# THE
# RADIO AND ELECTRONIC
# ENGINEER

## The Journal of the Institution of Electronic and Radio Engineers

## Improving Communication

THE INSTITUTION'S first *Journal* was published in 1926 primarily as a means of communication between members and to report papers read at Institution meetings. As the science and technology of radio emerged and ultimately developed into the wider field of electronics, the *Journal* grew into a permanent record of progress in electronic science and technology. As such it has attracted a world-wide readership and inclusion in international references which record initial work in the development of electronics.

Concentration on learned society activities gradually excluded from *The Radio and Electronic Engineer* news about members and their meetings, industrial activity and other topics of interest to the practising professional radio and electronic engineer. Indeed, this tendency to concentrate on highly specialized topics led to the following stricture from Mr. W. E. Miller in his Presidential Address to the Institution:*

'. . . However, the Institution is not completely fulfilling one of its major objects, the dissemination of knowledge, if it only provides papers and articles of such a high level that only a percentage of members are able fully to profit from them. . . .'

Mr. Miller then suggested

'. . . that in addition to the specialized papers which we must still continue to attract, we should also regularly publish a number of survey papers on the latest developments in various branches of radio-communications and electronics. . . .'

These constructive ideas led to the production of various Institution *Proceedings* individually designed to suit the members of the I.E.R.E. in many countries. Commencing with Canadian *Proceedings* in 1961, similar publications followed for members in India and Great Britain in 1963, and subsequently in France, Australasia and Israel. This arrangement has, however, certain drawbacks, not the least being the restriction of survey papers to national *Proceedings* and limitation of industrial and other news in *The Radio and Electronic Engineer*. Whilst the professional engineer must make the best use of the time he has available for keeping informed of new work, industrial news from far and wide and of Institution activities, it is a fact that the engineer's need to read an increasing number of publications has become an occupational hazard!

These are, therefore, some of the reasons which have prompted the decision to discontinue the *British Proceedings* although many of the features of that publication will in future be included in *The Radio and Electronic Engineer*. The opportunity is also being taken to make some typographical changes to the layout of the *Journal*, the most appropriate being a switch to the international A4 size (297 mm × 210 mm) which is about 17 mm deeper and 6 mm narrower than the present size. Resultant economies in cost—particularly in postage—will help to increase the publication of material of interest to a world-wide readership.

Thus *The Radio and Electronic Engineer* will continue especially to feature papers describing original research and development, and particularly encourage those papers which predict future technological trends. Special attention will also be given to securing reviews or survey papers on the lines already mentioned. Included in these will, of course, be the engineering design and development papers which describe problems and their solutions, and particularly emphasize the means whereby economic solutions have been achieved.

It is also intended to invite interpretive papers which present newly-developed theories and techniques in a form having immediate value in establishing design criteria, as well as 'tutorial' papers explaining recently established principles and techniques in various specialized fields. Finally, the very important 'short'

---

* *J. Brit. I.R.E.*, **13**, No. 1, p. 5, January 1953.

contribution will be encouraged. Typically, these are shorter than 2000 words in length and deal with device technology, circuit techniques, experimental methods and computer programs for electronic engineering. These communications will be given priority in publication.

Cessation of the British *Proceedings* will also eliminate duplication of reading matter on, for example, university research activities and the social implications of the continuous application of electronics to everyday life. This is the role of the National Electronics Council which since 1965 has published its own 'Review'. This publication continues to provide a bridge between the engineer and management—particularly the 'user' management. Problems of application or the development of new ideas as well as its regular feature on research in the universities, all contribute to making the *National Electronics Review* a useful reference to all readers of *The Radio and Electronic Engineer*. Indeed, an agreement has been reached with N.E.C. to enable all members and subscribers to the Institution's *Journal* to have a preferential subscription rate for the *National Electronics Review*.

The momentum of scientific change is increasing and innovation demands new approaches and constantly improving communication techniques. Nowhere, in the area of applied science, does this apply more than in the fast-moving field of electronics. The change in publishing policy enables all contributors to *The Radio and Electronic Engineer* to provide helpful 'feedback' in achieving an 'output' which is in accord with the needs of the end of the twentieth century.                                                                 G. D. C.

# INSTITUTION NOTICES

### Combined Programme of Meetings

The booklet containing details of Institution meetings in London, and meetings of the East Anglian, Kent and Thames Valley Sections for the second half of the current session (February to May 1971) is being sent to all members in South Eastern England with this issue of the *Journal*.

As previously, the booklet presents a combined programme of the above I.E.R.E. meetings together with the London area meetings of the Institution of Electrical Engineers. Members resident in other parts of the British Isles (or in other countries) who are able to attend any of these meetings may obtain a copy of the booklet on application.

Both Institutions wish to re-emphasize that all of their ordinary meetings for which no fee is charged are open to all members of the other Institution. This mutual arrangement extends to members of the Institute of Physics and the Physical Society and of the Institute of Mathematics and its Applications; meetings of these two bodies are also open to I.E.R.E. members without further formality under the agreement reached by the Standing Committee of Kindred Societies.

### Index to Volume 40

This issue of *The Radio and Electronic Engineer* is the last in the present volume, Volume 40. An Index will be sent out with the February issue.

Members who wish to have their *Journals* bound are asked *not* to send them to the Institution for this purpose before the Index is published.

### Fourth U.K.A.C. Control Convention

The fourth United Kingdom Automation Council Control Convention on 'Multivariable Control System Design and Applications' will be held at the University of Manchester, Owens Park, Manchester on 1st to 3rd September 1971. The Convention is being organized by the Institution of Electrical Engineers on behalf of the Technology Group of U.K.A.C. of which the I.E.R.E. is a member.

The increasing use of digital computers in control applications has led to a growing interest in problems of multivariable systems control. The main emphasis of the Convention will be on the development of design techniques for multivariable control systems and simulation of system interactions with particular reference to their applications to industrial processes. This will include:

Application studies of the use of specific design techniques on industrial plant.

Development of design techniques with specific reference to their underlying theory and use.

Analyses of multivariable control problems arising in industrial plant.

Comparative studies of different types of multivariable controllers, and practical surveys of the behaviour of existing industrial controllers in multivariable situations.

Intending authors should submit a 250 word synopsis to the Convention Secretariat without delay. Complete manuscripts will be required by 24th March 1971.

Further details and registration forms will be available in due course from the U.K.A.C. 1971 Convention Secretariat, Savoy Place, London WC2R OBL.

# Predicting Servomechanism Dynamic Performance Variation from Limited Production Test Data

By

**P. A. PAYNE,** (Graduate)†

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

and

**K. J. BAKER,** B.Eng.(Tech.)†

A probabilistic model describing the considerable variation in dynamic performance of a large family of electro-hydraulic servomechanisms is presented. Two fundamentally different techniques are employed to obtain the parameters of the model, both techniques giving very similar results. A third-order dominant transfer function (derived from the production test data) together with high frequency complex poles and a high frequency lag (both factors derived from field testing) form a sixth-order system. Variable parameters for the sixth-order system form the probabilistic model and are presented as a table for use in Monte Carlo simulation. The model is then used to derive experimentally verified production test limits in a new test domain, equivalent to the original frequency domain gates, showing how the use of a novel filtering technique, recently described elsewhere, may be employed to simplify the setting of test limits. The model is also shown to be of use in the prediction of dynamic errors in many domains (whilst only requiring data in one domain) and the initial production test data are shown to predict accurately the dynamic error variation for the complete family.

## List of Principal Symbols

| | |
|---|---|
| $V_o$ | system output voltage |
| $V_i$ | system input voltage |
| $s$ | Laplace operator |
| $a, A$ $b, B$ | transfer function coefficients |
| $\omega_0$ | time scaling factor |
| $f$ | frequency (Hz) |
| $\phi$ | phase lag (degrees) |
| $\omega$ | angular frequency (radians/second) |
| $f_{\phi 1, 2, 3}$ | frequency at which phase lag is $\phi_1, \phi_2, \phi_3$ |
| $T(j\omega) = \dfrac{N(j\omega)}{D(j\omega)}$ | fitted transfer function |
| $e(j\omega)$ | error between fitted transfer function and measured data |
| $T_1$ | time constant |
| $\zeta$ | damping factor |
| $\omega_n$ | undamped natural frequency |
| $\sigma$ | standard deviation |
| $\hbar_{t_i}$ | estimated impulse response at time $t_i$ |

## 1. Introduction

In recent years the authors and their colleagues have been investigating the application of various dynamic test techniques to production, acceptance and field testing of servomechanisms. Many problems have arisen in this work which are not solvable simply by considering only the nominal system created by the designer. Typical problems are:

† Department of Mechanical Engineering and Engineering Production, University of Wales Institute of Science and Technology, Cathays Park, Cardiff CF1 3NU.

(a) If the dynamic acceptance test is formulated in one domain how are equivalent tolerances best set in a different domain perhaps implicit in a new method of test?

(b) Can doubts on the validity of a test performed in one domain only be resolved by providing a model of the system applicable in all domains of interest as distinct from that domain of interest to the designer?

(c) What confidence can be placed in the response of the system to actual inputs based on standard acceptance tests such as harmonic testing?

(d) In a noisy environment is the variability between systems of the same family likely to be large or small compared with the uncertainties in the measurements themselves?

An early field study on a small sample of instrument servos revealed wide performance variation between acceptable servos and significant measurement uncertainties in the presence of noise.[1] Furthermore, if a toleranced performance specification used by the designer is then theoretically translated without regard for manufacturing feasibility into an envelope of possible responses in a test domain of particular interest,[2] it is possible unconditionally to tolerance this envelope in such a way that a system failing to meet the designer's specification is accepted.[3] This ambiguity is more likely to be avoided if the variation in the performance of real systems is better understood.

## 2. Concepts and Application of the Probabilistic Model

On the evidence of reference 1 the authors were granted access to the production test records of a large family of electro-hydraulic servomechanisms. These servomechanisms are essentially linear under

the stipulated test conditions, and are theoretically of eighth order. In practice a sixth-order model is adequate for performance description over a wide frequency range. However, the production test records are limited in range and have therefore been used to establish the dominant transfer functions of all servomechanisms in the family, firstly using a sensitivity method, and secondly using a least squares error-curve fit method of general applicability. From these predictions a sample of thirty transfer functions was chosen to form the basic probabilistic model which is the term used for a statistical model describing the performance variation between one servomechanism and another.

A limited number of field experiments have been performed in order to establish the variation in high frequency complex poles and an equivalent lag representing all other terms. The resultant probabilistic model is of sixth order and is made available in tabular form ready for Monte Carlo simulations. Philosophically, the probabilistic model is related to the design, development, and production phases in the manner of Fig. 1. Once established, the probabilistic model is used to

(i) establish simpler frequency domain production tests,

(ii) tolerance system impulse response equivalent to the original frequency domain limits,

(iii) show improvement in impulse domain tolerancing achieved by filtering,

(iv) correlate system dynamic errors with frequency domain properties.[4]

An outstanding feature of the model is the manner in which early production test results may be used to predict the behaviour of the family indicating that, with updating, the model may be used as an accurate monitor of system performance in all domains, during the production phase.

## 3. Modelling the Fast Response Electro-Hydraulic Servomechanism

The electrohydraulic servomechanism made available for this investigation is a typical fast response device as used for machine tool drives, transfer machinery and weapon systems. A simplified block
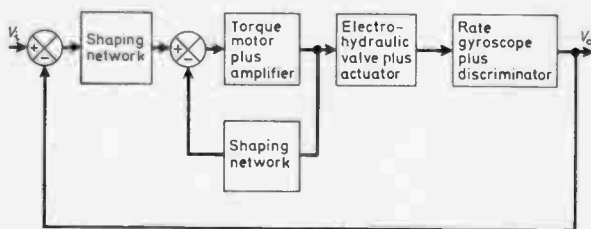


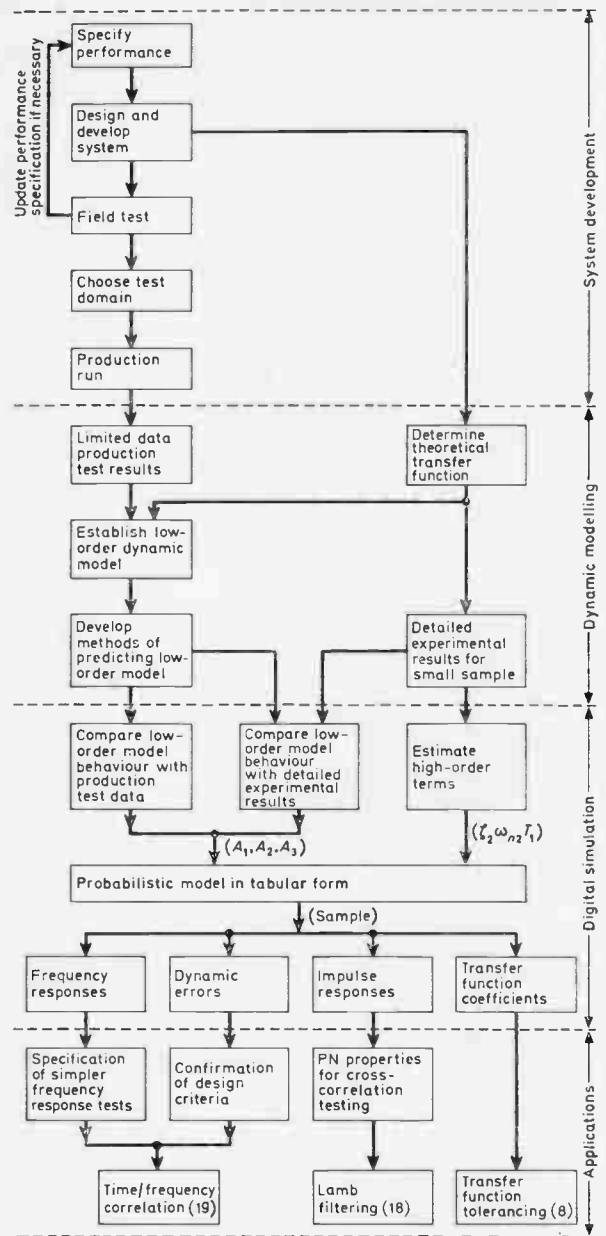Fig. 2. Block diagram of electro-hydraulic servo.



Fig. 1. Flow diagram for whole project.

diagram is shown in Fig. 2, and the open-loop Bode plot with all components at nominal values is shown in Fig. 3. There is a secondary resonance of $\omega_n = 300$ radians/second and $\zeta = 0.2$ due to oil resilience in the hydraulic drive part of the servo. From theoretical considerations, the nominal system transfer function is

$$\frac{V_o}{V_i}(s) = \frac{\sum_{j=0}^{j=1} b_j s^j}{\sum_{j=0}^{j=8} a_j s^i} \qquad \ldots\ldots(1)$$

or, in factored form,

$$\frac{V_o}{V_i}(s) = \frac{(1+s/18.5)}{\left(1+\dfrac{2\times0.61}{26.6}s+\dfrac{s^2}{26.6^2}\right)\left(1+\dfrac{2\times0.24}{292}s+\dfrac{s^2}{292^2}\right)\left(1+\dfrac{s}{256}\right)\left(1+\dfrac{s}{500}\right)\left(1+\dfrac{s}{770}\right)\left(1+\dfrac{s}{1730}\right)} \quad\dots\dots(2)$$
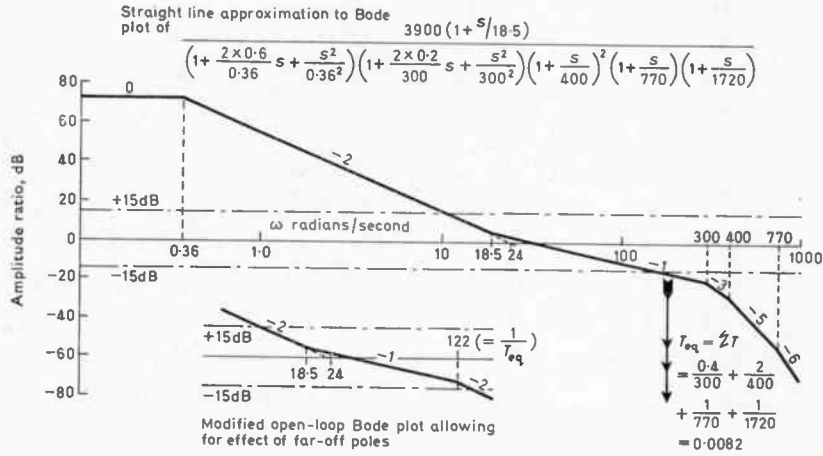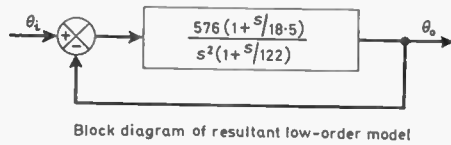
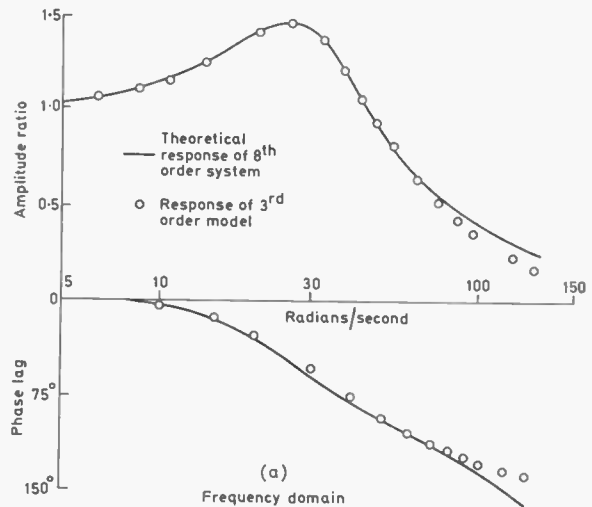**Fig. 3.** Derivation of low-order model from open loop Bode plot of high-order system.

Block diagram of resultant low-order model

**Fig. 4.** Comparison of I.3 model with I.8 system (below).

It is possible using the method of Towill and Mehdi,[5] to establish a third-order model of the system which adequately predicts system performance over a limited frequency range, and which, in addition, adequately predicts dynamic errors, thus confirming the Biernson argument[6] that dynamic errors are dependent only on behaviour in the 0 dB region of the open-loop Bode plot. Application of the Towill and Mehdi method, the working for which is shown in Fig. 3, gives for nominal conditions.

$$\frac{V_o}{V_i}(s) = \frac{1+5.38\left(\dfrac{s}{100}\right)}{1+5.38\left(\dfrac{s}{100}\right)+17.4\left(\dfrac{s}{100}\right)^2+14.2\left(\dfrac{s}{100}\right)^3} \quad\dots\dots(3)$$

The model is zero velocity lag, and physical reasoning[7] suggests that any dynamic model of this system must be so. Frequency responses for the full system and the third-order model are compared in Fig. 4. Up to a frequency of twice bandwidth, agreement in amplitude and phase is excellent. The equivalent dominant pole-zero array is also shown.

In order to predict variations in the impulse response or in the frequency response over a wide frequency range the model must be increased in order. It has already been shown[8] that a (I.6) model is adequate for describing performance in the frequency domain over a very wide range, including the secondary resonance. Therefore the simplest adequate frequency domain model for all purposes is defined as

$$\frac{V_o}{V_i}(s) = \frac{1 + \left(A_1 + T_1 + \frac{2\zeta_2}{\omega_{n_2}}\right)s}{(1 + A_1 s + A_2 s^2 + A_3 s^3)} \\ \times (1 + T_1 s)\left(1 + \frac{2\zeta_2}{\omega_{n_2}}s + \frac{s^2}{\omega_{n_2}^2}\right) \quad \ldots\ldots(4)$$

where the zero velocity lag property has been retained.

The coefficients $A_1$, $A_2$, and $A_3$ dominate the low-frequency response (and dynamic errors), whilst the additional lag and the secondary resonance shape the high-frequency response appropriately.

In the time domain, the interaction between the terms in equation (4) leads to an apparent attenuation of the secondary mode. This can be avoided and a better fit to the impulse domain obtained[9] by utilizing the time delay $e^{-T_1 s}$ giving

$$\frac{V_o}{V_i}(s) = \frac{\left[1 + \left(A_1 + T_1 + \frac{2\zeta_2}{\omega_{n_2}}\right)s\right]e^{-T_1 s}}{(1 + A_1 s + A_2 s^2 + A_3 s^3)} \quad \ldots\ldots(5) \\ \times \left(1 + \frac{2\zeta_2}{\omega_{n_2}}s + \frac{s^2}{\omega_{n_2}^2}\right)$$

Basically, the production test records (low-frequency information only, six data points only, but many systems) are used to establish the variation in $A_1$, $A_2$, and $A_3$, whilst limited field tests (wide frequency range, many data points, few systems) are used to establish the variation in $T_1$, $\zeta_2$, and $\omega_{n_2}$.

## 4. Available Production Test Records

The production test records were accumulated over a period of several years, the total number manufactured being regarded as infinite for statistical purposes. At the time of system development, possible test techniques were restricted to those described by Sumray,[10] and the system user therefore opted for frequency domain testing, with tolerances as shown in Table 1. In the light of Fig. 4 it is evident that this test corresponds to establishing a third-order dominant transfer function valid up to 20 Hz.

Figure 5 shows histograms constructed from the available production test data. Five of the distributions are approximately Gaussian, whilst the sixth (quadrature volts recorded at 90° phase lag) is bimodal. Superimposed, but to a different vertical



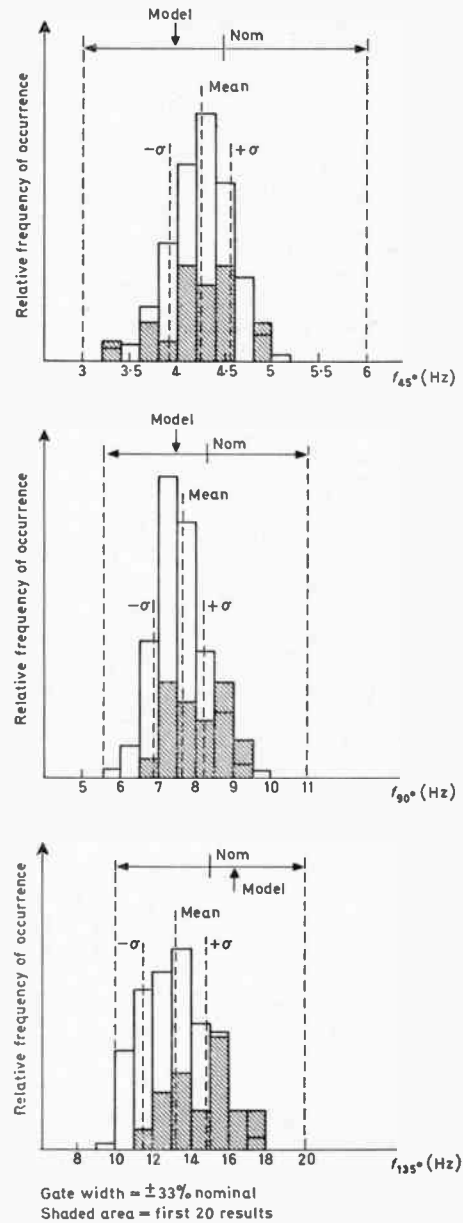Gate width = ±33% nominal
Shaded area = first 20 results

Fig. 5 (a). Production test results.

scale, are the results for the first twenty servos manufactured, and in some cases the ultimate range is

**Table 1.** Production test limits

| Phase lag | Frequency | | Amplitude ratio | |
|---|---|---|---|---|
| | Min. | Max. | Min. | Max. |
| 45° | 3·0 Hz | 6·0 Hz | 1·05 | 2·05 |
| 90° | 5·5 Hz | 11·0 Hz | 0·75 | 1·50 |
| 135° | 10·0 Hz | 20·0 Hz | 0·35 | 0·70 |

Gate = ± 33% nominal
Shaded area = first 20 results
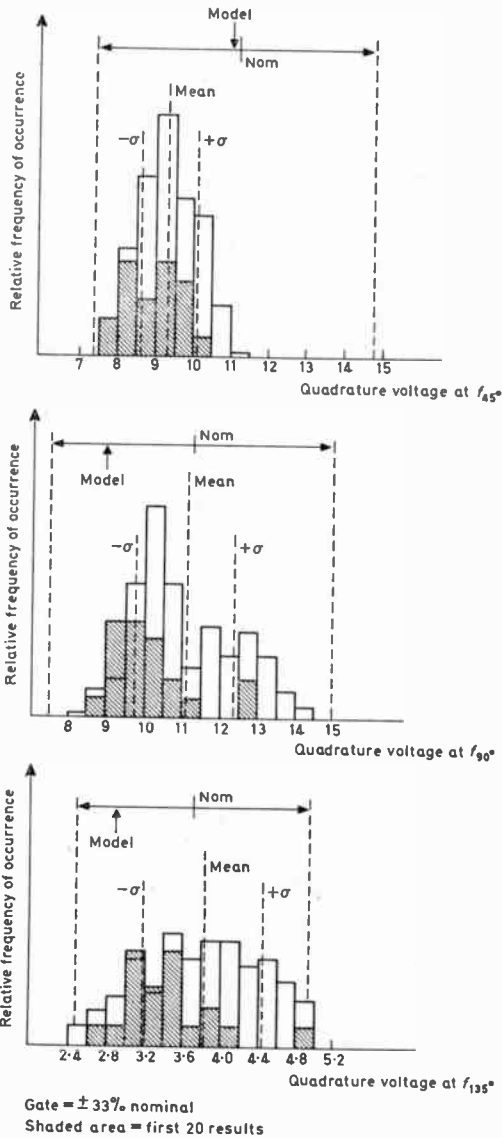
**Fig. 5 (b).** Reduction test limits.

fundamentally different transfer function prediction techniques have been developed in parallel, agreement between results obtained using the two techniques giving confidence to the probabilistic model. The curve fitting method is much more general in application, whilst the sensitivity method often leads to an insight into system variability. The two techniques are described in detail in Appendix 1.

Figure 6 shows the distribution of coefficients $A_1$, $A_2$, $A_3$ obtained from the production run summarized in Fig. 5. The predictions using the sensitivity technique and the least-squares error curve fitting technique are in pleasing agreement as quantified in Table 2, whilst the ranges are in broad agreement with the bounds determined experimentally by varying simulator transfer function coefficients one at a time
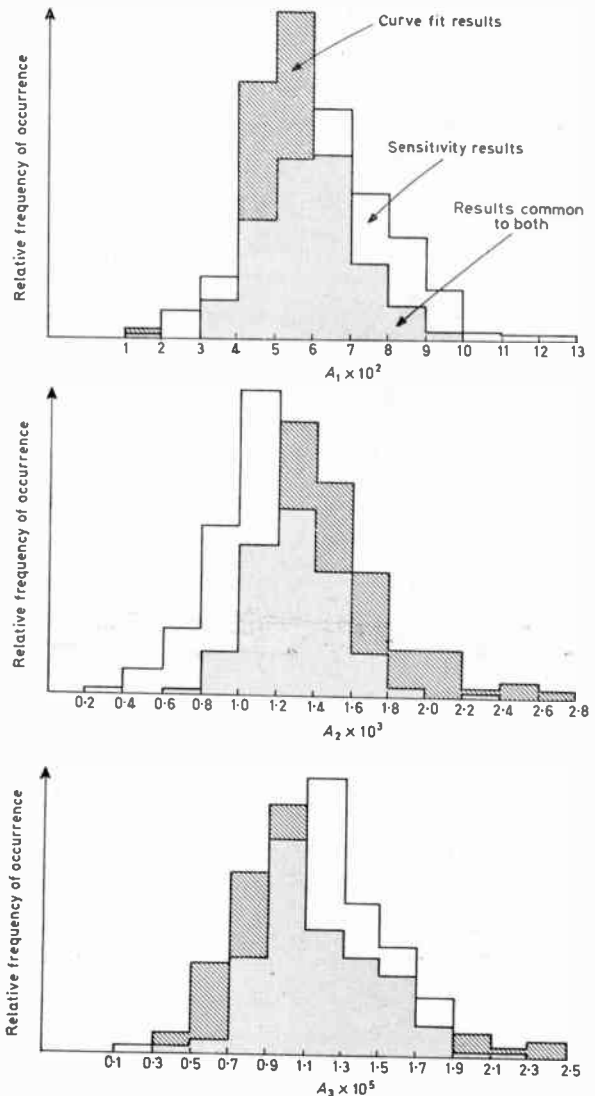


**Fig. 6.** Distribution of coefficients $A_1$, $A_2$, and $A_3$.

already apparent. In two cases (quadrature volts recorded at 45° phase lag, and $f_{45°}$), the experimental ranges for the large family are significantly smaller than the tolerance ranges, thus it is not possible for a practical servo to approach the higher tolerance in either case without falling outside tolerance at other frequencies.

## 5. Variation in the Low-order Model

Due to the limited nature of the production test results, both with regard to frequency range and number of data points, the six pieces of data are best used to estimate the low-order model coefficients $A_1$, $A_2$, $A_3$. To make best use of the results two

**Table 2.** Comparison of prediction of transfer function coefficients

| Coefficient | Curve fit model | | | Sensitivity model | | |
|---|---|---|---|---|---|---|
| | Mean | $\sigma$ | $100\sigma/$ mean | Mean | $\sigma$ | $100\sigma/$ mean |
| $A_1 \times 10^2$ | 5·61 | 1·24 | 22% | 6·40 | 1·68 | 26% |
| $A_2 \times 10^4$ | 14·60 | 3·74 | 26% | 11·40 | 2·75 | 24% |
| $A_3 \times 10^6$ | 11·12 | 3·96 | 35% | 12·15 | 2·78 | 23% |

until the tolerances of Table 1 were exceeded. All distributions except that of $A_3$ obtained from the curve fitting technique are approximately Gaussian. The skew distribution for $A_3$ (curve fit) accounts for the difference between the two resultant standard deviations for $A_3$. The ranges, however, are in close agreement.



Fig. 7. Distribution of system poles and zeros.

Each dominant transfer function predicted by the curve fit technique was factorized, the resulting system pole-zero arrays being shown in Fig. 7 which should be compared with the general sketch of Truxal.[11] It is thus seen that the tolerancing procedure of Table 1 results in about 70% of the complex poles having damping ratios between 0·5 and 0·9, and undamped natural frequencies between 28 and 40 radians/second. However, the transfer function coefficients are conditional, i.e. $A_1$, $A_2$, and $A_3$ must form particular patterns. The probabilistic model is therefore compressed into a standard random sample of 30, distributions for which are superimposed on Fig. 7. This sample is subsequently tabulated in Section 8 in a form suitable for Monte Carlo simulations.

## 6. Ability to Predict Performance Spread from Early Production Results

Production test results for the first twenty servos manufactured have already been indicated on Fig. 5. It is of considerable interest to compare the performance spread predicted from these early results with those drawn from the complete family as sampled in the probabilistic model. This comparison is shown in Table 3 together with results obtained from recent field tests. In all cases, for consistency the performance criteria are estimated from the appropriate low-order model, and the all-round agreement must be considered excellent. Since the servos have been in use for several years before being made available for field testing, differences between standard deviations for field tests results and the others may be in part at least attributable to wear and ageing effects.

## 7. Alternative Frequency Domain Test Procedure for Dynamic Error Acceptance Testing

Existing frequency domain go/no-go testing requires that the tester search for frequencies at which the real and quadrature components of the return signal are,

    case (1) equal magnitude, quadrature negative;
    case (2) real component zero, quadrature negative;
    case (3) equal magnitude, real and quadrature negative.

He then compares the measured frequencies and amplitude ratios with the limits given in Table 1.

It would clearly be simpler, and less time-consuming, to choose three fixed test frequencies. Using the probabilistic model described in this paper, Towill and Payne[4] have recently shown how three such test frequencies are readily chosen, the frequencies and new test limits being as shown in Table 4.

When tested against typical faulty systems, this new procedure *fails such systems at least three times in only six data points*, a very high degree of confidence thereby being achieved.
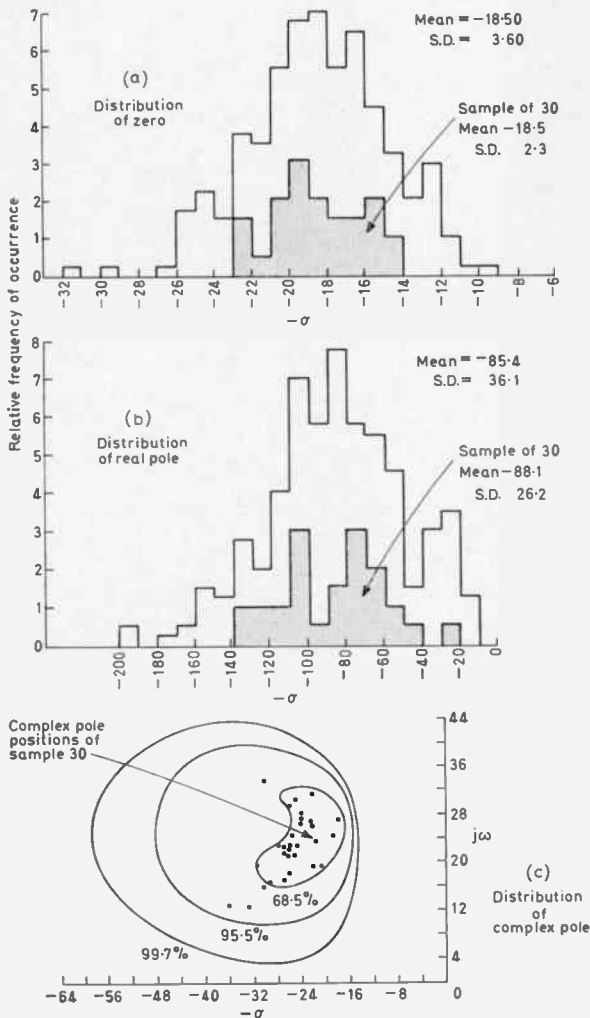
**Table 3.** Comparison of early production test results with those predicted from low-order part of probabilistic model and with recent field tests

| | Initial production | | Probabilistic model (Table 5) | | Field tests | |
|---|---|---|---|---|---|---|
| | Mean | $\frac{\sigma}{\text{mean}} \times 100\%$ | Mean | $\frac{\sigma}{\text{mean}} \times 100\%$ | Mean | $\frac{\sigma}{\text{mean}} \times 100\%$ |
| Bandwidth $\omega_b$ (rad/s) | 71·2 | 5·6 | 70·3 | 6·3 | 65·5 | 9·2 |
| Peak amp. ratio $M_p$ | 1·40 | 6·3 | 1·40 | 5·7 | 1·38 | 4·5 |
| Peak freq. $\omega_p$ (rad/s) | 27·6 | 9·1 | 27·4 | 8·8 | 25·2 | 8·8 |
| Impulse response peak | 35·3 | 6·0 | 34·5 | 6·4 | 33·9 | 7·8 |
| Step percent. overshoot | 29·2 | 15·0 | 29·6 | 13·6 | 27·2 | 9·2 |
| Ramp error peak | 0·019 | 7·3 | 0·019 | 7·4 | 0·019 | 6·6 |
| Sample size | 20 | | 30 | | 18 | |

**Table 4.** Spot frequencies and test limits for new dynamic error test procedure

| Frequency rad/s | Amplitude ratio limits | Phase lag limits |
|---|---|---|
| 18 | 1·10–1·50 | 0 – 30° |
| 50 | 0·70–1·30 | 45°–115° |
| 100 | 0·25–0·60 | 90°–150° |

## 8. Extending the Probabilistic Model to Predict Spread in Other Domains

Pseudo-noise stimuli together with input-output cross-correlation is well known as an impulse response estimator,[12,13] and it has been advanced as a possible contender for the production testing of servomechanisms. We must therefore establish the acceptable spread, equivalent to the frequency domain acceptance pattern, of the impulse response as estimated using cross-correlation.

It has already been established that lightly damped secondary modes greatly influence system impulse response, and that the far-off poles and zeros appear as a pure time delay.[9] Hence for impulse response prediction the variation in $\zeta_2$, $\omega_{n_2}$ and $T_1$ of equation (5) must be estimated, there being no guidance on such variation available from the production test results.

As outlined in Appendix 1, in establishing the dominant transfer function, $T_1$, $\zeta_2$, and $\omega_{n_2}$ were reckoned to sum to a constant contribution to low-frequency behaviour. To estimate $T_1$, $\zeta_2$, and $\omega_{n_2}$,

much more comprehensive tests are required, hence full frequency response tests were carried out on eighteen servos made available from time to time. Results for three typical systems are shown in detail in Fig. 8, the rapid change of amplification and phase due to the secondary complex poles being clearly discernible. Using curve fitting techniques based on standard forms,[14] supplemented where appropriate by the normalized phase slope procedure of Levinge,[15] the limits of $T_1$, $\zeta_2$, $\omega_{n_2}$ have been estimated from the 18 servos as follows:

$$0.0032 \geqslant T_1 \geqslant 0.0048 \quad \text{seconds}$$
$$0.05 \geqslant \zeta_2 \geqslant 0.15$$
$$282 \geqslant \omega_{n_2} \geqslant 346 \quad \text{radians/second}$$

It should be noted that the damping ratios observed for the secondary mode under the standard test conditions are somewhat smaller than those predicted theoretically in equation (2). However, it is well known that damping in hydraulic servos is much affected by oil temperature and coulomb friction, both somewhat variable quantities. Clearly, variations in $\omega_{n_2}$, $\zeta_2$, $T_1$ are not estimated with the same level of confidence as for the dominant transfer function coefficients, but additional systems made available since this analysis have indicated that the estimates are realistic.

Gaussian distributions were assumed for $T_1$, $\zeta_2$, and $\omega_{n_2}$ with $6\sigma$ values equal to the ranges quoted above. For each of the thirty dominant transfer functions randomly chosen from the production run, appropriate values of $T_1$, $\zeta_2$, and $\omega_{n_2}$ were allocated using random number tables. The complete probabilistic model is therefore as shown in Table 5.
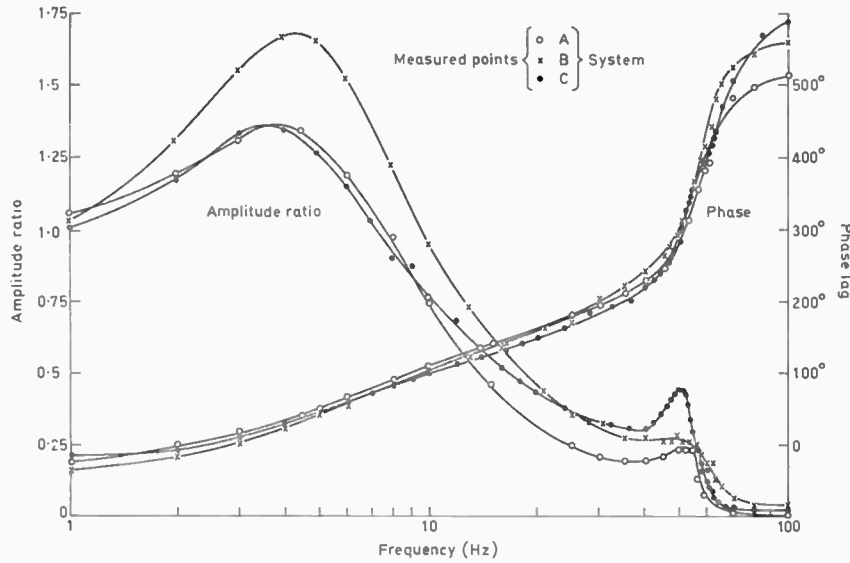
Fig. 8.
Three typical frequency responses.

**Table 5.** Characteristics of the probabilistic model

| | $A_1 \times 10^2$ | $A_2 \times 10^3$ | $A_3 \times 10^6$ | $\omega_{n2}$ (rad/s) | $\zeta_2$ | $T_1$ (seconds) | | $A_1 \times 10^2$ | $A_2 \times 10^1$ | $A_2 \times 10^6$ | $\omega_{n2}$ (rad/s) | $\zeta_2$ | $T_1$ (seconds) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 4·536 | 1·073 | 8·672 | 319 | 0·10 | 0·0044 | 16 | 5·837 | 1·512 | 8·506 | 308 | 0·12 | 0·0043 |
| 2 | 5·526 | 1·158 | 6·960 | 324 | 0·10 | 0·0039 | 17 | 5·109 | 1·229 | 7·319 | 285 | 0·09 | 0·0039 |
| 3 | 6·683 | 1·657 | 13·056 | 309 | 0·10 | 0·0039 | 18 | 5·150 | 1·169 | 7·286 | 294 | 0·11 | 0·0041 |
| 4 | 5·126 | 1·343 | 10·987 | 295 | 0·12 | 0·0043 | 19 | 4·383 | 1·074 | 9·183 | 317 | 0·13 | 0·0045 |
| 5 | 5·542 | 1·404 | 10·839 | 309 | 0·06 | 0·0033 | 20 | 5·201 | 1·320 | 11·324 | 328 | 0·11 | 0·0041 |
| 6 | 6·137 | 1·412 | 7·732 | 314 | 0·10 | 0·0039 | 21 | 4·817 | 1·368 | 8·410 | 331 | 0·09 | 0·0039 |
| 7 | 6·495 | 1·430 | 9·391 | 325 | 0·07 | 0·0036 | 22 | 6·174 | 1·506 | 9·707 | 323 | 0·11 | 0·0042 |
| 8 | 5·144 | 1·346 | 8·882 | 317 | 0·08 | 0·0037 | 23 | 4·792 | 1·123 | 9·401 | 309 | 0·08 | 0·0048 |
| 9 | 6·110 | 1·726 | 11·410 | 327 | 0·09 | 0·0039 | 24 | 6·365 | 1·493 | 16·334 | 320 | 0·10 | 0·0040 |
| 10 | 4·965 | 1·267 | 10·410 | 315 | 0·11 | 0·0042 | 25 | 6·326 | 1·645 | 15·830 | 303 | 0·08 | 0·0037 |
| 11 | 5·065 | 1·175 | 6·960 | 319 | 0·12 | 0·0044 | 26 | 5·725 | 1·402 | 10·234 | 316 | 0·11 | 0·0042 |
| 12 | 5·417 | 1·252 | 7·797 | 318 | 0·08 | 0·0037 | 27 | 5·475 | 1·330 | 9·843 | 300 | 0·11 | 0·0042 |
| 13 | 4·616 | 1·118 | 8·033 | 308 | 0·09 | 0·0039 | 28 | 6·717 | 1·649 | 13·218 | 329 | 0·09 | 0·0038 |
| 14 | 4·767 | 1·420 | 12·740 | 293 | 0·16 | 0·0050 | 29 | 6·562 | 1·494 | 11·306 | 302 | 0·09 | 0·0039 |
| 15 | 5·555 | 1·449 | 14·594 | 319 | 0·07 | 0·0035 | 30 | 4·880 | 1·130 | 6·247 | 319 | 0·10 | 0·0040 |

## 9. Predicting Performance Variation in Other Domains

The full model is readily employed to predict the spread in system response in any domain corresponding to the original frequency domain specification. Figure 9 shows the impulse responses of the probabilistic model of sample size 30, obtained by digital solution of equation (5). Provided the autocorrelation function of the pseudo-noise stimulus suitably matches the system, Fig. 9 is a guide to the variation in impulse estimates measured using a cross-correlator.

In Fig. 9 it is readily observed that the secondary mode has in some instances an amplitude approaching 50% of the 'dominant' part of the response, and this
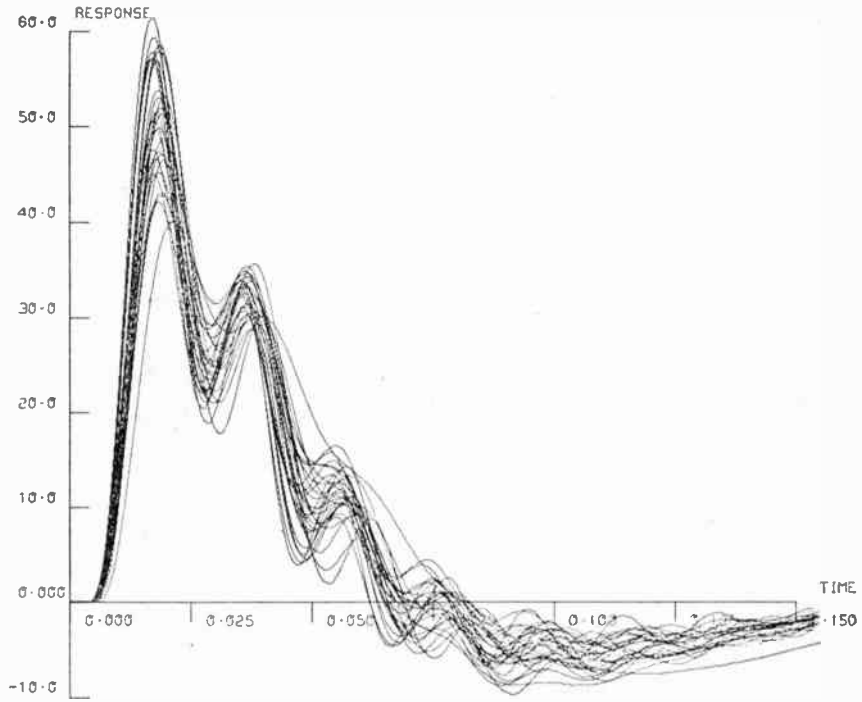
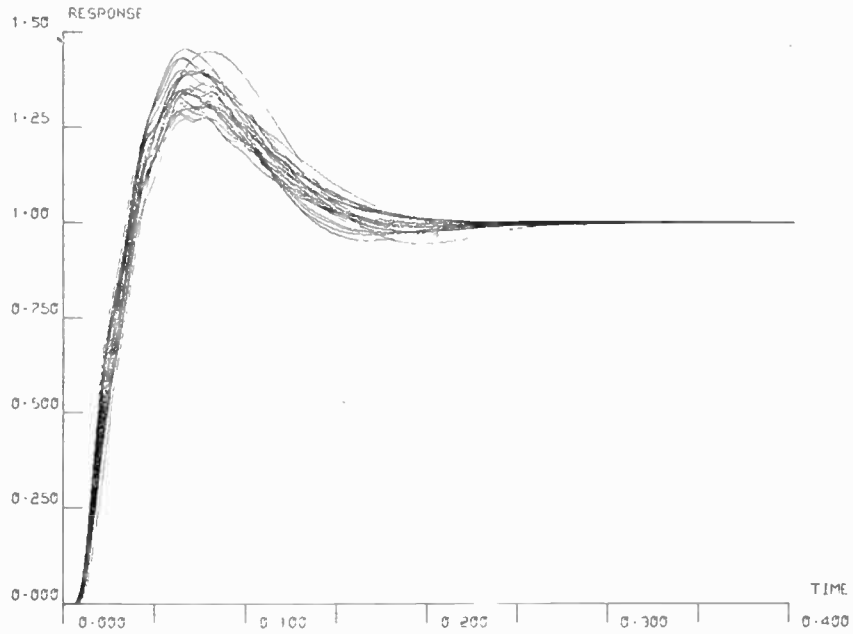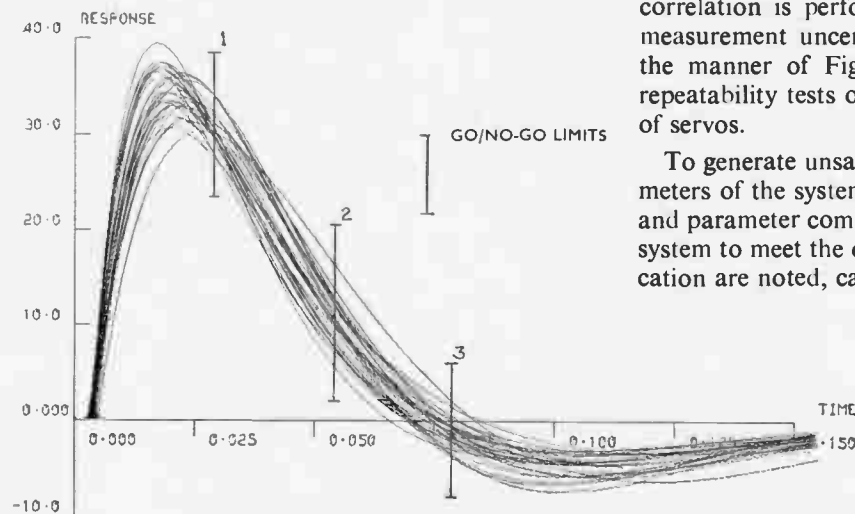Fig. 9. Spread in unit impulse response (Time in seconds).



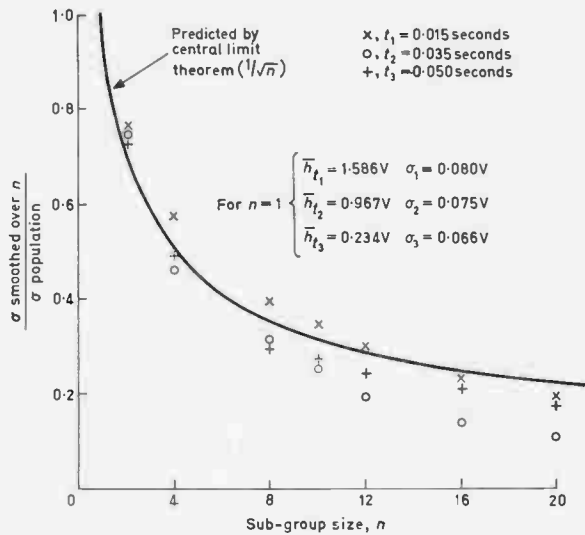Fig. 10. Spread in unit step response (Time in seconds).

secondary mode is therefore of considerable significance in the ability of the system to reject noise and disturbances. However, command signals can usually be represented by a series of steps, ramps and parabolas.[16] Reference to the spread in step response shown in Fig. 10 indicates that the secondary mode is attenuated to an extent where it is virtually undetectable, and will be even further filtered in the ramp and parabolic responses. The secondary mode is therefore of little consequence in assessing dynamic errors following the command signal, which is predictable on theoretical grounds,[17] since modal amplitudes of secondary resonances relative to primary modes are reduced by a factor $\lambda$ every time the excitation stimulus is integrated, where $\lambda$ is the ratio of secondary pole undamped natural frequency to primary pole undamped natural frequency. In this case, $\lambda$ is of the order of 10, so that the secondary mode amplitude in the step response is reduced to the order of 5%.

## 10. Filtering the Impulse Response

If only dynamic errors to steps, ramps and parabolas are important in system go/no-go testing, the secondary mode effects observable in the system impulse response serve only to complicate the judgement without adding to useful information. For circumstances in which the use of cross-correlation techniques for go/no-go testing are desirable, an alternative to integrating the impulse response and hence tolerancing the step response (see Fig. 10) is to use the novel Lamb filtering technique[18] in which the characteristics of the pseudo-noise stimulus are adjusted to filter out the secondary resonance, leaving a much smoothed impulse response as shown in Fig. 11. The relative ease of tolerancing the smoothed response will be demonstrated in Section 11.

Fig. 12. Reduction in measurement uncertainty. Cross-correlation at 200 Hz clock speed.

## 11. Tolerancing Impulse Response Estimates

Figures 9 and 11 give the basic data for setting up tolerances on go/no-go tests using cross-correlation techniques. Further information required includes a knowledge of measurement spread due to noise, and a knowledge of the impulse response of systems failing the original frequency response test. Measurement uncertainty depends on the test signal/environmental noise ratio, and clearly depends on the test environment. In fact, it is the authors' experience that signal/noise ratio varies enormously from one servomechanism to another within the same batch. For the case of wide-band random noise superimposed on the test signal, the measurement spread is inversely proportional to the square root of the number of sequences of the pseudo-noise stimulus over which correlation is performed.[12] Measurement time and measurement uncertainty can therefore be traded in the manner of Fig. 12 which is based on detailed repeatability tests on a typical member of this family of servos.

To generate unsatisfactory impulse responses, parameters of the system model (eqns. 4 and 5) are varied and parameter combinations resulting in failure of the system to meet the original frequency response specification are noted, care being taken to ensure adequate

Fig. 11.
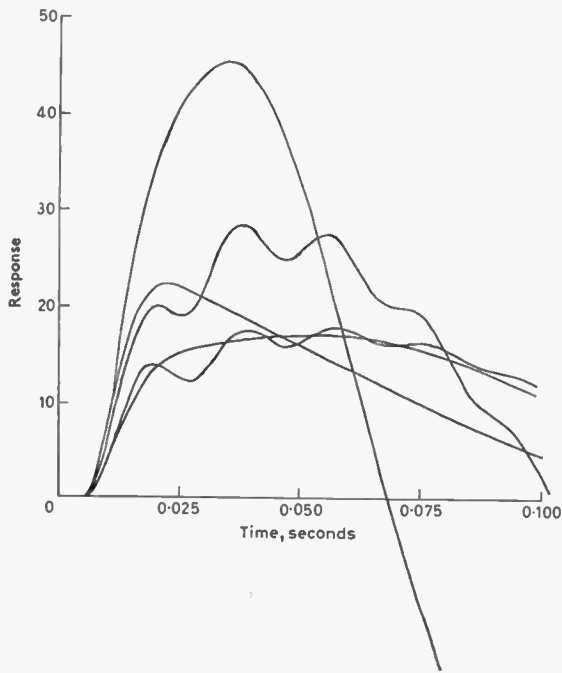Lamb-filtered unit impulse responses (Time in seconds).

Fig. 13. Simulated unit impulse responses for some failed systems.

coverage of both catastrophic failure and drift causes. Corresponding impulse responses such as those shown in Fig. 13 are then generated and critical regions of the impulse response identified by comparison with Fig. 9. Tolerances are then set and tested to determine the effectiveness of the test procedure.

Figure 14 shows the performance envelope derived from the probabilistic model sample of 30, together with the ranges recorded from field measurements made on a sample of 13 using a commercial cross-correlator. Allowing for sampling errors, the two methods of predicting suitable tolerances are in good agreement, and fully justify the theoretical model of the system based on equation (5). Six gates are shown, and incorporate an allowance for sampling errors. These gates are set on the basis of maximizing the chance of failure of bad systems whilst maintaining a marginal pass/fail characteristic for any system which has a marginal behaviour in the frequency domain.

Since there is a 100 : 1 range over which noise level may vary throughout the family an adaptive technique adjusting the measurement time to the noise conditions present would minimize test times. For example, the noise standard deviation for the system of Fig. 12 may be reduced to less than $2.5\%$ of any gate width by averaging over 25 sequences (128 seconds).

We define a measure of confidence, $M$, in the tolerancing procedure as follows:

$$M = \frac{\sum \text{number data points at which bad systems fail}}{\sum \text{data points}}$$

$$\ldots\ldots(6)$$

summed over all gates and over all simulated bad systems. Thus for the unfiltered impulse response with six gates and tested against twenty 'failed' systems, sum of data points $= 6 \times 20 = 120$. $M$ is summarized in Fig. 15 for three different procedures. It is seen that tolerancing Lamb-filtered responses gives a very high confidence measure for only three data points, in fact most bad systems are failed at least twice in three measurements. A further comparison between the Lamb filtered and the unfiltered techniques is shown in Fig. 15 where the unfiltered approach is seen to lead to acceptance of two out of twenty 'failed' systems (taking the three most successful gates for the unfiltered technique). This discrepancy may be overcome by taking all six gates shown in Fig. 14 but the measure of confidence, $M$ is then reduced. Since the philosophy of Lamb filtering implies knowledge (and therefore measurement) of the secondary resonance and also as shown leads to a higher degree of confidence in the measurements it must be concluded that this is a powerful technique for utilization in checkout of system dynamic errors.
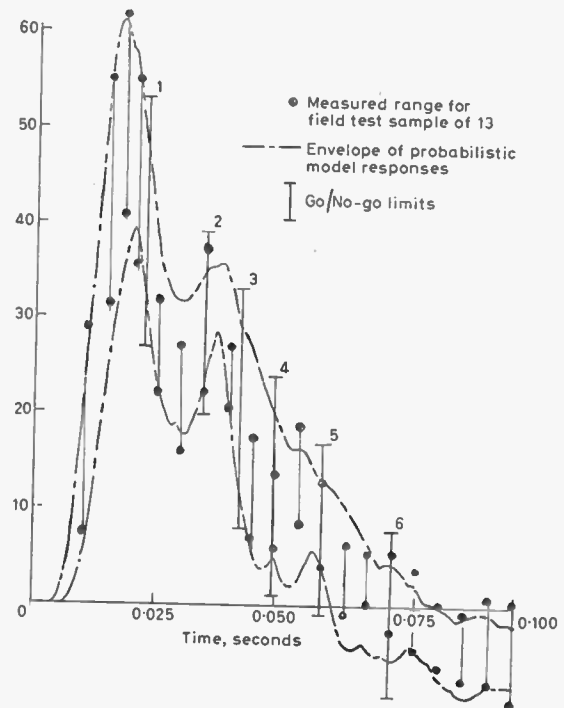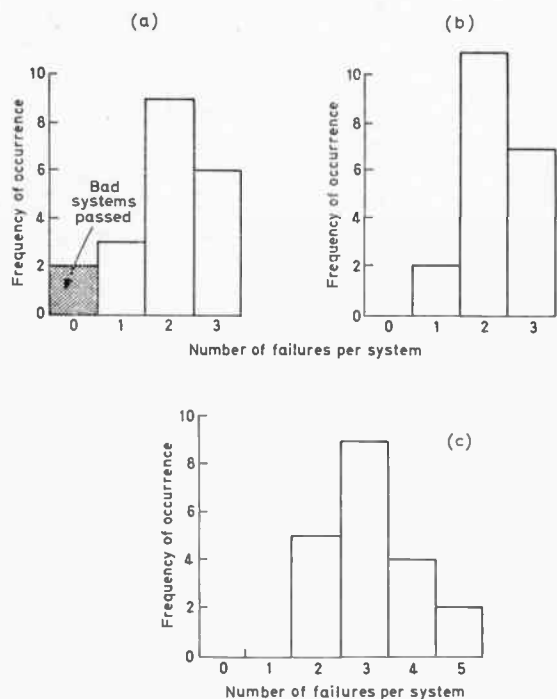


Fig. 14. Unit impulse response predicted and measured envelopes.

(a) Failure pattern for unfiltered case using three most effective gates. (1, 5 and 6 of Fig. 14.) $M = 0.65$.

(b) Failure pattern for filtered case using gates shown in Fig. 11. $M = 0.75$.

(c) Failure pattern for unfiltered case using all gates in Fig. 14. $M = 0.525$.

**Fig. 15.** Failure patterns for three test procedures.

## 12. Probabilistic Model as a Standard Test for New Identification Techniques

The system described by the probabilistic model of Table 5 is representative of many practical servo-mechanisms in that it contains passive, electrical, electronic and hydraulic components. Frequency and impulse responses for the system are of complex character due to the lightly damped secondary mode. In addition the low-order model transfer function coefficients coincide with those for a wide variety of control systems.[19] All these factors combined suggest that the model may be used in simulator studies including suitable noise and disturbance signals to establish the ability or otherwise of identifying practical systems using various proposed methods.[20, 21] Furthermore, the simulation approach may save much unnecessary expense involved in undertaking field studies to prove new test equipment.

## 13. Conclusions

The paper has quantified the meaning of production spread for a large family of electro-hydraulic servos, and established a probabilistic model which may be used to predict adequately performance variation in many domains.

It has also shown that the performance variation of the family is accurately predicted from the initial production test data.

Due to the general nature of the transfer function describing the electro-hydraulic servo, many of the conclusions are of wide applicability. Practical applications of the probabilistic model are shown to include the introduction of simpler frequency domain tests and the derivation of go/no-go gates on the system impulse response.

Two techniques suitable for deriving transfer function coefficients from very limited frequency domain data are also described.

## 14. Acknowledgments

## 15. References

1. Towill, D. R. and Lamb, J. D., 'First Probabilistic Model Describing Performance of In Service Instrument Servo-mechanisms', University of Wales Institute of Science and Technology, DAG TN 12, January 1968.
2. Marshall, J. J., 'Estimation of Parameters of a Type (I, 2) System Satisfying a Toleranced Deterministic Specification', University of Wales Institute of Science and Technology, DAG TN 8, August 1967.
3. Lamb, J. D., 'The Use of Pseudo-Random Binary Signals for the Production Testing of Dynamic Systems', Ph.D. Thesis, University of Wales Institute of Science and Technology, 1970.
4. Towill, D. R. and Payne, P. A., 'Frequency domain approach to the automatic testing of control systems', Proceedings of Conference on Automatic Test Systems, Birmingham, April 1970, pp. 529–48 (I.E.R.E. Conference Proceedings No. 17.)
5. Towill, D. R. and Mehdi, Z., 'Prediction of the transient response sensitivity of high order linear systems using low order models', *Trans. Inst. Meas. Control*, **3**, No. 1, T1, 1970.
6. Biernson, G., 'Estimating transient responses from open loop frequency response', *Trans. Amer. Inst. Elect. Engrs*, *Applications and Industry*, **75**, pp. 308–402, January 1956.
7. Towill, D. R. and Baker, K. J., 'Use of Sensitivity Functions to Generate a Probabilistic Model for a Family of Rate Servos', University of Wales Institute of Science and Technology, DAG TN 31, September 1969.
8. Payne, P. A. and Towill, D. R., 'Digital Simulation Applied to the Evaluation of System Dynamic Test Procedures', Proc. 1970 Summer Computer Simulation Conference, Denver, Colorado, pp. 462–69.
9. Towill, D. R. and Payne, P. A., 'Adequacy of Low Order Modelling for a Complex Electro-Hydraulic Servomechanism', University of Wales Institute of Science and Technology, DAG TN 36, August 1970.
10. Sumray, E. J., 'A survey of servo system test equipment', *Brit. Commun. Electronics*, **8**, pp. 198–204, March 1961.

11. Truxal, J. G., 'Automatic Feedback Control System Synthesis', p. 306 (McGraw-Hill, New York, 1955).
12. Hughes, M. T. G. and Noton, A. R. M., 'The measurement of control system characteristics by means of a cross-correlator', *Proc. Instn Elect. Engrs*, 109, Part B, No. 43, January 1962.
13. Cummins, J. D., 'A Note on Errors and Signal to Noise Ratio of Binary Cross-Correlation Measurements of System Impulse Response', AEEW Report No. R329, 1964.
14. Truxal, J. G., *Ibid*, p. 362.
15. Levinge, R. W., 'Measurement of resonances in high order transfer functions', *Trans. Inst Meas. Control*, 2, No. 3, pp. 102–5, March 1969.
16. Naslin, P., 'Dynamics of Linear and Non-Linear Systems', p. 128 (Blackie, London, 1965).
17. Towill, D. R., Cooper, J. S. and Lamb, J. D., 'Dynamic analysis of fourth order feedback control systems', *Intl. J. Control*, 10, No. 2, pp. 121–46, August 1969.
18. Lamb, J. D., 'The detection and filtering of system response modes of low damping using PRBS and cross-correlation', Proceedings of Conference on Automatic Test Systems, Birmingham, April 1970, pp. 521–7. (I.E.R.E. Conference Proceedings No. 17.)
19. Brown, J. M., Towill, D. R. and Payne, P. A., 'Generalisation of Dynamic Error Constraints Imposed by Frequency Response Measurements', UWIST DAG TN 35, September 1970.
20. Shuttleworth, J. E., 'Transient analysis using quasiperiodic forcing functions', *Electronics Letters*, 3, pp. 203–4, May 1967.
21. Kekre, H. B., 'Determination of linear system parameters by using polynomial inputs', *Intl J. Control*, 8, No. 3, pp. 211–19, 1968.
22. Payne, P. A., 'An improved technique for transfer function synthesis from frequency response data', *I.E.E.E. Trans. on Automatic Control*, AC-15, No. 4, pp. 480-3 August 1970.

## 16. Appendix: Data Processing to Predict Dominant Transfer Coefficients from Production Test Records

### 16.1 *The Sensitivity Method*

The sensitivity method involved considerable prior simulator study of the system[7] and utilized the production test phase data in the following manner.

Let $f_{\phi_1}$, $f_{\phi_2}$ and $f_{\phi_3}$ be the recorded frequencies at which the system phase lag is $\phi_1$, $\phi_2$, and $\phi_3$, respectively, and if first-order changes in transfer function coefficients are assumed then

$$|\Delta f_\phi| = \left|\frac{\partial f_\phi}{\partial A/A}\right|\left|\frac{\Delta A}{A}\right| \qquad \ldots\ldots(7)$$

where



Fig. 16.
Data processing to establish dominant transfer function part of probabilistic model.

$$|\Delta f_\phi| = \begin{vmatrix} \Delta f_{\phi_1} \\ \Delta f_{\phi_2} \\ \Delta f_{\phi_3} \end{vmatrix} \qquad \left|\frac{\partial f_\phi}{\partial A/A}\right| = \begin{vmatrix} \dfrac{\partial f_{\phi_1}}{\partial A_1/A_1} & \dfrac{\partial f_{\phi_1}}{\partial A_2/A_2} & \dfrac{\partial f_{\phi_1}}{\partial A_3/A_3} \\ \dfrac{\partial f_{\phi_2}}{\partial A_1/A_1} & \dfrac{\partial f_{\phi_2}}{\partial A_2/A_2} & \dfrac{\partial f_{\phi_2}}{\partial A_3/A_3} \\ \dfrac{\partial f_{\phi_3}}{\partial A_1/A_1} & \dfrac{\partial f_{\phi_3}}{\partial A_2/A_2} & \dfrac{\partial f_{\phi_3}}{\partial A_3/A_3} \end{vmatrix} \qquad \left|\frac{\Delta A}{A}\right| = \begin{vmatrix} \dfrac{\Delta A_1}{A_1} \\ \dfrac{\Delta A_2}{A_2} \\ \dfrac{\Delta A_3}{A_3} \end{vmatrix}$$

Clearly, if $|\Delta f_\phi|$ is available from test results and if $|\partial f_\phi/(\partial A/A)|$ is available from simulator or theoretical considerations, then $A_1$, $A_2$, and $A_3$ may be estimated. The technique is illustrated in the flow diagram, Fig. 16(a).

The sensitivity functions were determined for perturbations about the average values of $A_1$, $A_2$, and $A_3$ found to best match the means of the amplitude ratios of the production test data. Despite considerable experimentation, it was not found possible to balance the average production test amplitude ratio and phase data simultaneously, and the compromise adopted requires that the remnant phase errors implicit in this 'average' transfer function are allocated as the effect of the high frequency poles.

### 16.2 Curve Fitting Method

The curve fitting method used was that of Payne[22] in which for this application a zero velocity lag transfer function was fitted to the available data in a least error squared sense.

From equation (3), at frequency $\omega$ we may write

$$T(j\omega) = \frac{1 + A_1(j\omega)}{1 + A_1(j\omega) + A_2(j\omega)^2 + A_3(j\omega)^3} = \frac{N(j\omega)}{D(j\omega)} \quad (8)$$

If the measured data is $F(j\omega)$, then the error at a frequency $\omega_k$ is

$$e_k(j\omega_k) = F(j\omega_k) - T(j\omega_k) \qquad \ldots\ldots(9)$$

As indicated in reference 22, in order to minimize the sum of the errors squared, the equations to be solved become

$$\begin{vmatrix} U_2 - 2S_2 + \lambda_2 & T_3 & S_4 - U_4 \\ T_3 & U_4 & 0 \\ S_4 - U_4 & 0 & U_6 \end{vmatrix} \begin{vmatrix} A_1 \\ A_2 \\ A_3 \end{vmatrix} = \begin{vmatrix} 0 \\ U_2 - S_2 \\ T_3 \end{vmatrix} \quad (10)$$

where

$$\lambda_i = \sum_{k=1}^{n} \omega_k^i \times \frac{1}{\left|D(j\omega_k)_{M-1}\right|^2}$$

$$S_i = \sum_{k=1}^{n} \omega_k^i R_k \times \frac{1}{\left|D(j\omega_k)_{M-1}\right|^2}$$

$$T_i = \sum_{k=1}^{n} \omega_k^i I_k \times \frac{1}{\left|D(j\omega_k)_{M-1}\right|^2}$$

$$U_i = \sum_{k=1}^{n} \omega_k^i (R_k^2 + I_k^2) \times \frac{1}{\left|D(j\omega_k)_{M-1}\right|^2}$$

$n$ = total number of measurement frequencies

$R_k$ = real part of fitted transfer function at $\omega_k$

$I_k$ = imaginary part of fitted transfer function at $\omega_k$

$M$ = current iteration number

Thus previous values of $A_1$, $A_2$, and $A_3$ are used to evaluate $D(j\omega_k)$ for the current iteration. The iterative process is halted when differences become insignificantly small, as shown in Fig. 16(b). Based on a theoretical analysis, the phase lags of 45°, 90°, and 135° were reduced to 36°, 78°, and 109° respectively in order to allow for the effects of high-frequency poles. An attenuation factor of 0·93 was also allowed at $f_{109°}$.

## STANDARD FREQUENCY TRANSMISSIONS—November 1970

*(Communication from the National Physical Laboratory)*

| Nov. 1970 | Deviation from nominal frequency in parts in 10¹⁰ (24-hour mean centred on 0300 UT) | | | Relative phase readings in microseconds N.P.L.—Station (Readings at 1500 UT) | | Nov. 1970 | Deviation from nominal frequency in parts in 10¹⁰ (24-hour mean centred on 0300 UT) | | | Relative phase readings in microseconds N.P.L.—Station (Readings at 1500 UT) | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | GBR 16 kHz | MSF 60 kHz | Droitwich 200 kHz | *GBR 16 kHz | †MSF 60kHz | | GBR 16 kHz | MSF 60 kHz | Droitwich 200 kHz | *GBR 16 kHz | †MSF 60 kHz |
| 1 | − 300·2 | + 0·1 | + 0·1 | 683 | 623·5 | 17 | − 300·0 | − 0·1 | 0 | 671 | 619·7 |
| 2 | − 299·8 | 0 | + 0·1 | 681 | 623·6 | 18 | − 299·8 | + 0·1 | − 0·1 | 669 | 619·2 |
| 3 | − 299·8 | 0 | + 0·1 | 679 | 623·4 | 19 | − 299·9 | − 0·1 | − 0·1 | 668 | 620·1 |
| 4 | − 299·9 | + 0·1 | + 0·1 | 678 | 622·9 | 20 | − 300·1 | 0 | 0 | 669 | 620·4 |
| 5 | − 299·9 | + 0·1 | + 0·1 | 677 | 622·3 | 21 | − 300·1 | 0 | 0 | 670 | 620·6 |
| 6 | − 300·0 | + 0·1 | + 0·1 | 677 | 621·6 | 22 | − 299·9 | + 0·1 | 0 | 669 | 620·0 |
| 7 | − 299·9 | + 0·1 | + 0·1 | 676 | 620·5 | 23 | − 300·0 | 0 | 0 | 669 | 625·8 |
| 8 | − 299·8 | 0 | + 0·1 | 674 | 620·1 | 24 | − 300·0 | 0 | 0 | 669 | 626·6 |
| 9 | − 300·1 | + 0·1 | + 0·1 | 675 | 619·5 | 25 | − 300·0 | 0 | + 0·1 | 669 | 626·8 |
| 10 | − 299·9 | 0 | + 0·1 | 674 | 619·3 | 26 | − 300·0 | 0 | + 0·1 | 669 | 627·1 |
| 11 | − 299·9 | 0 | + 0·1 | 673 | 618·9 | 27 | − 300·0 | − 0·1 | + 0·1 | 669 | 627·7 |
| 12 | − 299·8 | + 0·1 | + 0·1 | 671 | 618·1 | 28 | − 300·1 | 0 | + 0·1 | 670 | 627·7 |
| 13 | − 300·1 | 0 | 0 | 672 | 618·5 | 29 | − 300·0 | 0 | 0 | 670 | 627·9 |
| 14 | − 299·8 | 0 | 0 | 670 | 618·1 | 30 | − 299·9 | − 0·1 | 0 | 669 | 628·6 |
| 15 | − 300·2 | 0 | 0 | 672 | 618·4 | | | | | | |
| 16 | − 299·9 | − 0·1 | 0 | 671 | 619·1 | | | | | | |

Note: The frequency offset for 1971 will be −300 × 10⁻¹⁰.

All measurements in terms of H.P. Caesium Standard No. 334, which agrees with the N.P.L. Caesium Standard to 1 part in 10¹¹.

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

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

# The Use of Cathode-Ray Tubes in Professional Equipment

By

## A. B. McFARLANE,

M.Sc. (Eng.), A.C.G.I., C.Eng., M.I.E.E.†

The vast range of professional equipment which uses a cathode-ray tube as the display device is considered in terms of the following classification:

(1) oscillography and precision measurements, (2) visual display computer terminals, (3) avionic displays, (4) radar, (5) medical electronics, and (6) automated and control systems.

In each case the general function and performance of relevant equipment is outlined, and the way in which these determine the type of cathode-ray tube used is discussed. Special aspects of tube design are dealt with in more detail, and desired future improvements indicated.

## 1. Introduction

The cathode-ray tube has held an unchallenged position as a display device for over thirty years, and will continue to do so for many more to come. Its unique, virtually instantaneous response to electronic signals accounts for this fact. The availability of many different types of transducer gives rise to their application in a very wide range of professional equipment. In this paper, the field is divided into the following major classifications:

 (i) Oscillography and precision measurements.
 (ii) Visual display computer terminals.
(iii) Avionic displays.
(iv) Radar.
 (v) Medical electronics.
(vi) Automated and control systems.

From early beginnings as an indicating device, the present-day tubes are capable of high precision measurements. In view of this, their complexity has of necessity increased. New uses are being evolved, especially in the avionics field and that of man/computer communications. In keeping with these requirements modern tubes are more rugged and reliable, and improved technology is being applied constantly to designs so that better performance is attained.

## 2. Oscillography and Precision Measurements

The basic operation of an oscilloscope instrument is to present to a viewer a visible graph of the variation of some electrical signal as a function of real time. This time interval may be as short as a few nanoseconds (i.e. $10^{-9}$ second), and the magnitude of the electrical signal may be only a few millivolts.

The geometrical configuration and mode of operation of present-day instrument tubes has been discussed in detail elsewhere,[1,2,3] and so Fig. 1 acts as

a brief reminder. It is common practice to specify the useable bandwidth of an oscilloscope instrument as the frequency range from d.c. up to the frequency at which the deflexion sensitivity is 3 dB less than the d.c. response. Since useable bandwidth is an important application parameter it has been used to summarize the characteristics of relevant cathode-ray tubes (Table 1). Here, care must be taken not to confuse real-time bandwidths with those quoted for sampling techniques which may seem to be fifty times wider, but can only be used on repetitive signal waveforms.
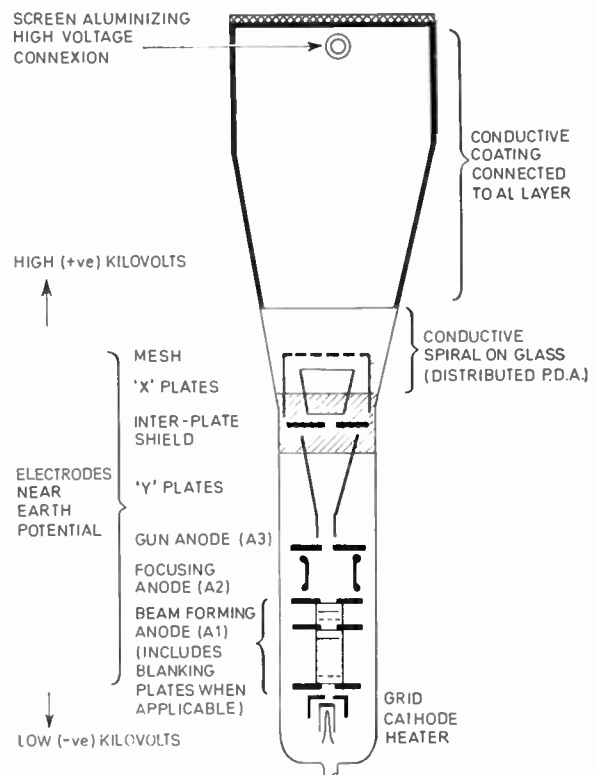


Fig. 1. Tube configuration.

† The M-O Valve Company Ltd., Brook Green Works, Hammersmith, London, W.6.

**Table 1.** Tube parameters for instrument bandwidth

| | Bandwidth (—3dB on d.c.) MHz | Typical 'Y' Amplifier Gain | Rise †Time (s) | Signal 'Y' Deflexion System | Scan Volts for Full 'Y' Scan | Post Deflexion Acceleration System | Typical Operating Screen Potential (kV) | Scan Volts for Full 'X' Scan | Maximum Overall Length (mm) |
|---|---|---|---|---|---|---|---|---|---|
| **(a) Single Trace.** | ≥200 | 300 | 1·5 | Continuous helical transmission line | 30 (8 cm) | mesh/distributed (ratio 10:1) | ‡ 20 (pws 4000 cm/μs) | 60 (10 cm) | 500/550 |
| | ≥100 | 500 | 3·5 | Lumped transmission line (sectioned plates) | 30 (8 cm) | mesh/distributed (ratio 10:1) | 15 | 120 (10 cm) | 450/500 |
| | ≥50 | 500/1000 | 7·0 | Conventional deflector plates | 40 (8 cm) | mesh/distributed (ratio 10:1) | ‡ 12 (pws 1750 cm/μs) | 110 (10 cm) | 350 |
| | ≥25 | 500/1000 | 15·0 | Conventional deflector plates | 30 (6 cm) | mesh/distributed (ratio 10:1) | 10 | 120 (10 cm) | 350/400 |
| | ≥10 | 500/1000 | 35·0 | Conventional deflector plates | 80 (8 cm) | distributed (ratio 4:1) | 4 | 180 (10 cm) | 350/400 |
| | <10 | 1000/2000 | >35·0 | Conventional deflector plates | 70/120 (6 cm) | Single band (ratio 2:1) or mono accelerator | 2 | 150/300 (8 cm) | 250/400 |
| **(b) Dual-Trace.** | ≥50 | 500 | 7·0 | Two gun conventional deflector plates | 50 [ 6 cm / 5 cm common ] | mesh/distributed (ratio 10:1) | ‡ 15 (pws 3000 cm/μs) | 200 [ 10 cm / separate X ] | 350/400 |
| | ≥25 | 500 | 15·0 | Split beam conventional deflector plates | 30 [ 6 cm / 6 cm common ] | mesh/distributed (ratio 10:1) | ‡ 10 (pws 500 cm/μs) | 80 [ 10 cm / common X ] | 400/450 |
| | ≥10 | 500 | 35·0 | Two gun conventional deflector plates | 40 [ 6 cm / 4 cm common ] | distributed (ratio 4:1) | 4 | 200 [ 10 cm / common X ] | 350/400 |
| | ≥5 | 500 | 60·0 | Split beam conventional deflector plates | 70 [ 6 cm / 6 cm common ] | distributed (ratio 4:1) or mono accelerator | 4 | 180 [ (10 cm) / common X ] | 350/400 |

† Fastest time for signal to rise from 10% to 90% of its amplitude.
‡ Photographic writing speed. (A.S.A. 10 000; lens $f$/1·2; demagnification 2:1; aluminized GH screen.)

The upper range listed starts at 200 MHz. Such tubes have reasonable deflexion sensitivity, which has been achieved by using the mesh/spiral p.d.a. (post-deflexion acceleration) principle. This was pioneered and first produced in England during 1957. A mesh or frame grid mounted close in front of the electron gun acts as an electrostatic shield between the deflexion system of the gun and the p.d.a. electrode. Consequently the deflexion system can be designed to give the required frequency response, and then with the minimum of interaction, the potential of the luminescent screen raised to give adequate trace luminance. For bandwidths greater than 200 MHz the deflexion system consists of a continuous transmission line in the shape of a wound ribbon helix. Thus, the deflecting voltage wave travels along the tube axis at the same velocity as the electrons themselves and avoids transit time effects.

For frequencies in the next range down to 100 MHz, a 'lumped' line formed by small inductances interconnecting several pairs (usually four to six) of short deflector plates in sequence gives satisfactory results.

The mesh type design (e.g. M-O Valve Co. 1400B) is still necessary for lower bandwidths down to about 25 MHz, but now a conventional pair of signal deflector plates can be designed to give satisfactory deflexion sensitivity, sufficiently short transit time (1·5 ns), and adequately high resonance frequency not to effect the overall level response.

Below 25 MHz, adequate photographic writing speeds can be achieved with 4 kV screen potential, hence reduced p.d.a. ratios (i.e. < 4/1) can be used and a distributed lens system without the mesh shielding is quite adequate (e.g. M-O V 1400A).

Right at the bottom end, monoaccelerator or single-band p.d.a. (ratio 2/1) tubes give satisfactory designs to operate up to 10 MHz.

Display presentation and measurement accuracy have been improved over the past few years to the instrument user's advantage, in the following ways:

(i) The widespread introduction of rectangular face plates which enables the front panel space of the instrument to be better utilized.

(ii) An increase in scan 'windows', particularly in the signal ($Y$) direction from 6 cm up to 8 cm.

(iii) The choice of having internal graticules (either 'black' or edge-lit) to avoid parallax errors, particularly when photographing transient traces.

(iv) The reduction of pattern distortion to $\leqslant \pm 1\%$.

(v) The improvement in deflexion linearity, especially in the time-base ($X$) direction, to < 5% for any two individual centimetres of scan.

(vi) Improved deflexion defocusing ratios to < 1·5, especially on mesh-type tubes where total scan angles of 40° are common.

(vii) The availability of luminescent screens which have an extremely thin aluminium backing layer. This eliminates 'hand capacitance' effects on the lower voltage p.d.a. tubes without incurring a loss of luminance.

Two important design features to the advantage of circuit designers are:

(viii) The availability of low wattage heaters (from 2 W to less than 1 W) for portable, transistor, and battery operated instruments.

(ix) The inclusion of deflexion blanking electrodes in the electron gun itself for on/off modulation of the electron beam. This means that bright-up pulse circuits are near earth potential and do not have to be highly insulated.

So far single trace tubes have been considered, and although two traces may be obtained by the technique of 'beam switching', such instruments are not so versatile as ones using dual-trace tubes of the split-beam (e.g. M-OV 1300P) or double gun variety (M-OV 1000H). Typical performances of the latter are indicated in section (b) of Table 1, the ultimate choice depending on instrument cost considerations.

For most purposes the GH(P31) green luminescent screen is used, but the GM(P7) type with its blue flash and much longer yellow afterglow has advantages when slower time base speeds are required. Table 2 compares the performance of these and other types of screen. For very long persistences, storage tubes[1] are available. Generally speaking, these are special versions of an existing conventional tube and contain a storage target.

For storage, the maximum writing speed depends on the electron charging characteristic of the permeable insulator target. In practice this is limited to a velocity of about 5 cm/μs, corresponding roughly to a frequency of 1 MHz at a peak-to-peak trace amplitude of 2 cm.

## 3. Visual Display Computer Terminals

In contrast to the long period of evolution of oscilloscope instrumentation, the last five years has seen the growth of a new application for cathode-ray tubes with the advent of digital computers. Here the vital need is for human communication with the computer to exploit fully its vast capabilities. This communication must take place through an interfacing equipment to which both parties can react. On the human side a visual display of graphical or alpha-numeric format in conjunction with a keyboard 'transmitter' is preferable. Figure 2 indicates the basic

**Table 2.** Luminescent screen data

| Phosphor Designation | | | Colour | | Persistence Classification† |
|---|---|---|---|---|---|
| European | EIA | M-O.V. | Fluorescence | Phosphorescence | |
| GH | P31 | 24 | green | green | medium-short |
| GM | P7 | 46 | purplish-blue | yellowish-green | long |
| LB | P38 | 27 | orange | orange | long |
| W | P4 | 18 | white | white | medium-short |
| GR | P39 | 29 | green | green | long |
| GJ | P1 | 01 | yellowish-green | yellowish-green | medium |
| LC | P26 | 19 | orange | orange | very long |
| LD | P33 | 23 | orange | orange | very long |
| — | P22R | — | red | red | medium |
| GE | P24 | 15 | green | green | short |
| BA | P16 | 22 | violet (u.v.) | violet (u.v.) | very short |
| KA | P20 | — | yellow-green | yellow-green | medium-short |
| — | P37 | — | blue | blue | very short |
| BE | P11 | 08 | blue | blue | medium-short |

† Persistence classification (decay time to 10%) as follows:

| very short | short | medium-short | medium | long | very long |
|---|---|---|---|---|---|
| $<1\,\mu s$ | $<10\,\mu s$ | $<1$ ms | $<100$ ms | $<1$ s | $>1$ s |

units schematically. For maximum efficiency, the equipment must be able to present data within microseconds to be consistent with 'calculation' times of the computer itself. The cathode-ray tube with its associated circuitry is fulfilling these requirements with great success.

Usually an alpha-numeric display is composed of several lines, possibly up to 60, each containing a given number of character positions (e.g. 50). The design of tube used depends on the manner in which the characters are to be generated, and this may be one of the following:

(i) Cursive stroke writing, in which fixed vector voltages are selected from an available range, depending on the character or symbol to be written, and then added sequentially to deflect the electron beam.

(ii) By the requisite selection of five pulses per line on a small seven-line scanned raster to give a possible $5 \times 7$ dot matrix.

(iii) The monoscope tube in which a preformed target character is scanned to produce a secondary emission current signal which is subsequently 'transferred' to the viewing tube in synchronism.

(iv) The extruded beam tube, which contains a stencil mask of the character repertoire. On selection of a particular character, the electron beam is positioned to pass through the character outline in question and thus produce a replica on the viewing screen.

(v) A tube which has a matrix electrode through which electron beams pass on their way to the screen. By suitably biasing the conductive surface inside the apertures certain beams may be prevented from passing through. Thus characters are formed by a dot matrix on the viewing screen.

(vi) By the suitable brightness modulation of a standard television scanning raster.

It should be noted that methods (iv) and (v) combine the character generation function within the viewing tube, but that (iii) still requires a separate viewing tube, as do (i), (ii) and (vi) where the character generation is achieved with solid-state circuitry. Precise details of these methods may be found elsewhere.[4-7] Many factors influence the techniques selected, for example, the cost of the electronic circuitry; the desired quality and legibility of the characters; the luminance of the symbols; the versatility of changing the character repertoire; or the number of viewing units required to operate from one terminal equipment.

In practice, electromagnetically deflected tubes are most commonly used for several reasons. Large beam currents are required to obtain sufficient luminance. Electrostatically-deflected tubes having reasonable deflexion sensitivities can only pass some 50 μA because of beam width restrictions through the deflector plates. Moreover on such tubes the deflexion angle is limited to 40° whereas angles up to 110° are not unreasonable magnetically. Finally screen potentials of 10 to 20 kV are necessary, depending on the display
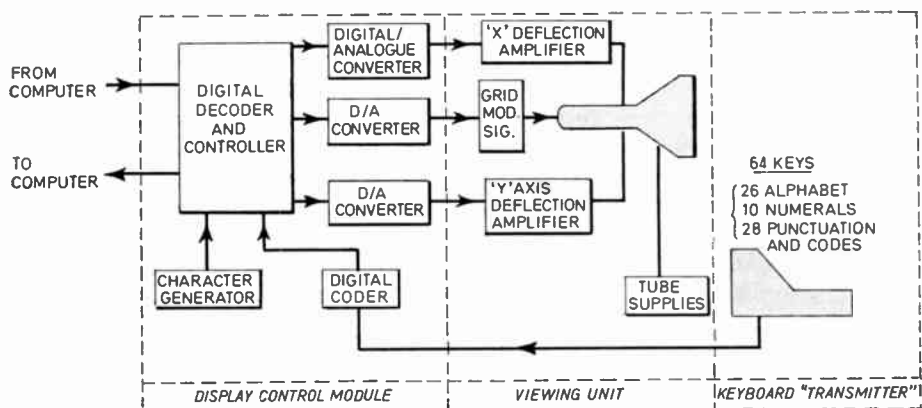
**Fig. 2.** Visual display terminal (schematic).

size, the larger sizes up to 61 cm diagonal being used mainly for graphical displays.[8] Electrostatically-deflected tubes would need a mesh/distributed p.d.a. system (see previous section) for tolerable scan voltages. Thus almost invariably the electron beam is positioned magnetically and the character 'written' magnetically. The time to relocate the beam to a new position is about 15 μs, which is usually a little longer than that required to write the character. It is possible to use a hybrid tube in which the 'writing' only is achieved by electrostatic deflexion. Since the character size and hence the scan amplitude is very small (e.g. 5 mm × 5 mm), poor deflexion sensitivity can be tolerated and p.d.a. is not necessary.

Having already outlined the methods of character generation, and the format of the display, it will be realized that the visible trace on the screen must be periodically refreshed to give an observer time to comprehend the text and act on it. Hence whilst phosphors are selected on physiological grounds to minimize operator fatigue, the choice does effect this refresh rate since the display must be free from tiresome flicker. Thus for the higher luminance screens, e.g. television white W (P4) and green GH (P31) which have short persistence the rate must be at least 50 Hz. For the longer persistence but more easily-burnt fluoride orange screen LB (P38) it may be safely reduced to some 16 Hz. As a compromise the green GR (P39) has been developed which has an intermediate persistence giving an adequate refresh rate of 30 Hz. The use of a storage tube[1] overcomes the refresh problem but such a system is severely limited for writing speed (see previous section), is not readily amenable to selective erasure, and the tube is very much more expensive.

Likewise, because of its complexity, the extruded beam tube is very costly. Moreover it suffers from the lack of versatility in that a particular tube has a fixed character repertoire and can only be used for a single viewing unit.

Visual computer terminals are in constant operation throughout the day and in bright ambient lighting. To obtain good contrast and visibility neutral tinted implosion panels are frequently attached to the bulb face, reducing the overall light transmission to some 40%. Such tubes offer operator protection without the need for other safety precautions. Conventional electrostatically-focused electron guns give characters of reasonable legibility but to improve this aspect of a display so-called laminar-flow guns are available.† Figure 3 indicates the essential electrode differences between the two types of tube. It has now been shown that an electron crossover does still occur in the beam-forming section of the new gun design, and yet it is capable of giving a cleaner-edged spot on the screen.
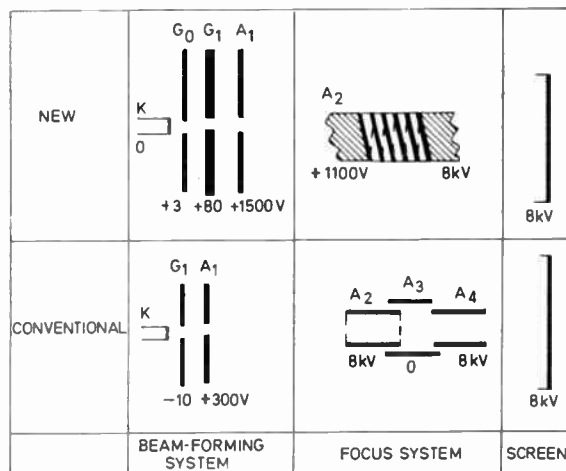


**Fig. 3.** Comparison of electron gun systems.

† Described in unpublished work (Thompson, R., Wreathall, M. and Ball, R).

The explanation for this is being determined. Such a system is invariably used in conjunction with a distributed focus lens, deposited on the inside of the neck of a tube in such a position as to minimize the image–object distance. Furthermore since the potential ratio is some 8/1, there is additional demagnification of the spot imaged on the luminescent screen. The focus voltage on such designs is more sensitive to the grid drive, which itself is greater than that normally required for a given beam current.

## 4. Avionic Displays

Aircraft instrumentation is another field in which the cathode-ray tube is being used increasingly to good effect. Again a digital computer processes volumes of data resulting from the many devices which monitor the plane's performance and progress on a flight. Certain aspects of this information have to be displayed on command to some member of the crew, and a multiple tube display offers the required flexibility for doing just this. For instance there are many advantages in projecting symbolic flight information into the pilot's view of the outside world, especially during take-off and landing. Such a display is termed 'head-up' (h.u.d.),[9, 10] and its action depends on locating the luminous cathode ray tube trace at the focal plane of an optical lens system, which then projects an image of it into the line of sight via a half silvered mirror. Thus a virtual image appears at infinity and the light beam is perfectly collimated.

The required properties of h.u.d. tubes are:

(i) They should be compact and lightweight.

(ii) Have narrow line-width since the projection optics will involve some magnification.

(iii) Yield high luminance since the display has to be viewed in very high ambient illumination.

(iv) The electrical insulation of the base and high voltage leads must be good especially at high altitudes.

(v) The glass face must have a precise geometry, in conjunction with a very accurate locating mask to position the tube with respect to the optical system.

(vi) The tube must be exceptionally rugged in two ways:

(a) by its envelope and internal contents being able to withstand accelerations up to 40 $g$ over some 2000 Hz of the audio frequency range without excessive movement or disintegration, and

(b) by being held sufficiently rigid in the locating attachment to prevent any movement with respect to auxiliary components (e.g. focus magnet) which would cause spot aberrations.

(vii) Withstand severe environmental conditions without deterioration.

With regard to (i), available tubes usually have a face diameter about 75 mm, are less than 250 mm in overall length, and are lighter than 1 kg. As with the computer terminal display tubes, magnetic deflexion is employed, but to obtain good resolution the focusing in these cases is also performed magnetically. For instance in a h.u.d. equipment having a 20° viewing angle, and assuming that the limiting resolution of the eye is 2′ of arc, 600 line pairs are required across 75 mm. This means a line width of 0·1 mm corresponding to a spatial frequency[11] of around 50 cycles/cm. To achieve this narrow line width, a special GJ(P1) luminescent screen of fine particles (~2 μm) has to be settled on the tube face. This material is used to give sustained high luminance. The screen operating potential needs to be as high as possible consistent with good environmental insulation. Thus 20 kV is a typical operating value and the e.h.t. lead and frequently the base as well are encapsulated in silicone rubber mouldings. Figure 4(a) shows a photograph of such a tube which also illustrates the locating mask specified under requirement (v) above. Regarding precision aspects of the tube, the glass bulb itself has to be fabricated with extreme accuracy otherwise the electron gun, which locates off the inside of the neck, may cause the electron beam to be deflected off-centre with respect to the focusing magnet and scan coils with resultant spot distortions.

Obviously, ruggedness is a very vital parameter contributing to the reliability of the tube performance. In this respect, the requirement for avionic systems exceed any other application. To this end, attempts have been made to use ceramic cone/neck assemblies for the tube envelope instead of glass. In such instances the simple triode gun electrodes are formed
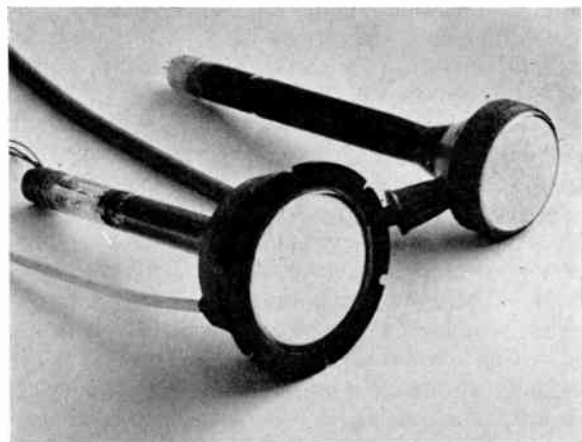


Fig. 4. (a) Head-up display tube (M-OV LD750).
(b) Radar projection tube (M-OV LD706).

by conductive coatings deposited on the inside surface of these envelopes. However it has been firmly established and confirmed that glass envelopes will safely withstand these $40\,g$ accelerations, and that gun structures can be adequately supported therein to minimize any relative movement between gun and envelope which can cause apparent elongation of the visible spot on the luminescent screen.

Turning to the head-down (h.d.) display which is usually mounted in a desk type console and presents data other than that associated with the flight path, ruggedness is again very important, but the electrical performance of such tubes is similar to those previously discussed for computer terminals. One major difference, however, is that simple neutral density filters do no longer give adequate contrast due to the exceedingly high level of incident illumination (around 30 000 nits) above the clouds. Even with an etched surface their specular reflection is unsatisfactory. Circularly-polarized filters, or special mesh-type filters, both having bloomed anti-reflecting surfaces offer improvements but this problem needs further attention. A 21 cm diagonal display seems a reasonable compromise with the restricted space available for electronics in aircraft.

In this Section, it is appropriate to consider airborne radar displays. For reason (i) above, projection principles have also been used to produce a large flat display from a small diameter tube. It is possible to achieve this with a Schmidt 'folded optical' system. The most used radar display is the plan-position indicator (p.p.i.) and for reasons to be given in Section 5 a fluoride, long persistence orange screen LC(P26), is necessary. The main problem with projection systems is to obtain adequate luminance of the image, and therefore a directional viewing screen must be used. This restricts the angle of view, but is not serious in an aircraft. With optical magnification of 10, a high screen operating potential up to some 30 kV is also necessary and therefore adequate electrical insulation precautions have to be adopted as well as x-ray shielding. Figure 4(b) shows a photograph of such a tube. It will be appreciated that requirements (i) to (vii) apply equally well to this type of tube.

Rectangular-faced tubes of both the conventional and storage[1] varieties, are also used for weather radar equipments. Again the question of trace visibility under high ambient lighting is important and television scan convertor systems have been used to improve the luminance of the display. This approach is discussed further in the next section.

With regard to the reliability of electronic equipment[12] it would appear that non-redundant avionic systems have a mean time between failures (m.t.b.f.) of some 1000 hours. Thus for the cathode-ray tube

itself to have 95% reliability for no faults during 1000 hours life, the m.t.b.f. has to be at least 20 000 hours which corresponds to a failure rate of $5 \times 10^{-5}$ per hour.

## 5. Radar

Cathode-ray tubes have been used as the display device in radar equipments for almost as long as their use in oscilloscopes. In this Section their application to marine and airport installations is considered. With the p.p.i. presentation, the electron beam is deflected electromagnetically[13] to traverse a radius from the centre to the edge of the tube face, in keeping with the instantaneous position of a very directional transmitting aerial. In actual fact, the aerial rotates continuously at about 20 revolutions per minute. Thus during one second the angle described is $2\pi/3$, during which time some 2000 radii have been traversed. As a particular radius is about to be 'drawn', a short pulse of electromagnetic energy which can be as short as $0.05\,\mu s$ duration starts off into space. Should a reflexion occur from an object $x$ nautical miles away, then an echo will return after $12.3x\,\mu s$. During all the scan sequences so far described, the electron beam in the tube has been biased off so that no light was visible from the luminescent screen. The echo signal is therefore used to switch on the beam current and hence produce a visible spot on the screen. The particular radius or radii affected depends on the bearing of the reflecting object. Thus by knowing the radial velocity of the spot movement, the distance from the face centre at which the echo spot occurs can be calibrated directly in nautical miles. For a 31 cm diameter tube and a maximum range of 6 nautical miles, the sweep time (i.e. deflexion time for one radius of the c.r.t.), is about $75\,\mu s$, whilst the sweep repetition interval may be typically $500\,\mu s$ or more dependent upon the parameters of the radar. Hence there is a long 'non-operational' time which may be used for inter-scan activities, such as the writing of alphanumerics. Furthermore since the elemental screen area under the beam spot ($0.5$ mm diameter) is excited for only $0.25\,\mu s$ at a time, it will be appreciated that to produce an observable 'picture' the luminescent screen must have a long persistence. A permanent echo will excite the same screen spot several times and hence build up the luminance. Transient echoes will not be repeated and their excited spots will 'fade away' very quickly. In practice, fluoride phosphors with a persistence of several seconds are used, depending on the range covered by the particular equipment. Alternatively the GM(P7) screen with its superimposed layers of two different phosphor powders may be used to give a blue flash for transient echoes, and a yellow trace for permanent ones.

Table 3 indicates the type of tubes available. Generally, the smaller sizes are used for river radar, where

**Table 3.** Types of radar tube

| Face Diameter (cm) | Face Curvature (cm) | Deflexion Angle | Maximum Neck Diameter (mm) | Maximum Overall Length (mm) | Typical Screen Potential (kV) | Focusing | Luminescent Screen | Special Features |
|---|---|---|---|---|---|---|---|---|
| ≤21 | 450/850 | 35°/50° | 30/38 23 | 300/450 | ≤14 ≤25 | electrostatic electromagnetic | LD(P33) ⎫ | ⎫ |
| ≤31 | 1000 | 40°/50° | 36 23 | 450/600 | ≤16 15/25 | electrostatic electromagnetic | LC(P26) or LD(P33) ⎬ | conventional radar types |
| ≤41 | 700/1400 | 50°/70° | 36 38 | 500/650 | 15 15/25 | electrostatic electromagnetic | LC(P26) ⎭ | ⎭ |
| ≤31 | 1000 | 50° | 36 | ≤650 | ≤12 | electromagnetic | P22 R ⎫ P1 ⎬ | bi-colour |
| ≤13 | optical flat | 30° | 36 | <500 | 25 | electromagnetic | GE(P24) BA(P16) | high resolution |
| ≤28 | optical flat | <40° | 36 | <600 | 10 | electrostatic | KA(P20) | direct view storage |
| ≤48 | >850 | 50° | 36/38 | <700 | 18 | electrostatic or electromagnetic | W(P4)† LC(P26) | rear-ported tube |

† Scan conversion system.

distances involved are much shorter than for ocean going equipments where ranges up to 50 nautical miles are common. Such tubes need a shorter persistence LD(P33) screen instead of the LC(P26) type, to avoid unecessary cluttering of the display. Assuming a 0·5 mm spot size, objects 5 metres apart can be detected. Noticeable tube features are the low deflexion angles, comparable with those pertaining to electrostatically deflected oscilloscope tubes. This is to keep the required scan power down as well as maintain good deflexion defocusing performance. Any power saving for the smaller neck sizes is questionable, and tubes with these tend to give poorer performance, due to the necessary reduction in the diameter of the electrodes in the gun. Faces are reasonably flat consistent with mechanical strength, but the long overall length is dictated by the bulb cone length for the low deflexion angles. The majority of electron guns are of the *einzel*, electrostatically-focused type, in which the spot size is very insensitive to changes in focus potential.

Present-day equipments tend to use transistors throughout which means that grid drive circuits are highly susceptible to permanent damage by an unusually large current pulse. As an additional safeguard, since tubes do operate at a high voltage, a transistor protection device (t.p.d.)[14] can be incorporated internally.

Here as described in the previous Sections, the problem of contrast and adequate luminance frequently arises. Direct view storage tubes may be used or a scan-conversion system. This uses a special non-viewed storage tube which has the radar picture 'written' on one side of a storage target in the usual polar co-

ordinates. This stored image is then 'read' by an electron beam scanning the surface in television raster fashion. The signal generated is displayed in synchronism on a viewing monitor tube, which may be one of the types discussed in Section 3. Since the whole reading process is repeated at 50 Hz, a very bright and easily viewable display results. Such a system is naturally more expensive than conventional equipment.

So far screens used in radar tubes have given a single colour, but there are advantages in having at least two colours available, so that echoes may be classified and easily recognized by the observer. For example, on airfield surveillance incoming craft might be one colour and outgoing ones a different colour. A 31 cm bi-colour tube has been reported,[15] in which two superimposed phosphor layers, one emitting red and the other green light, are bombarded in turn by switching the final anode potential between 6 and 12 kV.

For some radar installations it is useful to have the p.p.i. display superimposed on a background such as a chart or map. Rear ported tubes of up to 48 cm screen diameter are available for this purpose. A high-quality optical window is sealed into the cone of the tube so that a projector can image film or slides on the luminescent screen. Naturally this screen cannot be aluminized and some other form of overcoming the screen sticking potential[16] such as the use of a conductive transparent film on the glass face under the screen has to be adopted. An alternative method of producing a background to the display is that of using the focused spot (0·025 mm or 150 cycles/cm) of a extremely high resolution tube as a light source to scan

a film of the required data. An electrical signal is then obtained from a photomultiplier viewing the transmitted light, and this can be synchronously injected into the p.p.i. display. The reverse operation may also be utilized in which the 'live' p.p.i. display is produced on the high resolution tube, photographed, the film developed *in situ*, and subsequently projected on a large optical viewing screen together with other geographical data.

Such high resolution tubes[17] must be magnetically focused and deflected, have a very fine grain uniform luminescent screen of low noise content, and operate at high potentials. A very short persistence screen such as BA (P16) or P37 is necessary for these applications.

## 6. Medical Electronics

Nowadays electronics is playing an ever increasingly important part in medicine. This arises from the fact that electrical signals generated inside the body can be detected by electrodes suitably positioned on the surface of the skin. These signal waveforms are repeated continuously and therefore instruments like the electro-encephalograph (brain), electro-cardiograph (heart) and electro-myograph (muscles and nerves) are designed to display them as a function of time on a cathode ray tube

Such equipments are used both for monitoring purposes during surgical operations and also for surveillance on 'intensive care' units. In the former case, mobility is essential so that it can be transferred from one location to another with the patient. Hence the emphasis is on portability and self-contained battery supplies. Figure 5 illustrates the type of complex waveform (e.c.g.) due to the action of the
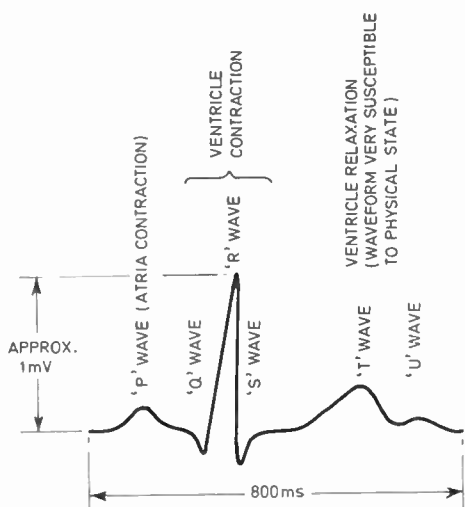
heart. Here the duration and amplitude of different sections result from definite sequential and coordinated functions of the heart beat. Changes in any part of the curve can be related to malfunction in a specific section of the heart. Since the heart-beat rate is around 75 per minute, the time scale in Fig. 5 is 0·8 seconds. For e.c.g. measurements the recognized standard scan speed is 2·5 cm/s, and the detected electrical signal is about 1 mV. Clearly a long persistence screen is necessary (e.g. GM) to obtain a readable trace. An instrument bandwidth of 1 kHz is ample, and therefore provided it has a low power consumption, a low voltage distributed p.d.a. tube operating at 3·5 kV will give perfectly satisfactory performance. In practice a 0·7 W cathode/heater assembly is acceptable (11·0 V and 0·068 A) and this voltage can be obtained from a series of Cd-Ni rechargeable batteries. Tube size is a matter of choice but to keep weight and dimensions down to a minimum for easy portability, scans of $5 \times 7$ cm are a reasonable compromise.

Since the required bandwidth is small, it is possible to use ordinary magnetically deflected monitor (television type) tubes, but these are more suitable for larger displays in the constant surveillance installations. Again time base speeds are up to 10 cm/s and several channels are desirable so that more than one parameter can be displayed simultaneously.

In cases where permanent records need to be taken over a considerable period of time, recording on direct print paper gives immediate results due to the fact that the print is developed automatically by a few seconds exposure to ultra-violet light. Other advantages are that the paper is much cheaper than film and the trace can be viewed by the unaided eye. Normally in photography, a lens is used to focus the object on the recording medium and this is a very inefficient 'light coupler'. To overcome the problem a cathode-ray tube with a fibre-optic faceplate[18] can be used (e.g. M-O.V. 700F). By this means the luminescent trace in the phosphor is 'guided' to the front surface of the glass and a contact print made. The spectral response of the phosphor has to match that of the recording paper. In practice BE (P11) screens are quite suitable. Tubes are now available with a full $7·5 \text{ cm} \times 12·5 \text{ cm}$ fibre-optic stack, but are expensive due to the cost of the latter. For the best optical coupling the numerical aperture (n.a.) of the stack should be high, and values from 0·45 up to 0·86 are practical. The diameter of the individual fibres is usually 50-100 μm, but for high resolution tubes, this must be reduced considerably (e.g. 5 μm) to avoid degrading the spot size. A typical 'photographic' writing speed for fibre-optic tubes recording on paper is 0·15 cm/μs. This is low compared with results for fast film (Table 1) but nevertheless is preferable in this application for the reasons already given. Fibre-optic tubes can of course be used



Fig. 5. Electrocardiographic waveform.

with film to obtain the high photographic writing speeds.

With regard to contact printing, attempts have been made to eliminate the luminescent screen of a tube altogether and allow the electron beam to penetrate the face and excite a photographic emulsion directly†. The face now has to be a very thin vacuum tight membrane. Whilst the principle worked satisfactorily for stationary film, the practical application to moving film, introduced difficulties in maintaining the film close enough to the mica window to prevent the electrons from being scattered in the air and at the same time not to graze the mica and fracture it. These problems however, can be overcome by careful mechanical design.

A completely different technique for medical diagnosis is that of using very short (microsecond) pulses of ultrasound (1·5–5 MHz) on the body, similarly to the way in which electromagnetic pulses are used in radar (see Sect. 5). A small radiating head transmits the pulse and receives echoes from tissue boundaries, bone, or foreign bodies. The signals are displayed on an oscilloscope as an amplitude–distance graph ('A' scan presentation). Alternatively by scanning an area of the body and using the return signal to modulate the intensity of a cathode ray tube beam, a form of p.p.i. display ('B' scan presentation) may be achieved. Again an electrostatically-deflected tube, with a shortish persistence screen (e.g. GH) operated at a medium voltage (6kV) is satisfactory.

Another means of detecting local changes inside the human body depends on the physiological effect that the surface temperature of the body in that particular region becomes a few degrees different (e.g. 4 deg C) from its surrounding areas. These positions can be detected by an infra-red camera and the resulting signals displayed either in 'A' or 'B' scan format on a cathode-ray tube. In the latter case, for a 13 cm tube having an LC screen, a 100 line raster refreshed twice per second is used. Thus a low voltage electrostatically-deflected tube meets the requirement.

## 7. Control and Automation Systems

Many of the equipments already discussed in earlier Sections can be used for control and automation purposes in industry. The relevant oscilloscope (Sect. 2) can be used to monitor a machine operation, or to check the quality of electrical components, whilst visual terminal units (Sect. 3) may be used to communicate with a computer which is controlling an industrial process. Already this type of installation is working in electricity generating plants, steel works, and large chemical plants. The basic requirement is to present information on a large number of concurrent and widely scattered processes to a central control position on command. Hitherto, simple indicators such as lamps or meters, which are invariably limited to monitoring single operations, have been used, but the cathode-ray tube display system is highly flexible and can cope with it all. Moreover, the data are presented in an easily assimilated manner. Even in automated control systems the facility of human intervention is necessary. Thus the visual display which supplies this facility is bound to stay.

Naturally in any automated control system the television display cannot be ignored, and closed circuit installations are frequently used to control motor traffic and railway marshalling yards, as well as many other surveillance duties. There is a very wide range of tubes from 21 cm diagonal up to 65 cm, mostly electrostatically-focused and magnetically-deflected, and employing the W(P4) luminescent screen. Deflexion angles up to 110° are available, but for the better overall presentation required for picture monitors, deflexion angles of 70° or less are preferred. Colour tubes are available, the main sizes being 48 cm and 63 cm. For large screen television there are colour as well as black and white projection systems. The former employ three separate tubes, with individual red, green, and blue luminescent screens. In order to obtain satisfactory luminance such tubes have to operate at potentials up to 50 kV. High quality television monitor tubes are also used in a variety of purposes such as for guided weapon control.

A particular control application not yet discussed is that of the train describer system used by British Railways in their centralized signal boxes.[19] Here a small rectangular tube (M-O V 700E) is used for presenting a four-digit display, which represents a particular train on a given section of line. The digits are drawn by selecting requisite pulses on a 5 × 7 matrix (see Sect. 3). Essentially such a device needs to give a very bright trace with good contrast, since it often has to be viewed in bright sunlight. For safety reasons it is desirable to keep the final operating potential low (e.g. 2 kV) and therefore a large beam current is necessary (e.g. 1 mA). This will give rise to an appreciable line width (1·5 mm) but in this context such a value is quite acceptable. Even under these conditions, lives exceeding 10 000 hours are achieved using the GJ (P1) screen.

Another application in which the cathode-ray tube has enabled a vastly improved performance to be attained is that of high-speed printing. A programmed computer indicates the text to be printed, and the characters are selected on a monoscope tube (see Sect. 3). The requisite signals are then transferred to modulate the grid of a special printing tube which is capable of producing a very high resolution spot.

† Described in unpublished work on electronography by C. J. Emberson.

The face of this tube consists of an array of wires sealed into it at a a density of 10 000 per cm$^2$. The beam current arriving at any particular wire depends on the actual character shape being 'written', and this current deposits a charged trace on a recording paper moving across the surface. This is subsequently 'developed' by dusting the paper with a special powder and then 'fixed' by gentle heating. The cathode-ray tube needs a spot size of about 0·2 mm and therefore is of the magnetically focused and deflected type operating at a potential of 25 kV.

In the previous Section the application of ultrasonics to medicine was discussed. However, the same technique can be used to control the homogeneity of precast materials, and to determine the presence of unseen, internal faults. The display equipment is basically the same as that listed in Section 6.

## 8. Concluding Remarks

The foregoing considerations attempt to cover the more important applications for the cathode ray tube. It is obviously necessary to continue technological improvements in their performance. In this context, the advent of the 'mesh tube' was a major breakthrough. In another area it was once thought inevitable that fine particle phosphors would be less efficient than coarser versions. This has now been disproved, and successful practical results achieved. Long persistence screens are associated with high 'burn' rate but now the question of the 'sodalites', a form of dark trace screen, is under investigation.[20] Already there have been some novel approaches to the contrast problem.[21] These are but a few examples, indicating how technology has and may improve. Thus, the future of the cathode-ray tube as a display device of mass data is exciting and expanding.

## 9. References

1. McFarlane, A. B., 'Instrument cathode ray tubes', 'Electronic Data Library', Vol. 3, p. 4 (Morgan-Grampian, London, 1969).

2. McFarlane, A. B., 'Recent advances in the design and use of cathode-ray tubes', Colloquium, I.E.E. Electronics Division, April, 1964.

3. Allard, L. S., 'Instrument cathode-ray tubes', G.E.C. Journal, 27, No. 3, p. 145, 1960.

4. 'Computer Display Review' Vol. I, II & III (Adam Associates).

5. Atkins, E. S., 'Alphanumeric displays', 'Electronic Data Library', Vol. 3, p. 21 (Morgan-Grampian, London, 1969).

6. 'Cathode-Ray Tube Displays', I.E.E. Colloquium. May, 1969.

7. Poole, H., 'Fundamentals of Display Systems' (Spartan Books, Washington, 1966).

8. Parslow, R. D., 'Computer Graphics, Techniques & Applications', (Plenum Press, New York, 1969).

9. Freeman, M. H., 'Head-up displays', Optics Technology, 1, p. 3, February 1969; p. 175, August 1969.

10. U.K. Symposium on Electronics for Civil Aviation. London, September, 1969.

11. Woodley, W. A. and Roger, D., 'High resolution and fibre-optic cathode-ray tubes', Brit. Commun. Electronics, 10, No. 9, p. 696, September, 1963.

12. Cluley, J. C., 'The reliability of electronic systems', The Radio and Electronic Engineer, 31, No. 2, p. 110, February 1966.

13. Fryer, S. J. and Wanless, D., 'Radar tubes and driving circuits', 'Electronic Data Library', Vol. 3, p. 45 (Morgan-Grampian, London, 1969).

14. British Pat. Application No. 29755/68: also Styles, D., 'Flashover protection', The Radio and Electronic Engineer, 38, No. 5, p. 292, November 1969 (Letter).

15. Boyland, D. A. and Norman, D. J., 'A two-colour cathode-ray tube', G.E.C. Journal, 35, No. 2, p. 77, 1968.

16. McFarlane, A. B., 'The Secondary Emission of Luminescent Screens'. M.Sc.(Eng.) Thesis, London University, 1955.

17. Allard, L. S., 'High resolution cathode-ray tubes,' Medical & Biological Illustration, 14, No. 3, p. 184, July 1964.

18. Allard, L. S., 'The application of fibre-optic faceplates in cathode-ray tubes', Industrial Electronics, 2, No. 6, p. 273, June 1964.

19. Bental, L. J., Evans, G. S., and Tomkins, A. J., 'A computer controlled train describer', The Radio and Electronic Engineer, 38, No. 12, p. 361, December 1969.

20. Taylor, M. J., Marshall, D J., Forrester, P. A. and McLaughlan, S. D., 'Colour centres in sodalites and their use in storage displays', The Radio and Electronic Engineer, 40, No. 1, p. 17, July 1970.

21. Davis, J. A., 'Recent Advances in Cathode-Ray Tube Display Devices', Proceedings Symposium sponsored by the Electronics Research Centre, N.A.S.A., Cambridge, Mass, 1967. p. 25.

# Multiple Stable State Merging
## A Practical Approach to the Design of Asynchronous Sequence Detectors and Similar Circuits

*By*

**D. T. MARKEY,** B.Sc. (Eng.), M.Tech., C.Eng., M.I.E.R.E.†

A practical method for reducing and merging the primitive flow table (or state table) of sequential machines is detailed. The method is particularly suited to more complex problems that are to be solved manually. By reducing the size of the primitive flow table at a very early stage in the design, the problems of reduction and merging are considerably simplified for certain types of machine. An example is included to show the use of the method and for comparison with the standard merging technique. This paper only considers the application to asynchronous machines.

## 1. Introduction

A sequential machine may be represented by the schematic diagram of Fig. 1, in which a general input to the system is designated as $x_i$ and a general output is designated as $Z_i$. A feedback, or secondary state variable, is shown as $Y_i$ (the next state) and as $y_i$ (the present state). In the general case, which is the Mealy model,

$$Y_i = f_1(x_1, x_2, \ldots x_n, y_1, y_2, \ldots y_m)$$

and

$$Z_i = f_2(x_1, x_2, \ldots x_n, y_1, y_2, \ldots y_m)$$

A simpler sequential machine is the Moore model in which

$$Y_i = f_3(x_1, x_2, \ldots x_n, y_1, y_2, \ldots y_m)$$

and

$$Z_i = f_4(y_1, y_2, \ldots y_m).$$

For a synchronous implementation of a sequential machine the feedback paths must contain storage and the inputs (x) and feedback variables (y) are clocked or strobed. In an asynchronous implementation of a sequential machine no storage is normally used and the feedback paths are not clocked but are continuously connected.

With a synchronous machine all the secondary states are stable; this is ensured by the storage and clocking used. An asynchronous machine, however, has stable states and unstable or transient states which occur during transition. A stable state is defined by the secondary state coding $(y_1, y_2, \ldots y_m)$ which it has been assigned. If states are 'merged' then a 'total state' is obtained, which is defined both by the secondary state coding and by the input coding $(x_1, x_2, \ldots x_n)$.

The standard manual approach to asynchronous logical design begins with the construction of a primitive flow table. This is a table displaying the required stable states of the design, one only per line, in terms of the input conditions; the output states may

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

also be shown on the table. It is usually helpful, in deciding on the stable states and the transitions, to consider simultaneously a state diagram and a primitive flow table. The state diagram is a pictorial version of the primitive flow table and is useful in developing the design. This paper shows that it may also be useful in considering possible mergers that may be made in the number of stable states.
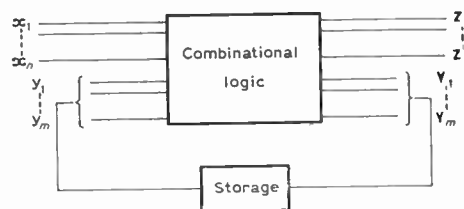


**Fig. 1.** Schematic of sequential machine.

When the primitive flow table is completed the next stage is to reduce this table, where possible, by removing redundancies. Next, rows are merged, where possible, so that the total number of rows is reduced to a minimum. Then a secondary state assignment is made and finally equations are extracted for the secondaries and for the output functions.

## 2. Flow Table Reduction

The primitive flow table has as many rows as there are stable states in the design. For a flow table having $r$ rows the number of secondary variables $s$ required is given by $r \leqslant 2^s$. The object of flow table reduction is to reduce the number of secondary variables, and hence the amount of storage and/or gating required to implement the design. The reduction is accomplished by locating redundancy in the number of stable states; this redundancy is caused by the fact that some of the stable states may be identical (or are equivalent in the case of incompletely specified flow tables). Reduction due to this cause has been

extensively covered in the literature.[1,2] In general, asynchronous designs have incompletely specified flow tables, since only one input change is normally allowed at a time and thus there are a great many 'don't cares' in the primitive flow table. Equivalent states are determined in an asynchronous machine (i) by having the same output states, (ii) by having the same input conditions (i.e. being in the same column), and (iii) by having next states which are the same for all possible input changes.

## 3. Flow Table Merging

The object of merging, as for reduction, is to reduce the number of rows in the flow table. Merging of any two rows is possible if there are no conflicting states, stable or unstable, in the rows to be merged. Unless a Moore model design is being attempted merging may take place regardless of the output states; 'don't care' conditions cannot cause confliction and so aid possible mergers. One method of determining the best merger is to use a merger diagram[3,4] but with designs having more than say 12 stable states, the number of possible mergers between rows may become very large, and the problem of which mergers to make may be too complex for the merger diagram method to solve. It is also possible to determine the best merger by using Boolean algebra, but a computer aided approach is required for larger designs.[5]

## 4. Multiple Stable State Merging

A multiple stable state is one which is defined by the secondary state coding and also by *more than one* input coding.

This method of merging and reduction uses the possibility that some total states, though distinct, may not need to be distinguished in a particular design. Thus two or more total states in the primitive flow table may be allocated one multiple stable state and this may be shown as a single state in the state diagram. Using this method the state diagram may often be greatly simplified.

There are many examples of design where this approach can be of use, but particularly the method is useful in sequence detectors and similar circuits where many conditions which are not required as outputs must be held. Circuits such as counters and shift registers normally require that all states be held distinctly and so multiple stable states cannot be formed in these cases.

Sequence detector circuits typically have a state diagram which clearly shows a flow path giving the required response, and in addition many other paths which either return to the starting state or else loop back to previous states. Frequently there is also a direct flow path for resetting the detector. The direct

detecting and resetting flow paths are the *main flow paths*. All the other paths are between directed or waiting states and are *subsidiary paths*. Sequence detector circuits often have a great number of subsidiary paths and states and it is between these that multiple stable states may often be formed. It is usual in designing sequence detectors to complete the main flow path first since this corresponds to the correct input sequence being applied. This path should be brought out clearly in the state diagram and then the subsidiary paths should be completed (see Figs. 5 and 6).

### 4.1 Preferred Conditions for Forming Multiple Stable States

(i) The stable states to be eliminated must have no vital function in the implementation, i.e. the states must be subsidiary states and must not be on the main flow path of the state diagram.

(ii) States to be grouped should have a bidirectional (for two states) or a looped (for more than two states) flow path connecting them.

(iii) The output states must be the same, this requirement is the same as for a Moore model machine.[3,4]

(iv) Multiple unstable states must also exist in the flow table, where necessary, so that inaccessible states do not occur.

(v) Multiple grouping of states may allow more than one input change to occur simultaneously where the transitions are included in a single multiple state.

Note that the conditions (i) to (iii) above are not mandatory, they only help in obtaining 'good' mergers that will not cause the final merging to be far from optimum. The only mandatory condition is that no conflicting states may be merged into one row.

### 4.2 Recognition of Possible Multiple Stable States

Clearly the easiest way of recognizing multiple states is to use the state diagram and to form the reduced flow table from the state diagram. In many designs it is possible to eliminate a great many states and the question arises how far should the reduction go. If extreme pre-merging is applied, very little final merging, if any, will be possible and the machine will often be designed as a Moore model. Very often the final design equations will then be the same as for
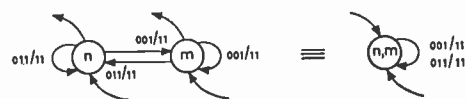


Fig. 2. Combination of two states.

D. T. MARKEY

optimum Moore merging. It is possible to combine any two states such as those in Fig. 2, which are not on the main flow path and which have no other conflicting unstable conditions. The greatest advantage from the method is obtained when states connected in triangles and 'boxes' are assigned a single state (Fig. 3).

All the states to be grouped need not be off the main flow path, e.g. in Fig. 4, states 5 and 6 may often be grouped into state 3 if the outputs in states 3, 5 and 6 are the same.
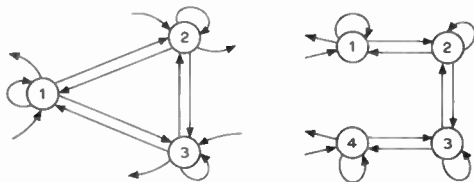


Fig. 3. Single state assignment.
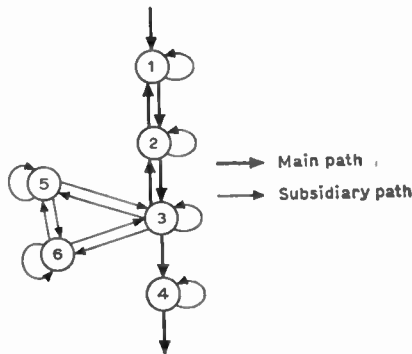


Fig. 4. Introduction of subsidiary flow path.
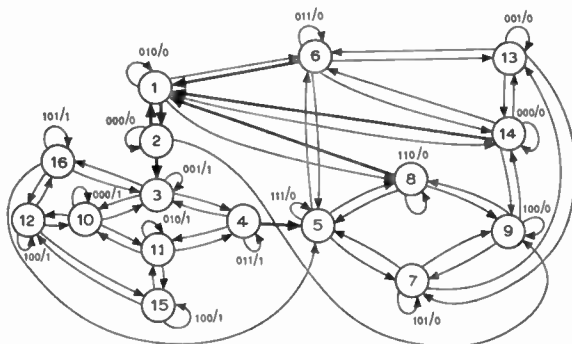


Fig. 5. Standard state diagram for sequence detector.

## 5. Example of Reduction Using the Multiple Stable State Technique

*Problem*. Design a sequence detector, such that with an input sequence 010→000→001 the output Z becomes 1. An input of 111 is required to reset the output to 0.

No equivalent state reduction is possible with this example which has 16 stable states, three input variables and one output variable. In Fig. 5 is shown the standard state diagram for this problem with the main paths brought out. The corresponding primitive flow table is shown in Table 1.

**Table 1.**

| 000 | 001 | 011 | 010 | 110 | 111 | 101 | 100 | Z |
|-----|-----|-----|-----|-----|-----|-----|-----|---|
| 2 | — | 6 | ① | 8 | — | — | — | 0 |
| ② | 3 | — | 1 | — | — | — | 9 | 0 |
| 10 | ③ | 4 | — | — | — | 16 | — | 1 |
| — | 3 | ④ | 11 | — | 5 | — | — | 1 |
| — | — | 6 | — | 8 | ⑤ | 7 | — | 0 |
| — | 13 | ⑥ | 1 | — | 5 | — | — | 0 |
| — | 13 | — | — | — | 5 | ⑦ | 9 | 0 |
| — | — | — | 1 | ⑧ | 5 | — | 9 | 0 |
| 14 | — | — | — | 8 | — | 7 | ⑨ | 0 |
| ⑩ | 3 | — | 11 | — | — | — | 12 | 1 |
| 10 | — | 4 | ⑪ | 15 | — | — | — | 1 |
| 10 | — | — | — | 15 | — | 16 | ⑫ | 1 |
| 14 | ⑬ | 6 | — | — | — | 7 | — | 0 |
| ⑭ | 13 | — | 1 | — | — | — | 9 | 0 |
| — | — | — | 11 | ⑮ | 5 | — | 12 | 1 |
| — | 3 | — | — | ⑯ | 5 | — | 12 | 1 |

In Fig. 6 six stable states have been eliminated by making stable states 3, 4, 6 and 7 multiple, using the extra limitation that only one change may occur at a time inside a stable state. In Fig. 7 reduction has been
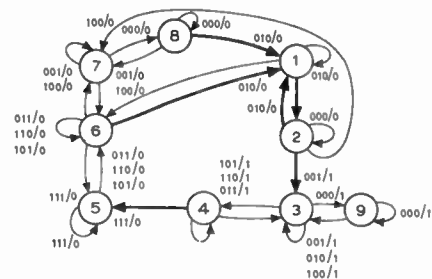


Fig. 6. Elimination of six stable states from Fig. 5.
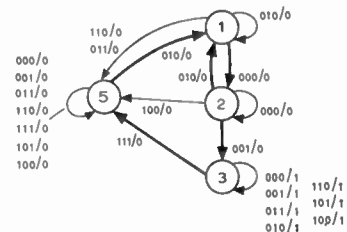


Fig. 7. Final reduction of Fig. 5.

302

carried to the maximum possible, only states on the main path remaining. The flow tables for these reduced state diagrams are given in Tables 2 and 3 and the corresponding merger diagrams in Figs. 8 and 9 respectively.

## Table 2.

| 000 | 001 | 011 | 010 | 110 | 111 | 101 | 100 | Z |
|---|---|---|---|---|---|---|---|---|
| 2 | — | 6 | ① | 6 | — | — | — | 0 |
| ② | 3 | — | 1 | — | — | — | 7 | 0 |
| 9 | ③ | 4 | ③ | 4 | — | 4 | ③ | 1 |
| — | 3 | ④ | 3 | ④ | 5 | ④ | 3 | 1 |
| — | — | 6 | — | 6 | ⑤ | 6 | — | 0 |
| — | 7 | ⑥ | 1 | ⑥ | 5 | ⑥ | 7 | 0 |
| 8 | ⑦ | 6 | — | 6 | — | 6 | ⑦ | 0 |
| ⑧ | 7 | — | 1 | — | — | — | 7 | 0 |
| ⑨ | 3 | — | 3 | — | — | — | 3 | 1 |

## Table 3.

| 000 | 001 | 011 | 010 | 110 | 111 | 101 | 100 | Z |
|---|---|---|---|---|---|---|---|---|
| 2 | — | 5 | ① | 5 | — | — | — | 0 |
| ② | 3 | — | 1 | — | — | — | 5 | 0 |
| ③ | ③ | ③ | ③ | ③ | 5 | ③ | ③ | 1 |
| ⑤ | ⑤ | ⑤ | 1 | ⑤ | ⑤ | ⑤ | ⑤ | 0 |
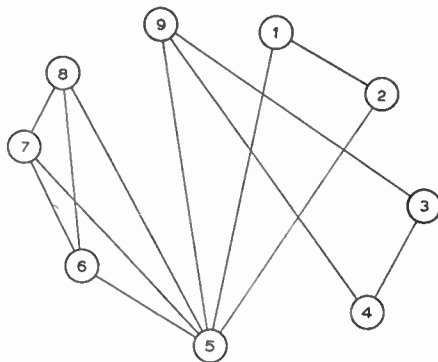


Fig. 8. Merger diagram.

From Fig. 8 the partition giving the best merger and assignment is (3, 4, 9) (1, 2, 5) (6, 7, 8); from this partition the merged flow table is shown in Table 4 and the assigned flow table in Table 5.

## Table 4.

| ⑨ | ③ | ④ | ③ | ④ | 5 | ④ | ③ |
|---|---|---|---|---|---|---|---|
| ② | 3 | 6 | ① | 6 | ⑤ | 6 | 7 |
| ⑧ | ⑦ | ⑥ | 1 | ⑥ | 5 | ⑥ | ⑦ |

## Table 5.

| $y_1 y_2$ | $x_1 x_2 x_3$ 000 | 001 | 011 | 010 | 110 | 111 | 101 | 100 |
|---|---|---|---|---|---|---|---|---|
| 0 0 | 00 | 10 | 01 | 00 | 01 | 00 | 01 | 01 |
| 0 1 | 01 | 01 | 01 | 00 | 01 | 00 | 01 | 01 |
| 1 1 | × | × | × | × | × | × | × | × |
| 1 0 | 10 | 10 | 10 | 10 | 10 | 00 | 10 | 10 |
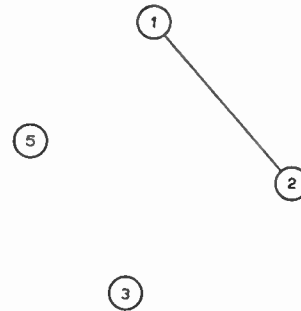


Fig. 9. Final merger diagram.

When the hazards have been removed the implementing equations simplify to:

$$Y_1 = \bar{x}_1 y_1 + \bar{x}_2 y_1 + x_2 \bar{x}_3 y_1 + \bar{x}_1 x_2 x_3 y_2$$
$$Y_2 = \bar{x}_2 y_2 + x_1 \bar{x}_2 \bar{y}_1 + x_1 \bar{x}_3 \bar{y}_1 + \bar{x}_1 x_2 x_3 \bar{y}_1 + \bar{x}_1 x_3 y_2$$
$$Z = y_1.$$

These equations require 11 NAND gates (plus 5 invertors) to implement. For the extreme case of state removal (Fig. 7) the merger diagram is extremely simple and is shown in Fig. 9, giving the partition (1, 2) (3) (5). However, a better assignment is possible for (1) (2) (3) (5), which is the unmerged partition. The assigned flow table is shown in Table 6.

## Table 6.

| $y_1 y_2$ | $x_1 x_2 x_3$ 000 | 001 | 011 | 010 | 110 | 111 | 101 | 100 |
|---|---|---|---|---|---|---|---|---|
| 0 0 | 01 | × | 10 | 00 | 10 | × | × | × |
| 0 1 | 01 | 11 | × | 00 | × | × | × | 10 |
| 1 1 | 11 | 11 | 11 | 11 | 11 | 10 | 11 | 11 |
| 1 0 | 10 | 10 | 10 | 00 | 10 | 10 | 10 | 10 |

After simplification the hazard-free implementing equations are as follows:

$$Y_1 = x_3 + x_1 + y_1 y_2 + \bar{x}_2 y_1$$
$$Y_2 = \bar{x}_1 y_1 y_2 + \bar{x}_1 \bar{x}_2 \bar{y}_1 + \bar{x}_2 y_1 y_2 + x_2 \bar{x}_3 y_1 y_2 + x_1 \bar{x}_3 y_1 y_2 + \bar{x}_1 \bar{x}_2 y_2$$
$$Z = y_1 y_2.$$

These equations again require 11 NAND gates and 5 invertors to implement.

## 6. Discussion

Merging by means of assigning multiple stable states is often possible by examination of the state diagram. The indication of the possibilities of multiple stable state merging is more clear if the main flow paths on the original state diagram are brought out. If multiple stable states are formed then a considerable reduction in the size of the flow table and in the merger diagram is possible for certain types of problem. The reader will be convinced of this if the merger diagram for the original primitive flow table (Table 1) is constructed and the result is compared with Figs. 8 or 9. It will be seen that there are 55

mergeable row-pairs in this merger diagram, a problem not easily soluble by the merger diagram method.

If care is taken in the choice of multiple stable states, i.e. the conditions outlined in Section 4.1 are strictly applied, then the final implementing equations are optimum or near optimum.

After partial merging by assigning multiple states, the residual problem of merging is enormously simplified, enabling much more complex problems to be tackled manually, although it is not possible to be sure of obtaining an optimum solution by the method. In the example given there are 16 states, with three stable states in column 1, thus an absolute minimum of three rows and two secondaries will be necessary after merging; this solution is rapidly obtained by the methods used. If the merging on the flow diagram is not carried out carefully then a final partition which is far from optimum may be obtained. Merging states on the main flow path is likely to give an incorrect solution but if incorrect mergings have been used in forming multiple states, these should become apparent when the merged state diagram or flow table is constructed.

Any redundancy in the original primitive flow table is very likely to be eliminated during the process of multiple stable state merging, as redundant states would be very likely to fulfil the conditions of Section 4.1. This further simplifies the design process as state reduction may not then be necessary and it is often possible to determine this from the merged state diagram. For example in the state diagram of Fig. 7, there could not be any redundancy as every state is on the main flow path.

## 7. Conclusions

To summarize, the method of multiple stable state merging is as follows:

(i) Form the state diagram.
(ii) Outline the main flow paths on the state diagram.
(iii) Consider possible multiple states under the conditions of Section 4.1.
(iv) Form the partially merged state diagram.
(v) Form the partially reduced and merged flow table.
(vi) Continue with merging and state assignment.

The type of problem where this method is likely to be of some use is for sequence detectors and similar circuits which have:

(i) a main flow path which only includes a small proportion of the total number of states in the design, and
(ii) where the number of outputs is small compared with the number of inputs.

This method is proposed as a practical manual approach to a fairly complex design of this type.

For very complex designs a more drastic pre-merging may be necessary in order that any solution may be obtained in a reasonable time. For less complex designs the rules may be followed more closely so that a solution very near optimum is found. Thus an approach suited to the particular design being attempted is called for enabling a practical solution to be obtained.

The most convenient method of designing asynchronous logical systems is to use computer-aided design algorithms where these are available. For a summary of these possibilities see references 5 and 6. One of the practical disadvantages of using c.a.d. is that the designer often loses the 'feel' for the problem being tackled and thus a minimal solution may be lost. Unlike c.a.d. methods the multiple stable state merger method requires considerable 'feel' from the designer before its value becomes apparent. It is in the *preparation* of the original state diagram that the method may best be used to save a considerable amount of work, i.e. states that do not conflict and which are not on the main flow path may be allotted a single state whilst the problem is being considered and before the primitive flow table is prepared. With this approach a partially merged flow table is the next step after the state diagram.

The main limitations of the method are:

(i) It is only applicable to a certain class of sequential circuits.
(ii) A considerable 'feel' for this method and for asynchronous design is required.
(iii) It can only extend the range of manual problems slightly.
(iv) It is not applicable if the asynchronous version of full reduction is to be used instead of merging.[2]

## 8. References

1. Paull M. C. and Unger, S. H., 'Minimizing the number of states in incompletely specified sequential switching functions', *I.R.E. Trans. on Electronic Computers*, EC-8, No. 3, pp. 356–67, September 1959.
2. Hill, F. J. and Peterson, G. R., 'Switching Theory and Logical Design', (Wiley, New York, 1968).
3. Lewin, D. W., 'Logical Design of Switching Circuits', (Nelson, Edinburgh 1968).
4. Kreiger, M., 'Basic Switching Circuit Theory', (Macmillan, New York, 1967).
5. Markey, D. T., 'A Critical Study of Asynchronous Logic Design'. M. Tech. Dissertation, Brunel University, 1970.
6. Lewin, D. W. and Waters, M. G., 'Computer aids to logic system design', *Computer Bull.*, 13, No. 11, pp. 382–8, November 1969.

# An Ultrasonic Position Sensor for Automatic Control

*By*

**W. P. WILLIS,** B.E., M.E.†

and

**Professor L. KAY,**

Ph.D., C.Eng., F.I.E.R.E.‡

An instrument has been designed and experiments carried out to determine the resolution which may be possible using airborne ultrasonics as a sensing medium. The experiments were conducted using one transmitter and two receivers from which it is shown that a resolution of the order of a 3 mm cube is practicable using a single wide angle radiator and a surrounding cluster of 3 wide angle receptors.

## 1. Introduction

In an automated factory it is often necessary to know the moment a component passes through a specified position in space. Sometimes it is required that this be known fairly accurately without physically touching the component and optical systems are in common use for the purpose. An alternative method which may at times have advantages is ultrasonic echo-location using pulses which are sufficiently short to control azimuth as well as range resolution. This is an improvement on an earlier instrument described by one of the authors.[1] Physically small radiating and receiving transducers which have a wide-angle polar response under continuous wave operation may then be used.

## 2. Range Bounds

With the use of very short transmission pulses (5 µs) it is possible to operate a range gate (5–30 µs say) which passes an echo only if this originates from a very short annulus of range (1·5–9 mm). This annulus is in the form of an ellipse having an inner and outer bound given by:

inner bound

$$\frac{\left(x - \frac{L}{2}\right)^2}{\left(\frac{Dc}{2}\right)^2} + \frac{y^2}{\left(\frac{Dc}{2}\right)^2 - \left(\frac{L}{2}\right)^2} = 1$$

outer bound

$$\frac{\left(x - \frac{L}{2}\right)^2}{\left\{\frac{(D+d)c}{2}\right\}^2} + \frac{y^2}{\left\{\frac{(D+d)c}{2}\right\}^2 - \left(\frac{L}{2}\right)^2} = 1$$

both equations being of the form

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$$

---

† New Zealand Post Office, Wellington; formerly of the Department of Electrical Engineering, University of Canterbury, Christchurch.

‡ Department of Electrical Engineering, University of Canterbury, Christchurch.

where $D$ = echo delay
  $d$ = annulus width
  $L$ = spacing between transmitter and receiver
  $c$ = velocity of propagation in air
$x$ and $y$ = range and azimuth co-ordinates.

Short sections of these bounds are sketched in Fig. 1 for each receiver relative to the common transmitter.

## 3. Azimuth Bounds

Each receiver channel receives pulses of duration $k$. Passing these through an AND gate imposes the restriction that there will be an output only if the time for an echo pulse to reach one receiver differs from the time for it to reach the other receiver by less than $k$, since the distance from the transmitter to the reflecting object is common. The path difference must therefore be less than $kc$. The boundary thus imposed is the locus of a point which moves so that its distances from two fixed points (T and R) differ by $kc$, a constant. This is the bifocal definition of a hyperbola whose equation has the general form
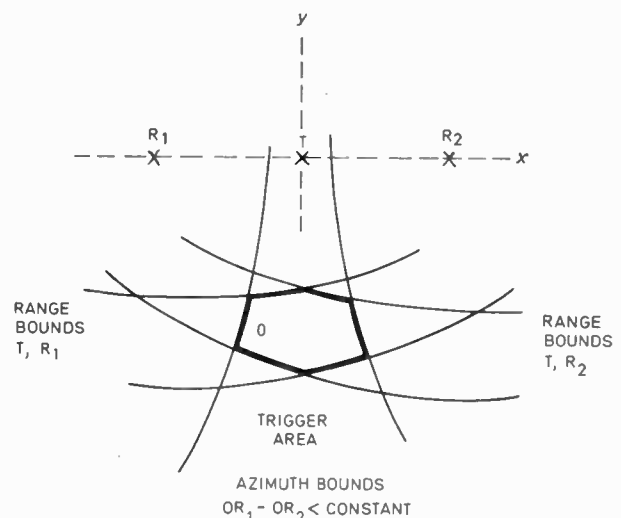


Fig. 1. Range and azimuth bounds.

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$$

The equation defining the system azimuth bounds is found to be

$$\frac{x^2}{\left(\frac{kc}{2}\right)^2} - \frac{y^2}{L^2 - \left(\frac{kc}{2}\right)^2} = 1$$

where $x$ and $y$ are as before.

Figure 1 also shows the two hyperbola resulting from the AND gate. It will be seen that the ellipsoidal bounds together with the two hyperbola form what may be called a 'trigger area'.

Any echo originating from within this area will pass through both the range gates in the two receivers and the common AND gate as shown in Fig. 2.

Thus both range and azimuth resolution are unambiguously determined by the pulse duration and the transducer spacing alone.
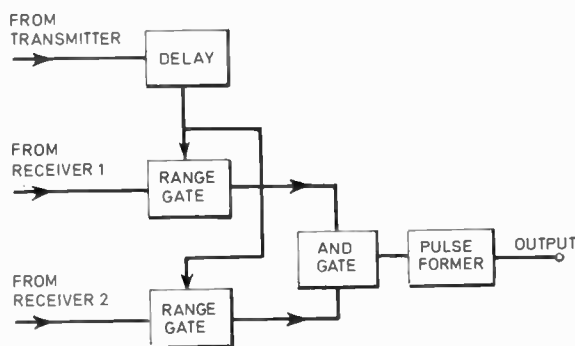
Fig. 2. Signal processor functional diagram.

## 4. Instrumentation

The key element in the system is the transducer which, through its impulse response, determines both the range and azimuth bounds of the resolved area. Both receiving and transmitting transducers were based on the design by Kuhl et al.[2] Recent work by Clark at Canterbury on further development of the solid dielectric transducer has shown how the response can be controlled in manufacture and the thickness reduced to approximately 2 mm. In the equipment being described the diameter was effectively 1 cm producing a beam width of 20° at 100 kHz. Differences in sensitivity of less than 1 dB are possible without excessive selection of pairs of units and the maximum absolute sensitivity of the unit as a receiver is of the order of −57 dB relative to 1 V/μbar at 80 kHz falling to −65 dB at 150 kHz. Operating as a transmitter with a d.c. polarization of 150 V, the maximum peak radiated power from these units was 1 mW.
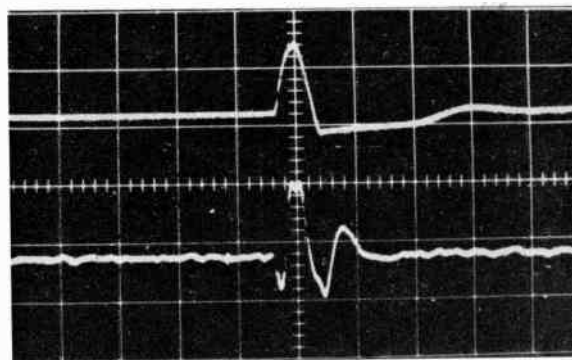
Fig. 3. (a) Electrical pulse at transmitter.
(b) Echo received from small sphere.

Due to the good coupling and the consequential wide bandwidth the transmitted acoustic waveform can be a pulse of only $1\frac{1}{2}$ cycles at 100 kHz with a peak power of 1 mW. Figure 3 shows the received echo from a small smooth sphere which is compared with the waveform applied to the transmitter.

The electronics consisted essentially of high gain band-limited amplifiers (50 kHz bandwidth centred on 100 kHz) to limit the effect of ambient noise in the air. Each of these was followed by a rectifier detector. The signal processor converted the rectified echo pulses into constant amplitude constant duration pulses which were fed to the range gates. Echoes coinciding with the opening of the range gates (range bounds) were fed to the AND gate (azimuth bounds). Thus only echo pulses originating from within the trigger area were seen at the AND gate output. These were further formed into a pulse suitable for operating control mechanisms such as a relay.

## 5. Resolution—Theoretical and Measured

We are here defining resolution as that area from which an echo may be received which would operate a mechanism, i.e. the trigger area. A simple computer program was written to determine the effect of varying parameters such as transducer spacing, range gate
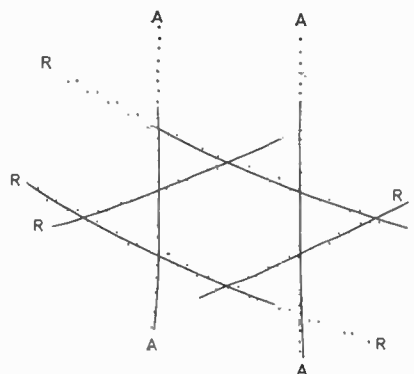
Fig. 4. Specimen output of program.

width, polar response, AND gate pulse length, and range, from which the trigger area shown in Fig. 4 is a typical print-out. An experimental measurement with the equipment is shown in Fig. 5. The effects of independently varying pulse duration, range gate width, the spacing of transducers and the range itself are shown in Fig. 6. The curves are the plots of experimental measurements but the results from the computer were within a few percent, and the error could be accounted for by the difference between the actual and the assumed velocity of propagation. Where a parameter is fixed its value was as follows:

| | |
|---|---|
| range from transmitting transducer | 15 cm |
| transducer spacing | 15 cm |
| range gate width | 30 μs |
| AND gate input pulse length | 30 μs |

It is simple to convert to another value. For example, to determine the resolution in azimuth for an AND gate pulse duration of 10 μs at a range of 10 cm, the value of 12 mm given by the azimuth curve (varying range) in Fig. 6 at 10 cm should be divided by 3, since the pulse duration for the figure is 30 μs. This gives an azimuth resolution of 4 mm.
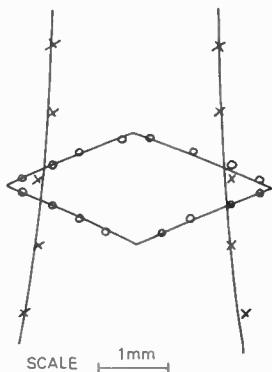


Fig. 5.
Experimental trigger area.

SCALE ├── 1mm ──┤

## 6. Three-dimensional Resolution

A third receiving transducer may be used to feed a third range gate, and the output of this can be applied to a second AND gate together with the output of the first AND gate. This produces the means for restricting the range and azimuth bounds in three dimensions, although the trigger volume will not be a simple shape. It will be self-evident however from Figs. 5 and 6 that a volume of the order of 3 mm cube will not be difficult to realize using the 5 μs pulses obtainable with the solid dielectric transducers.

## 7. Errors

The system is not error free; noise, air flow, atmospheric conditions, etc., will introduce errors. These, however, need not be significant under controlled conditions.

Ambient noise in the low-frequency range is reduced by filtering without seriously affecting the pulse duration. High frequency ambient noise is usually
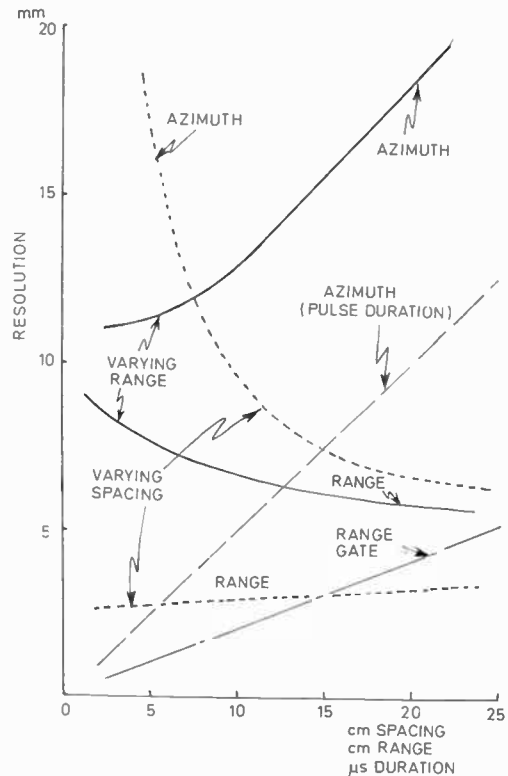


Fig. 6. Resolution vs. parameters.

directional and thin plastic guards can be placed between the system and the noise source. System self-noise is not significant since the signal level is so high. A 32 km/hr wind can cause a range error of up to 0·15 mm or an azimuth error of 4·5 mm. A change in temperature from 0°C to 50°C varies the velocity by approximately 9%.

Thus with reasonable precautions quite accurate and repeatable results can be obtained using the system.

## 8. Acknowledgment

## 9. References

1. Kay, L., 'Ultrasonic probe for sensing the position of objects', U.K. Patent No. 3,283,292, 1966.
2. Kuhl, W., Schodder, G. R. and Schroder, F. K., 'Condenser transmitters and microphones with solid dielectric for airborne ultrasonics', *Acustica*, 4, pp. 519, 1954.

# Contributors to this issue

**Mr. P. A. Payne** (G. 1965) received his technical education at Willesden Technical College and Borough Polytechnic, London. He has held appointments as an electronic engineer with G. and E. Bradley Ltd., E.M.I. Electronics Ltd. and K and N Electronics Ltd. In 1967 he joined the research staff of the dynamic analysis group in the Department of Mechanical Engineering and Engineering Production of the University of Wales Institute of Science and Technology. Mr. Payne is the author and co-author of a number of papers in the field of dynamic analysis and is currently completing the requirements for a Ph.D.

**Professor D. R. Towill** (F. 1970) was appointed to the Chair of Engineering Production in the Department of Mechanical Engineering and Engineering Production in the University of Wales Institute of Science and Technology at the beginning of 1970. He has recently been awarded the Clerk Maxwell Premium for a previous paper which was considered to be the most outstanding published in the *Journal* in 1969. A fuller note on Professor Towill's career appeared in October 1969.

**Professor Leslie Kay** (F. 1965) took his B.Sc. at King's College, Durham, in 1948, following which he spent 10 years with the Royal Naval Scientific Service on the development of sonar systems. In 1958 he joined the Department of Electrical Engineering at Birmingham University, where he continued research on underwater sound, and also commenced studies on air sonar for the blind and ultrasonics in solids. He took his Ph.D. in 1962 at Birmingham and in 1963 he became Head of the Electrical Engineering Department at Lanchester Polytechnic. In 1966 he was appointed to the Chair in Electrical Engineering at the University of Canterbury, Christchurch, New Zealand, where his studies of sonar systems in air, water and solids are continuing. Professor Kay is chairman of the Institution's New Zealand Advisory Council and he has contributed several previous papers to the *Journal*.

**Mr. Warren Willis** obtained his B.E. degree in electrical engineering at the University of Canterbury, New Zealand, in 1968. He gained his M.E. degree at the same University in the following year for his work on an ultrasonic sensing system. He is now working in the Engineer-in-Chief's Office of the New Zealand Post Office where he is mainly concerned with inductive interference between power and telecommunications systems.

**Mr. J. H. Gunton** graduated in physics at Hertford College, Oxford, in 1967. Since then he has been working in the Tube Investments Research Laboratories at Hinxton Hall near Saffron Walden, Essex. His main research has been concerned with the visualization of ultrasound and other problems relating to non-destructive testing.

**Mr. K. J. Crook** obtained a B.Sc. at Nottingham University and in 1956 he joined EMI Electronics Ltd. working on the design of EMIDEC 2400. When the EMI Computer Division was merged with International Computers and Tabulators Ltd. he became joint logic designer for the 1902/3 computer. He was in charge of logic design for the 1901 computer and mathematical attachments for smaller 1900 processors. Mr. Crook is now managing a group of projects in ICL at Stevenage including design automation development, software associated with testing of logic assemblies, and graphic aids for use in engineering development.

**Miss Joan Blythin** graduated in 1963 from Liverpool University with a B.Sc. degree in mathematics and has since been employed at the development laboratories of ICL at Stevenage. She has worked on various software projects associated with the design and manufacture of computers. These include investigations of wiring optimization techniques, applications of computer graphics and techniques for testing ogic arrays.

# A Computer Controlled Tester for Logic Networks and a Method for Synthesizing Test Patterns

*By*

**K. J. CROOK,**

B.Sc.†

and

**Miss J. BLYTHIN,**

B.Sc.†

This paper describes a total test system for logic networks as manufactured for digital electronic equipment. The system provides a means of physically testing the manufactured item together with a method for preparing test data automatically using a computer. The method uses sensitive path techniques with forward and reverse simulation of a software model of the network which is held in the computer store. The justification for automatic test synthesis lies in the large number of network types which are designed.

## 1. Introduction

The complexity of logic networks in digital equipment has long since precluded manual testing. The justification for automatic test equipment is summed up simply as 'the only reasonable test method'.

The paper describes one approach used within the authors' company for production testing of complex logic arrays, the prime features being a fully automatic test facility (excepting actually handling the printed circuit boards) and a method of synthesizing test patterns requiring minimum action on the part of the board designer. It is relevant to note that the output of one division of that company is of the order 150 board types annually and that the manufactured quantities run to many hundreds of thousands annually.

## 2. Logic Networks

The logic to be tested is formed by mounting dual-in-line bipolar integrated circuits (DILICS) on multilayer printed circuit boards (macroboards). Four sizes of board are available, the largest having 96 logic pins at its edge and holding up to 75 DILICS, i.e. about 180 logic elements, although the trend to medium-scale integration is raising this figure. In general the logic array is a random mixture of sequential logic (flip-flops, delays, pulse gates etc.) and passive gating logic. The singular case of only gating logic is a rare event and in any case is trivial so far as testing and synthesis are concerned.

## 3. The Test Equipment

To test a logic network it is necessary to check that for any pattern of signals presented at the inputs of the network, the network will produce the correct outputs according to the logic function of the board. Later sections of this paper deal with methods of generating

a sufficient set of such input patterns; here we are concerned with the facilities required in the tester, known as the TE110.

In logic generally, the signals used fall into two broad classes:

(1) Quasi-d.c. 'levels' such as, for example, the inputs to an adder and

(2) Pulses used for creating the sequential situation which characterizes digital equipment. The fact that the 'levels' would appear to an untrained observer as rather narrow pulses is irrelevent so long as the real pulses appear even narrower.

Since it is virtually impossible to constrain the board designer to use only certain pins for level inputs and certain pins for pulse inputs it follows that the tester must be capable of treating each pin on the macro-board as either an input for one of the above types of signal or as an output. In an earlier equipment this selection was achieved by providing an adaptor with each board type but with the proliferation of board types and test stations the economics of this approach became suspect. In the TE110, any one of the 96 signal channels can be set independently of the others to be either

(a) *a transition*, i.e. a signal which switches to a given state at a certain time and remains in that state for some relatively long time, or

(b) *a positive- or negative-going pulse* occurring at some time later than the transition, or

(c) treated as an *output* from the network.

Figure 1 represents a typical logic network. If the inputs A are activated at a time $t = 0$, then a time $t_1$ will elapse before the signals have propagated to the point C. This is the appropriate time to clock the flip-flops using pulses on inputs B. A further time elapses before this event reaches the network outputs at time $t_2$ which is the appropriate time to sample the

---

† Computer Systems Development, International Computers Limited, Stevenage, Herts.

**Fig. 1.** A typical logic network.
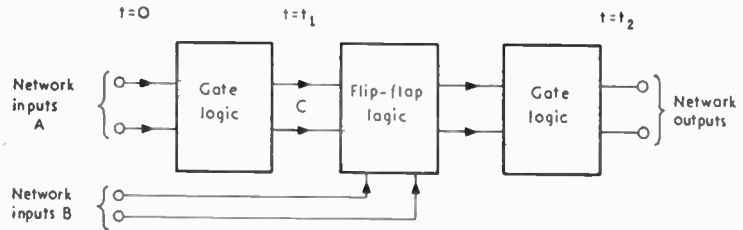
outputs. Ideally one would like independent control of the times $t_1$ and $t_2$ for each channel but this is a very expensive luxury. In TE110, all transitions occur at the same time $t = 0$ which is deemed to start the test. All pulses occur at a time $t_1$ which can be set from the computer at a value from 10 to 630 ns in steps of 10 ns. The pulse width is separately but similarly defined. The outputs are sampled at a time $t_2$ which is set from the computer at a value from 0 to 40 ms in steps of 10 ns. This control is achieved by loading a binary value of the time into a counter which is counted down at 100 MHz using a mixture of conventional counters and delay techniques.

Each channel (Fig. 2) is provided with five flip-flops. FF1 to FF4 are set up by the computer at the start of a test sequence. D is set up by the computer at the start of each test within a sequence. Four signal 'rails' are provided as shown. If FF1 is set, then the input is to be a transition whose phase is determined by the value of D. If D is set then switch S1 is closed, if not S2 is closed so that the correct transition activates the signal generator SG. Similarly if FF2 is set a positive pulse is routed via switch S3 if D is set; FF3 performs the same selection for negative pulses. Where the channel is an output from the network FF4 is set and disconnects the signal generator from the channel by opening S5 leaving the channel connected to the comparator C which can assess the value of the signal which appears from the network. Table 1 indicates the signal which is presented at an input channel according to the state of the flip-flops.

**Table 1**

| | $D = 1$ | $D = 0$ |
|---|---|---|
| FF1 set | +going transition | −going transition |
| FF2 set | positive pulse | negative level |
| FF3 set | negative pulse | positive level |

The comparator C is a limiting differential amplifier which compares the output signal from the network against a high and low reference signal, providing one of three answers; either 'good high', 'good low' or 'invalid' (Fig. 3). At the sampling time $t_2$ the comparator output is clocked into two flip-flops S and V (Fig. 2), S indicating valid status if set and invalid status if unset, V indicating the value 1 or 0 for a valid signal.

Summarizing then, each of the 96 channels can be set to provide one of six signals as an input or to assess the value of an output.

## 4. The System

The tester described has no ability to make any decision as to the correctness of the response from the board. Any piece of test equipment must include this ability and various strategies are possible, of which the simplest to conceive is a free-standing tester containing all facilities necessary. The problem with such a strategy lies in changing the testing method if necessary. In the case of TE110 it is made entirely passive and embedded in a computer system, the computer
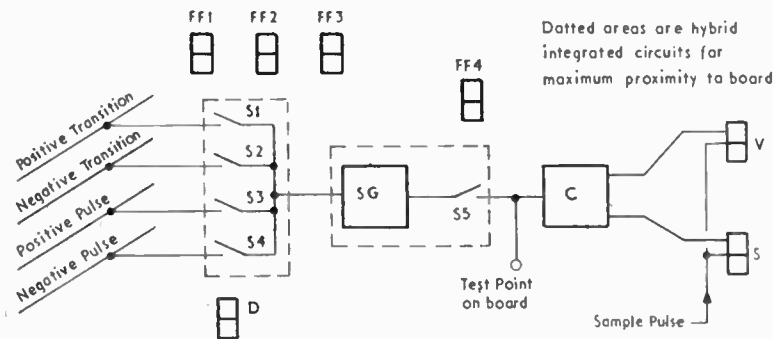


**Fig. 2.** Channel switching control.
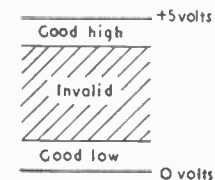


**Fig. 3.** Valid states.

Fig. 5. Arrangement of tester.
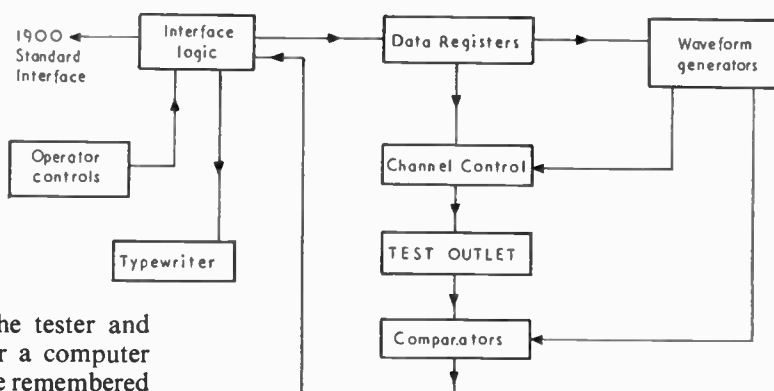
being responsible for the running of the tester and evaluation of the results. The need for a computer may seem a large overhead but it has to be remembered that in a computer manufacturing organization many of the production and test centres will already carry a computer, e.g. for testing peripherals.

In order to gain maximum flexibility, the tester connects as a normal peripheral via the peripheral Standard Interface and can be connected to most ICL 1900 computers. The minimum system consists of the computer with 8192 words of core store, a tape reader, an optional tape punch for logging and a TE110 tester. This system can be enhanced (Fig. 4) by more store and by adding testers up to the number of peripheral channels available. The tester (Fig. 5) contains the logic to interface to the computer, the electronics associated with the facilities described in Section 3, a typewriter to communicate with the operator and a control panel through which the operator controls the tester.

The computer contains the normal *Executive* program which controls the whole system via the main console typewriter. The console is used to tell the computer to load via the tape reader the tester control program and the various test data for each board type being tested, and generally to control the whole system. When the operator fits a board into the tester he dials in the board type to the computer. The computer inspects the store to see if the required test data is in and if so, it sets up the channels as specified therein. The operator next dials in the serial number of the board and the computer begins the test sequence.
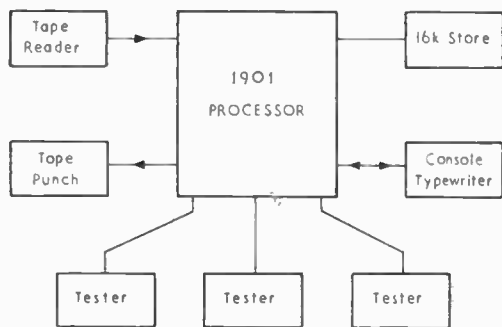


Fig. 4. Typical system configuration.

For each test it finds from the test data the input pattern and timing which it sends to the tester. The tester applies these inputs to the macroboard, assesses the output from the board and sends the data to the computer which decides from the test data supplied whether or not the response is correct. If it is, the computer proceeds to the next test and if the sequence is complete it punches out a short logging message, prints the same message on the local typewriter and lights an 'Accept' light. If a test fails, the computer proceeds to an analysis routine which attempts, very successfully, to identify the nature of the fault. It does this by repeating the failed test a number of times varying some of the timing parameters and observing the consistency of the results. From this it classifies the fault and prints one of the following:

(a) BROKEN—The board consistently fails the same test.

(b) LATE—The test passes if the sample is set later.

(c) JUST LATE—The test sometimes passes and sometime fails but always passes if the sample is set later.

(d) MARGINAL—The output is never 1 or 0, i.e. the output is a bad output.

(e) INCONCLUSIVE—The results are not consistent. The fault is probably an intermittent one. Fortunately this report occurs infrequently.

This together with all the information available about the failing test is monitored on the local typewriter, a short logging message is punched out and the 'Reject' light is illuminated. This board is then passed to the diagnostic and repair facility which uses a special mode of the control program to allow test repeat facilities in order to locate the fault.

The punched-out log is filed on magnetic tape for later Q.A. statistics analysis and is examined once per week to give the week's throughput figures for that test station. At no time is the operator required to write anything as the logging is entirely automatic.

Where more than one tester is connected, the control program deals with them in turn. The tester

operator is not aware of this except perhaps to notice that tests take a little longer as more testers are added to the system. Other incidental facilities of the control program are:

(a) Calibration of testers automatically.

(b) Tape copying facility.

(c) Automatic rearrangement of store when it fills up with test data.

(d) Amendment facility to enable timing faults in the test data to be overcome temporarily.

(e) Operator control of the number of test data cycles to be applied to each board.

The breadth of these facilities was not foreseen initially and they have indeed grown over the years. This appears to indicate that without great foresight and with greater complexity emerging, there is much to support the strategy of total 'soft' control of a test system. It is certainly much easier to adapt the environment as circumstances demand, when only one program has to be altered, than when a dozen testers have to be modified. Against it is the cost of the supporting system and the fact that it is slower than it could be if the equipment was more autonomous.

Consideration is being given to the possibility of adding a computer-controlled probe to the tester to enable fault finding to be done automatically. At present this is done manually and takes an undesirably long time. Any automatic aids in this area promise very worthwhile returns.

## 5. Manual Preparation of Test Patterns

At the time when the work on test pattern synthesis described in this paper was started, a comprehensive system for their manual preparation was already in use. This required the engineer to specify a list of the tests needed for a particular macroboard in a flexible format using simple English expressions. The list was then compiled by a program into a set of binary test patterns to be presented to the test equipment. For the more complex logic networks, the preparation and checking of these tests lists was a considerable task and it was found that an average of 3 weeks skilled engineering effort was needed to produce a correct set of test patterns for one macroboard. This clearly placed a great strain on the resources of design projects to the point of requiring extra staff to be employed entirely on testing. Some form of computing aid seemed most desirable and various methods of synthesizing
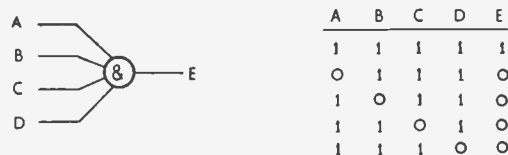


Fig. 6. Example of standard test list for a 4 input AND gate. 5 tests are sufficient.

test patterns from a definition of the logic network were investigated.

The aim was not to replace the manual system by a completely automatic one but to provide an automatic extension capable of dealing with the more routine macroboards.

## 6. The Synthesis Process

In testing a macroboard we must look for two things, firstly that each component works within its specified tolerances, and secondly that the network of components and wiring corresponds logically to the engineer's design.

In devising the test procedure we must also take into account the nature of the equipment used to do the actual testing in particular that only values on input and output connectors of the macroboard are accessible to the tester.

The approach used is to check that each logic element in the network can assume all its possible states. For each type of element in our standard set is held a list of applied inputs with expected results forming a sufficient test (Fig. 6). Each element in the network is taken in turn and a complete test sequence is evaluated using these standard lists.

In order to set up the required values on the element-under-test (e.u.t.) it is necessary to trace back through the logic network to find a consistent pattern of inputs to apply to the macroboard. Similarly in order to inspect the resulting signal a 'sensitive path' is found from the e.u.t. to one of the output connectors. The critical signal is guided along this path without being masked by the values of other signals incidental to it (Fig. 7).

In fact both these trace processes must be carried out simultaneously as the route chosen for the sensitive path must necessarily affect the choice of input pattern.

This method is implemented in the simple form just described and proves quite adequate for dealing with

Fig. 7. A sensitive path from e1 to OP
(a) with e1 giving the correct result and
(b) with e1 broken.
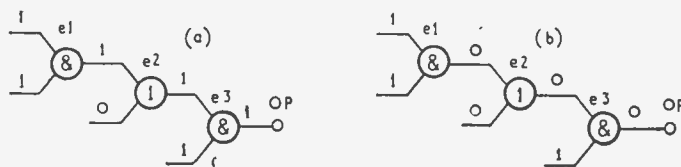
networks containing only logic gating elements. Unfortunately not many of the macroboards to be tested fall into this category: most include sequential logic.

These elements pose considerable problems as the timing and sequence of occurrence of the applied values have a considerable effect on the result achieved. In some cases stable conditions can only be obtained by the use of pulsed rather than level signals. The real significance of this is that it is not always possible to transmit a value through the network from input to output in one test cycle (Fig. 8). More than this, when working from some point in the network (the e.u.t.) it is not possible to tell how many cycles are necessary until actual values have been computed.

In order to test these more general networks a more complex strategy is needed. Each proper test, that is a complete operation on one element as already described, is divided into test fragments where each fragment corresponds to one test cycle. In each cycle, values are transmitted one stage from input pins or the outputs of bistables to output pins or holding bistables.

The outputs of these holding bistables are then picked up in the next cycle and passed on a further stage until critical signals have reached output pins where they may be tested. This is illustrated in Fig. 9. Each test fragment is thus only concerned with a section of the network composed of gating logic with input and output pins and bistables on the boundary. The method described for gating logic can therefore be applied locally with modifications for dealing with boundary conditions. The main task of the synthesizer is really to build-up proper tests from these test fragments.

Working outwards from the e.u.t., fragments will be added to the forward edge to carry critical signals one step nearer to the outputs and to the trailing edge to establish setting up conditions on the inputs (see Fig. 9).

In most cases the sub-networks corresponding to fragments of a proper test will not be discrete and some elements may be common to all, their values changing in each cycle.

Having established a general method, there remain some matters of practical detail to be considered. Some of these are by no means trivial and have not as yet been satisfactorily solved for all cases.
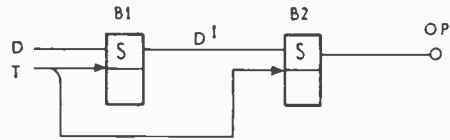


**Fig. 8.** Cycle 1 D clocked into B1. Cycle 2 $D^1$ clocked into B2 and tested at OP.

Before any real tests can be carried out on a network, bistable elements must be set initially to some known state. In fact, this process must be repeated every ten to twenty test cycles to provide a restart point on which the tester may cycle during fault diagnosis. The initializing procedure, therefore, must itself occupy as few cycles as possible.

The necessity for distinguishing between pulsed and level signals makes the evaluation of gating logic more complex as there are four, rather than two, valid states for any signal, namely: positive level, negative level, positive pulse and negative pulse. In most instances, the polarity of the signal is of prime importance but where, for example, it triggers a bistable the only distinction is between a pulse and a level.

### 6.1. Details of Implementation

In order to synthesize test patterns for a network two sets of information are needed. Firstly, the particular network must be described in terms of elements in the standard set with the interconnexions between them. Secondly, the interface between the given network and the tester must be given. Signals on input or output connexions are assigned channel numbers corresponding to sockets on the tester. Input signals are also categorized as levels, positive pulses or negative pulses.

The standard element set is defined partly within the program and partly by data. A limited set of logic functions is programmed in, but timing rules and lists of sufficient tests are supplied as data to allow for some flexibility.
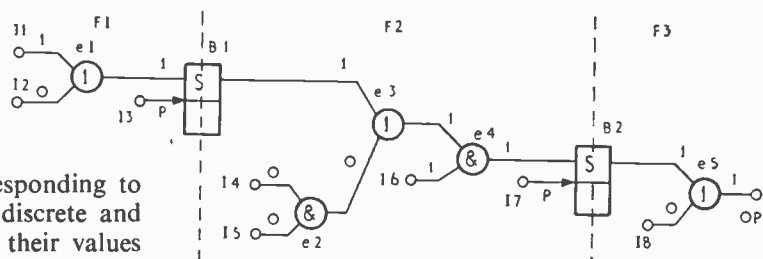


**Fig. 9.** A test of element e3. Sensitive path e3, e4, B2, e5, OP.
Cycle 1 fragment F1 set B1.
Cycle 2 fragment F2 clock path value into B2.
Cycle 3 fragment F3 carry value of B2 to output OP.

From the information supplied, a notional model of the network is built up. For each element there is a data cell including element type, forward and backward pointers to preceding and succeeding elements, a map of tests performed and bit patterns showing current and initial (start of cycle) states. In addition, a list of the source and load elements for each signal is kept.

The process is described for a general network of gating and sequential logic. The first step is to explore the network and mark some properties of the elements into the data cells. Working from inputs to outputs, those signals which may be pulses are noted. Working from outputs to inputs each element is given a distance number $d_i$ equal to the number of logical steps in the shortest path to an output pin, and a pointer $p_i$ to a connected element of distance $d_i - 1$ (Fig. 10). The next step is to initialize bistable elements to some known state. Where both set and reset lines are provided care must be taken not to pulse both in the
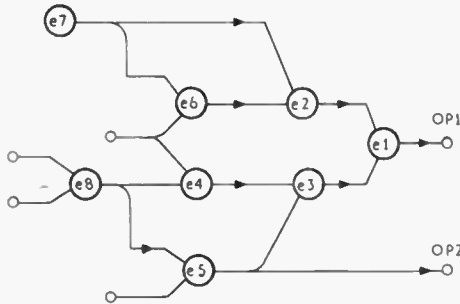
this is compared against the standard list for that type and sufficient tests generated to complete the map. To perform a proper test, input conditions must be established such that the result of the e.u.t. reaches an output pin without being masked by other signals. Starting at the e.u.t., a sensitive path segment is built up using pointers $p_i$ until an output pin or bistable is reached. If a bistable is reached, this is used to hold the critical value over to the next test cycle when the path will be continued. Appropriate test values are marked on the inputs of the e.u.t. and elements in the path segment (Fig. 7).

A trace is made from these path values towards the input pins attempting to achieve a consistent input pattern.

Where a bistable output is encountered, the current trace on that line is terminated and the conditions required held in a list Q. The trace ends when a failure condition is detected or all lines have reached input pins or bistable outputs. If a fail condition occurs, a



| i | $p_i$ | $d_i$ |
|---|---|---|
| e1 | OP1 | 1 |
| e2 | e1 | 2 |
| e3 | e1 | 2 |
| e4 | e3 | 3 |
| e5 | OP2 | 1 |
| e6 | e2 | 3 |
| e7 | e2 | 3 |
| e8 | e5 | 2 |

Fig. 10. General network.

same test, or the resulting state of the bistable will be unpredictable. Otherwise an input pattern consisting of all 1's (each channel supplies a pulse or level at logical 1) is applied to the network. The effect of this pattern is computed, and if some bistables still remain in an unknown state, the pattern is successively re-applied until their condition can be known. There are cases where this procedure fails. A more subtle method is needed, but has not yet been found. Having established initial conditions, proper tests can be computed. Elements are taken in decreasing order of distance $d_i$ and sufficient tests generated according to element type. In this way elements with the longest sensitive paths are tested first. This has the advantage of reducing the overall number of test patterns as tests may incidentally be performed on elements along this path (Fig. 11). The data cell for each element contains a bit map of tests performed. For the e.u.t.,

monitor is given and the current test abandoned. If list Q is not empty the current test fragment is held in abeyance list A and a new fragment computed with the aim of satisfying conditions in list Q. The Q list elements now take the place of the path and Q is cleared. If the new list Q is again found to be non-empty, this fragment is added to A and so on until a Q-empty fragment is achieved. This means that there are no more input conditions to be satisfied (Fig. 12).

If our original sensitive path segment ended in a holding bistable rather than an output pin, the above process must be repeated with a new section of path from the holding bistable B1 to another B2 or an output, and again until the elusive output pin is finally reached (Fig. 13).

Having concentrated on the problems of finding input patterns for the network, no mention has been
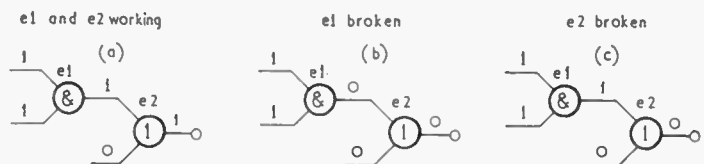


Fig. 11. Typical test patterns under fault conditions.
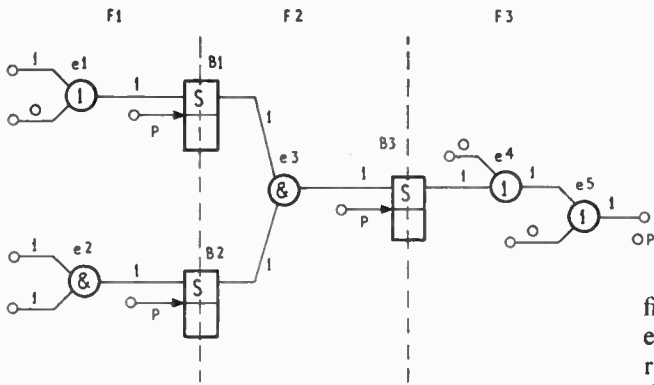
| Fragment | Aim | Start A | Q | End A |
|----------|--------|---------|-------|-------|
| F3 | test e4 | — | B3 | F3 |
| F2 | set B3 | F3 | B1,B2 | F3,F2 |
| F1 | set B1,B2 | F3,F2 | — | F3,F2 |

**Fig. 12.** Example of a test on e4. Sensitive path e4, e5, OP. Test fragments performed in the order F1, F2, F3.

made of computing corresponding output patterns and times. When the list A of test fragments, mentioned above, is emptied, a complete scan is in fact made from inputs emulating the action of the tester in order to find the true state of the network at the end of each test cycle. This is required for several reasons: to obtain an output pattern and sample time, to provide a monitor of the state of mid-board elements for the engineer and to find the initial state of the network for the next cycle. The monitor of mid-board points is to aid fault diagnosis. While the tester is performing a test of the macroboard, the engineer may examine mid-board signals with an oscilloscope and compare their values with those on the monitor chart.

As the standard element set allowed by the synthesizer is restricted a macro-definition facility has been provided to enable engineers to describe m.s.i. elements as sub-networks containing standard set elements. When these macro-elements occur in the network description data they are replaced by an equivalent sub-network. For example, a double AND-OR gate as shown performing the function E = A & B u C & D can be replaced by the equivalent network of three standard gates shown in Fig. 14.



| Segment | Path |
|---------|------------|
| S1 | e1, e2, B1 |
| S2 | B1, e3, B2 |
| S3 | B2, e4, OP |

**Fig. 13.** Segmented path for a test on e1.
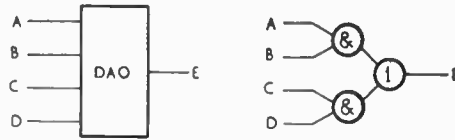


**Fig. 14.** Network of three standard gates equivalent to double AND-OR gate.

The program suite comprises four programs. The first is used to update the file of standard and macro-element descriptions. The second and third programs read the tester channel definition data and network definition data respectively, forming files used by the fourth program which synthesizes test patterns for the network and writes them on magnetic tape. The second program on the second call punches these patterns on a paper tape, for use by the tester. It also prints charts. The system is run on a 1900 computer, and uses 17k words of core store, two magnetic tape decks and an E.D.S. The average run time using a 1903A for a network of 150 elements is 25 minutes.

### 6.2. Results

Test patterns have been synthesized under experimental conditions for 14 macroboards averaging 150 elements each. The number of test patterns generated was a little over 2 per element. The time spent by an engineer is about 2 to 3 days per board, compared with 2 to 3 weeks writing test schedules by hand. However, the percentage of boards for which tests have been fully synthesized is sadly rather low, for two reasons. First, the synthesis process is somewhat sensitive to 'unusual' logic arrangements on which the simulation fails. Many of the problems associated with this can be overcome by making more element outputs available to the tester, but this approach is resisted by designers who already have insufficient pins available for design purposes. Secondly, the 'macro' technique noted in Section 6.1 is not very satisfactory and the proper approach for new elements and for medium-scale integration is to create a new procedure for each new package. Unfortunately, the rate at which new packages are being produced by the semiconductor industry implies a large software team simply to keep the synthesizer up to date.

### 7. References

1. Roth, J. P., 'Diagnosis of automata failures, a calculus and a method', *IBM J. Res. Development*, 10, No. 4, pp. 278–91, July 1966.
2. Boyce, A. H., 'Computer generated diagnosing procedures for logic circuits', Proceedings of Conference on Automatic Test Systems, p. 333 (I.E.R.E. Conference Proceedings No. 17).
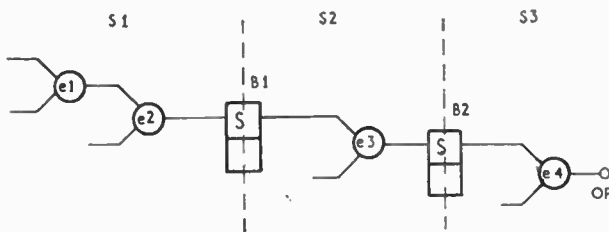
# The Visualization of Ultrasound in Solids

By

**J. H. GUNTON**, B.A. †

and

**D. M. MARSH**, M.A. †

In order to improve ultrasonic testing methods, a highly sensitive Schlieren apparatus has been built to visualize the propagation of ultrasound in solids. The design of the apparatus is described and some of the initial results are shown.

## 1. Introduction

Ultrasonic testing is used extensively in the steel tube industry with considerable success. However, the knowledge available on the exact path taken by the pulsed ultrasound inside the tube and the degree of correlation between defect size and orientation and the returning signal remains limited. We decided therefore that a direct visualization of the propagation of ultrasound in solids could not only promote a more efficient and predictable use of the ultrasonic energy, but also increase the information obtained in ultrasonic testing.

We considered three methods of visualization: photoelasticity, interferometry and Schlieren visualization. The photoelastic method is insensitive to ultrasound in liquids and interferometry is not as sensitive to ultrasound as the Schlieren system. Consequently, we decided to build a Schlieren system which could easily be adapted for photoelastic studies, or for interferometric work.[1] Further, the Schlieren system has been used extensively for the visualization of ultrasound in liquids[2] and there seemed no reason why the principle should not be extended to visualize both continuous wave and pulsed ultrasound in solids.

The final aim of the project is to use pulsed ultrasound synchronized with a strobed light source and to follow its path in a transparent model of a tube. The tube dimensions will be scaled up together with the wavelength of the ultrasound (perhaps using 200 kHz ultrasound instead of the normal 5 MHz) so that the effect of defects as small as 0·025 mm can be examined. It is also possible to deliver more ultrasonic power from lower frequency transducers which improves the visualization. Initially, however, work has been concerned with continuous wave ultrasound in order to test the principle of Schlieren visualization. Propagation in water was first studied before going on to the relatively unexplored field of ultrasound in glass. Glass rather than a plastic was used because its elastic properties are similar to those of steel.

† Tube Investments Research Laboratories, Hinxton Hall, near Saffron Walden, Essex.

## 2. Theory of Schlieren Visualization

The Schlieren principle is illustrated in Fig. 1. The first mirror and the light source produce a collimated beam of light through the working section. The second mirror serves a double purpose. If the working section is empty, it focuses the parallel beam of light to a point. The Schlieren stop then cuts off this undeflected beam. If, however, an object is placed in the working section so as to scatter or refract the light beam, the second mirror acts to image the point at which the deflexion originates as shown in the lower diagram. This deflected light is not impeded by the Schlieren stop, provided the stop is small enough. Thus, the image produced by the Schlieren system is a dark field with any deflecting regions showing up light.

Since ultrasound alters the density (and hence the refractive index) of any medium through which it



(a) Undeflected light beam.



(b) The effect of a disturbance in the working section.
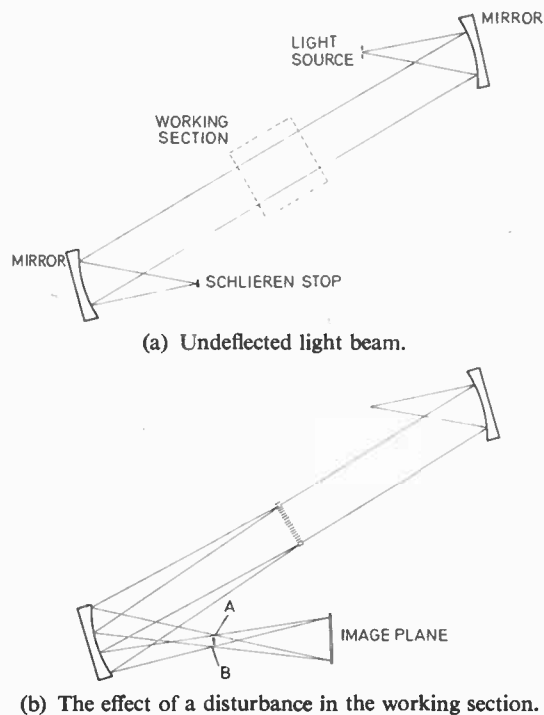
**Fig. 1.** The Schlieren principle.

propagates, light is refracted on passing through it. A Schlieren system can therefore, in principle, be used to visualize the passage of ultrasound through transparent media.

The ultrasonic beam presents to the light a periodic variation in refractive index. The effect of this is that the focus of the Schlieren system is deflected to a series of discrete positions as shown in Fig. 5(b). This is because the ultrasound acts as a phase diffraction gating and the light only interferes constructively when it is deflected through certain specific angles, given by

$$\alpha_n = \frac{n\lambda}{\Lambda}$$

where $n = 0, 1, 2, \ldots$

$\lambda$ = wavelength of light

$\Lambda$ = wavelength of ultrasound.

$n$ is the diffraction order, the zero order ($n = 0$) representing the undeflected beam. As shown in Fig. 1(b), light deflected through a given angle is brought to a focus (e.g. A and B) in the plane of the Schlieren stop at a distance $\alpha f$ from the undeflected focus ($f$ is the focal length of the second mirror). Hence, with ultrasound present, a series of foci is obtained in the focal plane, displaced by $\alpha_n f$. In order that all the deflected light shall pass the Schlieren stop, it follows that this stop must be smaller than $\alpha_1 f$, i.e.

$$D = \frac{\lambda f}{\Lambda} \qquad \ldots\ldots(1)$$

where $D$ = diameter of Schlieren stop.

Thus, if we decide to work with 200 kHz ultrasound in glass using 6 m (20 ft) focal length mirrors, $D$ must be less than 0·1 mm. This means that any deflexions of light from its true path due to inaccurately polished mirrors, intrinsic mirror aberrations, relative movement of light source, mirrors and Schlieren stop, and variations in the refractive index of the medium in the light path must not be greater than 2 seconds of arc.

There are two intrinsic aberrations in the mirror Schlieren system illustrated in Fig. 1. Firstly, there is spherical aberration due to the fact that the mirrors are spherical rather than parabolic and secondly, there is astigmatism because the mirrors are being used off-axis. The symmetry of the set-up removes all other aberrations. For a given mirror diameter, both of these aberrations are reduced as the focal length is increased. We decided to use 6 m focal length 35 cm (12 in) diameter mirrors (made by Optical Works Limited) which allowed us to offset the mirrors by only 1°35′. If the polishing errors are $\lambda/25$ per cm of the mirror surface, the following performance is predicted:

| | |
|---|---|
| error due to diffraction | 0·25 seconds of arc |
| error due to polishing | 0·5 seconds of arc |
| error due to spherical aberration | 0·4 seconds of arc |
| error due to astigmatism | 2·0 seconds of arc |

Astigmatism of this order can be corrected using a simple cylindrical lens, reducing it to less than 0·1 seconds of arc. Hence, using the mirrors described, one should obtain the desired performance given in Section 2 of less than 2 seconds of arc deflexion, provided that no further errors are introduced by the rest of the apparatus.

## 3. The Apparatus
### 3.1 The Main Structure

The above requirement implies that the mirrors must be located to a fraction of a second of arc relative to one another and to the light source and Schlieren stop. This is achieved by mounting them at either end of a steel box 9 m × 0·92 m × 0·46 m (30 ft × 3 ft × 1½ ft) made, for convenience, in three sections out of 3·2 mm (⅛ in) thick steel. This is mounted on two sets of coil springs to isolate it from building vibrations. The whole box is lagged with 2·5 cm (1 in) thick polystyrene foam and this is covered with an outer skin of steel. This eliminates air movement in the apparatus, which proved to be very troublesome in preliminary experiments.

The layout of the components in this structure is shown in Fig. 2, and the complete structure can be seen in Fig. 3. The mirrors (D) are situated at two opposite corners of the box, with the working section (E) in the centre. The 'input optics'—light sources (A), condenser lenses, aperture (C), etc.—are mounted on
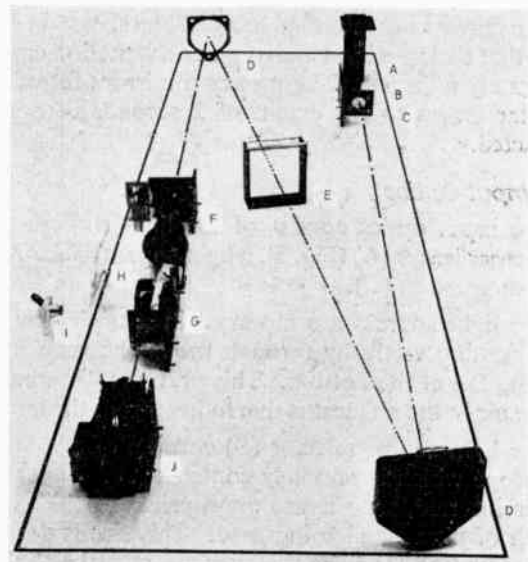


Fig. 2. The components of the Schlieren system removed from the apparatus to show the optical layout.
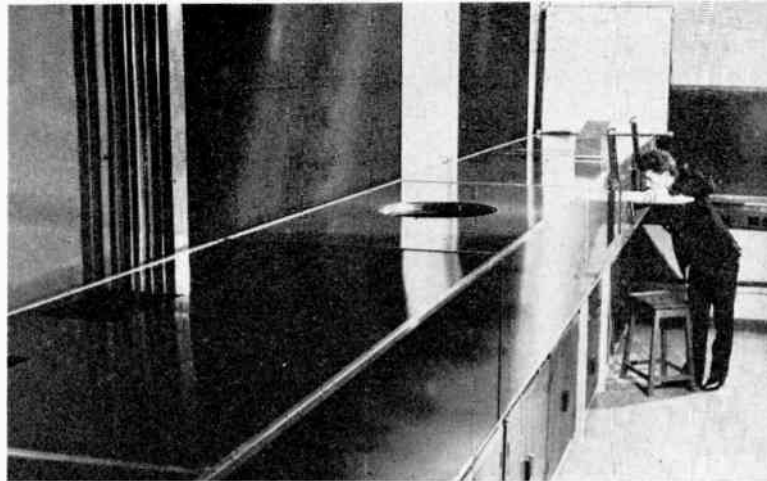
Fig. 3.
The completed
Schlieren system.

an optical bench parallel to one side of the box and the 'output optics'—Schlieren stop (F), camera (J) etc.—are mounted on another optical bench at the other end of the apparatus. The inclusion of optical benches not only facilitates adjustment of components, but also means that additional components can be included at a later stage.

### 3.2 Mirror Mounting

It was found that the method of mounting the mirrors could have a significant effect on their performance. It was shown during Foucault tests on the mirrors at the centre of curvature that they must be mounted loosely so that the minimum constraints are placed upon them. We were surprised to find also that rotation of the mirror about its optical axis had an appreciable effect on the mirror performance.

The mirrors are mounted in a cell on three steel bars such that their position and angular orientation can be accurately adjusted. The mounted mirrors introduce angular errors to the extent of 2 seconds of arc as predicted.

### 3.3 Input Optics

The input optics consist of the light sources and condenser lenses (A) (Fig. 2), a light source selector (B) and an aperture (C).

The light sources are mounted in a triple-skinned box, passing vertically through the apparatus with a cooling fan at the bottom. This prevents disturbance of the air in the apparatus due to heat from the lamps.

The light source selector (B) consists of a rotating periscope which is remotely controlled. An aperture (C) is inserted, with a size appropriate to the wavelength of ultrasound being used. This size is derived from equation (1) since the aperture must be slightly smaller than the Schlieren stop because of aberrations in the system.

### 3.4 Output Optics

The Schlieren stop (F) (Fig. 2) has its position finely adjustable in three dimensions independently. This is important so that the undeflected light can be accurately stopped off. The stop can also be rotated, which is useful when a slit light source and line stop are used. A system of prisms and mirrors (G) then sends the light either on to a small back-projection screen (H), on to a large external screen (Fig. 3) or simultaneously to an eye-piece (I) and camera (J). This provides adequate facilities both for experimental work and for demonstration.

## 4. Visualization of Ultrasound in Water

Initially, it was decided to assess the performance of the apparatus using ultrasound propagating into water. Since the wavelength of ultrasound in water is a quarter of that in glass, this experiment is not as demanding on the sensitivity of the apparatus as shown from equation (1). Apart from the desire to test the apparatus in a relatively well-tried application, it is of considerable interest to study the beam-patterns in water from various transducers examining, for example, the effect of perspex lenses on the pattern. A water-tank was constructed to study such effects. The sides were of $1·25$ cm ($\frac{1}{2}$ in) plate glass to avoid excess bowing due to the weight of the water.

Figure 4 is a Schlieren photograph of ultrasound propagating into water from a plane unbacked transducer. It shows the reflexion and transmission caused by a perspex block. It also shows pronounced side-lobes and an asymmetrical beam pattern, even though the transducer is symmetrical.

In order to gain quantitative information from this sort of picture, it is necessary to relate the intensity of light deflected into the various diffraction orders to the intensity of ultrasound. This relationship is given by the Raman-Nath equations:
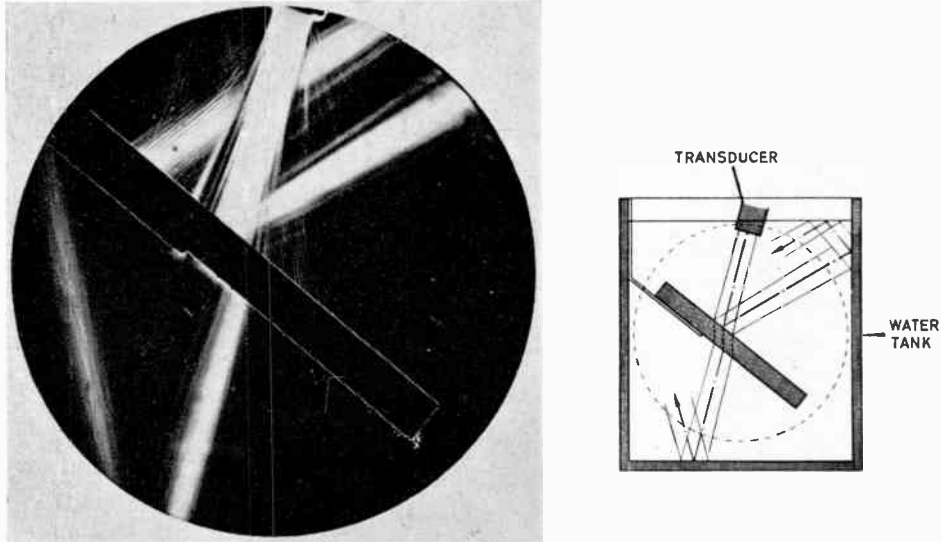
**Fig. 4.** Schlieren visualization of a 2 MHz beam of ultrasound in water, reflected and refracted by a 'Perspex' block as shown in the diagram.

$$\frac{\partial A_n}{\partial x} = \frac{-\rho n^2}{2} B_n + \frac{A_{n-1}}{2} - \frac{A_{n+1}}{2}$$

$$\frac{\partial B_n}{\partial x} = \frac{\rho n^2}{2} A_n + \frac{B_{n-1}}{2} + \frac{B_{n+1}}{2}$$

$$I_{n(x)} = A_n^2(x) + B_n^2(x)$$

where $I_n$ = light intensity in the $n$th order.

$A_n$ and $B_n$ are the real and complex parts of the light amplitude.

$\rho = \lambda^2 / \mu k p \Lambda^2$.

$x = 2\pi k p d / \lambda$.

$\mu$ = refractive index.

$p$ = acoustic pressure.

$k$ is a constant for a given medium.

$d$ = length of ultrasonic beam in the direction of the light path.

These equations can be solved on an analogue computer as shown by Berry.[3] Our method gave the relationship between $I_n$ and $x$ ($x$ is proportional to the acoustic pressure $p$) for a given transducer and a given medium. This was achieved by producing graphs for constant $\rho x$, i.e. for constant value of $\lambda d / \mu \Lambda^2$.

The graphs for the first six orders of diffracted light are shown superimposed in Fig. 5(a). They indicate that, as expected, all the light goes into the zero order when $p = 0$, but as $p$ is increased, more and more light is deflected into the other orders, until at $A$ there is virtually no light in the zero order. Figure 5(b) shows how a slit light source is deflected as the ultrasonic energy is increased. Since the light source is not monochromatic, there are several lines in each order. The light in the zero order is never extinguished completely, but reaches a minimum. This is both because the

intensity of the ultrasound is not completely uniform along the optical path at the place where it is sampled, and because of the spread in light wavelength. It is found that, despite the limitation just mentioned, the ultrasound energy can be estimated to within 20% which is quite adequate for our purposes.

## 5. Visualization of Ultrasound in Glass

From the Raman-Nath graphs, we estimated that 5 W of ultrasonic power would be necessary from a 2·5 cm diameter transducer to visualize ultrasound in a glass block. However, using a 40 W 2 MHz oscillator and coupling the ultrasound from the transducer to the glass block with a layer of water, we found that the cavitation in the water, and hence the scattering, was such that ultrasound was scarcely visible in the glass. Further, when mercury was used to improve the ultrasonic coupling and to eliminate cavitation, results were even more disappointing. We think this is because of the non-wetting properties of mercury. However, if—as shown in Fig. 6(b)—the transducer is arranged so that the reflected beam does not interfere with the incident beam at the transducer surface, visualization of ultrasound in the glass block is possible. Figure 6(a) shows the visualization of longitudinal and shear waves generated in the glass block by this arrangement. In this initial photograph, the speckled background is due to the imperfections in the plate glass. However, it is reasonable to deduce from this photograph that, if the ultrasonic power can be increased either by using a higher power oscillator or by using pulsed ultrasound, propagation of ultrasound in tubes and the reflexion from defects can be studied.
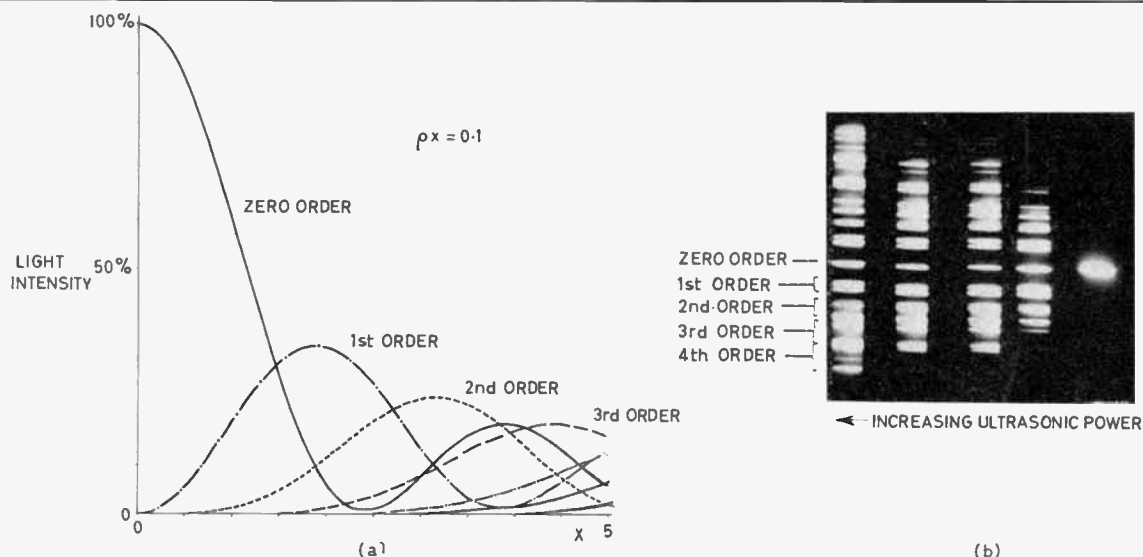
Fig. 5. Graph showing the relationship between light intensity and acoustic pressure and a photograph of the diffracted image of a slit showing this relationship in practice.
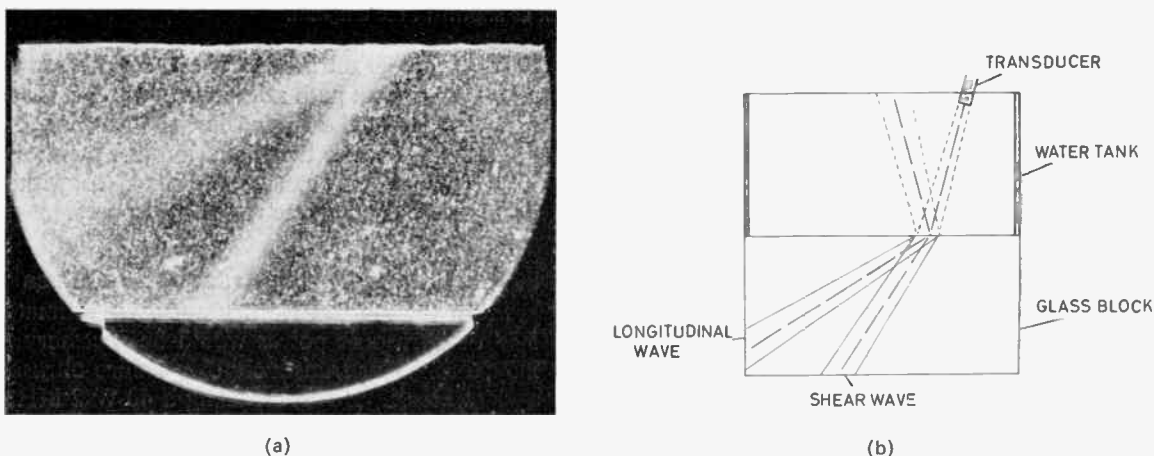


Fig. 6. Schlieren visualization of shear and longitudinal waves generated in a glass block. The diagram shows the arrangement of the transducer and glass block.

## 6. Conclusions

The apparatus as it has been described, has been shown to be capable of the visualization of ultrasound in solids. Its potential for the demonstration and improvement of ultrasonic testing of steel tubes has therefore been proved, and work is at present in progress on the building of a pulsed light source in order to realize this potential. This should provide a better understanding of present ultrasonic testing methods and could lead to significant advances in the non-destructive testing of steel tubes.

*Note added in October 1970.* Since this paper was written significant advances have been made. Using a stroboscopic light source it has been possible to form a stationary image of pulsed ultrasound in water and glass. The light source is a development of the argon jet spark discharge lamp producing light pulses of 200 ns duration, which has allowed the resolution of individual waves of sound at 2 MHz. This will be the subject of a future publication.

## 7. Acknowledgment

This paper is published by permission of the Chairman of Tube Investments Limited.

## 8. References

1. Oppenheimer, A. K., Urtiev, P. A. and Weinberg, F. J., 'On the use of laser light sources in Schlieren-interferometer systems', *Proc. Roy. Soc.*, 291, 1966.
2. Whaley, H. L., Cook, K. V., McClung, R. W. and Snyders, L. S., 'Optical methods for studying ultrasonic propagation in transparent media'. Paper presented at Fifth International Conference on Non-destructive Testing, Montreal, 1967.
3. Berry, M. V., 'The Diffraction of Light by Ultrasound'. (Academic Press, New York, 1966).

# A Cellular 8421 B.C.D. Multiplier

*By*

**G. WHITE**, B.Sc.†

An iterative array of macro-cells is proposed to multiply together two $N$ digit decimal numbers with a propagation delay of $(2N - 1)$ macro-cell delays. The array will also add two decimal numbers to the product without adding to the complexity.

Recent papers[1,2] have described iterative arrays for multiplication of binary numbers. Decimal multiplication can be achieved using similar techniques and with the introduction of large-scale integration the larger cells can now be manufactured. A $3 \times 3$ array is shown in Fig. 1 with the primary and intercell numbers appropriate to the identity $385 \times 428 + 419 + 216 = 165415$, the intercell lines each consisting of four leads, since they carry binary numbers from 0 to 9.

The array operates by forming the partial products in columns of the correct weight and by summing these partial products into the final product. Therefore the required arithmetical function of the macro-cell is:

$$R = A$$
$$Q = B$$
$$P = \frac{AB + C + D}{10}$$

$$S = \text{remainder of } P$$

There are $N^2$ partial products; hence the number of macro-cells required is $N^2$ giving a propagation delay of $2N - 1$ cell delays.

The macro-cell function can be realized by multiplying with a $4 \times 4$ binary multiplier array,[2] and then dividing by 5 to convert the result back into 8421 b.c.d. as used in a b.c.d. adder/subtractor system.[3]

This gives 11 cell delays per macro-cell. The structure of the macro-cell is shown in Fig. 2 with the primary and intercellular numbers appropriate to the identity $3 \times 8 + 2 + 4 = 30$, which corresponds to the bottom cell in Fig. 1.

The binary multiplier operates similarly to the decimal multiplier described above; the partial products being straightforward to generate, since the multiplier can only be 0 or 1. The arithmetic function of a cell is

$$V = \frac{XY + W + Z}{2}$$

$$U = \text{remainder of } V$$

which gives the Boolean relationship

$$V = XYW + XYZ + WZ$$
$$U = XY \oplus W \oplus Z$$

The result of the multiplication is in binary, which has to be divided by 10 to convert into 8421 b.c.d. Division by 2 is accomplished by the equivalent of shifting one place, leaving the cells to divide by 5. The arithmetic function of the cell is:

$$K = \frac{E2^3 + F2^2 + G2^1 + H2^0}{5}$$

$$L2^2 + M2^1 + N2^0 = \text{remainder of } K$$

which gives the Boolean relationship:



Fig. 1. Cellular array for b.c.d. multiplication.

† Department of Engineering Technology, Twickenham College of Technology, Twickenham, Middlesex.
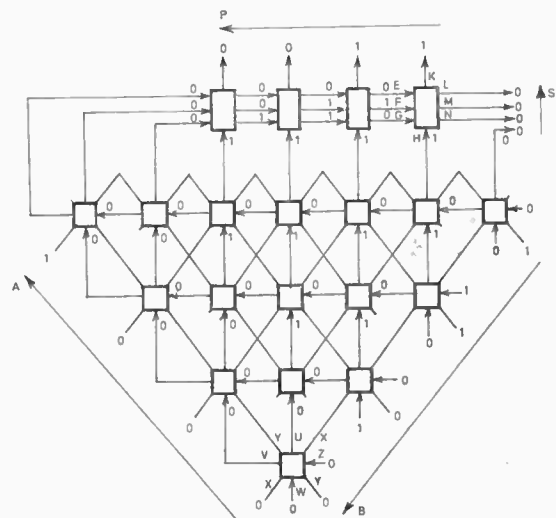
Fig. 2. Cellular structure of macro-cell.

$$K = E + FG + FH$$
$$L = EH + F\bar{G}\bar{H}$$
$$M = E\bar{H} + GH + \bar{E}\bar{F}G$$
$$N = \bar{E}\bar{F}H + FG\bar{H} + E\bar{H}$$

The decimal multiplier array has been built at Twickenham using standard TTL integrated circuits to fabricate the cells. With the propagation delay of one cell being in the order of 50 ns, a 6 decimal digit output is produced in $2\frac{3}{4}$ μs. The macro-cells would be candidates for large scale integration, because of the high complexity of gates per pin.

These could be used in such applications as desk calculators or special-purpose computers.

## References

1. Hoffmann, J. C., Lacaze, B. and Csillag, P., 'Multiplieur parallèle à circuits logiques itératifs', *Electronics Letters*, **4**, p. 178, 1968.
2. Guild, H. H., 'Fully iterative fast array for binary multiplication and addition', *Electronics Letters*, **5**, p. 263, 1969.
3. White, G., 'Design of a parallel b.c.d. adder/subtractor', *Electronic Engineering*, **41**, pp. 229–30, February 1969.

# Contributors

# to

# this issue



**Mr. G. White** obtained his B.Sc. degree in electrical engineering from the University of Surrey in 1967 and then joined the National Coal Board's Mining Research Establishment, working on the development of a data collection system for use underground. From March 1968 to January 1969 he was a research assistant at Letchworth College of Technology investigating cellular arrays for arithmetic units. He then joined Muirhead Ltd., working in their digital systems and data communications department on the development of a special purpose computer for typesetting. He continued his research on cellular arrays at Twickenham College of Technology part-time and since July 1970 he has worked full-time at the College on industrially-sponsored research on the subject.



**Professor William Gosling** (F. 1967) graduated with a B.Sc. degree in physics from Imperial College in 1953. He then worked with de Havilland Propellers Limited for five years on various aspects of the design of electronic test systems for aircraft and missiles. In 1958 he was appointed to a lectureship in the School of Engineering Science of the University College of Swansea and subsequently became a senior lecturer. In 1966 he was appointed to the Chair of Electrical Engineering. Professor Gosling is Chairman of the Technical Committee and was recently elected to serve on the Council of the Institution.



**Mr. A. B. McFarlane** studied at Imperial College of Science and Technology and graduated in electrical engineering (radio and telecommunications) in 1945. He has worked for 20 years on research and development projects concerning cathode ray tubes, including pioneer work on the now universally accepted basic mesh p.d.a. tube design, and on the secondary electron emission of luminescent screens. A thesis on the latter subject was accepted for an M.Sc. (Eng.) degree in 1955. Since 1965 Mr. McFarlane has been manager of cathode ray tube development at the M-O Valve Company.

# Measurements on a Solid-State Noise Source

By

**J. A. ROBERTS,** M.Sc.†

and

**Professor W. GOSLING,**
B.Sc., C.Eng., F.I.E.R.E.‡

Some planar silicon junctions with a breakdown voltage of 5-6V produce noise with good waveform up to 1GHz. Flicker noise is absent.

Among the solid state noise sources presently in use are avalanche diodes, which have breakdown voltages typically between 12 and 50 V. Diodes which break down in the region 5 to 6 V are probably subject to both Zener and avalanche breakdown processes, since the temperature coefficient of the breakdown voltage can be positive, negative or zero dependent on the magnitude of the breakdown current. No experimental observations of the noise properties of diodes of this type appear to have been published hitherto.

As a consequence of measurements on some 120 units, it has been found that certain planar silicon junctions with a breakdown voltage of 5 to 6 V produce noise with exceptionally good waveform and spectrum characteristics. Complete freedom from multi-state (so-called 'popcorn') and flicker noise may be obtained with good repeatability. The noise level is typically 35 to 40 dB above $kT_0$, and a temperature coefficient of amplitude of 0·01 dB per degree C is observed about room temperature. For diodes of this type, noise level remains consistently within $\pm 1$ dB over the frequency range 20 Hz to 50 MHz. High levels of noise can be detected at frequencies up to 1 GHz, with suitable circuit arrangements.

The variation of noise voltage with bias current is of particular interest since an inverse square variation of noise voltage $V_n$ with bias $I_B$ is invariably obtained. The graph shown in Fig. 1 illustrates the noise voltage change with bias current of a particular unit. The dashed curve traces the relationship

$$V_n = \frac{K}{(I_B)^{\frac{1}{2}}}$$

where $K$ is a constant for a particular unit.

The precise relationship together with the exceptionally low temperature coefficient suggests that the device is a suitable replacement for the temperature-

limited diode in much noise generating equipment. The absence of flicker noise observed, which allows the device to be used at audio frequencies, suggests that the current density is considerably lower than in avalanche devices.

It is of interest to compare the inverse square noise/current relationship obtained with that for diodes operating well into the avalanche breakdown mode. For example, for Read type avalanche diodes, Hines§ has developed a relation for the mean-square noise current.

$$\bar{I}_n^2 = 2qI_B B \left\{ \frac{C_T^2 V_a^2/I_B^2}{2m^2 \tau_x^2} \right\}$$

where $q$ = electronic charge
$B$ = noise bandwidth
$\tau_x$ = delay time
$C_T$ = total depletion zone capacitance
$V_a$ = voltage drop across the avalanche zone
$m$ = constant for breakdown, in silicon $m \simeq 6$
$I_B$ = direct current bias

The Read diode, as described, is a true avalanche diode ($V_{Br} = 80$ V). The noise voltage will be given by an expression of the form
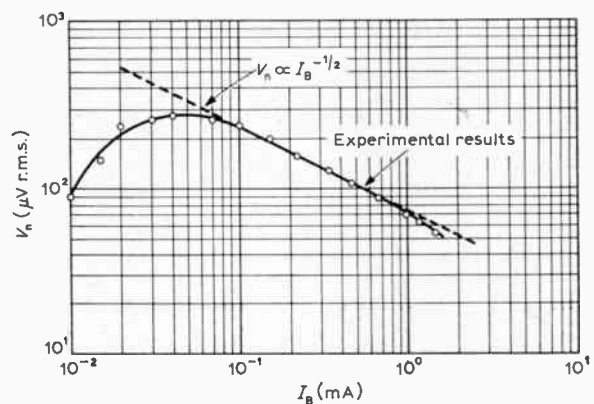


**Fig. 1.** Variation of r.m.s. noise voltage ($V_n$) with diode current ($I_B$). $B = 1$ MHz.

---

† Formerly at the University College of Swansea; now with Marconi-Elliott Microelectronics Ltd., Witham, Essex.

‡ Electrical and Electronic Engineering Department, University College of Swansea, Swansea, SA2 8PP, Glamorganshire.

§ Hines, M. E., 'Noise theory for the Read type avalanche diode', *I.E.E.E. Trans. on Electron Devices*, ED-13, pp. 158–63, January 1966.

$$V_n \propto \left(\frac{qB}{I_B}\right)^{\frac{1}{4}}$$

This is in agreement with Fig. 1 and would thus support the hypothesis that noise generation in avalanche-Zener diodes is principally an avalanche phenomenon. An explanation for the relatively favourable noise statistics and spectrum might be sought in terms of more detailed characteristics of the breakdown process, for example relative infrequency of microplasmas. Work on this aspect continues.

The inverse square law relationship breaks down at small values of $I_B$ where $I_n$ is large; the noise current fluctuations then become of the same order as the bias current. Operation in such a manner is not desirable since the peaks of the noise waveform will be clipped.

# Conference on 'Electronic Control of Mechanical Handling'

The worldwide quest for increased productivity has provided an impetus to the development and use of mechanical handling both in industry and in public service. In many cases the use of these techniques has been made possible only by the employment of electronic control, while in others the use of electronics has led to the evolvement of complex and sophisticated systems for the automatic handling of goods and materials.

The Institution believes that a conference on the theme 'Electronic Control of Mechanical Handling' would be most valuable at this stage. It will take place at the University of Nottingham from Tuesday, 6th to Thursday, 8th July 1971, and will have the association of the Institutions of Mechanical, Electrical and Production Engineers, the Institute of Materials Handling and the National Materials Handling Centre, Cranfield.

While many of the electronically controlled systems of mechanical handling are peculiar to their application the techniques employed frequently have far wider potential use. Thus it is intended that the papers presented at this Conference, and the ensuing discussions will provide 'cross-fertilization' between users in different industries as well as between users and manufacturers of electronic equipment for controlling mechanical handling systems. By definition, the common denominator of all papers presented will be 'electronic control'; it is not, however, intended that they should deal with detailed electronic circuitry.

The Organizing Committee, which is under the Chairmanship of Mr. Laurence Clarke, Chairman of the Instrumentation and Control Group of the I.E.R.E., and includes representatives from Industry, Government Research Establishments and Universities, now invites papers for the Conference. In the first instance synopses of proposed contributions are requested. Preference will be given to papers describing original work, new systems or new applications.

Papers discussing the economic advantages of electronically controlled systems will also be welcome. For guidance a non-exclusive range of subjects is suggested:

*Overall Management of Mechanical Systems:* including—Inventory control by computer; Computer direction of stacker cranes; Positioning devices; Computer control of order picking; Automatic warehouses.

*Routing Systems:* including—Automatic routing of cartons, cases, pallets, etc.; Postal letter and parcel sorting machines; Driverless tractors, etc.; Airport baggage handling.

*Cranes, Stackers, Palletizers and Robots:* including—Electronics associated with component handling; Liner trains, etc.; Control of dockside cranes; Steel slab yards; Programmable industrial robots, and automatic assembly devices.

*Components and Sub-systems:* including—Code reading devices; Electronic power control fork-lift trucks and stacker cranes and other power machines; Control accessories; Photoelectronic counters; Ultrasonic counters; Closed circuit television, etc.

A synopsis (typically about 200 words) should be long enough to enable the Committee to assess the scope of the proposed paper, and should be sent as soon as possible to the Conference Secretariat, Institution of Electronic and Radio Engineers, 8–9 Bedford Square, London, WC1B 3RG.

Final papers may be either in short form or full length, i.e. containing either 2000 words or 4000 words approximately. Papers will be pre-printed and are therefore required in final form by 31st March 1971.

Further information and registration forms for the Conference will be available in due course from the Institution.

# SUBJECT INDEX

Papers and major articles are denoted by printing the page numbers in bold type

# INDEX OF PERSONS

Names of authors of papers published in the volume are indicated by bold numerals for the page reference.

Biographical references are denoted by в.

# INDEX OF PERSONS