text{......(24)}$$

The $g(\theta)$ which maximizes eqn. (23) is given in Fig. 6.

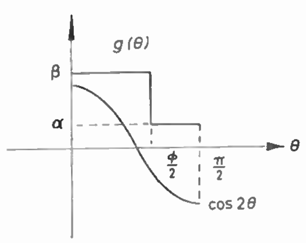


Fig. 6. Plot of $g(\theta)$.

Thus

$$C_1 = \frac{2\beta}{\pi}\int_0^{\phi/2} d\theta - \frac{2\alpha}{\pi}\int_{\phi/2}^{\pi/2} d\theta$$

$$= \left(\frac{\alpha-\beta}{\pi}\right)\phi - \alpha \qquad (0 \leqslant \phi \leqslant \pi). \qquad \text{......(25)}$$

Similarly

$$\rho = \frac{2\beta}{\pi}\int_0^{\phi/2} \cos 2\theta\,d\theta + \frac{2\alpha}{\pi}\int_{\phi/2}^{\pi/2} \cos 2\theta\,d\theta$$

$$= \left(\frac{\beta-\alpha}{\pi}\right)\sin\phi \qquad (0 \leqslant \phi \leqslant \pi). \qquad \text{......(26)}$$

The equation of the circle is then given by

$$\left[h_1 + \frac{(\beta-\alpha)}{\pi}\phi + \alpha\right]^2 + h_2^2 < \left(\frac{\beta-\alpha}{\pi}\right)^2 \sin^2\phi. \qquad \text{......(27)}$$

The boundary curve can be found by considering simultaneously the equation

$$f(h_1^*, h_2^*, \phi) = \left[h_1^* + \frac{(\beta-\alpha)}{\pi}\phi + \alpha\right]^2 + (h_2^*)^2 - \left(\frac{\beta-\alpha}{\pi}\right)^2 \sin^2\phi \qquad \text{......(28)}$$

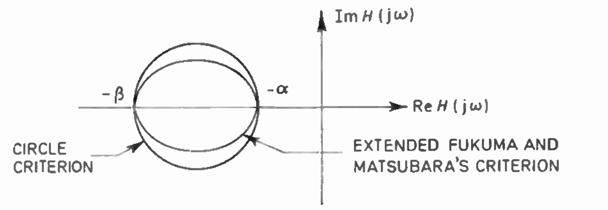


Fig. 7. Critical contours for circle criterion and extended Fukuma and Matsubara's criterion and jump resonance.

and its partial derivative with respect to ϕ

$$f'(h_1^*, h_2^*, \phi) = 2\left[h_1^* + \frac{(\beta-\alpha)}{\pi}\phi + \alpha\right]\frac{(\beta-\alpha)}{\pi} - \left(\frac{\beta-\alpha}{\pi}\right)^2 \sin^2\phi$$

$$= 0. \qquad \text{......(29)}$$

Hence the envelope of the family of circles is given by

$$h_1^* = -\left[\frac{(\beta-\alpha)}{2\pi}(2\phi - \sin 2\phi) + \alpha\right] \qquad (0 \leqslant \phi \leqslant \pi) \qquad \text{......(30)}$$

$$h_2^* = \pm\frac{(\beta-\alpha)}{\pi}\sin^2\phi. \qquad \text{......(31)}$$

Given a non-linearity bounded in a sector with slope lines α and β, the contour described by eqns. (30) and (31) will lie inside the critical circle dictated by the circle criterion as shown in Fig. 7.

6. Worked Example

The merit of Theorem 2 of the circle criterion and the extended Fukuma's criterion is demonstrated by the following example.

Consider the feedback system of Fig. 1 with the non-linearity characteristic shown in Fig. 3 and the transfer function of the linear plant given by

$$G(s) = \frac{K}{S(S+1)}. \qquad \text{......(32)}$$

Assuming that the system meets all the requirements for application of circle criterion and the describing function method of analysis. It is desired to compute the range of gain, K of $G(s)$ for the non-existence of jump resonance.

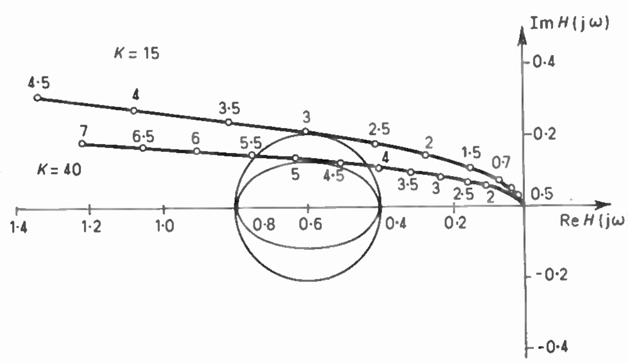


Fig. 8. Jump resonance investigation using circle criterion and extended Fukuma and Matsubara's criterion.

It is noted from eqn. (5) or eqn. (20) that the incremental gain of the non-linearity given in Fig. 3 be bounded in a sector with slope lines

$$\alpha = \frac{2}{5}, \qquad \beta = \frac{4}{5}. \qquad \text{......(33)}$$

By circle criterion the critical circle lies symmetrically on the real axis of $G(j\omega)$ plane and passes through the points $(-1/\alpha)$ and $(-1/\beta)$. This circle is mapped onto the inverse polar plane by the transformation (see Appendix),

$$\left(h_1 + \frac{\beta+\alpha}{2}\right)^2 + h_2^2 = \left(\frac{\beta-\alpha}{2}\right)^2$$

or

$$\left(h_1 + \frac{3}{5}\right)^2 + h_2^2 = \left(\frac{2}{5}\right)^2. \qquad \text{......34)}$$

By the extended Fukuma and Matsubara's criterion, the critical ellipse as given in eqns. (30) and (31) is

$$-\frac{2}{5} - \frac{1}{5\pi}(2\phi - \sin 2\phi) \pm j\frac{1}{5\pi}(1 - \cos 2\phi)$$

$$(0 \leqslant \phi \leqslant \pi). \quad \text{......(35)}$$

Equations (34) and (35) are plotted in Fig. 8. Also plotted in Fig. 8 are the loci of $H(j\omega)$ for various K values.

$$H(j\omega) = \frac{j\omega(1+j\omega)}{K}. \qquad \text{......(36)}$$

It is seen that the circle criterion assures the non-existence of jump resonance if the gain of the linear plant $G(s)$ lies in the range

$$0 < K < 15$$

while the extended Fukuma and Matsubara's criterion assures the non-existence of jump resonance for the range of gain of $G(s)$

$$0 < K < 40.$$

The system was simulated on an analogue computer as arranged schematically in Fig. 9. The input applied to this simulated system was a sine-wave having an amplitude of 1 V. The closed-loop gain of the system was

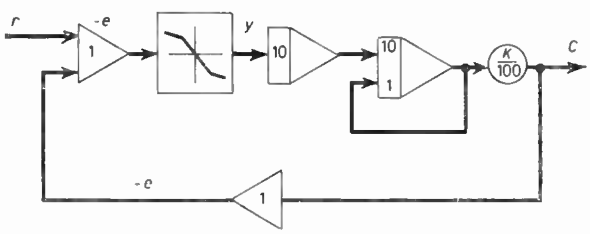
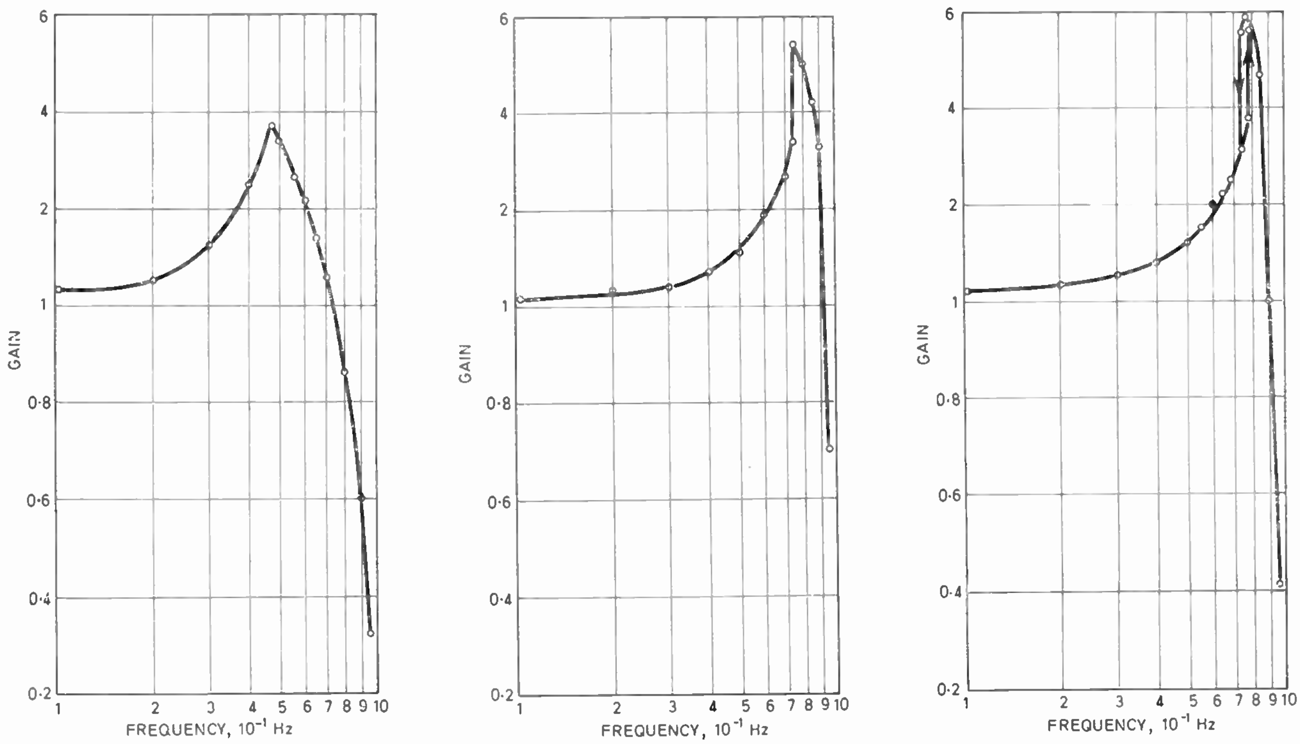


Fig. 9. Analogue simulation of a feedback system.

recorded for each incremental change of frequency and was found free from jump resonance for $K = 15$ and 40 as dictated by the two criteria respectively. The measurement was carried out for various K values to search for jump resonance. Laboratory observation revealed that for $K = 45$, discontinuities in the gain characteristic occur at $f = 0{\cdot}75$ Hz and $f = 0{\cdot}77$ Hz. The frequency responses for $K = 15$, 40 and 45 are shown in Fig. 10



(a) $K = 15$ (b) $K = 40$ (c) $K = 45$

Fig. 10. Closed-loop frequency responses for various K values.

7. Conclusion

Fukuma and Matsubara obtained the boundary curves inside which jump resonance can occur by normalizing the non-linearity characteristic. Their discussion has been extended by permitting non-linearities bounded in a sector based on the concept of circle criterion. For the conditions which assures the non-existence of jump resonance it shows that the extended Fukuma and Matsubara's work is less conservative than the circle criterion. Nevertheless, one should not jump to the conclusion that the circle criterion is less useful. It must be pointed out that the Fukuma and Matsubara's criterion was developed through the describing function method. The systems to which this criterion may be applied represent a rather restrictive sub-class of the class of systems to which circle criterion applies. For this sub-class of systems the circle criterion is expected to be more conservative. The choice of a certain criterion lies in the conditions of the problem at hand. However, the sector-bounded-non-linearity concept is very useful in the system design particularly when it is more desirable to alter the non-linearity than to compensate the linear part of the system. The stability conditions dictated by the extended Fukuma and Matsubara's criterion are dependent upon the sector which bounds the non-linearity and not upon the non-linearity characteristic. Thus it represents a significant improvement of the existing theory.

8. References and Bibliography

1. West, J. C., Douce, J. L. and Livesley, R. K., 'The dual-input describing function and its use in the analysis of nonlinear feedback systems', *Proc. Instn Elect. Engrs*, **103**, Pt. B, pp. 463–74, 1956.
2. Hatanaka, H., 'The frequency responses and jump resonance phenomena of nonlinear feedback control systems', *Trans. Amer. Soc. Mech. Engrs, J. Basic Engineering*, Ser. D, **85**, pp. 236–42, June 1963.
3. Fukuma, A. and Matsubara, M., 'Jump resonance criteria of nonlinear control system', *Trans. I.E.E.E. on Automatic Control*, **AC-11**, No. 4, pp. 699–707, October 1966.
4. Zames, G., 'On the input-output stability of time-varying nonlinear feedback systems'—Part I: conditions derived using concepts of loop gain, conicity and positivity', *Trans. Inst. Elect. Electronics Engrs on Automatic Control*, **AC-11**, pp. 228–39, April 1966.
5. Zames, G., 'On the input-output stability of time-varying nonlinear feedback systems—Part II: Conditions involving circles in the frequency plane and sector nonlinearities' *Trans., I.E.E.E.*, **AC-11**, pp. 465–76, July 1966.
6. Sandberg, I. W., 'A frequency-domain condition for the stability of feedback systems containing a single time-varying nonlinear element', *Bell Syst. Tech. J.*, 43, pp. 1601–8, July 1964.
7. Murphy, G. J., 'A frequency-domain stability chart for nonlinear feedback systems', *Trans. I.E.E.E.*, **AC-12**, pp. 740–3, December 1967.
8. Moore, J. B., 'A circle criterion generalization for relative stability', *Trans. I.E.E.E.*, **AC-13**, pp. 127–8, February 1968.
9. Narendra, K. S., 'A geometrical criterion for the stability of certain nonlinear nonautonomous systems', *Trans. I.E.E.E. on Circuit Theory*, **CT-11**, pp. 406–8, September 1964.
10. Gibson, J. E., 'Nonlinear Automatic Control' (McGraw Hill, New York, 1963).
11. Levinson, E., 'Some saturation phenomena in servomechanisms with emphasis on the tachometer stabilized system', *Trans. Amer. Inst. Elect. Engrs*, **72**, Pt. II, pp. 1–9, 1953.

9. Appendix: Transformation of the Critical Region of the Circle Criterion in the Polar Plane into the Inverse Polar Plane

Since the polar plane and the inverse polar plane are complex, let the points in the polar plane and the inverse polar plane in rectangular co-ordinates be represented respectively by

$$Z = X + jY \qquad \ldots\ldots(37)$$

and

$$W = U + jV. \qquad \ldots\ldots(38)$$

Any point Z may be transformed into a point W in a one to one correspondence by the transformation

$$W = \frac{1}{Z} \qquad \ldots\ldots(39)$$

except for $Z = 0$ which has no image. This transformation maps circles and lines into circles and lines.

Consider the following equation which represents any circle or line in the polar plane

$$A(X^2 + Y^2) + BX + C + D = 0 \qquad \ldots\ldots(40)$$

where A, B, C and D are real numbers.

From eqns. (37), (38) and (39)

$$X = \frac{U}{U^2 + V^2}$$

$$Y = \frac{-V}{U^2 + V^2} \qquad \ldots\ldots(41)$$

$$U = \frac{X}{X^2 + Y^2}$$

$$V = \frac{-Y}{X^2 + Y^2}. \qquad \ldots\ldots(42)$$

Therefore, eqn. (40) becomes by the transformation

$$A\left[\frac{U^2}{(U^2 + V^2)^2} + \frac{V^2}{(U^2 \div V^2)^2}\right] + B\frac{V}{U^2 + V^2} - \\ - \frac{C}{U^2 + V^2} + D = 0$$

or

$$D(U^2 + V^2) + BU - CV + A = 0. \qquad \ldots\ldots(43)$$

Now, consider the three critical regions dictated by the circle criterion.

Case (i) ($\alpha > 0$)

The critical region shown in Fig. 11(a) is given by

$$(X + X_0)^2 + Y^2 = r_1^2$$

where

$$X_0 = \frac{1}{2}\left(\frac{1}{\alpha} + \frac{1}{\beta}\right)$$

$$r_1 = \frac{1}{2}\left(\frac{1}{\alpha} - \frac{1}{\beta}\right). \qquad \ldots\ldots(44)$$

Then from eqn. (40), it is seen that

$$A = 1$$

$$B = \left(\frac{1}{\alpha}+\frac{1}{\beta}\right)$$

$$C = 0$$

$$D = \frac{1}{4}\left[\left(\frac{1}{\alpha}+\frac{1}{\beta}\right)^2-\left(\frac{1}{\alpha}-\frac{1}{\beta}\right)^2\right] = \frac{1}{\alpha\beta}. \qquad \ldots\ldots(45)$$

Equation (43) becomes

$$U^2+V^2+(\beta+\alpha)U+\alpha\beta = 0. \qquad \ldots\ldots(46)$$

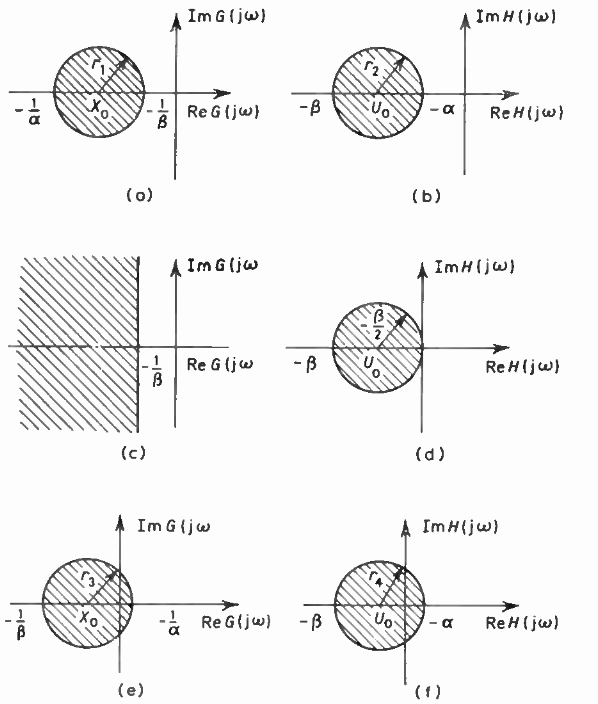


Fig. 11. Transformation of critical regions of circle criterion in the polar plane into the inverse polar plane.

Equation (46) may be written as

$$(U+U_0)^2+V^2 = r_2^2$$

where

$$U_0 = \tfrac{1}{2}(\beta+\alpha)$$

$$r_2 = \tfrac{1}{2}(\beta-\alpha). \qquad \ldots\ldots(47)$$

The transformation of the critical region in the polar plane into the inverse polar plane is shown in Fig. 11(b).

Case (ii) ($\alpha = 0$)

The critical region shown in Fig. 11(c) is given by

$$X = \frac{-1}{\beta}. \qquad \ldots\ldots(48)$$

Then from eqn. (40),

$$A = 0$$

$$B = 1$$

$$C = 0$$

$$D = \frac{1}{\beta}.$$

Equation (43) becomes

$$U^2+V^2+bU = 0. \qquad \ldots\ldots(49)$$

It may be written as

$$\left(U+\frac{\beta}{2}\right)^2+V^2 = \left(\frac{\beta}{2}\right)^2. \qquad \ldots\ldots(50)$$

The transformation is shown in Fig. 11(d).

Case (iii) ($\alpha < 0$)

The critical region shown in Fig. 13(e) is given by

$$(X-X_0)^2+Y^2 = r_3^2$$

where

$$X_0 = \frac{1}{2}\left(\frac{1}{\alpha}+\frac{1}{\beta}\right)$$

$$r_3 = \frac{1}{2}\left(\frac{1}{\beta}-\frac{1}{\alpha}\right). \qquad \ldots\ldots(51)$$

From eqn. (40)

$$A = 1$$

$$B = -\left(\frac{1}{\alpha}+\frac{1}{\beta}\right)$$

$$C = 0$$

$$D = \frac{1}{4}\left[\left(\frac{1}{\alpha}+\frac{1}{\beta}\right)^2-\left(\frac{1}{\beta}-\frac{1}{\alpha}\right)^2\right] = \frac{1}{\alpha\beta}. \qquad \ldots\ldots(52)$$

Thus eqn. (43) becomes

$$(U+U_0)^2+V^2 = r_4^2$$

where

$$U_0 = \tfrac{1}{2}(\beta+\alpha)$$

$$r_4 = \tfrac{1}{2}(\beta-\alpha). \qquad \ldots\ldots(53)$$

The transformation is given in Fig. 11(f).

Manuscript received by the Institution on 19th May 1972. (Paper No. 1489/IC69.)