

THE RADIO AND ELECTRONIC ENGINEER

The Journal of the Institution of Electronic and Radio Engineers

Over-Production — of Engineers ?

RECSSION in trade is caused or conditioned by natural or man-controlled factors. A deliberate cutting back of much new enterprise—whether in advanced work such as space investigation or in expanding and improving consumer products—may be required to meet the economic capability of both public and private sectors in withstanding the effects of recession. Currently there is world-wide difficulty in reconciling both public and private expenditure with income whether that is to be provided through direct or indirect taxation or through private investment. Over-simplified, the inability to balance costs with resources leads to inflation which must be arrested if a healthy state of economic affairs is to be achieved.

Although this reasoning is simplified it is hard to appreciate why the engineering profession and industry should so often be among the first casualties of the measures taken to attempt to cure a depressive economic condition—the more so because there is almost invariably a concomitant reaction affecting not only personnel in the industry itself but also the prosperity of dependent companies in other industries. How can such measures be reconciled with the undeniable fact, emphasized in a recent Presidential Address to this Institution,* that it is engineers who create wealth? This is of course the wealth that is based on the production and effective use of both capital and consumer goods.

Changing patterns in political economy can, however, be equally applied to changing patterns in technology. Whilst it is true that the domestic products of the radio and electronic engineer can—and are—subject to the fluctuations of politically-influenced national economy, there is nevertheless a rising demand for all consumer products—not least television, radio and 'hi-fi' equipment—as was pointed out in the recent manpower survey.† There is also a growing demand for electronic products which can increase and improve production in other industries, to say nothing of the vital need for stepping up the introduction of more and more electronic systems and devices to provide increased safety in travel by land, sea or air. There is thus ample justification for reorientation of training so that younger people can be directed toward progressive industries such as electronics rather than older industries which, for economic, social, or other reasons, are no longer able to offer prospects of future employment.

It is distressing to read in, for example, the *Indian Proceedings* of the Institution‡ of one country apparently contemplating trying to resolve its economic problems by cutting back on the training of engineers of the future. Distressing because whether the country is underdeveloped or supposedly advanced, ultimate wealth and the exploitation of nature's resources can only be achieved by technological development and this demands employment of qualified engineers with the ability to exploit natural resources and thereby create the wealth to ensure economic viability of any community. The engineer's training today, which throws greater emphasis on industrial management, enables him to meet these challenges.

It seems, therefore, that a change of emphasis would be more productive than an economic clamp. Is the engineer able to meet that economic and political challenge? For its part, this Institution encourages engineers to think more of management problems through its own Management Techniques Group.

G. D. C.

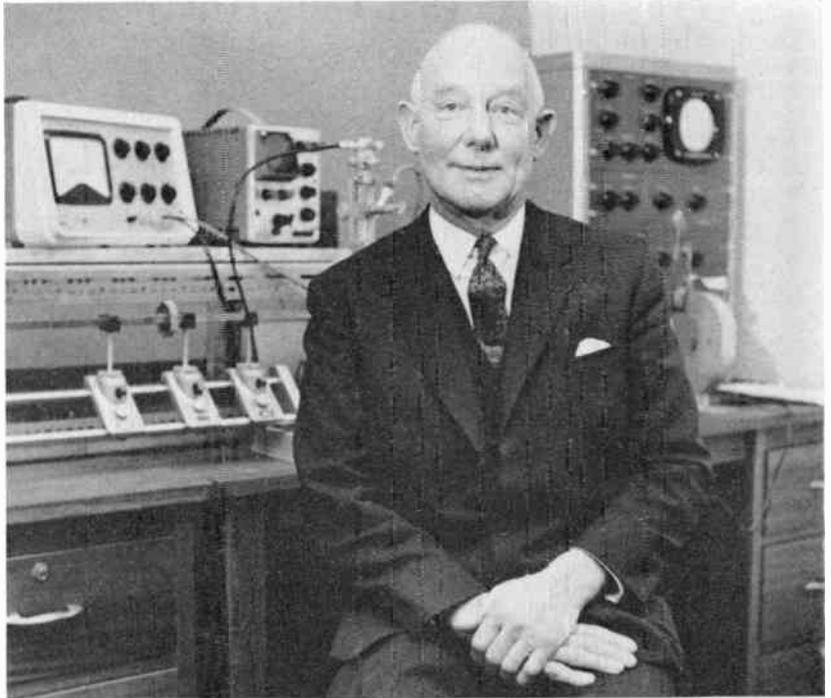
*Sir Leonard Atkinson, Presidential Address, *The Radio and Electronic Engineer*, 35, pp. 3–9, January 1968.

†'Qualified Manpower in the Electronics Industry', NEDO, January 1971 (see *The Radio and Electronic Engineer*, 41, p. 97, March 1971).

‡*I.E.E.-I.E.R.E. Proceedings—India*, 8, No. 4, p. 143, September/October, 1970.

The Seventh Clerk Maxwell Memorial Lecturer

Professor H. M. BARLOW
Ph.D., F.R.S., C.Eng.



Throughout most of his professional life, Professor H. M. Barlow's specialization has been microwave research and engineering and for this reason it was particularly appropriate that he should be invited to give the seventh Clerk Maxwell Memorial Lecture. It was appropriate too that, by kind permission of the Provost, the Lecture should be delivered in the Engineering Lecture Theatre of University College London, where Professor Barlow has worked continuously, apart from periods of practical engineering training and war service, for well over 50 years. For although he retired from the Pender Chair of Electrical Engineering in 1967 and was granted the title of Emeritus Professor, he is continuing his microwave work as an Honorary Research Associate in the Department.

Born in London in 1899 and educated at Wallington Grammar School and the City and Guilds Engineering College, Professor Barlow served in the 1914-18 War as a Sub-Lieutenant in the R.N.V.R. and on demobilization became a student at U.C.L. After graduating and spending two years with engineering companies, he was appointed to the Academic Staff of the College's Faculty of Engineering in 1925. On the outbreak of the Second World War he joined the Air Ministry and for four years was concerned with radar development at the Telecommunications Research Establishment; in 1943 he became Superintendent of the Radio Department at the Royal Aircraft Establishment.

He returned to U.C.L. in October 1945 and was appointed to a chair in Electrical Engineering; in the following year he was elected a Fellow of the College and he later became Dean of the Faculty of Engineering and a member of the U.C.L. Committee. Other academic appointments held by Professor Barlow have included membership of the Academic Council of the University of London, Governorship of Woolwich Polytechnic, and service on the London Regional Advisory Council for Higher Technological Education. He has also served on numerous official bodies including the Radio Research Board of D.S.I.R. and Advisory Boards of the Ministry of Supply and he is a member of the B.B.C. Engineering Advisory Committee. He represents the Royal Society on the Board of the

National Electronics Council and his industrial associations have included a directorship of Marconi Instruments Ltd.

Professionally Professor Barlow has been a leading member of the Institution of Electrical Engineers, having served on its Council for several years. As well as the Faraday Medal of the I.E.E., he has received a number of its Premiums for outstanding papers. He was elected a Fellow of the Royal Society in 1961, a Foreign Member of the Polish Academy of Sciences in 1966 and a Fellow of the Institute of Electrical and Electronics Engineers in 1956. A notable international honour was the award of the Dellinger Gold Medal of the International Scientific Radio Union in 1969; he has been closely concerned with the work of U.R.S.I. and served as chairman of the British National Committee. In July of this year Heriot-Watt University, Edinburgh, will confer upon him an Honorary D.Sc. degree.

Author of several books and very many papers on numerous aspects of radio engineering, it is probably his pioneer investigations of different kinds of waveguide for communications that have attracted the widest interest. (One of his earliest papers on this subject was, in fact, read at the 1947 Convention of this Institution and he also contributed a paper more recently on applications in railway signalling and control.*) He has also directed highly significant research on r.f. measurements, particularly in the use of the Hall effect.

During the first four years of his retirement, Professor Barlow has continued much of his technical work, and he is developing a long-standing interest in the problems of assessing the response of people to technological innovation. He also acknowledges increasing enjoyment of detective stories which he finds are akin to the search for new scientific information.

*'The exploitation of microwaves for trunk waveguide multi-channel communications', *J. Brit.I.R.E.*, 7, No. 6, pp. 251-8, October 1947.

'High frequency guided electromagnetic waves in application to railway signalling and control', *The Radio and Electronic Engineer*, 33, No. 5, pp. 275-81, May 1967.

The Seventh Clerk Maxwell Memorial Lecture

GUIDED ELECTROMAGNETIC WAVES

By

Professor H. M. Barlow, B.Sc.(Eng.), Ph.D., F.R.S., C.Eng., F.I.E.E.†

Delivered at a Meeting of the Institution at University College London on 9th March 1971.

1. Introduction

May I first say how deeply I value the honour you have done me in inviting me to give this, the Seventh in the Series of Memorial Lectures to James Clerk Maxwell. The task sets a challenge, for which my limitations are obvious, but it also gives me the opportunity of paying homage to that great man whose researches have done so much to influence and direct our thinking in electrical science.

Many biographies have been written about Clerk Maxwell and the results of his investigations,¹ seemingly making it that much more difficult to find a fresh avenue of approach, but such was the versatility of his ideas that any possible limitation imposed in that way is quite inconceivable. Thus, the choice of 'guided electromagnetic waves' as the theme of the present lecture implying, as it does, consideration of boundary conditions at interfaces between different media gives, I feel, any amount of scope for discussion of a subject intimately bound up with Maxwell's work, and of particular importance today.

2. Highlights of Clerk Maxwell's Life and Work

Before considering the more specialized interest, perhaps I may recall, very briefly, some features of Maxwell's career and of the setting in which his creative-ness flourished so successfully.

Born in Edinburgh on 13th June 1831, he came of an old-established Scottish family. There was nothing particularly remarkable about his boyhood; it is said that he was naturally rather shy but he showed a lively imagination and the usual boyish inquisitiveness. As he developed in his early 'teens, his exceptional powers became more apparent and before going up to Cambridge he had achieved a distinguished place at his school, the Edinburgh Academy, gaining special prizes in Mathematics and English Verse.

As an undergraduate, first at Peterhouse and later at Trinity College, he was soon acknowledged as one of the intellectual set but, at the same time, he was sought after as a genial and amusing companion, able to talk well on a wide variety of subjects. His contemporaries speak too of that characteristic dry humour, which many people found so attractive in his later life. He took the Mathematical Tripos in 1854 and was declared

Second Wrangler. After a period as a Fellow of Trinity he went, in turn, to Chairs of Physics at Marischal College, Aberdeen, King's College, London, and then back to Cambridge. It is generally agreed that the most fertile part of his career was the five years he spent at King's College, when he was in his early thirties.

Most of us think of Clerk Maxwell, first and foremost, as a great mathematician and although this cannot be wrong, it does seem that the key to his remarkable success lay, more particularly, in a rather different direction. William Hopkins, Maxwell's Tutor at Cambridge, says 'it was impossible for Maxwell to think wrongly on any physical subject but in analysis he was far more deficient'. Sir James Jeans assesses Maxwell's mathematical technique as 'adequate though not outstanding' and it has been pointed out that some of his algebraic calculations, such as those on viscosity of gases, were not very reliable. Whatever his status as a mathematician, there can be no doubt that Maxwell had that very special gift, invariably shown by our greatest scientists, of being able to see beyond the mathematical symbols and their formal relationships, to the wider consequences of the analysis. Above all, there is much in evidence an intuitive understanding which dominates his work.

In Maxwell's day there was a distinct emphasis on 'physical reality' and he went along with Michael Faraday and Kelvin in attempts to construct physical models to illustrate basic concepts of electrical behaviour. Maxwell was, however, rather more wary in accepting some of the consequences which such models suggested and he was more prepared to discard deductions where, in his view, they did not fully meet the requirements. It can be said that Maxwell was the kind of mathematician to whom every symbol was alive and represented, in his mind, a clear physical picture, although not always in terms of tangible things. Faraday's influence was obviously tremendous and after absorbing the full implications of his experiments, Maxwell set about giving mathematical expression to some of the ideas that came out of them. Thus, he translated the intensity of Faraday's 'electrotonic state' into what he called magnetic vector potential. The possibility of waves with transverse electrical vibrations was really inherent in Faraday's concept of 'lines of force' but it was left to Maxwell to establish this as a natural consequence of his theory. Einstein, writing more recently,¹ remarks that the 'physical reality' of the Maxwell era tends now to be thought of 'as represented by continuous fields

† Honorary Research Associate, Department of Electronic and Electrical Engineering, University College London, Gower Street, London, W.C.1.

governed by partial differential equations and not capable of any mechanical interpretation'. He instances quantum mechanics, in which the quantities that appear make no claim to describe 'physical reality' itself, but only the probabilities for the appearance of a particular physical reality on which attention is fixed.

Our gratitude to Maxwell rests primarily on what he did for electromagnetic theory, bringing together the mathematics of the earlier philosophers, like Ampère, Gauss and Poisson, with the physical concepts of Faraday, thus enabling relationships to be formulated that led, among other things, to the electromagnetic theory of light. His paper on 'The dynamical theory of the electromagnetic field', quoted so often, is undoubtedly one of the greatest productions of the human mind, and without it the essential directive and coherent influences we value so much today would not have been achieved. No statement of basic principles could have been more far-reaching, more precise in describing what happens to an electromagnetic field in any defined environment. The unsurpassed genius of that achievement has, perhaps, tended to obscure some of the other great contributions that Maxwell made.

Looking at the two Volumes of Maxwell's Scientific Papers, one cannot help being struck by the wide scope of the topics represented there. In fact, electromagnetic theory occupies but a small part of his scientific writings, which include studies ranging from Saturn's Rings, dynamical theory of gases and fluids, to geometrical problems such as the delineation of surface contours. Indeed, it was Maxwell who was largely responsible for laying down the principles governing contour mapping, as they are applied today. Maxwell's remarkable intuition is perhaps nowhere better demonstrated than in his work on 'The dynamical theory of gases'. Of part of this memoir Jeans writes 'Maxwell, by a train of argument which seems to bear no relation at all to molecules, or to the dynamics of their movements, or to logic, or even to ordinary common sense, reached a formula which, according to all precedents and all the rules of scientific philosophy, ought to have been hopelessly wrong. In actual fact it was subsequently shown to be exactly right, and is known as Maxwell's law to this day'.

In the latter part of his life, as the first Professor of Experimental Physics at Cambridge, Maxwell was responsible for the setting up of the Cavendish Laboratory, opened in 1874. He held strong views about the need for careful measurements and the rewards these could bring in the development of new ideas. He encouraged a group of Fellows and Post-graduates, including W. D. Niven, Chrystal, J. A. Fleming, Glazebrook and Poynting, to make the fullest possible use of the facilities at the Cavendish but, strange as it may seem, he was never himself a dedicated experimentalist. In this connexion, it is perhaps significant that Maxwell, with all the resources of the Cavendish Laboratory behind him, never seems to have attempted to obtain experimental proof of the existence of electromagnetic waves, so brilliantly predicted by him but left to Heinrich Hertz to demonstrate eight years later.

To sum up then, this very short biographical note, we can say that Maxwell contributed most, by bringing to bear on the problems he tackled, a kind of intuitive understanding far greater in depth and breadth than all but very few attain. In his comparatively short lifetime he established a re-orientation of thinking on the whole subject of propagation of electrical disturbances, and it is quite impossible to contemplate what he might have achieved had he not died at the early age of 48.

3. Guided Waves

My particular concern on this occasion relates to electric waves that are constrained to travel along a pre-determined path.² To ensure this, some mechanism is required either continuously or periodically tending to concentrate the electromagnetic field of the wave and its associated energy, within the ambit of the chosen path. Thus, a wave is said to be guided when the transverse spread of the field, which accompanies uninhibited propagation, is deliberately restricted. Within this rather broad definition are included waves transmitted by cables, along the inside of hollow tubes, or indeed those to which an iterative system of lenses or irises is applied along the path (Fig. 1). In the context of the present discussion iterative arrangements will not be considered and only continuously applied restraint to

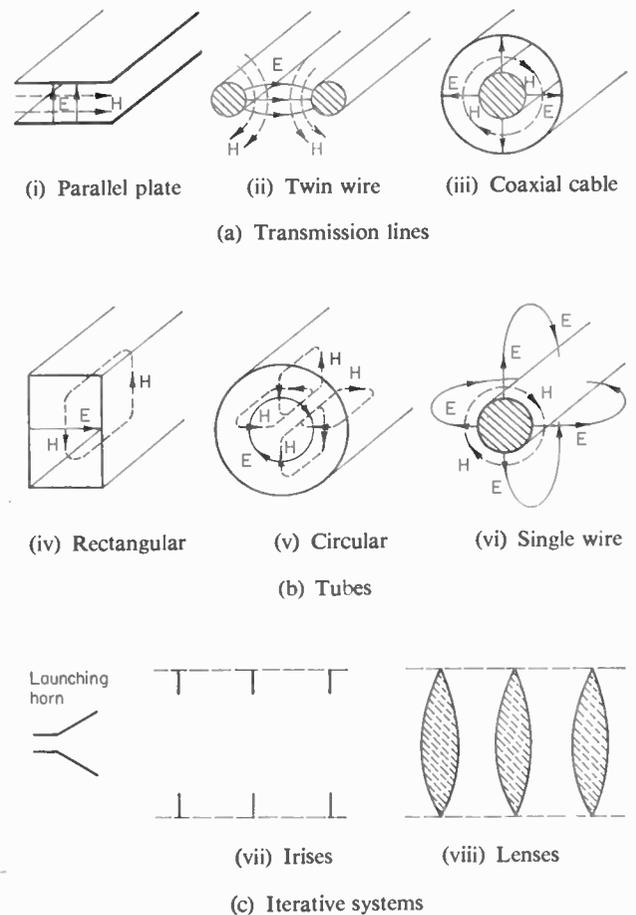


Fig. 1. Different forms of waveguide.

extension of the field in the transverse plane will be dealt with. Thus by implication, to provide the guidance required, either a transverse standing-wave must be brought into operation, as in hollow tubular waveguides, or an evanescent field distribution must be set up at right-angles to the direction of propagation, as in the case of waves exhibiting no cut-off at the lower frequencies.

A guidance system of this kind must incorporate at least two different media providing an interface between them, along which the wave can be propagated. As is well known, Maxwell's equations define precisely the behaviour of an electric wave in a medium of given permeability, permittivity and conductivity so that, knowing the requirements for continuity of the electric and magnetic field components across an interface and the electrical properties of the two media concerned, it becomes a comparatively simple matter to interpret what goes on at the boundary. The truly remarkable thing about Maxwell's theory is that, for homogeneous media, it is practically all-embracing, applicable to the most complex arrangements of conductors and dielectrics. It is true, of course, that in some circumstances the relationships become unmanageable, especially when there are a number of simultaneous variables, but this is not a limitation arising from the way in which the equations have been formulated; rather is it a reflection of the need to investigate structures of conductors and dielectrics, presenting almost unlimited variations. In ordinary practical applications Maxwell's equations have never been known to fail and it is only when relativity or quantum considerations become important that significant limitations appear.^{3,4}

4. Requirements for the Transmission of Power and Modes of Operation

In any electromagnetic field the Poynting Vector gives a measure of the power flow. This is calculated in terms of the product of transverse electric and magnetic field components at any given point and yields the power-density normal to the plane containing the field components (Fig. 2). The total power is then derived by integration over the cross-sectional area of interest.

To obtain finite net power crossing a surface, the electric and magnetic field components over the surface must co-exist to some extent in time and space. Thus, a travelling wave, normal to the surface, provides the requisite conditions for power to pass through it, while a pure standing-wave does not. A perfect reflecting surface cancels the tangential electric field at the surface and consequently consumes no power.

Interwoven with this very helpful and convenient way of interpreting the rate of transport of energy in an electromagnetic field, is the concept of 'lines of force' and whatever their shortcomings in a physical sense, there can be no doubt of their value in providing a visual picture of the mechanism satisfying the requirements. For this Clerk Maxwell was one of the first to acknowledge his indebtedness to Faraday.

Out of an analysis of guided wave propagation, using Maxwell's equations, come particular solutions repre-

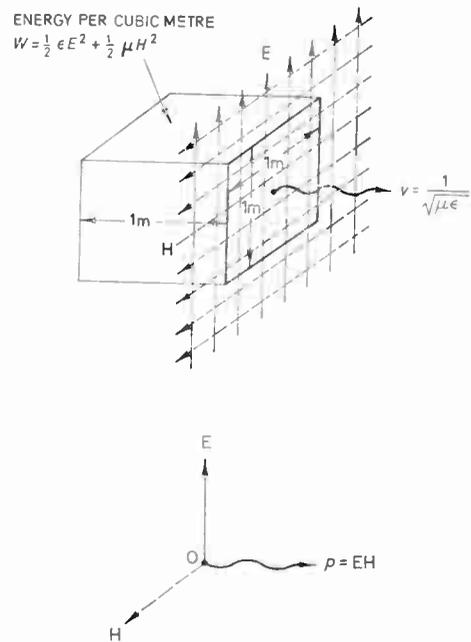


Fig. 2. Energy and power carried by a plane wave.

sending modes of operation appropriate to the boundary conditions, and these are readily identified in relation to their characteristic field distributions. Sometimes very complicated patterns of the electric and magnetic field quantities are found, especially in the case of hybrid waves, but the Poynting Vector always enables power in any given direction to be evaluated.

5. Twin-wire Lines and Coaxial Cables as Waveguides

These structures are perhaps the most familiar ones for the transmission of electrical energy. Since the electromagnetic field, employed in the dominant mode, is very nearly purely transverse to the direction of propagation of the wave and has a configuration that remains practically unchanged whatever the frequency, its useful application extends over the whole spectrum from d.c. upwards to microwaves. Calculations on the basis of so-called 'lumped' parameters of the guide, resistance, inductance, capacitance and leakage, each per unit length, have always been popular because the circuit concept is thereby maintained and, as a rule, high accuracy is offered. The alternative approach, applying Maxwell's equations, takes into consideration progressive change of field distribution from point-to-point along the guide and, in general, when using parallel conductors of high conductivity separated by a homogeneous dielectric, the results are not significantly different. There are, however, cases of growing importance in which the dielectric is stratified and the surfaces supporting the wave, normally represented by conductors, may have impedances that are by no means negligible. In these circumstances Maxwell's equations are required to explain the behaviour satisfactorily. While lumped circuit theory is properly described as the counterpart of field theory, it is only strictly equivalent at relatively low frequencies when the wavelength is

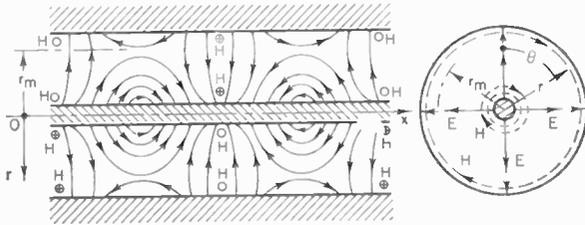


Fig. 3. Hybrid TEM-dual surface wave in a coaxial guide.

large compared with the transverse dimensions of the guide. Thus, we find that the dominant wave set up between parallel conductors and having no frequency cut-off, always includes to some extent surface waves, one associated with each of the two supporting surfaces. In most cases the wave is still dominated by TEM content, but it is necessarily a hybrid with a component of the electric field in the direction of propagation.⁵ (Fig. 3). At high frequencies this may have important consequences, particularly in its effect on dispersion characteristics when the line is used for telecommunications. It happens that the change of phase-velocity with frequency for a pure TEM wave is in the opposite sense to that of a surface wave, and consequently judicious adjustment of the two components in the actual hybrid wave can be expected to bring about almost constant phase-velocity over a very wide frequency band. The technique requires some arrangement for varying the reactance of one of the surfaces supporting the field, without adding significantly to the attenuation, and this is readily achieved by a thin coating of low-loss dielectric whose permittivity is higher than that of the surrounding medium in which the power is transmitted.

The developing interest in other aspects of the use of stratified dielectrics, both with and without conductors, has been pursued recently by Clarricoats,⁶ Oliner⁷ and others. The fact that very good dielectrics suitable for application at high frequencies and having a range of relative permittivities from 2.4 upwards to about 10, such as can be obtained with titanium dioxide loaded polystyrene, are now commercially available, has done much to encourage these investigations. By grading the dielectric, using layers of different materials over the cross-section of a guide, re-distribution of the energy-density of the field can be obtained and if this is towards greater uniformity, lower losses are sometimes possible. Both the twin-wire line and the coaxial cable, with a homogeneous dielectric between the conductors, suffer from bad energy-density distributions in the transverse plane, and with the new dielectrics there appear to be good prospects of improving the situation.

It seems therefore, that there are at least two important objectives that might be profitably pursued in the further development of twin-wire and coaxial structures as waveguides for high frequencies. Because surface roughness of conductors is so important in such circumstances, we must be prepared to recognize more realistically the effect of the surface impedance of the metals employed. The surface-wave content of the hybrid that characterizes the propagation always gives a slight evanescence to the field over the cross-section, and when the outer conductor of a coaxial cable is removed to

infinite radius, the guide transforms into the familiar single-wire transmission line supporting an axially cylindrical surface-wave. It is of interest to observe that proximity of the outer conductor to the inner increases the radial decay of the field and thus the outer conductor, in acting as a screen, tends to compress the field within the boundary which it establishes.

6. Surface Waveguides

Bearing in mind that the single-wire transmission line is, in fact, a particular case of the coaxial structure for which the outer conductor is at infinite radius and that the field distribution in the radial plane has become purely evanescent, it follows that there can be no radiation from the guide so long as there are no bends (Fig. 4).⁸ With a field that extends to infinity, any obstacle outside the supporting surface tends to disturb the smooth propagation of the wave. In practical arrangements, however, the wave is made to hug the guide so closely that any disturbing object more than a few wavelengths away from the surface has relatively negligible effect. Thus the reactance of the wall of the guide is made sufficient to give a rapid radial decay of the field. Unfortunately, this implies strengthening the storage field within the guide itself and as a consequence

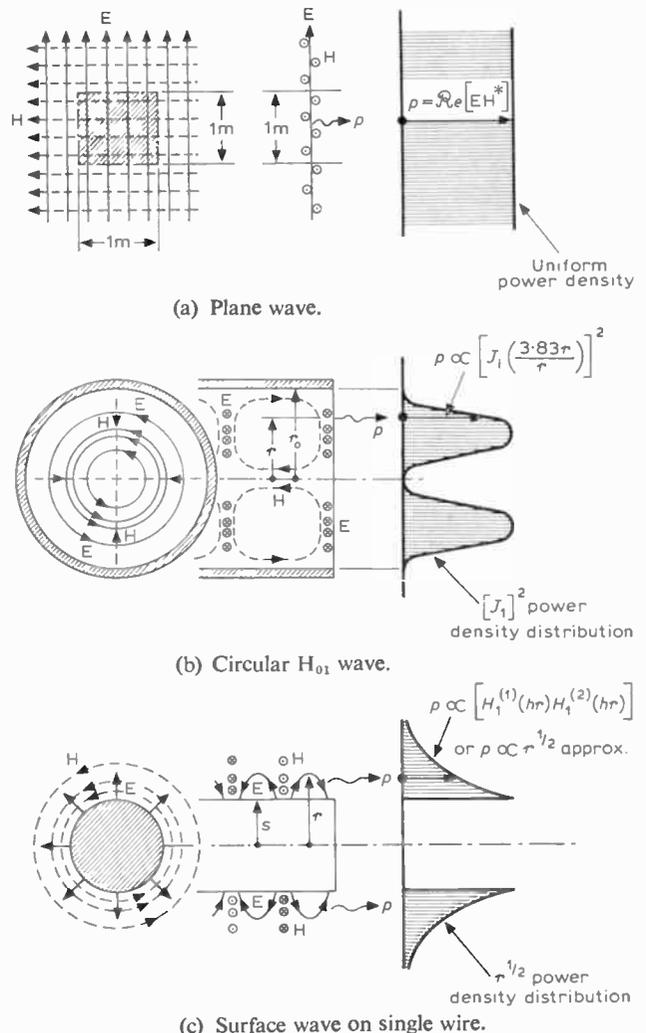


Fig. 4. Power distribution for various waves.

there is a tendency for losses to increase. Moreover, any departure by the guide from a straight line produces radiation and although this can be made small, it represents a handicap in telecommunications applications. To overcome the difficulty a screened surface waveguide may be employed, but this is reverting to the coaxial structure already discussed.

An alternative form of surface waveguide is a dielectric rod or fibre, without any metal core. For the transverse magnetic wave the surface requires to be inductive and therefore must have a sufficient diameter at the frequency of operation. It is found that a hybrid HE_{11} mode (or dipole mode) is also possible and that this has no cut-off. In all cases, as much of the field as possible should be kept in the air outside the guide to reduce losses, without at the same time making the field spread too far from the surface. At optical frequencies this requires a fibre of very small diameter. A coaxial metal shield around a dielectric surface waveguide rod is also feasible and such a structure appears to offer scope for further investigation.

7. Waveguides Employing Transverse Standing Waves

In this form of guide the wave is deliberately reflected backwards and forwards in the transverse plane, while allowing it to progress longitudinally. Thus, the wave may be regarded as including two components, each taking zig-zag paths crossing one another diagonally (Fig. 5). As the frequency is reduced there comes a situation in which it is no longer possible to accommodate the required half-wavelength within the cross-section of the guide and, at this point, cut-off occurs with complete inhibition of propagation. The field along the guide is then evanescent. For efficient operation a guide of this kind must have good reflecting surfaces and high-conductivity metals with a smooth profile are generally used for this purpose. Thus, some form of metal tube with a highly polished inside surface will meet the requirements and has found wide application. With a minimum transverse dimension of the order of a half-wavelength, microwaves are generally required to give manageable sizes.

The hollow tube, trapping within it a microwave, can have almost any cross-section provided it is large enough, but either a rectangular or circular shape is the most common. The corresponding field configuration or mode, supported by the guide, may vary widely according both to its shape and the frequency of operation. Maxwell's equations for the field, applying the appropriate boundary conditions at the surfaces, define

precisely the wave mode to be expected in any given circumstances. It is remarkable how elegantly this is done and how versatile is the theory in application to the most complex structures.

It was Lord Rayleigh⁹ who first drew attention to the possibilities of this kind of waveguide, and the subsequent investigations of a distinguished group of scientists at the Bell Telephone Laboratories¹⁰ in the early 1930s established the technique in practical form.

Today, there is a particular interest in the so-called low-loss guide for millimetre waves, potentially useful in long-distance telecommunications overland (Fig. 4). It so happens that the particular H_{0n} family of modes in a tubular guide of circular cross-section requires a progressively decreasing field strength at the boundary wall as the frequency is raised, and this tends to reduce the current in the surface. For frequencies of 35 GHz up to 100 GHz and a tube of about 5 cm diameter, the losses can be surprisingly small. Unfortunately, many modes are possible in such a guide and when conversion from one to another is freely permitted, not only is the dissipation increased but serious distortion is introduced when a signal is transmitted along the guide. Thus, the skill in application to telecommunications rests on the provision of arrangements for suppressing unwanted modes while allowing unimpeded progress to the H_{01} mode, normally employed. This is done by making the spurious waves travel along the guide more slowly, thereby severely discouraging transfer of energy to them during transit. A circular tube formed either of a wire helix, a bare metal surface transversely corrugated or coated with a thin layer of dielectric, can meet most of the requirements, especially when one of these waveguide structures is associated with mode-filters to clean up any remaining contamination of the desired H_{01} mode (Fig. 6). The engineering development of low-loss waveguide, for possible application to heavily-loaded trunk routes, is now being pursued by the British Post Office and by many telecommunications authorities overseas.¹¹ If the work goes forward according to plan there should be in this country a fully engineered prototype in operation by the mid-1970s.

Encouraged by the rising efficiencies with which high-power microwaves can now be generated, and the availability of suitable rectifiers for conversion to direct current, engineers are also turning their attention to circular H_{01} waveguide in large diameters (between 0.75 metre and 1.0 metre) for operation at about 10 GHz as a means of transmitting powers up to 1000 MW with relatively small losses.¹² This circularly symmetrical

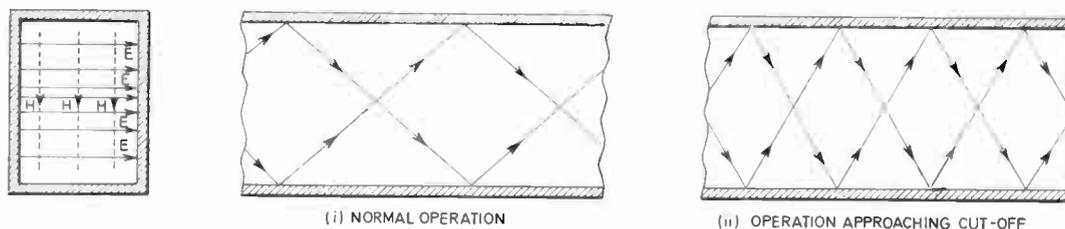


Fig. 5. Wave propagation in rectangular guide.

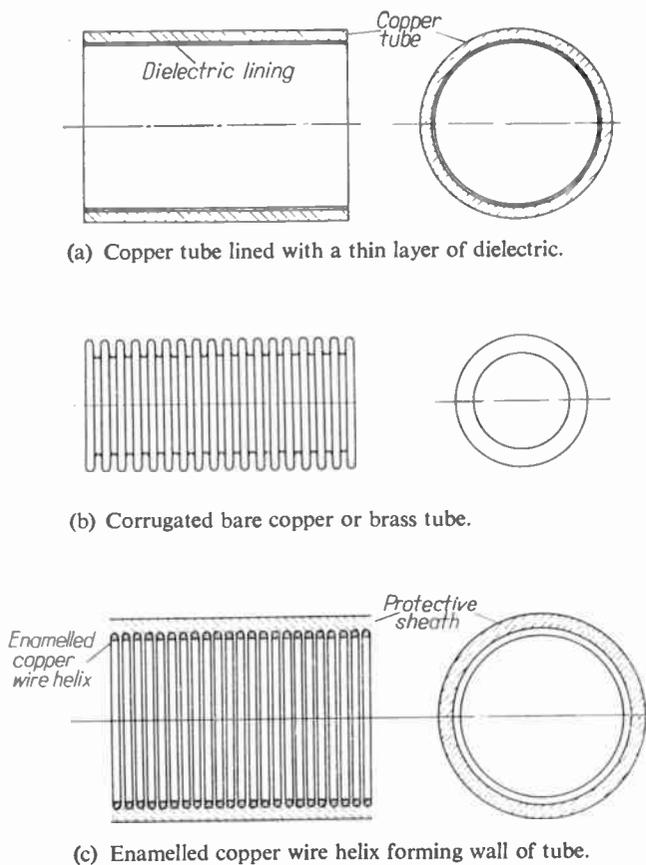


Fig. 6. Waveguide tubes for preserving the H_{01} mode.

mode is, in many respects, ideal for large power, not only because it possesses high electrical breakdown capability, but more particularly because transverse joints in the waveguide present no heating problem as they do in other cases. Recent work has been recorded on a taper for launching such a wave in a tube 75.4 cm (29.7 in) diameter.¹³

Another practical development of special interest is the proposal to use a waveguide, the wall of which is semi-transparent or partly cut away, running along the side of a railway track to provide continuous electrical coupling to trains in motion for signalling services. Among other guides, the circular H_{01} structure is under consideration for this purpose.

8. Conclusion

It is recognized today that guided electromagnetic fields first came into being with the original twin-wire direct current circuit, while the corresponding electric wave operation can be said to have developed when time variation of the electromagnetic field arose.

Maxwell's theory is not limited in any way by such considerations and although it happened that at the time he was formulating his equations, he was particularly concerned with very high frequencies, any part of the spectrum is equally amenable to his treatment. What wonder that we revere this man whose mind encompassed such an achievement, and who built an edifice that will stand for all time.

9. Acknowledgments

The author is much indebted to University College London for the facilities provided in presenting this lecture; to Professor Alex Cullen for making available the demonstration equipment, and especially to Mr. Ron Hillyer for setting up the instruments and operating them.

10. References

1. 'James Clerk Maxwell, Commemoration Volume 1831-1931', with essays by Sir J. J. Thomson, Max Planck, Albert Einstein, Sir Joseph Larmor, Sir James Jeans, William Garnett, Sir Ambrose Fleming, Sir Oliver Lodge, Sir Richard Glazebrook and Sir Horace Lamb.
2. Collins, R. E., 'Field Theory of Guided Waves' (McGraw-Hill, New York, Toronto and London, 1960).
3. Jeans, J. H., 'The Mathematical Theory of Electricity and Magnetism' (Cambridge University Press, 1925).
4. Cullwick, E. G., 'Electromagnetism and Relativity' (Longmans Green, London, 1957).
5. Barlow, H. M., 'Hybrid TEM-dual surface wave in a coaxial cable', *Proc. Instn Elect. Engrs*, 116, No. 4, pp. 489-94, 1969.
6. Clarricoats, P. J. B., 'Propagation along unbounded and bounded dielectric rods', *Proc. Instn Elect. Engrs*, 108C, pp. 170-86, 1961.
7. Oliner, A. A., 'A new class of reactive-wall waveguides for low-loss applications', *Advances in Radio Science, International Union of Radio Science, 16th General Assembly, Ottawa, August 1969*.
8. Barlow, H. M., and Brown, J., 'Radio Surface Waves' (Clarendon Press, Oxford, 1962).
9. Lord Rayleigh, 'On the passage of electric waves through tubes or the vibrations of dielectric cylinders', *Phil. Mag.*, 43, pp. 125-32, 1897.
10. Schelkunoff, S. A., 'Electromagnetic theory of lines and shields', *Bell Syst. Tech. J.*, 13, pp. 522-79, 1934.
11. Southworth, G. C., 'Hyper-frequency wave guides—general considerations and experimental results', *Bell Syst. Tech. J.*, 15, pp. 284-308, 1936.
12. Proceedings of the Conference on 'Trunk Telecommunications by Guided Waves', London, September 1970. (I.E.E. Conference Publication No. 71.)
13. Okress, E. (Ed.), 'Microwave Power' (Academic Press, New York, 1970).
13. Tomiyasu, K., '17.5 ft-long multiconical taper for TE_{01} mode in 29.7 in-diameter waveguide at X band', *Proc. Instn Elect. Engrs*, 116, No. 3, pp. 373-6, March 1969.

Manuscript received by the Institution on 10th September 1970. (Address No. 44.)

Influence of Noise in Automatic Measurements by Insertion Test Signals

By

G. BARBIERI, Dr.Ing.†

and

P. D'AMATO, Dr.Ing.†

Reprinted from the Proceedings of the Conference on Television Measuring Techniques held in London from 12th to 14th May 1970.‡

A theoretical study is presented on the noise affecting the automatic television measurements carried out by test signals in the field blanking interval (insertion test signals). On the basis of the results it is shown that special signals (national test signals) can be used, for automatic measurements, instead of the international signals and the measurement devices are defined so as to obtain measurements affected as little as possible by noise.

1. Introduction

In recent years the need has arisen to carry out continuous measurements during normal transmissions throughout the whole of a television distribution network. For this purpose, special test signals are inserted in the vertical blanking interval of the video signal and the distortion affecting these signals is measured at different points of the network.

The C.C.I.R. has defined the characteristics of the international test signals to be inserted in lines 17, 18,

330 and 331 of the video signal.¹ These signals, shown in Fig. 1, are particularly suitable for oscillographic measurements, but are not equally suitable for remotely-controlled automatic measurement.

The RAI, which is currently studying an automatic measurement system, has adopted special national signals.² This is in order to simplify the data acquisition devices as much as possible without neglecting, however, measurement accuracy, especially when noise is present. The national test signals cover lines 19, 20, 21 and

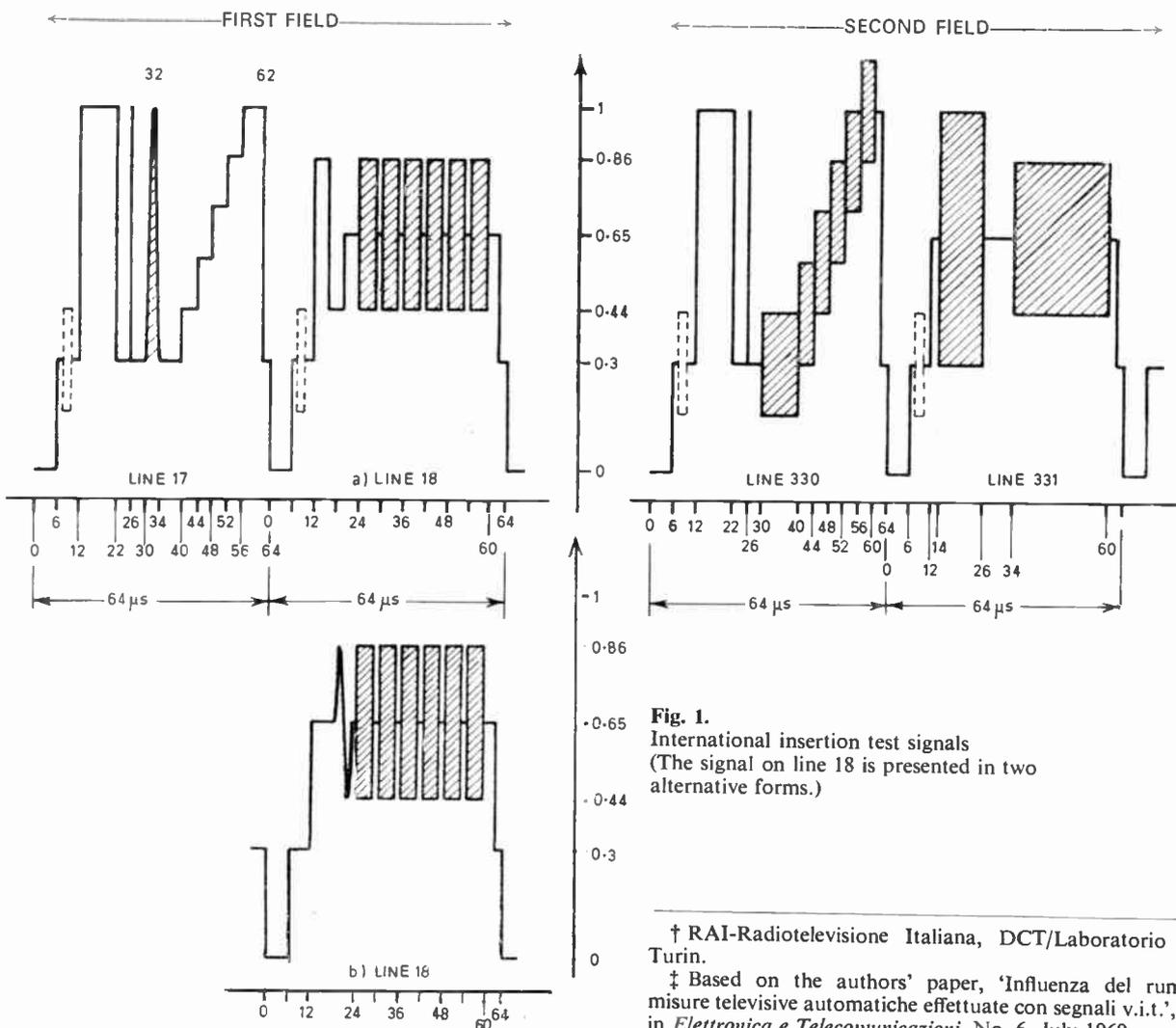


Fig. 1.
International insertion test signals
(The signal on line 18 is presented in two alternative forms.)

† RAI-Radiotelevisione Italiana, DCT/Laboratorio Ricerche, Turin.

‡ Based on the authors' paper, 'Influenza del rumore nelle misure televisive automatiche effettuate con segnali v.i.t.', published in *Elettronica e Telecomunicazioni*, No. 6, July 1969.

LINES	19 & 332 (BLACK)	20 & 333 (GREY)	21 & 334 (WHITE)	NUMBER OF FIELDS	MEASUREMENTS
LUMINANCE WAVEFORMS					
SUPER- IMPOSED SINE WAVES OF 0.2 V _{pp}	1 MHz	1 MHz	1 MHz	1	QUASI-STATIC NON-LINEARITY
	3 MHz	3 MHz	3 MHz	1	AMPLITUDE- FREQUENCY RESPONSE 1st VALUE
	4.43 MHz	4.43 MHz	4.43 MHz	1	AMPLITUDE- FREQUENCY RESPONSE (2nd VALUE) DIFFERENTIAL GAIN AND PHASE
				7	NOTHING

Fig. 2. National insertion test signals.

332, 333, 334 of the video signal as shown in Fig. 2. Three different luminance levels are always present on these lines, namely the black level on lines 19 and 332, the grey level on lines 20 and 333 and the white level on lines 21 and 334.

Sinewave signals with 0.2 V peak-to-peak amplitude and different frequencies are superimposed on these levels according to a 10-fields-cycle. In the three lines of the 1st field there is a 1 MHz burst; in the three lines of the second field there is a 3 MHz burst and in the three lines of the third field a burst at the colour sub-carrier frequency (4.433 6187 5 MHz). (In the remainder of the paper the colour sub-carrier frequency will be abbreviated to 4.43 MHz.)

In the next fields no signal is superimposed on the luminance levels.† The aim of this paper is to study the influence of noise on automatic measurements in order to explain the reason why national insertion signals were adopted and to define the characteristics of data acquisition devices, so as to obtain measurements affected as little as possible by noise. The measurements which can be carried out with the national test signals are:

- (a) Quasi-static linearity. This is made by using the 1 MHz bursts of the first field. The measurement result is the percent error between the burst which has got the maximum amplitude and the burst with minimum amplitude, namely:

$$\epsilon\% = \left(1 - \frac{V_{\min}}{V_{\max}}\right) \cdot 100$$

- (b) Differential gain. The measurement is made in the same way as the previous one, but with the 4.43 MHz bursts.‡
- (c) Amplitude-frequency response. The amplitude of the bursts at 3 MHz and 4.43 MHz is compared to the amplitude of the burst at 1 MHz:

$$\epsilon_{3/1}\% = \left(\frac{V_3}{V_1} - 1\right) \cdot 100 \quad \epsilon_{4.43/1}\% = \left(\frac{V_{4.43}}{V_1} - 1\right) \cdot 100$$

† The cycle of the signal has been varied with respect to that of ref. 2.

The choice of the sign is done so that the error is positive if $V_3 > V_1$ or $V_{4.43} > V_1$.

The bursts used are those superimposed on the d.c. level corresponding to grey (lines 20 and 333).

- (d) Differential phase. The phase shift between the bursts at 4.43 MHz superimposed on different luminance levels is measured. Two values are given: black-to-grey and grey-to-white differential phase.‡

As it may be noted, all the measurements, except for differential phase, are 'percent error' measurements between sinewave amplitudes on different levels and with different frequencies. It is therefore possible to make a unique theory for all the measurements, except for differential phase, which shall be studied separately.

2. Introductory Notes on the 'Percent Error Meter'

In order to carry out these kind of measurements, first of all it is necessary to separate the bursts to be compared from the video signal and to measure their mean or peak value. Since the bursts appear on the video subsequently, the results of these measurements have to be stored; two of them must then be chosen and sent simultaneously to the analogue-to-digital converter, which computes the percent error.³

First of all we want to examine the behaviour of the detector and store shown in Fig. 3. A band-pass filter has been inserted, adjusted to the burst frequency f_0 , which can be useful to eliminate not only the luminance levels, but also part of the noise added to the video signal. The filter is followed by a detector, the characteristics of which will be examined later, and by a store, which should be analogue for reasons of cost and simplicity.



Fig. 3. Block diagram of detector and store.

An operational amplifier, the input of which is suitably gated, is used as the store. This amplifier integrates the input signal from the instant in which the filter output has reached its steady state, until the end of the burst. Once the burst ends, the gate switches off and the output remains constant for the whole time necessary to complete measurement operations. Integration or 'averaging' tends to eliminate the a.c. components of the signal outcoming from the detector.

In order to establish the filter, detector and integrator characteristics most satisfactorily, it is necessary to determine the statistical properties of noise and signal/noise ratio in the various points of the chain shown in Fig. 3.

‡ These measurements differ from those laid down in the C.C.I.R. standard.—Editor.

3. Statistical Properties of Noise and Signal/Noise Ratio in the Various Points of the Circuit of Fig. 3

The useful signal fed to the circuit is assumed to consist of a continuous sinewave:

$$v_s = V_s \cos \omega_0 t$$

Hence at point A the signal power (apart from a constant equal to the inverse of the input resistance) is:

$$[P_s]_A = V_s^2/2$$

There is also at point A a noise voltage $n_A(t)$ with Gaussian distribution and with a rectangular spectrum between 0 and 5 MHz (video band limits). Let η_A be the power spectral density. The variance σ_A^2 , which gives the power associated with noise (apart from the same constant as above) is:

$$\sigma_A^2 = \eta_A f_c$$

with $f_c = 5$ MHz.

The ratio between signal power and noise power at point A is hence:

$$\left[\frac{P_s}{\sigma^2} \right]_A = \frac{V_s^2}{2\eta_A f_c}$$

The filter is assumed to have a response similar to that of a parallel resonant circuit. For such a filter the transfer function modulus is:

$$\frac{V_o}{V_u} = \frac{a}{\sqrt{a^2 + (f - f_0)^2}}$$

The value of a determines the filter selectivity, whereas f_0 is the resonance frequency, which coincides with the signal frequency. The filter response is assumed symmetrical with respect to f_0 ; this is acceptable if $a \ll f_0$ (Fig. 4). The filter does not alter the power associated with the signal, but it remarkably reduces noise power.

At point B there is in fact a noise power spectral density equal to:

$$\eta_B = \eta_A \frac{a^2}{a^2 + (f - f_0)^2}$$

The noise power is therefore:

$$\begin{aligned} \sigma_B^2 &= \eta_A \int_0^{f_c} \frac{a^2}{a^2 + (f - f_0)^2} df \approx \eta_A \int_{-\infty}^{+\infty} \frac{a^2}{a^2 + (f - f_0)^2} df \\ &= \eta_A a\pi \end{aligned}$$

The ratio between signal and noise powers at point B is

$$\left[\frac{P_s}{\sigma^2} \right]_B = \frac{V_s^2}{2\eta_A a\pi}$$

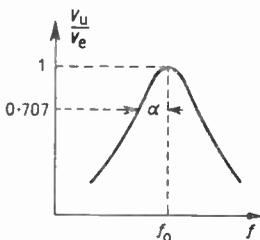


Fig. 4. Amplitude-frequency response of the band-pass filter.

Because of a well-known property of linear circuits, the statistical distribution of noise at the filter output is still Gaussian.

The filter is followed by the detector, which for the present can be assumed 'ideal'. At its output an ideal detector gives a signal proportional to the envelope of the incoming signal. Hence the study of the useful signal and noise properties at point C is connected with the study of the envelope of signal and noise at point B.

It is advisable to substitute a discrete spectrum for the continuous spectrum of noise at point B as shown in Fig. 5.

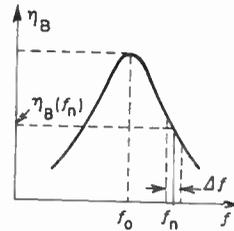


Fig. 5. Approximation of a continuous spectrum by a discrete spectrum.

The continuous spectrum is divided into strips covering frequency intervals Δf , centred on frequencies f_n . Every portion of the continuous spectrum is substituted by a sinewave with frequency f_n and an amplitude A_n , such that the mean power is equal to that of the corresponding continuous spectrum portion. Hence:

$$A_n = \sqrt{2\eta_B(f_n) \cdot \Delta f}$$

The phases φ_n are uncorrelated random variables with uniform distribution between 0 and 2π . The sum of the sinewaves:

$$n'(t) = \sum_n A_n \cos(\omega_n t + \varphi_n) \quad \dots\dots(1)$$

can well approximate to the noise voltage at point B, i.e. $n_B(t)$. It is possible indeed to prove^{4,5} that $n_B(t)$ is a random function with Gaussian distribution and variance:

$$\sigma^2 = \frac{1}{2} \sum_n A_n^2$$

This is a particular case of the central limit theorem.

By exploiting the approximate representation it can be proved that the noise at the filter output can take the form of a single sinewave voltage, namely

$$n_B(t) = V(t) \cos[\omega_0 t + \Phi(t)]$$

In other words, the noise $n_B(t)$ can be considered a sinewave, with frequency f_0 and random amplitude and phase. In fact, from (1) it may be derived as

$$\begin{aligned} n_B(t) &= \sum_n A_n \cos[\omega_n t + \varphi_n(t)] \\ &= \sum_n A_n \cos[(\omega_n - \omega_0)t + \varphi_n(t)] \cdot \cos \omega_0 t - \\ &\quad - \sum_n A_n \sin[(\omega_n - \omega_0)t + \varphi_n(t)] \cdot \sin \omega_0 t \\ &= x(t) \cos \omega_0 t - y(t) \sin \omega_0 t. \end{aligned}$$

The amplitudes $x(t)$ and $y(t)$, i.e. the components of vector $n_B(t)$ are random functions, the spectrum of which, because of the filter response symmetry, has the same shape of that of $n_B(t)$, but shifted towards the origin and with double ordinates.

Besides, since $x(t)$ and $y(t)$ are the sum of many uncorrelated sinewaves, with definite amplitudes and frequencies and random phases, their distribution is also Gaussian. It is also possible to prove that they are independent.⁵

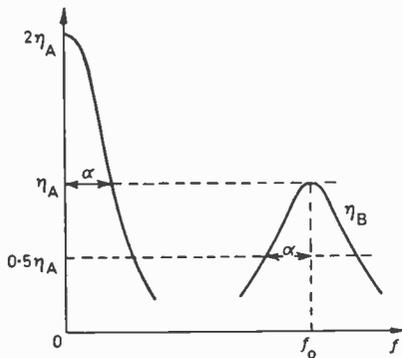


Fig. 6. Spectra of $x(t)$ and $y(t)$ in relation to the noise spectrum at the detector input.

Even though it is not necessary for the present work, it is useful to remember that the envelope $V(t)$ has a Rayleigh probability density $q(V)$, namely

$$q(V) = \frac{V}{\sigma_B^2} \exp\left(-\frac{V^2}{2\sigma_B^2}\right)$$

whereas phase $\Phi(t)$ has a uniform probability density $q(\Phi)$ in the interval $(0, 2\pi)$, i.e.:

$$q(\Phi) = \frac{1}{2\pi}, \text{ for } 0 \leq \Phi \leq 2\pi$$

$$q(\Phi) = 0, \text{ for } \Phi < 0; \Phi > 2\pi.$$

Once the properties of noise $n_B(t)$ alone have been defined, the properties of the whole signal can be derived. This is rather difficult in the general case.⁴⁻⁶

Anyhow, if the signal/noise ratio is high, the problem is very easily solved. In fact, it can be assumed that the representative vector of the useful signal is amplitude-modulated only by the in-phase component of noise $x(t)$, whereas the quadrature component $y(t)$ causes only a phase modulation. Hence the envelope of the useful signal and noise consists of a component V_s and a random function $x(t)$, the statistical properties of which have already been described.

Phase properties, connected with those of the function $y(t)$ will be examined under Section 7 dealing with differential phase measurement. At point C of the diagram of Fig. 3 the noise is Gaussian and has the following power spectral density:

$$\eta_C = 2\eta_A \frac{a^2}{a^2 + f^2}$$

and the noise power is:

$$\sigma_C^2 = 2\eta_A \int_0^{f_c} \frac{a^2}{a^2 + f^2} df \approx 2\eta_A \int_0^\infty \frac{a^2}{a^2 + f^2} df = \eta_A a\pi.$$

As may be noted, the noise power is equal to that before the detector. As to the useful signal, the detector doubles its power, i.e.

$$[P_s]_C = V_s^2 = 2[P_s]_B$$

Hence the signal/noise ratio, too, is doubled.

The theory is correct only in case of an ideal detector. The real detector can be an average- or a peak-detector. It is necessary to choose between the two kinds, considering the real and not the theoretical behaviour. It has been experimentally observed that the peak detector operates in a different way from the theory, since the capacitor in parallel to the load resistance normally charges to the value of noise peaks.

Therefore, by keeping the useful signal constant, at the detector output a d.c. voltage is obtained, which increases as input noise increases. This serious inconvenience does not occur for the average detector. This operates rather in conformity with the theory, since the mean value of a rectified signal is proportional to the envelope with the coefficient $1/\pi$ or $2/\pi$ according to whether rectification is half-wave or full-wave.

It must be added that at the output of such a detector there are, besides noise and the d.c. useful signal, periodical oscillations at multiples of f_0 . These oscillations, however, are considerably reduced by the circuits which follow, and hence will not be considered.

The detector is followed by the storage circuit, which in this analysis is substituted by an ideal device giving, at every instant, an output voltage equal to the average of the input signal, computed in a time interval T , immediately preceding the considered instant.

The output of such a device is a continuous function, the statistical properties of which have to be determined. In order to evaluate the transfer function it must be observed that the response to a generic sinewave input voltage

$$v_e = V_e \cos \omega t$$

is

$$\begin{aligned} v_u &= \frac{1}{T} \int_{t-T}^t V_e \cos \omega \theta d\theta \\ &= \frac{V_e}{\omega T} [\sin \omega(t-T) - \sin \omega t]. \end{aligned}$$

Hence the transfer function is

$$\frac{V_u}{V_e} = \frac{\exp(-j\omega T) - 1}{j\omega T}$$

The diagram of the squared modulus is shown in Fig. 7. The greater T is, the more marked is the filtering action of the device. The integrator, thus conceived, does not affect the useful signal, hence:

$$[P_s]_D = V_s^2 = [P_s]_C$$

The noise keeps a Gaussian distribution, but its

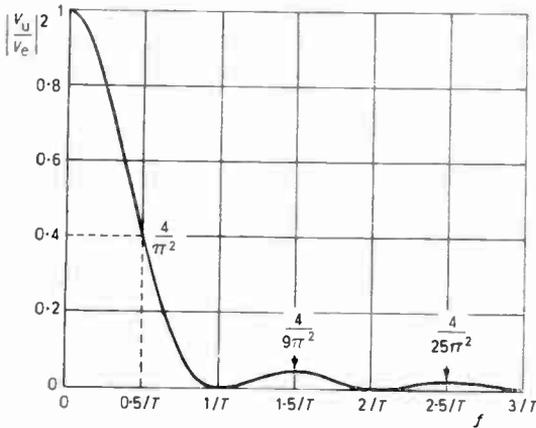


Fig. 7. Square of the transfer function modulus of an ideal integrator.

frequency spectrum is strongly narrowed. In fact:

$$\eta_D = 2\eta_A \frac{a^2}{a^2 + f^2} \cdot \frac{1 - \cos 2\pi T f}{2\pi^2 T^2 f^2}$$

and the noise power is:

$$\sigma_D^2 = 2\eta_A \int_0^{f_c} \frac{a^2}{a^2 + f^2} \cdot \frac{1 - \cos 2\pi T f}{2\pi^2 T^2 f^2} df$$

This integral can be computed only by numerical methods.

The signal/noise ratio is

$$\left[\frac{P_s}{\sigma^2} \right]_D = \frac{V_s^2}{2\eta_A \int_0^{f_c} \frac{a^2}{a^2 + f^2} \cdot \frac{1 - \cos 2\pi T f}{2\pi^2 T^2 f^2} df}$$

As a conclusion of this study, it may be interesting to consider how much the noise is reduced and the signal/noise ratio is improved, in the path from point A to point D.

For this purpose the factor ζ is introduced, defined as the ratio between noise power in D and A:

$$\zeta = \frac{\sigma_D^2}{\sigma_A^2} = \frac{2}{f_c} \cdot \int_0^{f_c} \frac{a^2}{a^2 + f^2} \cdot \frac{1 - \cos 2\pi T f}{2\pi^2 T^2 f^2} df$$

Factor ρ is also introduced as the ratio between the signal/noise ratios in D and A:

$$\rho = \frac{\left[\frac{P_s}{\sigma^2} \right]_D}{\left[\frac{P_s}{\sigma^2} \right]_A} = \frac{2}{\zeta} = \frac{f_c}{\int_0^{f_c} \frac{a^2}{a^2 + f^2} \cdot \frac{1 - \cos 2\pi T f}{2\pi^2 T^2 f^2} df}$$

4. Determination of the Filter and Integrator Characteristics for a 'Percent-Error-Meter' for National Insertion Test Signals

By using the relations of the preceding Sections, it is possible to design the filter and integrator in the best way. If we consider national insertion signals, the measurement of the mean voltage can be made for a maximum time $T = 48 \mu s$, in order to have a 2 μs safeguard before and after the measurement time. More precisely, if ΔT is the total available time and T the

integration time, the time reserved for transient phenomena produced by the filter is $\Delta T - T$. Transients die in about 5 time constants τ . Since in a resonant circuit there is the following connexion between τ and a

$$\tau = \frac{2L}{R} = \frac{2Q}{\omega_0} = \frac{1}{2\pi a},$$

between T and a there must be the following relationship:

$$T = \Delta T - 5\tau = \Delta T - \frac{5}{2\pi a},$$

or

$$a = \frac{5}{2\pi(\Delta T - T)}$$

The factors ζ and ρ depend on the characteristics of the filter and of the integrator, which are mutually related by the preceding equations. With a simple substitution it is therefore possible to express ζ and ρ as functions of the variable T :

$$\zeta = \frac{2}{f_c} \int_0^{f_c} \frac{25}{25 + 4\pi^2 (\Delta T - T)^2 f^2} \cdot \frac{1 - \cos 2\pi T f}{2\pi^2 T^2 f^2} df$$

$$\rho = \frac{f_c}{\int_0^{f_c} \frac{25}{25 + 4\pi^2 (\Delta T - T)^2 f^2} \cdot \frac{1 - \cos 2\pi T f}{2\pi^2 T^2 f^2} df}$$

ζ and ρ have been evaluated by means of a computer for different values of T , in order to find in which conditions ζ becomes minimum and maximum. The results are shown in the following table.

T (μs)	a (MHz)	ζ (dB)	ρ (dB)
6	0.018	-20.20	23.21
12	0.022	-20.70	23.71
18	0.026	-21.20	24.21
24	0.033	-21.82	24.83
30	0.044	-22.38	25.39
36	0.066	-22.90	25.91
42	0.132	-23.40	26.41
48	∞	-23.81	26.82

It may be inferred that it is advisable to increase the integration time as much as possible. The filter can also be completely eliminated. From the above considerations a percent-error-meter to operate with national insertion signals can be defined (Fig. 8). By eliminating the three bandpass filters (one for each frequency), it is possible to use a single average detector preceded by a differentiating network which eliminates the luminance levels. In this way the circuit is remarkably simplified, without any impairment of its noise immunity. The differentiator can have a 800 ns time-constant. In this case, transients virtually die in 4 μs . It must be noted, however, that if the average-detector is full-wave, it gives a nearly correct answer even though transients are not completely dead. Therefore the integration time can still be 48 μs , so that part of transients occur in the integration time and part in the guard interval at the beginning of the burst. The detector output is sent,

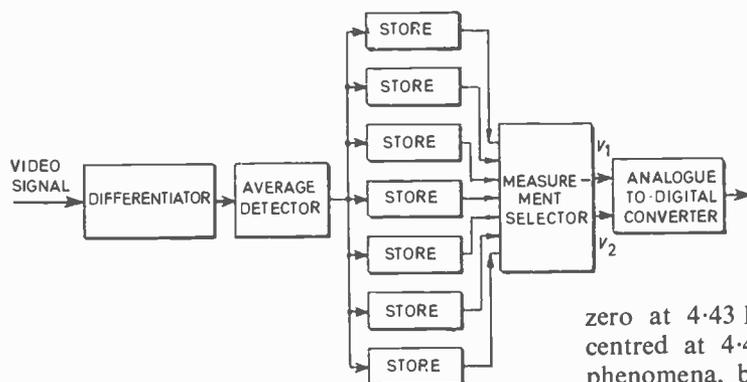


Fig. 8.
Block diagram of the
percent-error meter.

through suitable gates, to the storage devices. These devices store the amplitude of the bursts involved in the measurements (exactly 7, i.e. the three bursts at 1 MHz, the three bursts at 4.43 MHz and the one at 3 MHz superimposed on the grey level). The storage outputs are sent to a 'measurement selector'. According to the measurement to be performed, this circuit selects two of the inputs and feeds them into the analogue-to-digital converter which computes the percent error.

5. Percent-Error Measurement with International Insertion Test Signals

If international insertion signals are used, it is not possible to carry out quasi-static linearity measurements, with a method similar to that of national signals.

Therefore we only consider the measurements of amplitude-frequency response and differential gain. The amplitude-frequency response measurements are made by means of the 'multiburst' of line 18 (Fig. 1). It is advisable that the 'multiburst' signal be directly fed to an average-detector, avoiding filters, which would increase the complication of the circuit without being useful in eliminating noise. The detector should be followed by stores similar to those used with national signals. The influence of noise is more noticeable because of the shorter integration time ($5.5 \mu\text{s}$). The impairment is easily assessed by observing that the output noise power is inversely proportional to the integration time, when no filter is applied. Since the ratio between integration times in the two cases is 8.72, ζ and ρ are worsened by 9.4 dB. For a total evaluation of the performance of national and international test sets it is more interesting to compare the signal/noise ratios in the two cases, the signal/noise ratios of the input video signals being equal.

By remembering that the 'multiburst' peak-to-peak amplitude is 420 mV, whereas the peak-to-peak amplitude of the bursts used in national signals is only 200 mV, it may be easily deduced that the application of international signals causes a worsening by 3 dB of the output signal/noise ratio.

Differential gain measurement is made with the 4.43 MHz sinewave superimposed on the staircase of line 330.

During signal generation, the staircase is shaped by a Thomson filter with a transfer function having the first

zero at 4.43 MHz.¹ In this way a band-pass filter, centred at 4.43 MHz, does not give rise to transient phenomena, because of the luminance level variations. On the risers a transient occurs, due to the distortion which may affect the 4.43 MHz sinewave. This transient is assumed to die only after three time-constants. The transient at the black level, coinciding with the beginning of the sinewave, lasts longer, but the time available for the measurement is $10 \mu\text{s}$ instead of $4 \mu\text{s}$. For these reasons it is advisable to eliminate the d.c. levels by means of a band-pass filter, instead of a differentiating network. This filter may have a 500 ns time-constant, to which a Q of 7 corresponds. In this case $2.250 \mu\text{s}$ for integration and 250 ns for safeguard would be available. The integration time is so short that it must be multiple of the sinewave period: otherwise the integration result depends on the phase relation between the gate and the sinewaves.

Figure 9 shows the waveforms of the signal at the band-pass filter output with $Q = 7$, in case of a strong differential gain distortion.

In the situation above described, the following values of ζ and ρ are obtained:

$$\zeta = -11.6 \text{ dB}; \quad \rho = +14.6 \text{ dB}.$$

With respect to the national test signals ζ and ρ are worsened by 12.2 dB. It is useful to compare again the signal/noise ratios at the integrator output in the condition of identical signal/noise ratio in the input video signal, as it has already been done for the amplitude-frequency response measurement. As the sinewaves superimposed on the staircase have a peak-to-peak amplitude of 280 mV, the use of international signals causes a worsening in ζ and ρ by 9.3 dB.

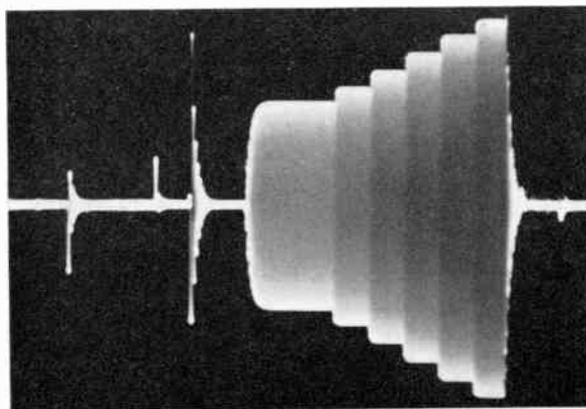


Fig. 9. Waveform of the signal at the band-pass filter output.

6. Statistical Properties of Percent-Error Measurements

We want to establish the statistical properties of the error:

$$\varepsilon\% = \left(1 - \frac{V_1}{V_2}\right) \times 100$$

once the statistical properties of voltages V_1 and V_2 are known.† From the results of Section 3 it may be inferred that V_1 and V_2 are random functions with Gaussian distribution, variance σ_{D1} and σ_{D2} , and mean value V_{10} and V_{20} , which coincides with the 'true value' of the measurement, i.e. the value expected if noise is absent.

It is also assumed that $\sigma_{D1} = \sigma_{D2} = \sigma_D$, since the sinewave bursts are picked up from the same video signal and are hence affected by the same noise.

The statistical distributions of V_1 and V_2 are therefore:

$$q_1(V_1) = \frac{1}{\sqrt{2\pi}\sigma_D} \exp\left[-\frac{(V_1 - V_{10})^2}{2\sigma_D^2}\right]$$

$$q_2(V_2) = \frac{1}{\sqrt{2\pi}\sigma_D} \exp\left[-\frac{(V_2 - V_{20})^2}{2\sigma_D^2}\right]$$

Since V_1 and V_2 are statistically independent, the error probability density can be derived by the formula:

$$q(\varepsilon) = \int_0^\infty q_2(V_2) \cdot q_1[V_2(1-\varepsilon)] \cdot \left|\frac{dV_2}{d\varepsilon}\right| \cdot dV_2,$$

i.e.

$$q(\varepsilon) = \int_0^\infty \frac{1}{2\pi\sigma_D^2} \times \exp\left[-\frac{(V_2 - V_{20})^2 + [V_2(1-\varepsilon) - V_{10}]^2}{2\sigma_D^2}\right] \times V_2 dV_2 \dots\dots(2)$$

The choice of integration limits has been done remembering that V_1 and V_2 are always positive.

The integral in (2) is of the kind:

$$\int_0^\infty \exp[-(ay^2 + by + c)] y dy, \dots\dots(3)$$

and can be solved by considering that:⁷

$$\int_0^\infty \exp[-(ay^2 + by + c)] dy = \frac{1}{2} \sqrt{\frac{\pi}{a}} \exp\left[-\frac{b^2 - ac}{a}\right] \cdot \left(1 - \operatorname{erf} \frac{b}{\sqrt{a}}\right)$$

The error function, $\operatorname{erf} z$ (where z is a general variable), has the following expression:

$$\operatorname{erf} z = \frac{2}{\sqrt{\pi}} \int_0^z e^{-t^2} dt$$

The solution of the integral (3) is hence:

$$\frac{1}{2a} e^{-c} - \frac{b}{4a\sqrt{a}} \exp\left[-\frac{b^2 - ac}{a}\right] \cdot \left(1 - \operatorname{erf} \frac{b}{\sqrt{a}}\right).$$

† The sign convention for error and the fact that it is expressed in relative value and not in percent is immaterial.

In our case we get:

$$q(\varepsilon) = \frac{1}{2\pi[1+(1-\varepsilon)^2]} \cdot \exp\left[-\frac{V_{10}^2 + V_{20}^2}{2\sigma_D^2}\right] + \frac{(1-\varepsilon)V_{10} + V_{20}}{2\sqrt{2}\cdot[1+(1-\varepsilon)^2]^{3/2}\sqrt{\pi}\sigma_D} \times \exp\left[-\frac{[(1-\varepsilon)V_{20} - V_{10}]^2}{2\sigma_D^2[1+(1-\varepsilon)^2]}\right] \times \left[1 - \operatorname{erf}\left(-\frac{(1-\varepsilon)V_{10} + V_{20}}{\sqrt{2\sigma_D^2}\cdot[1+(1-\varepsilon)^2]}\right)\right] \dots\dots(4)$$

This complicated expression is simplified remarkably when, for both V_1 and V_2 , the signal/noise ratio is high enough (at least 20 dB). In this case we have

$$\frac{1}{2\pi[1+(1-\varepsilon)^2]} \exp\left[-\frac{V_{10}^2 + V_{20}^2}{2\sigma_D^2}\right] \approx 0$$

$$\operatorname{erf}\left[-\frac{(1-\varepsilon)V_{10} + V_{20}}{\sqrt{2\sigma_D^2}\cdot[1+(1-\varepsilon)^2]}\right] \approx -1$$

We get then:

$$q(\varepsilon) \approx \frac{(1-\varepsilon)V_{10} + V_{20}}{\sqrt{2\pi}\cdot\sigma_D[1+(1-\varepsilon)^2]^{3/2}} \times \exp\left[-\frac{[(1-\varepsilon)V_{20} - V_{10}]^2}{2\sigma_D^2[1+(1-\varepsilon)^2]}\right] \dots\dots(5)$$

If the previous assumption is valid, the mean value or true value

$$\varepsilon_0 = 1 - \frac{V_{10}}{V_{20}}$$

can be substituted in (5) in the place of ε throughout, except in the numerator of the argument of the exponential function. Making the substitution we get:

$$q(\varepsilon) = \frac{V_{20}}{\sqrt{2\pi}\sigma_D\sqrt{1-(1-\varepsilon_0)^2}} \exp\left[-\frac{(\varepsilon - \varepsilon_0)^2 \cdot V_{20}^2}{2\sigma_D^2 \cdot [1+(1-\varepsilon_0)^2]}\right]$$

and hence:

$$q(\varepsilon) = \frac{1}{\sqrt{2\pi}\sigma_\varepsilon} \exp\left[-\frac{(\varepsilon - \varepsilon_0)^2}{2\sigma_\varepsilon^2}\right], \dots\dots(6)$$

where:

$$\sigma_\varepsilon = \frac{\sigma_D\sqrt{1+(1-\varepsilon_0)^2}}{V_{20}} \dots\dots(7)$$

The statistical distribution of the percent error is still Gaussian with mean value ε_0 and variance σ_ε^2 .

This approximate result can be obtained more easily by differentiating with respect to V_1 and V_2 :

$$d\varepsilon = -\frac{1}{V_{20}} dV_1 + \frac{V_{10}}{V_{20}^2} dV_2$$

By remembering that a linear combination:

$$z = ax + by$$

of two random variables x and y with Gaussian distribution has a Gaussian distribution, too, and that its variance is:

$$\sigma_z^2 = a^2 \sigma_x^2 + b^2 \sigma_y^2$$

it is possible to derive again the expressions (6) and (7) for $q(\varepsilon)$ and σ_ε^2 .

7. Differential Phase Measurement

Figure 10 shows the block diagram of an instrument which can be used for this measurement. The measurement is performed by the well-known diode phase detector, shown in Fig. 11.

The detector output is a function of the two input signals amplitudes and of their phase difference $\Delta\phi$. By limiting the two input signals it is possible to eliminate the influence of their amplitude on the detector output. In order to obtain a good linearity of the detector characteristic it is advisable to make equal the amplitudes of the signals outcoming from the limiters. Under this assumption, the detector output is:

$$u = K \sin \frac{\Delta\phi}{2}$$

where K is a proportionality constant. Since the phase error seldom exceeds 40° , the law:

$$u = K \frac{\Delta\phi}{2}$$

can be adopted with a maximum error of $48'$.

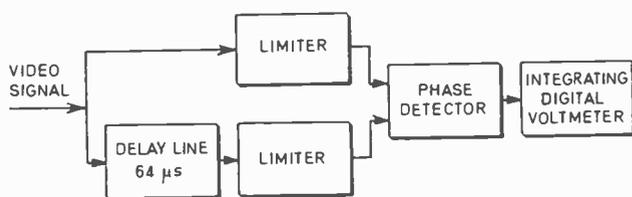


Fig. 10. Block diagram of the differential phase meter.

In the following analysis the limiters and the detector will be dealt with as an unique ideal device, giving an output proportional to the phase shift between the two input signals and independent of their amplitudes. The statistical properties of the phases of the input signals are easily determined under the simplifying assumptions of Section 3.

Only the quadrature component $y(t)$ of the noise vector introduces a phase modulation in the useful signal v_s . Hence the random component of phase is:

$$\phi(t) = \arctan \frac{y(t)}{V_s} \approx \frac{y(t)}{V_s}$$

The properties of $\phi(t)$ are directly derived from those of $y(t)$; hence $\phi(t)$ is a random function with Gaussian distribution and variance:

$$\sigma_\phi^2 = \sigma_B^2 / V_s^2$$

where σ_B^2 is the variance of $y(t)$.

The spectral density of $\phi(t)$, the physical meaning of which is not bound to any electrical power, is proportional to $y(t)$, with a proportionality factor $1/V_s^2$.

Consider now two signals

$$v_{s1} = V_{s1} \sin(\omega t + \phi_{10}), \quad v_{s2} = V_{s2} \sin(\omega t + \phi_{20}),$$

taken from two adjacent lines of the video signal. Because of their distance apart, the noise affecting v_{s1} and that affecting v_{s2} can be assumed to be uncorrelated. The phase shift $\Delta\phi$ between the two signals will be a

random function with Gaussian distribution, mean value $\phi_{10} - \phi_{20}$ and variance:

$$\sigma_\phi^2 = \sigma_{\phi_1}^2 + \sigma_{\phi_2}^2 \quad \dots\dots(8)$$

This derives from the known properties of the difference of two independent random functions with Gaussian distributions.

Equation (8) shows a connexion between the variance at the phase meter output and the signal/noise ratio of the two signals v_{s1} and v_{s2} .

The two signals are affected by the same noise, since they are taken from the same video signal. If they have equal amplitudes, their signal/noise ratio is equal, too; hence to obtain σ_ϕ^2 it is sufficient to worsen this ratio by 3 dB. Besides, since the two functions ϕ_1 and ϕ_2 are statistically independent, the spectral density of their difference is the sum of the spectral densities of ϕ_1 and ϕ_2 .

Let us now examine the measurement device more in detail. The two signals to be compared, extracted from the video signal in two adjacent lines, are fed to the 'limiter-detector' set. Obviously, the delay line should have a $64 \mu s$ delay in every condition. To achieve this, a close loop control of the line delay may be employed, but this does not affect our considerations about noise.

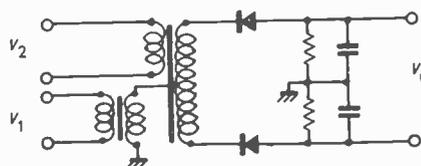


Fig. 11. Phase detector circuit.

In the case of national signals, the bursts on lines 19/20 and 20/21 of the even fields, and on lines 332/333 and 333/334 of the odd fields are compared.

With considerations analogous to those of Section 4 it can be proved that the integration time of the voltmeter after the detector must be extended as much as possible.

For international signals, the measurement is carried out by comparing the staircase on line 330 with the reference burst of line 331. In this case, the signals should be filtered and from the results of Section 5 it is possible to get immediately the value of σ_ϕ^2 .

The sinewave superimposed on the staircase and the reference signal have a peak-to-peak amplitude of respectively 280 mV and 420 mV peak-to-peak, whereas the burst of the national test signals are smaller (200 mV peak-to-peak amplitude). This introduces an improvement of 4.6 dB with respect to national signals, which is however overcome by the 12.2 dB impairment caused by the short duration of steps. Hence, for the same input video signal/noise ratio, the signal/noise ratio at the integrator output is worsened by 7.6 dB in the case of international signals.

The considerations of this Section are just a first approximation to the problem. The next step should be to analyse better the effect of limiters on noise, without assuming them as 'ideal'.

8. Statistical Properties of the Measurement Results of a Digital Instrument

From the considerations of the previous Sections it is possible to define completely the statistical properties of the measurements, if the instrument used is analogue. Since the instruments used in the automatic measurements are digital, it is necessary to make further studies in order to know the effect of the analogue-to-digital conversion. The following observations are valid for any kind of measurement.

As an example, we refer, however, to a percent-error-meter, which gives the results with the two integer digits alone. This instrument will indicate 0 if ε is between -0.5% and $+0.5\%$; +1 if ε is between $+0.5\%$ and $+1.5\%$, and so on. In general, calling the quantization interval $2\Delta\varepsilon$ and the centre value of the i th interval ε_i , the measurement will give the result ε_i , whenever the actual value is between $\varepsilon_i - \Delta\varepsilon$ and $\varepsilon_i + \Delta\varepsilon$.

The probability that the measurement result be ε_i is therefore:

$$Q(\varepsilon_i) = \int_{\varepsilon_i - \Delta\varepsilon}^{\varepsilon_i + \Delta\varepsilon} q(\varepsilon) d\varepsilon$$

$$= \frac{1}{2} \operatorname{erf} \frac{\varepsilon_i + \Delta\varepsilon - \varepsilon_0}{\sqrt{2}\sigma_\varepsilon} - \frac{1}{2} \operatorname{erf} \frac{\varepsilon_i - \Delta\varepsilon - \varepsilon_0}{\sqrt{2}\sigma_\varepsilon}$$

By knowing $Q(\varepsilon_i)$ it is possible to derive the mean value $\bar{\varepsilon}_i$, the mean square value ε_i^2 , the standard deviation σ_{ε_i} and the variance $\sigma_{\varepsilon_i}^2$. We get:

$$\bar{\varepsilon}_i = \sum_{-\infty}^{+\infty} \varepsilon_i Q(\varepsilon_i); \quad \varepsilon_i^2 = \sum_{-\infty}^{+\infty} \varepsilon_i^2 Q(\varepsilon_i);$$

$$\sigma_{\varepsilon_i} = \sqrt{\varepsilon_i^2 - \bar{\varepsilon}_i^2}$$

In the case of digital measurements, the true value ε_0 , the error more frequently indicated by the instrument ε_{i_0} , and the mean value $\bar{\varepsilon}_i$ are normally different.

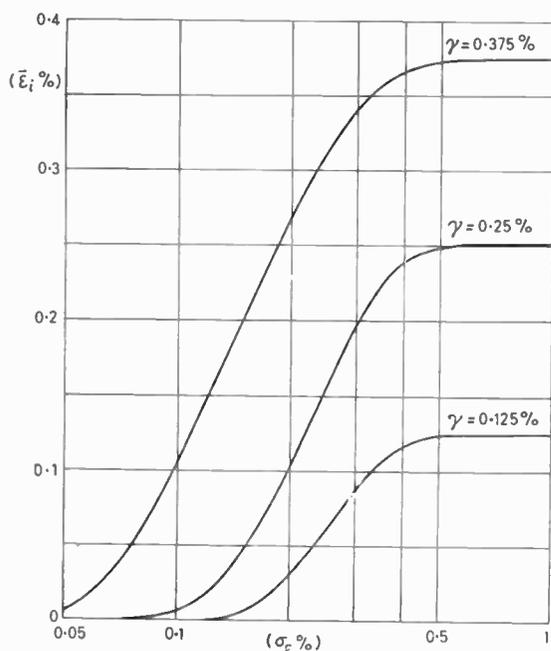


Fig. 12. Mean value of the digital measurement results as a function of σ_ε .

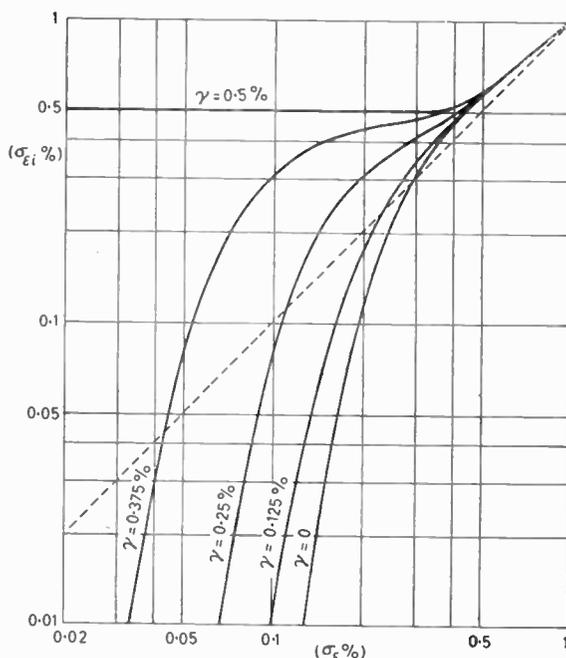


Fig. 13. Standard deviation of the results of digital measurements, as a function of σ_ε .

By means of an electronic computer the values assumed by $\bar{\varepsilon}_i$ and σ_{ε_i} in function of σ_ε have been computed. The shift γ between the true value ε_0 and the more probable value ε_{i_0} has been assumed as parameters. The results shown in the diagrams of Figs. 12 and 13 are particularly interesting for an assessment on the reliability of the measurements.

As to the mean value $\bar{\varepsilon}_i$ it may be noted that it coincides with ε_{i_0} when noise is absent, whereas it tends to ε_0 when the signal/noise ratio increases.

For $\sigma_\varepsilon > 0.5\%$, $\bar{\varepsilon}_i$ coincides virtually with ε_0 .

Because of the conversion into digital form it is not possible to know directly ε_0 . When noise is great, however, this value can be attained by averaging the results of a great number of measurements. That is to say, the presence of noise makes it possible to obtain information which otherwise would remain unknown.

From the diagram of Fig. 13 it clearly appears that for high signal/noise ratios, the dispersion of measurement results is remarkably reduced by the analogue-to-digital conversion. An exception is given by $\gamma = 0.5\%$, when there is a dispersion also if noise is absent, since the instrument indicates $\varepsilon_0 + \gamma$ for a 50% if the cases and $\varepsilon_0 - \gamma$ in the remaining 50%.

This is true, of course, only if the measurement result is given by two digits only.

As noise increases, the curves reach a field, in which the analogue-to-digital conversion leads to $\sigma_{\varepsilon_i} > \sigma_\varepsilon$ and finally all the curves tend to coincide with the straight line $\sigma_{\varepsilon_i} = \sigma_\varepsilon$. This occurs for $\sigma_\varepsilon > 0.5\%$, namely, where the dispersion of the measurement results is very great in respect to the quantization intervals. In this case, the analogue-to-digital conversion does not noticeably alter the statistical properties of the results.

9. Conclusions

(1) The presence of an analogue-to-digital converter in the measuring instrument does not prevent one from knowing the statistical properties of the results; in fact, the standard deviation of any measurement can be computed if the video signal/noise ratio is known. The latter is defined as the ratio between peak-to-peak amplitude of the picture signal (0.7 V) and r.m.s. noise value.⁸

Let us assume for instance that the input video signal/noise ratio is 50 dB and that a percent error measurement is being carried out by using national insertion signals.

Let us also assume that the distortion is zero. The signal/noise ratio for the burst with 0.2 V peak-to-peak amplitude is reduced to 30 dB. An improvement of 26.82 dB will occur at the integrator output, as shown under Section 4.

By applying relation (7) of Section 6, we find that the standard deviation of the measurement is equal to 0.2%. Finally, to take into account the analogue-to-digital conversion, the graph of Fig. 13 must be consulted and corresponding to values $\sigma_e = 0.2\%$ and $\gamma = 0$, we get $\sigma_{e_i} = 0.11\%$.

(2) The theories presented lead also to the specification of measuring instruments both for national and international insertion signals.

(3) In the case of analogue measurements, the video signal/noise ratios being equal, the use of international signals causes an increase in the standard deviation by 1.41 times for the amplitude-frequency response; by 2.9 times for differential gain measurement, and by 2.4 times for differential phase measurement.

Finally the analogue-to-digital converter modifies these values to an extent that can be determined by means of the diagrams of Figs. 12 and 13.

Because of the results given above and of the greater complexity and higher cost of the devices using international signals, the RAI has deemed it advisable to build up a system operating with national signals, as a first step before starting the construction of a more sophisticated system operating with international signals.

10. References

1. C.M.T.T. Document 1025/E, New Delhi, 1970.
2. Castelli, E. and Giugliarelli, G., 'Il controllo automatico delle reti di distribuzione televisive', *Alta Frequenza*, p. 633, July 1968.
3. Zetti, G., 'Convertitore analogico-numerico per il calcolo diretto dell'errore percentuale', *Elettronica e Telecomunicazioni*, No. 5, p. 154, 1969.
4. Panter, P. F., 'Modulation, Noise and Spectral Analysis', (McGraw-Hill, New York, 1965).
5. Bennet, W. R., 'Methods of solving noise problems', *Proc. Inst. Radio Engrs*, 44, pp. 609-38, May 1956.
6. Rice, S. O., 'Mathematical analysis of random noise', *Bell Syst. Tech. J.*, 23, pp. 282-332, July 1944; 24, pp. 46-157, January 1945.
7. Abramowitz, M. and Stegun, I. A., 'Handbook of Mathematical Functions', page 302, formula 7.4.2. (Dover, New York, 1965.)
8. Zetti, G. and D'Amato, P., 'Misuratore automatico del rapporto segnale-disturbo nell'intervallo di cancellazione di quadro di un segnale televisivo', *Elettronica e Telecomunicazioni*, No. 1, p. 2, 1969.

11. Bibliography

1. Papoulis, A., 'Probability, Random Variables and Stochastic Processes'. (McGraw-Hill, New York, 1965).
2. Davenport, W. B. and Root, W. L., 'Introduction to Random Signals and Noise'. (McGraw-Hill, New York, 1958).
3. Schwartz, M., 'Information, Transmission, Modulation and Noise'. (McGraw-Hill, New York, 1959).

Manuscript received by the Institution on 17th April 1970. (Paper No. 1376/Com. 40.)

© The Institution of Electronic and Radio Engineers, 1971

Rate Control for Vacuum Co-deposition of Thin Film Cermets

By

J. E. MORRIS,

B.Sc., M.Sc.,†

and

A. D. BOOTH,

D.Sc., F.Inst.P., C. Eng., F.I.E.R.E.†

In vacuum co-deposition of thin cermet resistor films it has been found that the electrical properties are sensitive to deposition rates, thus needing a two-channel rate controller. The requirements of such an instrument are outlined and alternative systems are briefly compared, resulting in the choice of a digital data processor and step feedback control. A detailed description of the logic and circuit design of the instrument built is presented along with a brief discussion of the operational limitations and difficulties. Possible extensions of the instrument to other needs are considered.

1. Introduction

In the course of the development of thin film micro-electronic circuits in this laboratory an investigation into thin film resistors is being pursued. Conduction mechanisms and deposition processes in cermet films are being considered using Au-SiO cermets as a simple system. Preliminary results indicated that the structures of the films, and hence their electrical properties, are extremely sensitive to deposition rates.¹ In addition, film structure is a function of substrate temperature which is, in turn, also dependent to some extent on deposition rate, due to the radiation heating effects of the SiO vapour source.² It has therefore become necessary to develop an instrument to control deposition rates within acceptable limits.

This paper will briefly review some of the alternative system designs considered and will describe the instrument developed to suit the needs of the present project. A few suggestions as to how the instrument's flexibility might be extended to the needs of others will be described in closing.

2. Basic Requirements

In this laboratory, the present rate monitoring equipment has been supplied by Sloan Instruments Corporation and includes quartz crystal oscillators, which are mounted inside the vacuum system as sensor heads, and two Sloan thickness monitors. The frequency of the crystal oscillator increases as material from the source is deposited on the exposed crystal in such a way that the change is proportional to the mass deposited. The signal from the oscillator is mixed with a variable reference frequency inside the monitor unit to yield a beat frequency which is accessible as an output. The beat frequency is converted to an analogue voltage displayed on a front meter and available as an output to a chart recorder.

Evaporation is accomplished by Joule heating of the sources using two Allen-Jones high current power supplies. These supplies may be regarded as essentially 2 kVA variable-ratio transformers ('variacs').

Rates used in the Au-SiO cermet depositions fall in the range of 1 to 100 Hz per second for the deposition

apparatus geometry when expressed in terms of crystal oscillator frequency changes. An arbitrary design requirement of detection of a 1% variation from the desired rate has been set. The instrument is required to display the deposition rate and the total deposited thickness.

It has been stated above that the crystal frequency change is directly proportional to the mass deposited on it, but there are two sources of rapid and erroneous frequency changes or crystal mode changes. Poor seating of the crystal itself in the sensor head may give rise to frequency jumps of the order of 100 Hz. These changes are not regarded as a normal operating difficulty and can be eliminated, for example by re-machining the crystal seating. Thermal effects in the crystal due to its exposure both to the hot vapour stream and line-of-sight radiation from the source are, however, a normal operational difficulty. In particular, a crystal exposed to the SiO vapour is sensitive to this type of effect since the heating produced by the vapour source is quite significant.² The problem may be reduced by either water cooling of the crystal head or the use of special circular cut crystals which have recently become available commercially. The latter solution has not been tried in this laboratory and the former, while very effective in reducing the problem, does not entirely eliminate it. In addition, provision of water cooling and associated feed-throughs represents an extra expense. So another requirement of the instrument is that it should be able to detect crystal mode changes and take appropriate action to compensate for the errors they produce, provided of course that such detection can be performed economically.

3. System Design

3.1 Power Control

The obvious method of controlling the variac supply is to provide a means of producing small rotations of the variac wiper shaft. Stepper motors of sufficiently high torque were considered too expensive and in any case the discrete winding nature of the variac is unsuitable for fine control.

A triac phase control system operating on the input line voltage can provide the appropriate fine control, and can, in fact, be arranged to provide whatever control delicacy is required. Variation of the source power is achieved by variation of the resistance in a phase lag

† Department of Electrical Engineering, University of Saskatchewan, Saskatoon, Canada.

trigger network, for example by means of a stepper motor driving a potentiometer. A version of this arrangement which is much cheaper than it may appear is a bi-directional stepper driving a wiper across contacts to which is connected a continuous chain of discrete resistors appropriately chosen to provide very fine control near mid-scale with control coarseness increasing toward the outer limit. Such an arrangement requires a digital trigger pulse to initiate a power correction rather than a continuous error signal.

3.2 Analogue Rate Control

The Sloan Instrument Corporation produces a commercial rate controller which operates on the analogue output from the Sloan monitor. The signal is differentiated and the resulting voltage level, (representing deposition rate), is compared with a set level, (representing the required rate), by means of a differential amplifier the output of which is proportional to the error in the deposition rate and is used to generate corrections to the source power.³⁻⁸

As a worst case consider a deposition rate of 1 Hz/second. On the Sloan monitor the most sensitive analogue output range is 100 kHz full scale corresponding to an analogue level of 3 V. The rate of change in the analogue level is then 30 μ V/s which gives a differential output level of 30 μ V for a network time-constant of 1 second. Design of a system to operate at a detection level of 1% of this value places very stringent requirements on stability. In addition some precautions are necessary to protect the amplifiers against saturation when crystal mode changes occur.

An alternative analogue system was briefly considered. This consisted of three sample-and-hold circuits followed by a differential amplifier to replace the differentiation network of the Sloan systems. The analogue Sloan output was then to be sampled by each sample-and-hold circuit in turn and two successive samples applied to the appropriate inputs of the differential amplifier the output of which was to be compared with a variable setting as before. It was concluded that this system could not be realized with comparable accuracy at comparable cost.

Crystal mode change errors may be eliminated in both the above systems by comparison of the rate or rate deviation levels with preset limit levels to provide inhibit signals to the power correction circuitry and to initiate overvoltage protection, if required.

3.3 Digital Rate Control

Since this project was started several digital deposition monitors have appeared on the market in competition with the Sloan monitor. Digital processing of the crystal frequency change is not only an obvious direct solution but also presents the usual advantages of digital data handling over analogue (e.g. no drift problems and reduced effect of component deterioration on efficiency). The final decision to build a digital system instead of an analogue system was based on comparison of estimated total costs, the determining factor being the anticipated (but unknown) stability problems involved in developing an analogue system from scratch and the time factor these problems may have created.

The basic system design is shown in Fig. 1. Input pulses are counted into the counter (C) for a defined gating period at G1. Either the crystal or the beat frequency output from the Sloan monitor may be used as the input source via appropriate pulse shaping networks. Since Sloan monitors are available in this laboratory already they are used as convenient sources. 50% of the cycle time is used as a convenient gating period, simplifying later interpretation of the displayed deposition data. For detection of a 1% deviation from a required rate of 1 Hz/s frequency change a 100 s gating period is required. This period is variable down to 1 second for the required 1% detection with 100 Hz/s. The response time of the deposition rate to a small step change in source power has been estimated from tests to be about 1 second so cycle times greater than this figure are not expected to produce instabilities in the feedback loop. System sensitivity could be increased for a given gating period by using crystals with a lower frequency constant,³ allowing higher sampling rates with low deposition rates.

After the count is complete the previous count stored in the buffer B1 is subtracted from C by means of S1. A negative result indicates either a mode change or possibly, in the case of an input derived from the Sloan monitor, a decreasing beat frequency. It is not good practice to use a decreasing beat frequency since the crystal and internal monitor oscillators tend to lock together as their frequencies approach, so a negative result is considered to be a mode change error. The result is compared with previously set limits (E) and if these limits are exceeded it is concluded that a mode change has taken place. With long gating periods it is conceivable that successive mode changes may occur within one cycle time in such a way that they partially cancel and are interpreted as a genuine result. However, long cycle times are used for low rates where heating effects are minimal and experience indicates that mode changes will probably not occur sufficiently often to give this problem with the materials in use. Should the problem arise, however, further reduction in mode change occurrences by the use of a water-cooled sensor head should eliminate it.

The output of the subtracting unit S1, which now represents the deposition rate, is compared (S2) with the required rate set up by the operator to produce a rate deviation count with associated sign. This result is entered into the buffered display (B2) by the rate display trigger pulse t_1 . If the earlier decision was that a mode change has occurred, the trigger pulse is inhibited (G2) and the previous result remains in the buffer (B2). In effect, then, it is assumed in the event of a crystal mode change that the rate has not changed since the previous cycle. In the event of mode change detection, an error light is turned on. If this light is on continuously, there is either a drastic error in the genuine rate, very frequent mode changes, unrealistic mode change detection settings or a negative beat frequency, all of which require the operator's attention. The error light therefore provides a check on system operation. (Note the possibility of a small error being introduced if a mode change occurs since the rate is not necessarily the same

as that in the previous cycle, particularly if the previous cycle initiated a power correction which will now be repeated.) The significance of the sign logic (SL1) will become apparent in the next section. At this stage the contents of the counter C1 can be entered by a pulse t_2 into the buffer B1 in readiness for the next cycle.

The rate deviation in B2 is now added to the total deviation in B3, the new total being entered into B3 where it is available as a visual output. If either the total

deviation or the instantaneous rate deviation is beyond preset acceptable limits then a correction signal is applied at t_4 to the bi-directional stepper of the power control unit previously described. A correction pulse arising from an excessive total deviation is inhibited if the sign of the instantaneous deviation is not the same as the sign of the total. At the end of the timing sequence the first counter is cleared (t_4) and the control timing is reset (t_5) in readiness for the next cycle.

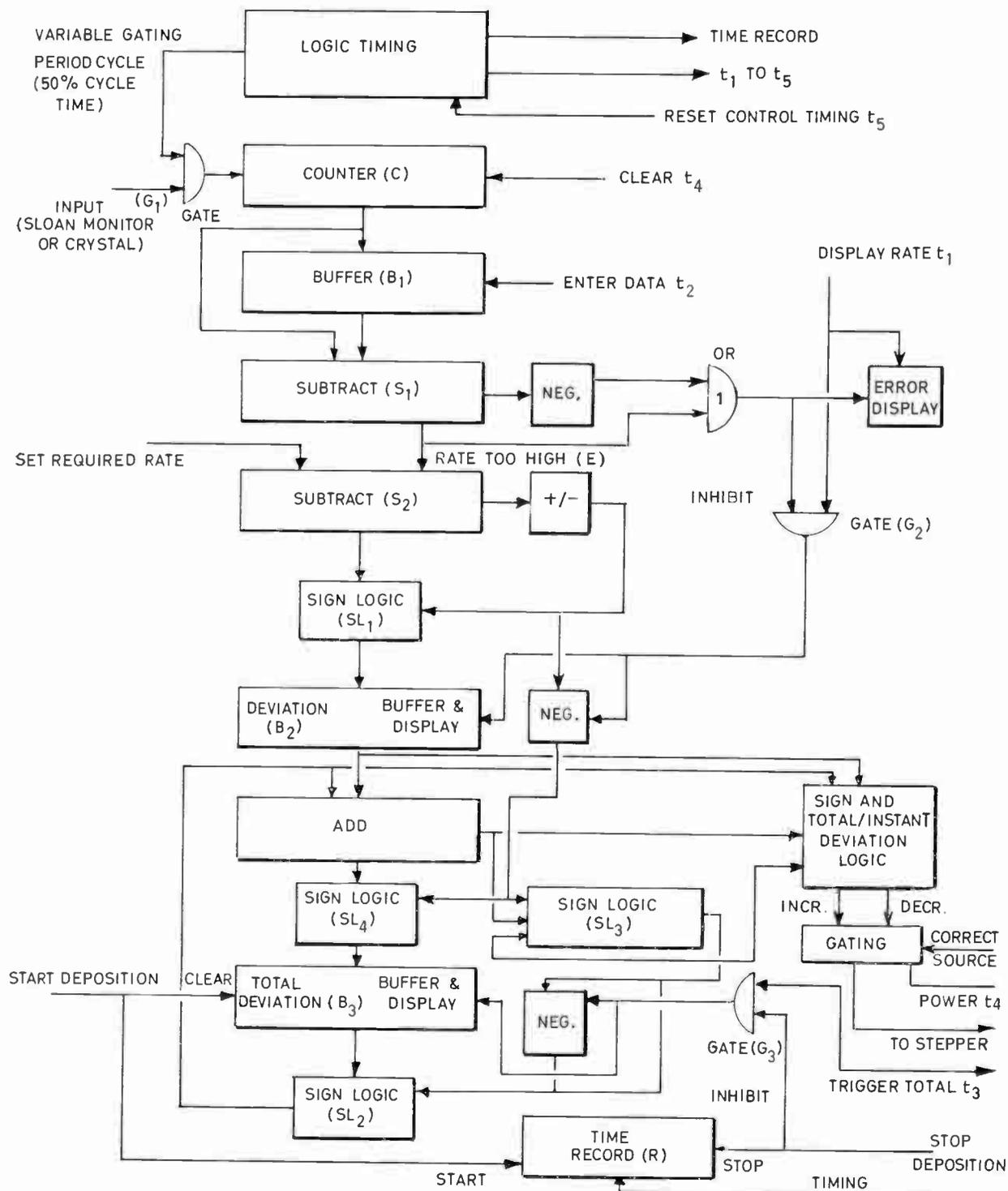


Fig. 1. Digital rate control system.

At the start of deposition itself, i.e. when the substrate shutter is opened, a pulse clears the total deviation buffer (B3) and starts a counter (R) driven by the timing circuit, to provide a record of deposition time. When the shutter is closed the total deviation count gate (G3) and the time record are stopped. From the set rate, deposition time and total deviation the total deposited thickness may be calculated. Note that the basic design philosophy is oriented towards the idea that deposition rate is more important than total deposited thickness in determining the final properties of the deposited film. This is true in the present context of co-deposition of Au-SiO cermet where errors introduce variations not only in final resistance values at a given thickness but also in the temperature coefficient of resistance. The latter parameter is particularly important for practical resistors, the resistance value itself being adjustable after deposition by a variety of trimming techniques.

4. Instrument Design

4.1 Power Control

Details of the power control circuits are shown in Fig. 2. Figure 2(a) shows the triac phase control with a conventional r.f. suppression network and a bypass switch which allows direct use of the high current power supplies without rate control. The trigger diode (MPT28) triggers the triac when the capacitor voltage reaches a nominal 28 V and r (330 Ω) limits the surge current to less than 100 mA.

The maximum current output from the Allen-Jones supply is normally 400 A. If the maximum current may be cut to 350 A for the immediate application then the conduction angle at the instant of triac turn-on (θ) is given by

$$\text{load power/max. power} = 1 - \theta/\pi + \sin 2\theta/2\pi = (350/400)^2$$

i.e. $\theta = 65^\circ 18'$, which is the phase lag which would be required of the phase network if there were no trigger circuit. The triggering requirement of 28 V introduces an inherent phase delay of $5^\circ 10'$, however, so the phase lag, required at C is $60^\circ 8'$, i.e.

$$\tan \zeta = \omega RC = 1.742$$

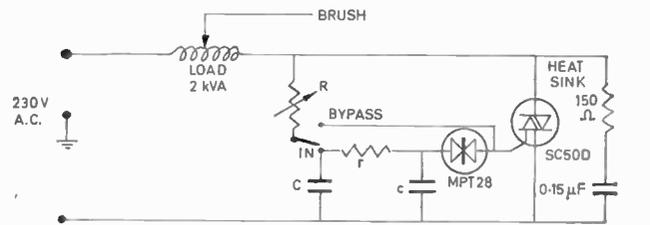
Therefore $RC = 4.62 \times 10^{-3}$. Now the maximum trigger current required for the triac is 50 mA which limits R to about 4 kΩ and hence C is of the order of 1 μF. C is selected to be 0.82 μF and c is 33 nF. The centre value of R is then approximately 5.5 kΩ. Minimum and maximum values of 4 kΩ and 9 kΩ respectively give an approximate 5 A variation about a centre at 100 A (Table 1) which experience indicates is an appropriate control limit. R1 and R2 (Fig. 2(b)) are determined by these outer limits for R3 open and short circuited. The centre value of R3 is then determined by the centre

Table 1. Control limits

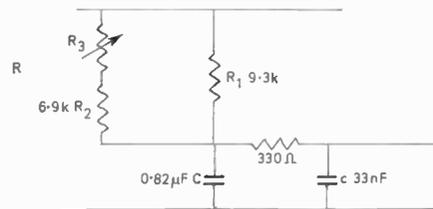
R	ωRC	θ	Power/Max	Current
4 kΩ	1.29	$57^\circ 20'$	0.83	105 A
5.5 kΩ	1.77	$65^\circ 50'$	0.75	100 A
9 kΩ	2.90	$76^\circ 05'$	0.66	94 A

value of R. R3 is composed of discrete resistors. The spring-loaded wiper makes before break, minimizing transients (Fig. 2(c)). Power resistors are used where appropriate.

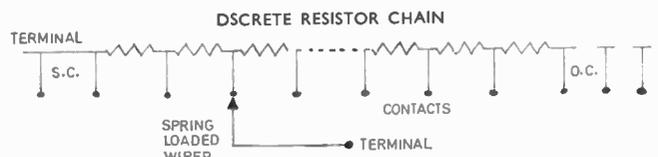
Figure 2(d) shows the drive circuits for the bi-directional steppers. Indicated component values are for RTL drive. The ganged hold switch allows inhibition of corrective action during initial set up procedures and the manual step facility provides for centring R3 in the hold mode.



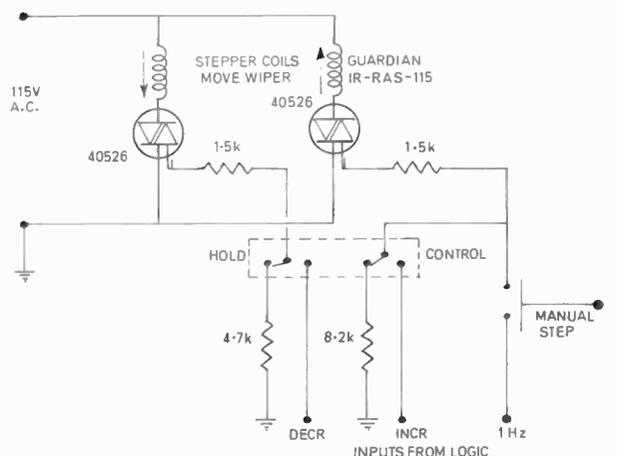
(a) Triac control.



(b) Phase network.



(c) Structure of R3.



(d) Stepper control.

Fig. 2. Power control

4.2 Logic

The choice of plastic pack RTL logic (supplied by Motorola) was made on the basis of ready availability of complex operations contained in one multiple package which allows parallel processing relatively cheaply. Noise immunity of RTL is not good but the expected problems did not materialize except on trigger lines which picked up transients from other parts of the logic system, primarily buffer outputs. Capacitors are shown on the diagrams where these were found to be necessary. Diagrams refer to the system as built and operating and include later modifications to the original design except those described in Section 5 below.

An Electrostatics 50-4 power supply was used and the logic was mounted on an Augat GP-13-120 wire wrap board mounted on hinges with the wire wrap terminals extending into a shielded enclosure. Neon binary output displays were used with an unregulated supply, the transformer also being used for isolation of the stepper drive supply. A key to the logic symbols and packages used is contained in Table 2 and Fig. 3.

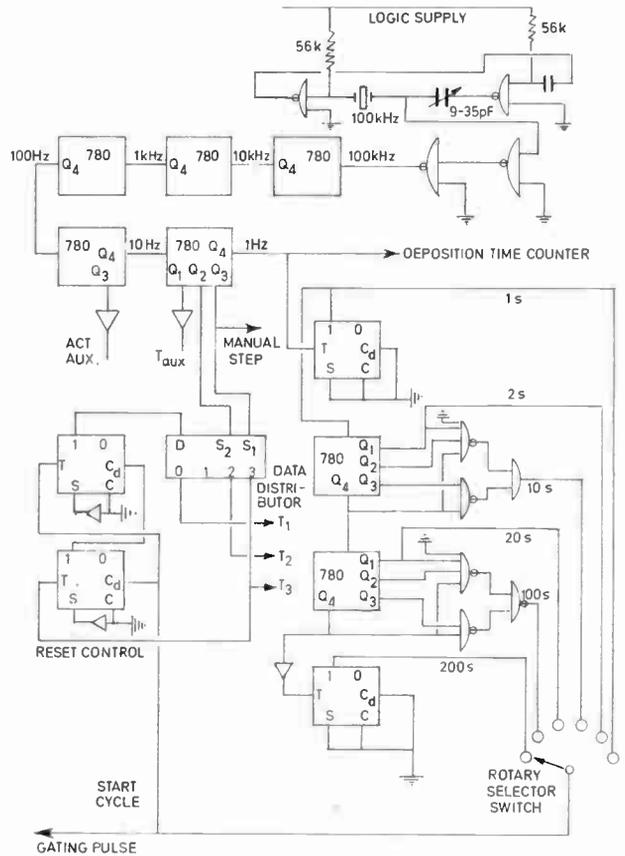


Fig. 4. Timing logic.

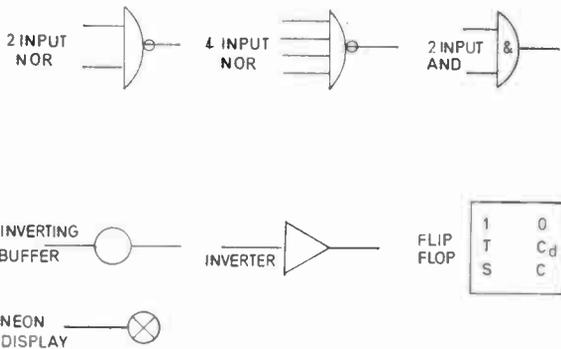


Fig. 3. Logic symbols (positive logic).

Table 2. Logic packages used

MC724P	Quad 2-input NOR gate
MC725P	Dual 4-input NOR gate
MC767P	Quad latch
MC771P	Quad exclusive-OR
MC777P	Binary up counter
MC780P	Decade up counter
MC789P	Hex inverter
MC790P	Dual flip-flop
MC799P	Dual buffer (inverting)
MC9704P	Quad parallel adder
MC9707P	Quad exclusive-OR
MC9713P	Quad exclusive-OR

4.2.1 Timing (Fig. 4)

A 100 kHz crystal oscillator is used for timing and provides stability far in excess of that required at reasonable cost. The 100 kHz signal is divided down to provide gating periods (half cycle times) of 1, 2, 10, 20, 100 and 200 seconds. In addition a signal is provided to drive the deposition time counter. Control timing is best

explained with reference to Fig. 4. The cycle is started by the turn off of the gating pulse which enables the data distributor at D. As S1 and S2 pass through the combinations (0, 0), (0, 1), (1, 0) and (1, 1) the 0, 1, 2 and 3 outputs go high in turn giving timing pulses T₁, T₂ and T₃. Two auxiliary timing pulses (Act Aux and T_{aux}) are required for gating with the main pulses to provide short pulses for certain operations. The fall of T₃ disables D for the rest of the control half cycle and the control timing is reset by the next gating pulse.

Master timing (oscillator and countdown) is common to both channels, but individual gating period selection and control timing are required for deposition of two materials at widely differing rates.

The complete control timing sequence is shown in Fig. 5. Edge and level triggering are indicated.

4.2.2 Deposition rate (Fig. 6)

Input pulses are counted into a string of binary scalars corresponding to counter C in the system diagram (Fig. 1). The previous count stored in the latches (buffer B1) is subtracted by means of adding the counter output to the complements of the stored count. A low carry out from the parallel adder chain indicates that the stored count is greater than the new and is used as one source of error detection.

The input may be derived from the Sloan beat frequency output, as in this case, or directly from the sensor head crystal oscillator which can be driven from the

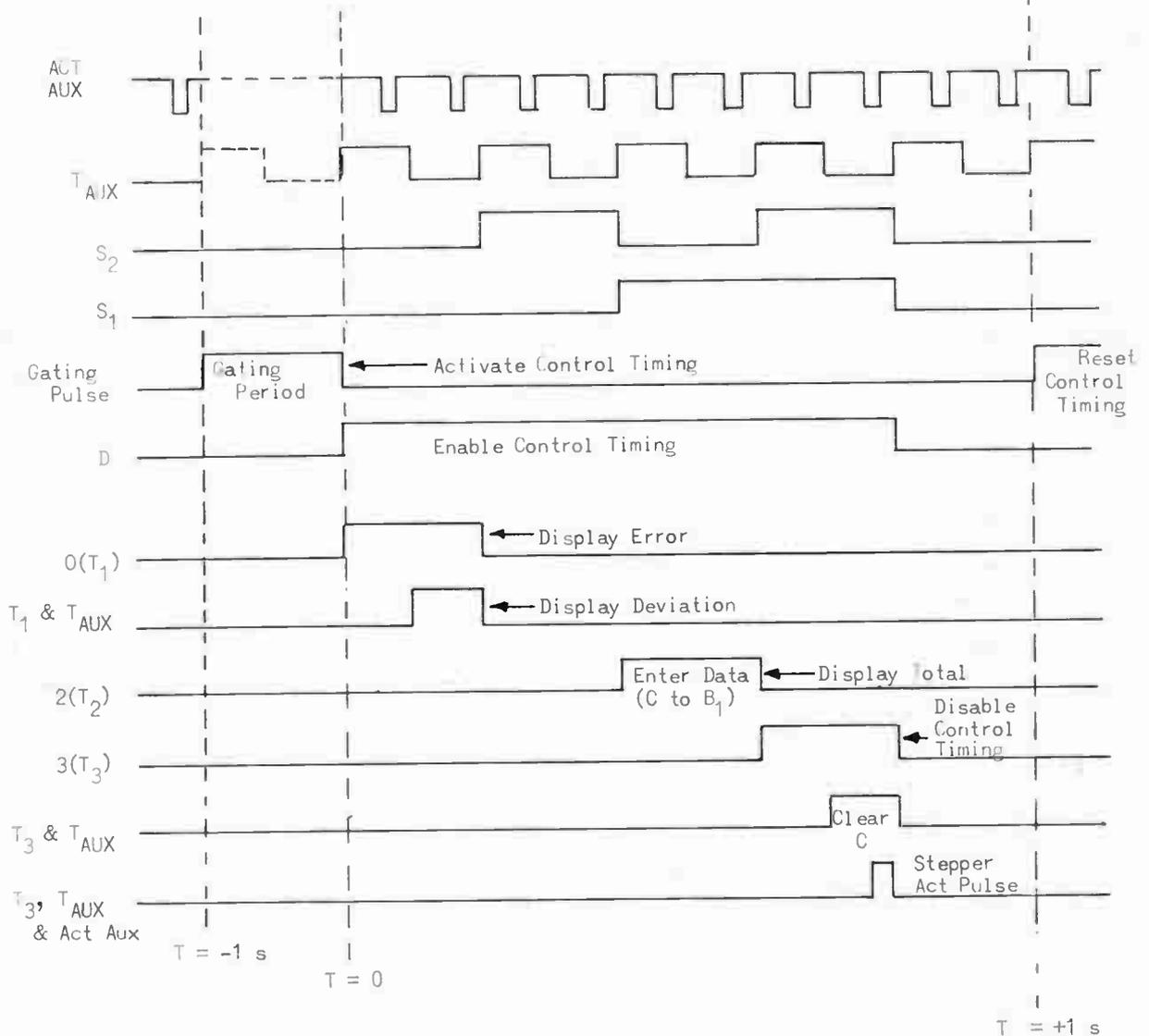


Fig. 5. Timing sequence.

instrument. In the latter case RTL logic may not be suitable for the first binary counting stage since typical frequencies (6 MHz) are beyond the specified limits of the MC777P. An alternative shaping network will also be required. The input circuit shown requires an additional Schmidt trigger shaping unit if the full 0–100 kHz frequency range of the Sloan output is to be used due to distortion of the pulses at frequencies below about 5 kHz.

4.2.3 Rate/deviation (Fig. 7)

If the rate count is positive the real value is presented to the inputs of the adders (S_2). The desired rate is therefore set up in complementary form on switches as shown. The carry level once again indicates the sign of the deviation and use of exclusive-OR modules (SLI) allows the presentation of the real result at the latch terminals. As previously described, the display trigger is inhibited (G_2) by mode change detection E (i.e. by a negative rate or an unrealistically high rate detected by partially variable switch settings as shown), which also activates the error light.

4.2.4 Total deviation (Fig. 8)

Addition of the deviation to the total deviation is performed with appropriate sign decisions. If the signs of the total and the rate deviations are the same then that sign will be the new total sign and real values are added. But if the signs are different the resultant sign will depend on which is the greater, indicated by the carry level of the adder operating (in this case) as a subtractor. If both are positive and b are both low and the total deviation input to the adders from the exclusive-ORs are the final inputs to the total flip-flop register. When the total is positive (negative) and the instantaneous deviation is negative (positive) point b (a) is high and the complement of the total is entered at the adders. In these cases a high carry level indicates the rate deviation magnitude is greater than that of the total and the new total sign is that of rate deviation. The opposite holds for a low carry level. Real values of the new total are obtained except when opposite signs and a low carry occur in which case the exclusive-OR input to the total

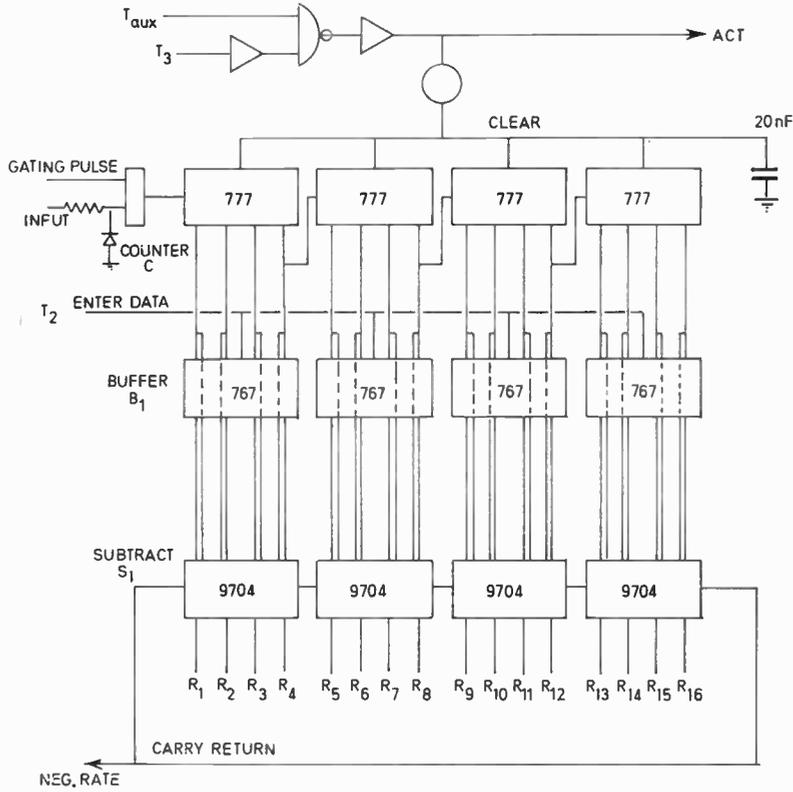


Fig. 6. Deposition rate logic.

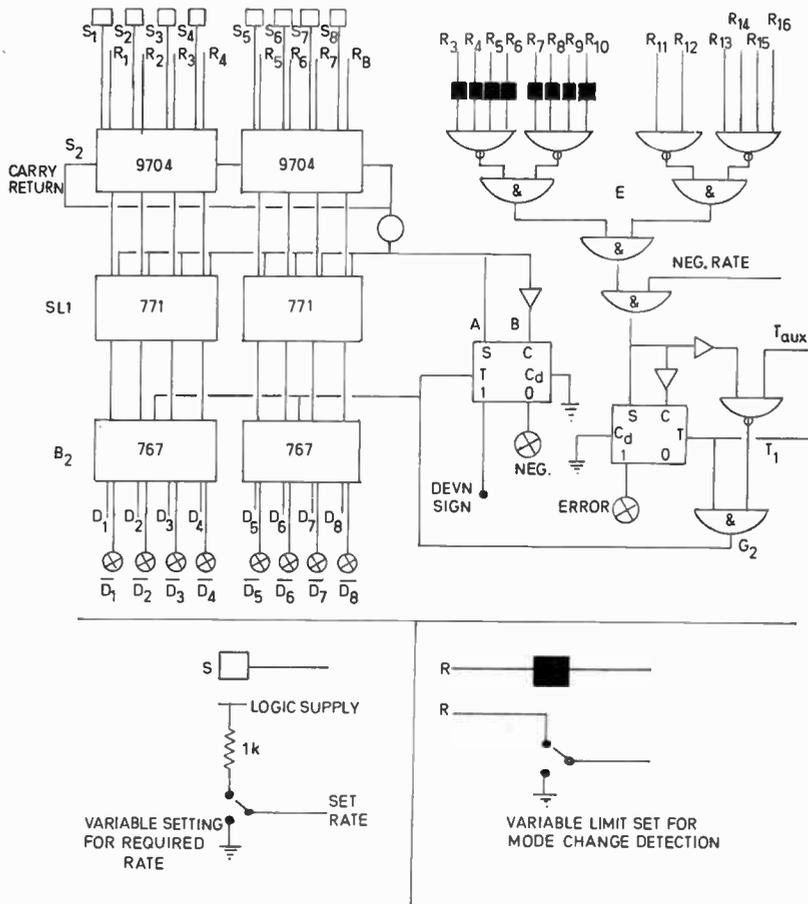


Fig. 7. Rate deviation logic.

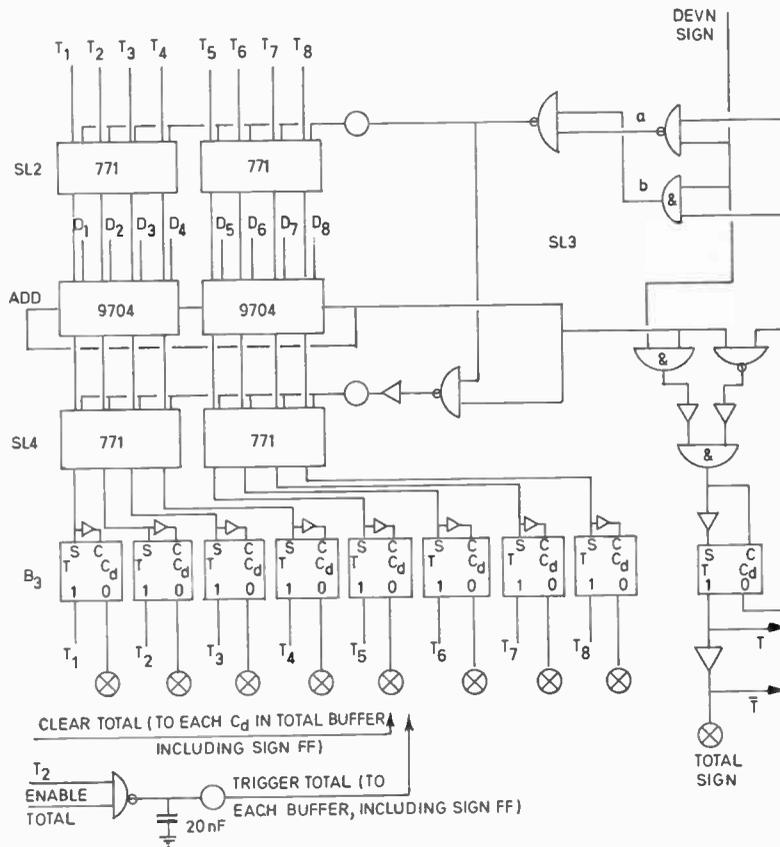


Fig. 8. Total deviation logic.

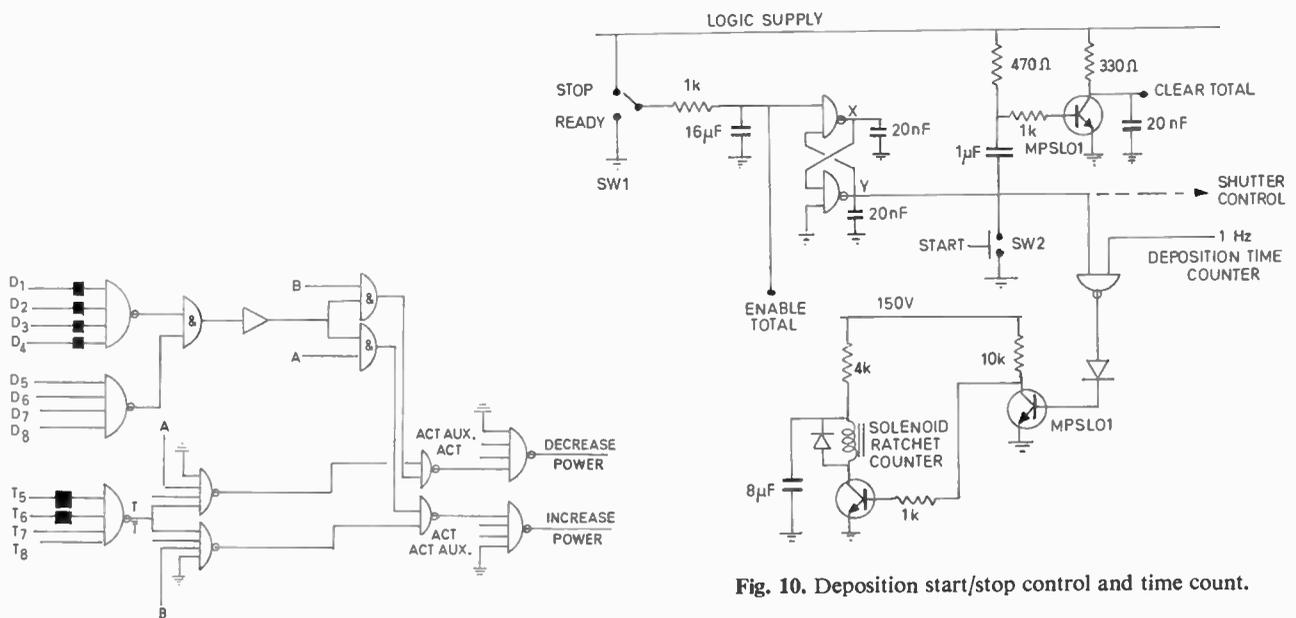


Fig. 9. Power correction logic.

register is required. The delay through the logical loop is much greater than the flip-flop trigger time so no intermediate buffer is required.

4.2.5 Power control logic (Fig. 9)

Correction limits for both rate and total deviations are partially variable. Signs of total and rate deviations are given by T, \bar{T} and A, B respectively.

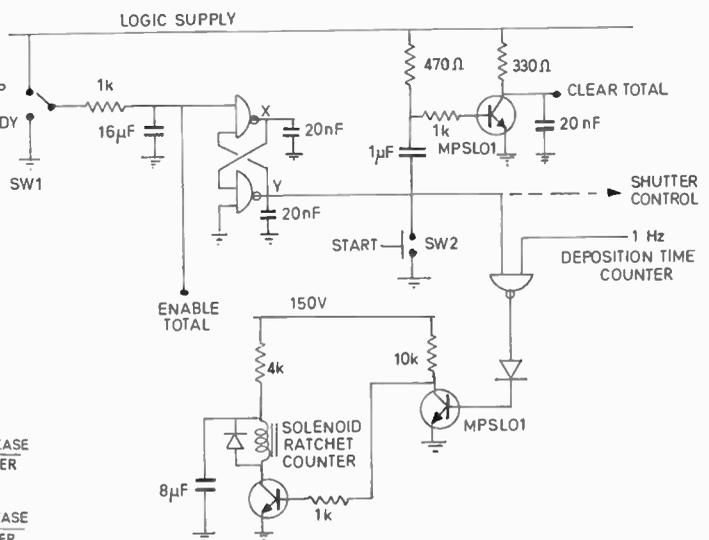


Fig. 10. Deposition start/stop control and time count.

4.2.6 Start/stop deposition control (Fig. 10)

The circuits shown are shared between the two channels. Before deposition is started switch 1 is in the stop position inhibiting the total deviation count and holding X and Y low and high respectively. Prior to deposition, switch 1 is transferred to ready enabling the total count but not affecting X or Y. At the start of deposition switch 2 is momentarily closed clearing the total register and locking X and Y high and low respectively. Y opens the time count gate and deposition is

recorded in seconds by the ratchet counter. Turning switch 1 to stop disables the total counting, sets X and Y low and high respectively and stops the time count.

5. Discussion

A two-channel rate controller has been built as described above and is operating successfully. Total cost of components used was very close to \$900 (approx. £400) on a unit price basis. This figure for a two-channel system compares very favourably with available commercial units. Basic differences from previous approaches to the problem include the error rejection facility and the adoption of an independent channel approach in preference to the master-slave concept.³ Extension to a multi-channel system by merely increasing the number of channels would be simple, but a more economical way would be to time-share the arithmetic operations in a central processing unit, with individual channels consisting only of data storage. Addition of a preset time shut-off may be achieved by using the timing frequency dividers (possibly extended) as a down counter with a start count entered by the deposition start signal. The start/stop deposition control may be conveniently connected to a solenoid-operated shutter in the vacuum system.

No conversion from frequency to equivalent thickness has been attempted within the unit for two reasons:

- (a) Such a conversion restricts the deposition geometry inside the system with particular reference to sensor head placement.
- (b) There is a significant variation in the frequency to thickness calibration factor with substrate temperature.

It has been found with the system in use that the $\pm 5\%$ variation designed for is greater than that required provided the rate is permitted to settle first. This would, however, vary considerably with the requirements. A more positive triac drive current has also been found to be desirable and is altered to two times the previously quoted figure.

A major problem remains with the accuracy of rate control at low rates. At a rate of 1 Hz/s the required gating period is 100 s for 1% accuracy. Mode changes may not be detected and significant rate deviation can occur within the gating period.

There are two possible solutions. Either an alternative crystal with a greater (frequency change)/(unit deposited mass) may be used or the crystal must be placed closer to the evaporant source. Present crystals have been chosen on the basis of minimal thermal mode change characteristics and decreasing the source to crystal distance increases the radiant energy incident on the crystal. Both solutions then, while allowing a decrease on the cycle time, increase the incidence of thermal mode changes. The figure 1 Hz/s corresponds, however, to a rate of 0.5 nm/s for Au at the substrate with present system geometry and a brief review of the commercial units indicates a similar loss of accuracy at low deposition rates.

6. Acknowledgments

Financial support was provided by NRC Grant No. A-1616 and DRB Grant No. 5501-32. The advice of D. A. Freestone was very much appreciated

7. References

1. Morris, J. E. (to be published).
2. Breitweiser, G., Varadarajan, B. N. and Wafer, J., 'Influence of film condensation and source radiation on substrate temperature', *J. Vac. Sci. Tech.*, **7**, pp. 274-7, 1970.
3. Behrndt, K. H. and Love, R. W., 'Automatic control of film-deposition rate with the crystal oscillator for preparation of alloy films', *Vacuum*, **12**, pp. 1-9, 1962.
4. Bath, H. H. A., English, J. and Steckelmacher, W., 'Automatic control and monitoring system for thin film deposition', *Semicond. Prod. and Solid State Tech.*, **10**, No. 1, pp. 27-34, 1967.
5. *Ibid*, No. 2, pp. 25-29, 1967.
6. Gerber, R., 'Use of simultaneous vaporizing of two substances to obtain thin films', *Rev. Sci. Instrum.*, **38**, pp. 77-8, 1967.
7. Chopra, K. L. and Randlett, M. R., 'Influence of deposition parameters on the coalescence stage of metal films', *J. Appl. Phys.*, **39**, pp. 1874-81, 1968.
8. Riegert, R. P., 'Simultaneous Control of Many Vacuum Deposition Sources', Application Note, Sloan Instruments Corp.

Manuscript first received by the Institution on 21st October 1970 and in final form on 11th January 1971. (Paper No. 1377/IC 41.)

© The Institution of Electronic and Radio Engineers, 1971

Stability Analysis of an Improved Mode Control Amplifier

By

S. MARJANOVIĆ, Ph.D.†

The stability and bandwidth characteristics of a mode control amplifier can be substantially improved by addition of a resistor in series with computing capacitor. It is shown that the effects of such a modification of the feedback path configuration on computing accuracy can easily be balanced out.

Notation

A_1	open-loop gain of switched amplifier SA
A_2	open-loop gain of computing amplifier A
K_{10}, K_{20}	gain-bandwidth product in rad/s of amplifiers SA and A, respectively
R_{o1}	open-loop output resistance of switched amplifier SA
Z_C	impedance of feedback capacitor C
Z_o	output impedance of switched amplifier SA
C	feedback capacitor
r	resistor for frequency response compensation
$T_o = CR_{o1}$	
$T_1 = Cr$	compensating time constant
R_1, R_2, C_0	error compensation network
V_{oT}	output voltage when system is in TRACK
V_{oH}	output voltage when system is in HOLD
i_f	current through rC branch
i_c	current through compensating capacitor C_0
s	complex frequency

1. Introduction

The use of mode control amplifiers in fast iterative hybrid computing systems has posed a serious problem on matching the speed requirements with the current handling capabilities of the switches employed. A solution has been found in the use of switched amplifiers (Fig. 1) instead of simple switches. Such a mode control amplifier consists of two amplifiers: a high gain operational amplifier (A) and switched amplifier (SA) which is a voltage follower whose output can be turned off.¹ The modes of operation are established in the following way: When in TRACK/STORE mode the switch S is left open disconnecting the corresponding resistive network from the rest of the circuit, while the switched amplifier SA can be either on or off. When SA is on, the output of the computing amplifier A is proportional to the sum of TRACK inputs. If SA is in the non-conducting state the input of amplifier A is connected only to its proper output via feedback capacitor C. Such a combination of an amplifier and capacitor operates as a storage unit which stores the output voltage corresponding to the instant of SA changeover.

Another combination of modes of operation of the arrangement of Fig. 1 is when switch S is on. The COMPUTE mode is established when the switched amplifier (SA) is non-conducting. The output voltage of the computing amplifier A is then proportional to the integral

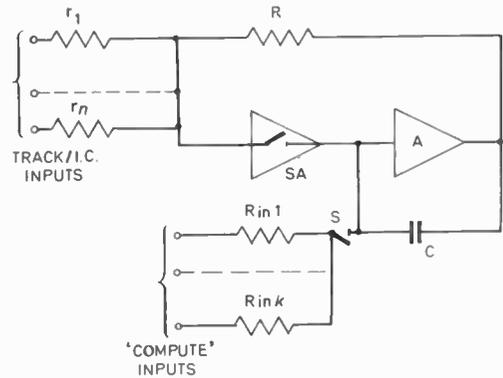


Fig. 1. Two-mode control amplifier operating either as a TRACK/STORE or I.C./COMPUTE unit.

of the COMPUTE inputs. If SA is conducting, the summing junction of amplifier A is connected to both the COMPUTE inputs and the output of SA amplifier. Since the output resistance of the SA amplifier is small (due to local feedback from output to input), the influence of amplifiers SA and A therefore operates as an inverting amplifier with output voltage proportional to the sum of signals on r_1 to r_n inputs. This mode of operation is thus identical to previously described TRACK mode; the only difference is that the voltage across feedback capacitor C is used to establish initial conditions (I.C.) when the system is switched into (COMPUTE) INTEGRATE. The details of operation of the mode control amplifier of Fig. 1 are described by Eckes.¹

The stability criteria necessary for the design of the mode control amplifier have been discussed by the present author elsewhere.² These criteria establish relationships between the parameters of the switched amplifier (SA) and computing amplifier (A) on one hand and overall performance of the mode control amplifier on the other hand.

2. Compensation

The use of mode control amplifiers in fast computing systems imposes stringent conditions on bandwidth characteristics of the amplifier if signal delays are to be kept within the permissible limits. For this reason such frequency compensation techniques which aim at reducing the bandwidth are not acceptable.

In this paper a method, based on results obtained with an analysis of the stability of the system, has been devised which can greatly improve the frequency response.

† Formerly with the Department of Electrical and Electronic Engineering, University of Birmingham; now with the Mihailo Pupin Institute for Automation and Telecommunications, P.O. Box 906, Belgrade, Yugoslavia.

2.1 Stability Analysis of the Non-compensated System

The two-mode control amplifier presents a feedback system having two feedback loops, one across the main amplifier and the other across both the switched and main amplifier. The complexity of these loops introduces the problem of ensuring stability of the system while still preserving its bandwidth. In performing an analysis the following assumptions will be made.

1. The amplitude characteristic rolls off at the rate 6dB/octave starting from d.c. unless otherwise stated.
 2. The gain-bandwidth product is finite.
- These two assumptions imply that the gain at d.c. must be infinite.

3. The values of the computing impedances are small compared with the input impedance of the amplifiers and the effect of the finite value of the input impedance can be neglected.
4. The output impedance of the switched amplifier is approximately resistive. This is because the integrated operational amplifiers have local feedback around the output stages for the purpose of reducing the value of the open-loop output impedance. The output impedance stays nearly constant since the impedance of the output stage (which is usually an emitter follower) and the loop gain both decrease with frequency.
5. The effect of the finite value of the output resistance of the amplifier A can be neglected.

For the analysis of the TRACK mode an equivalent circuit is necessary and is shown in Fig. 2. The switched amplifier of open-loop gain A_1 is replaced by a voltage source $A_1/(1+A_1)V$ whose output impedance is $R_{o1}/(1+A_1)$, R_{o1} being the open-loop output resistance of the switched amplifier.

The output voltage is then

$$V_o = \frac{A_1}{1+A_1} \frac{Z_c}{Z_o} \frac{1}{1 + \frac{1}{A_2} \left(1 + \frac{Z_c}{Z_o}\right)} V \quad \dots\dots(1)$$

where A_2 is open-loop gain of amplifier A, and

$$Z_c = \frac{1}{sC}, \quad Z_o = \frac{R_{o1}}{1+A_1},$$

$$A_1 = \frac{K_{10}}{s}, \quad A_2 = \frac{K_{20}}{s}$$

and K_{10} and K_{20} are the gain-bandwidth products in rad/s of the switched and main amplifier, respectively, giving the open-loop transfer function $G(s)H(s)$

$$\frac{K_{10}K_{20}/2}{s[T_o s^2 + s(1 + T_o K_{20}) + K_{10}]} \quad \dots\dots(2)$$

where $T_o = CR_{o1}$.

The closed-loop transfer function is then

$$\frac{V_o}{V_i} = \frac{K_{10}K_{20}/2}{T_o s^3 + (1 + T_o K_{20})s^2 + K_{10}s + K_{10}K_{20}/2} \quad \dots\dots(3)$$

The root loci of a system whose open-loop transfer function is defined by equation (2) cannot be drawn

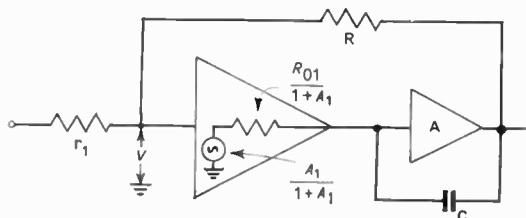


Fig. 2. Equivalent diagram of TRACK/STORE unit.

directly when either K_{10} or K_{20} is assumed to be a parameter. This is because each appears not only as a single multiplying factor, but also in the coefficients of the polynomial in the denominator. Therefore, the characteristic equation of the system $1 + G(s)H(s) = 0$ has to be rewritten so as to be in the form

$$1 + G_1(s)H_1(s) = 0$$

where G_1H_1 comprises either K_{10} or K_{20} , whichever is chosen to be the parameter. This will always be possible provided the parameters in the characteristic equation are of linear form.

The characteristic equation of the closed-loop gain, can be rewritten in two ways depending on whether K_{10} or K_{20} is allowed to vary.

The two new open-loop gain equations are therefore

$$G_1H_1 = \frac{K_{10}(s + K_{20}/2)}{s^2(T_o s + 1 + T_o K_{20})} \quad \dots\dots(4)$$

where $K_{20} = \text{constant}$

$$G_2H_2 = \frac{K_{20}(T_o s^2 + K_{10}/2)}{s(T_o s^2 + K_{10} + s)} \quad \dots\dots(5)$$

where $K_{10} = \text{constant}$.

The plot of equation (4), shown in Fig. 3, represents the root loci of the system whose open-loop gain function is given by equation (2) for a fixed value of K_{20} with K_{10} as a parameter.

Let us consider the root loci of Fig. 3. As the complex pair of poles is clearly predominant, the effect of the real pole can be neglected. It can be seen that there is

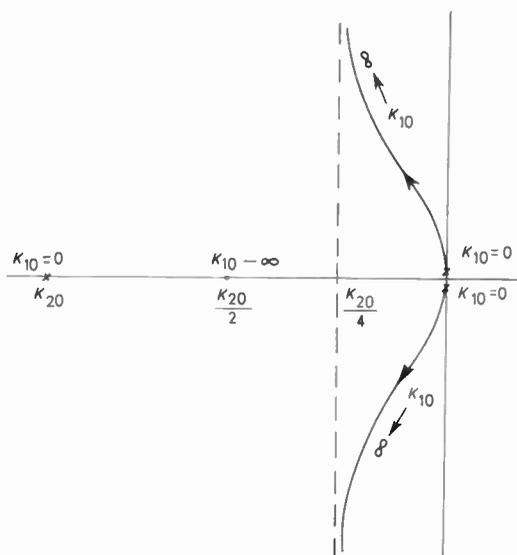


Fig. 3. Root locus diagram for uncompensated system.

a value of K_{10} for an optimum damping ratio at the point where the straight line drawn from the origin touches the root loci. For small values of K_{10} the poles are very near the imaginary axis and the system is, if not unstable in practice, then very underdamped. For large values of K_{10} the root loci move away from the imaginary axis towards an asymptote at $K_{20}/4$. In this case, the effect of the real pole cannot be neglected; in fact its influence becomes predominant and the whole system tends to become an idealized system whose characteristics depend solely on the main amplifier.

2.2 Stability Analysis of the Compensated System

The proposed compensation as shown in Fig. 4 consists of a small resistor r placed in series with the computing capacitor C between the summing junction of the computing amplifier A and common terminals of integrating input resistors. It is clear that the resistor r will not have any effects when in the COMPUTE mode while when in TRACK the voltage across C will not be equal to the output voltage. The error is proportional to the rate of change of the output voltage at the moment when the amplifier is switched from TRACK/I.C. to either COMPUTE or HOLD.

It will be shown in Section 3 that the unwanted features of the compensation with r can be easily balanced out.

Consider the loop gain transfer function of the compensated system which can be obtained by substituting Z_c with

$$Z_c = r + \frac{1}{sC} = \frac{1 + T_1 s}{sC} \dots\dots(6)$$

where T_1 denotes compensating constant rC .

Since the feedback ratio $H(s) = \frac{1}{2}$, the loop gain transfer function is

$$GH = \frac{K_{10} K_{20}(1 + T_1 s)/2}{s[(T_o + T_1)s^2 + (1 + T_o K_{20} + T_1 K_{10})s + K_{10}]} \dots\dots(7)$$

When compared with the loop gain transfer function of the noncompensated system, (see equation (3)), a marked difference is the presence of the zero $s_z = 1/T_1$ in the GH function. No new poles are added, but the positions of the existing ones are altered with respect to the locations in the original system.

For the purpose of analysing the effects of the added zero, the characteristic equation $1 + GH = 0$ should be

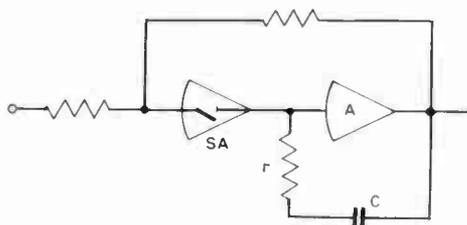


Fig. 4. Compensated system.

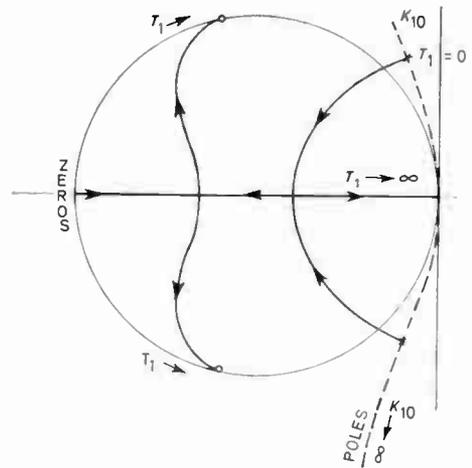


Fig. 5. Root locus diagram of compensated system.

rearranged so as to have $G_1 H_1$ in the form in which T_1 appears as a multiplying factor:

$$G_1 H_1 = \frac{T_1 s(s^2 + K_{10}s + K_{10}K_{20}/2)}{T_o s^3 + (1 + T_o K_{20})s^2 + K_{10}s + K_{10}K_{20}/2} \dots\dots(8)$$

The start of the root contours, when T_1 varies from zero to infinity, is at the poles, which in this case lie on the root loci of the non-compensated system. The root contours end at the zero placed on the root loci defined by

$$\frac{K_{10}(s + K_{20}/2)}{s^2} + 1 = 0 \dots\dots(9)$$

The completely drawn root contours are shown in Fig. 5 for fixed values of K_{20} and K_{10} and with T_1 as a parameter.

The following conclusions can be reached after the study of the root contours:

(1) The compensating resistor has improved stability because when T_1 increases, the predominant complex pair of poles moves away from the imaginary axis along an approximately circular path, and after a certain value of T_1 , the pair breaks into two real poles, thus transforming a heavily underdamped system into overdamped one. Apart from bringing the poles closer to the real axis, the increase of the compensating time constant T_1 brings the zeros closer to the origin, thus improving stability even further.

(2) The compensating resistor has increased the bandwidth. While for very small values of T_1 , the bandwidth of the system is hardly affected, the main contribution to the widening of the frequency response comes from the zero approaching the origin. This is because the modulus of the complex pair does not change much when T_1 rises.

(3) There is an optimum value of the compensating resistor above which the system again assumes the oscillatory response. In the diagrams this is indicated by the appearance of the complex pair of poles for larger T_1 .

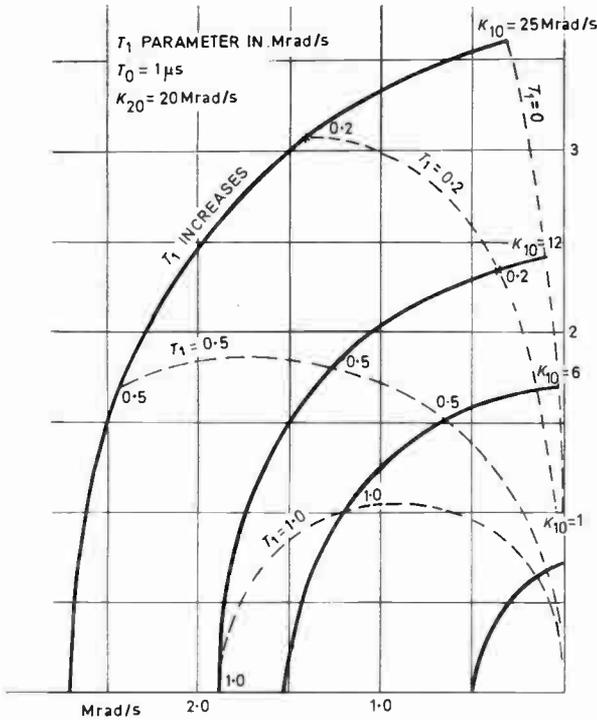


Fig. 6. More detailed representation of predominant poles location of Fig. 5.

(4) Less compensation is required when a switched amplifier of higher bandwidth is used. The root contours of constant compensation ($T_1 = \text{const}$, Fig. 6) show that when K_{10} is increased the bandwidth as well as the stability, is improved.

It is interesting to investigate the stability of the system for different values of the holding capacitor. This can be done by analysing the location of the predominant poles and zero when the ratio of the time constants T_0/T_1 are kept constant while their magnitudes are being altered. An example is shown in Table 1.

Table 1

$T_0 = T_1$ (μs)	Zero z (Mrad/s)	Predominant poles (p_1, p_2) (Mrad/s)	Remaining pole (Mrad/s)
0.1	10	$-1.38 \pm j 5.02$	-55.23
0.2	5	$-1.43 \pm j 3.49$	-52.64
0.5	2	$-1.47 \pm j 1.93$	-51.06
1.0	1	$-1.48 \pm j 0.87$	-50.53
2.0	0.5	-0.64 - 2.35	-50.27
5.0	0.2	-0.22 - 2.78	-50.11

$K_{10} = 6 \text{ Mrad/s}$, $K_{20} = 100 \text{ Mrad/s}$, $T_0/T_1 = 1$.

It can be seen from the Table that if the system is compensated for the smallest value of the holding capacitor, the stability will not be impaired very much when the capacitor is increased. For larger values of $T_0 (= T_1)$ the two most predominant poles are separated by the zero placed close to the nearest pole to the origin.

The open-loop gain can be approximated with two poles, separated wide enough to be able to maintain a phase margin of 45° . For large values of the holding capacitor the two poles become complex conjugates (Fig. 5) and the system becomes slightly underdamped.

In concluding the effects of the compensation with a resistor in series with the holding capacitor in the feedback loop around the main amplifier, it can be said that the advantages regarding the stability and bandwidth are substantial to justify the complications which this method may entail.

3. The Balancing of the Effects of the Compensating Resistor r

The stability analysis carried out in Sect. 2.2 has shown that great improvement with regards to bandwidth and stability has been achieved with addition of resistor r in series with the feedback capacitor C . However, the presence of compensating resistor r introduces an error when switching from I.C. into COMPUTE or from TRACK into HOLD (STORE).

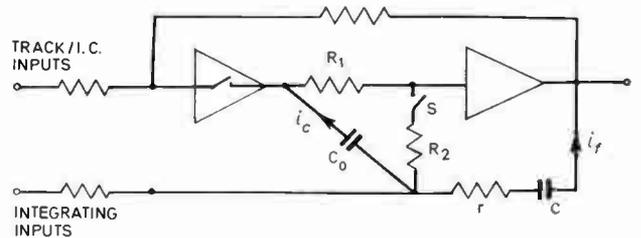


Fig. 7. Complete diagram of compensated system of Fig. 4.

Unwanted features of such a resistive compensation are due to current flow through r, C branch which produces a voltage drop across r . A compensation method designed to balance out the effects of the voltage drop on r is illustrated by an arrangement shown in Fig. 7. The compensation network consists of resistors R_1, R_2 , capacitor C_0 and switch S_0 .

The switch S_0 is on in TRACK and off in HOLD. If the current through the capacitor C_0 is neglected, the voltage across the same capacitor can always be made equal to the voltage drop across r . When the switch S_0 is turned off the voltage step across r is cancelled by the voltage stored on C_0 , which is switched in series with the voltage on feedback capacitor C . The conditions to be satisfied are that the output voltage when in HOLD, V_{oH} must be equal to the output voltage, V_{oT} , in TRACK.

$$V_{oT} = V_{oH} \quad \dots\dots(10)$$

These two voltages are given by

$$V_{oT} = i_f \left(\frac{1}{sC} + r + R_2 \right) \quad \dots\dots(11)$$

$$V_{oH} = i_f \left(\frac{1}{sC} + R_1 + R_2 \right) \quad \dots\dots(12)$$

where i_f is the current flowing through the feedback capacitor. The cancellation of the effects of the compensating resistor r , under the previously stipulated condition is achieved for

$$R_1 = r. \quad \dots\dots(13)$$

It is, however, impossible to achieve complete cancellation of the voltage drop across r because of the current which flows through the capacitor C_0 . Since the current is frequency dependent, the calculations must be done for the worst case when the current reaches the largest value. The current through the capacitor C_0 is approximately equal to

$$i_c = (R_1 + R_2)i_f s C_0. \quad \dots\dots(14)$$

For the small values of the compensating resistor r the current through the feedback capacitor C is approximately equal to

$$i_f = V_o s C. \quad \dots\dots(15)$$

The necessary condition to track the input voltage with 0.1% accuracy is therefore

$$(R_1 + R_2)V_o C C_0 \omega^2 r = 10^{-3} V_o \quad \dots\dots(16)$$

where $\omega = 2\pi f$.

On substituting $R_1 = r$, $C_0 = C$, from equation (16), the value of r is given by

$$r = \sqrt{\frac{R_2^2}{4} + \frac{10^{-3}}{(C\omega)^2}} - \frac{R_2}{2} \quad \dots\dots(17)$$

For $C = C_0 = 10\,000$ pF, $R_2 = 100 \Omega$, the maximum permissible value of $r = 26 \Omega$ can be used for the stability compensation.

Since the current that flows through C_0 is proportional to the value of this capacitor it is desirable to choose its value as small as possible. The minimum value is set by the permissible drift of the output voltage in HOLD which is inversely proportional to the series combination of C and C_0 .

The effect of capacitor C_0 on the stability can be best evaluated by considering the star network formed from the R_1 , R_2 , C_0 delta. The impedance of the arm in series with the input impedance of the amplifier is comparatively small and it can be neglected. The other two branches, whose impedances are also small, only slightly increase T_0 and T_1 producing overall little effect on the stability of the system.

4. Experimental Results

The mode control amplifier whose analysis was presented in this paper has been built and tested. The experimental results obtained during the testing of the circuit show that the analysis of stability and error has provided a good insight into the design procedure. The stability criteria have produced a useful background in searching for means of improving the performance of the mode control amplifier.

The difference between the compensated and non-compensated system is demonstrated in Fig. 8, where the small-signal pulse response of the system is shown: (b) when fully compensated, and (c) without compensation; (a) shows the signal at TRACE input.

The diagram (c) shows an oscillatory response typical for an underdamped system. This indicates presence of a pair of complex conjugate poles near imaginary axis, as shown in Fig. 6. The poles are located along the dashed

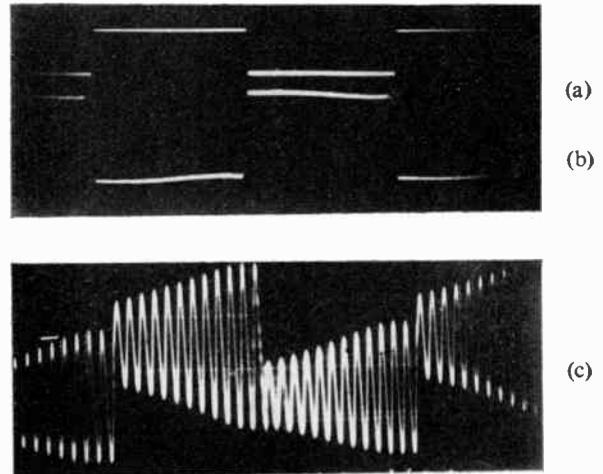


Fig. 8. Pulse response:
 (a) input signal, (b) output signal, compensated system.
 (c) output signal, uncompensated system.
 Horizontal scales: 2 μ s/div.
 Vertical scales: (a) 50 mV/div, (b) and (c) 20 mV/div.

curve drawn for compensating time-constant $T_1 = 0$. The diagram (b) of Fig. 8 clearly shows the absence of complex conjugate poles with T_1 chosen to have the smallest possible value necessary for the complex pair of poles to be transformed into two real poles.

5. Conclusions

The stability analysis of the mode control amplifier, presented in this paper, has shown that compensation is possible with a simple modification of the feedback loop around the main amplifier. By placing a resistor in series with the computing capacitor there is a zero added in the transfer function which, apart from improving the stability, increases the bandwidth. The value of the compensating resistor for optimum performance is relatively small so that the effects on computing accuracy can easily be balanced out. The experimental results clearly show the advantages of the proposed modifications.

6. Acknowledgment

The author wishes to express his gratitude to Professor H. A. Prime from the Electrical Engineering Department of the University of Birmingham for suggesting the topic of this research and to Dr. V. N. Kostić, the Director of the Mihailo Pupin Institute, for placing facilities for this work to be completed.

7. References

1. Eckes, M. R., 'A fast mode-control switch for iterative differential analysers', *Trans I.E.E.E. on Electronic Computers*, EC-14, pp. 946-50, December 1965.
2. Marjanović, S. and Noaks, D. R., 'Design criteria for the mode control of hybrid computer amplifiers', *Proc. Instn Elect. Engrs*, 117, No. 4, pp. 687-92, April 1970.

Manuscript first received by the Institution on 24th April 1970, in revised form on 29th September 1970 and in final form on 4th December 1970. (Paper No. 1373/IC 42.)

A Camera Tube Lag Meter

By

J. R. SANDERS, M.A.†

Reprinted from the Proceedings of the Conference on Television Measuring Techniques held in London from 12th to 14th May 1970.

The exposure of a camera tube to a brief flash of light and the measurement of the magnitude of the signal current after a known number of television fields can provide an indication of the speed of response of the tube to changing light levels. A meter which permits accurate lag measurements, in the presence of noise, to within an accuracy of $\pm 0.1\%$ of white level is described and its performance is discussed.

1. Introduction

The response of a television camera tube to changes of light reflected from the scene is not instantaneous and complete: the deficiencies of a tube in this respect are referred to as 'lag'.^{1,2}

Lag in a photoconductive camera tube results from inherent delays in the response of the photoconductive target material to changing light levels (photoconductive lag), and from incomplete discharge of the target by the beam during any one scan (capacitive lag).

With the now widespread use of photoconductive tubes, an accurate and rapid method of measuring overall lag characteristics is required. This is of special importance in the case of tubes used in colour cameras, since abnormal lag characteristics are one cause of coloured 'trailing' on moving objects.³

2. General Description

The equipment to be described was designed to satisfy two distinct applications with somewhat conflicting requirements. The first application is for the routine testing of camera tubes in a studio centre to a specification agreed between the camera tube manufacturer and the broadcaster. This specification is quite rigid and requires a simple set of tests, the results of which indicate whether or not the tube is acceptable.

The second application is in the laboratory for the examination of new camera tubes or for more detailed examination of existing tubes. In this case the equipment is required to be extremely versatile and provide facilities not required for routine measurements. These differing requirements are satisfied in the present equipment by making it versatile but extremely simple to operate, the complexity of setting the instrument for a particular measurement not being apparent to the operator who is provided with almost self-explanatory controls.

Until recently these tests have usually been performed by means of a mechanical shutter allowing light to fall on the camera tube target for a pre-determined time. A disadvantage of this method is that the precise time at which the illumination is switched on or off is difficult to control accurately. It has also been the practice to observe the magnitude of residual signals with a long persistence or storage oscilloscope; these signals inevitably have random noise added to them by the input stage of the

camera tube pre-amplifier and since they are typically of the order of 1% of the normal tube output, observations by means of an oscilloscope are of limited accuracy and fatiguing to the observer.

To overcome the disadvantage of a mechanical shutter, a small cathode-ray tube with a short persistence phosphor is used as the light source. No focusing or deflexion of the electron beam is provided, so that a comparatively large area of phosphor is uniformly excited to provide a source of light approximately 1 cm square. The light source is switched on and off at regular intervals and is arranged to illuminate a small area of the tube target using a normal camera lens to focus an image of the light source on the target; the remaining area of the tube face is in darkness.

Two types of measurement are possible. The first is a measurement of the speed of response of the tube when illumination is applied after a period of darkness, and is often referred to as 'build-up lag'. The second is the speed of response to the removal of illumination; this is often described as 'decay lag'.

During any single television field throughout the on/off cycle of the light source, samples may be taken of the magnitude of the video output from the tube. The samples correspond to two areas of the picture, one dark, and the other illuminated. By measuring the difference between these samples the magnitude of the tube output can be obtained, the value being independent of the black-level setting of the camera. The way in which the tube output builds up ('build-up lag') may be investigated by sampling after the light source has been turned on; 'decay lag' may be measured by sampling after the light source has been turned off. Any field during the light source cycle may be selected for sampling so that it is possible to examine, field by field, the response of the tube.

Samples of the tube output taken at the same point in every light source cycle will all be of identical magnitude apart from added noise. These successive samples are averaged in a circuit with a long time-constant so that the final measurement is almost independent of added noise.

The two areas of the picture to be sampled are defined by two internally generated 'white pulse' waveforms and, in order that one of these areas may be aligned with the illuminated area of the target, both waveforms and the tube video signal are fed to a picture monitor; the areas being sampled then appear as white rectangles (or

† B.B.C. Research Department, Kingswood Warren, Tadworth, Surrey.

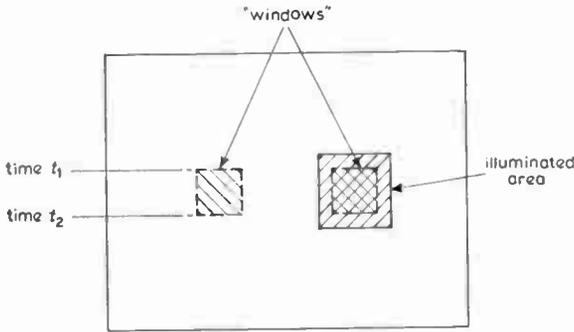


Fig. 1. Position of 'windows' relative to area of illumination.

windows) superimposed on the camera output. Typical positions of these 'windows' relative to the area of illumination are shown in Fig. 1.

The duration of the light pulse may be adjusted between the limits of 1 and 256 fields and the periods for which the light remains off between 64 and 1024 fields. The precise times at which the illumination is switched on and off is controlled by the 'window' waveforms. For decay-lag measurements the light source is switched off at time t_1 , shown in Figs 1 and 2(a); if, for example, 'FIELD 1' is to be measured, sampling starts immediately and ceases at time t_2 , but if 'FIELD 3' is selected, sampling still occurs between times equivalent to t_1 and t_2 but exactly three fields after the light source switched off.

In the case of build-up lag measurements, the illumination is switched on at time t_2 and sampling again takes place between times t_1 and t_2 (Fig. 2(b)) on whichever of the ensuing fields is selected.

3. Circuit Description

3.1 The Window Generator and Video Mixer

A block diagram of the complete equipment is shown in Fig. 3. Video signals are provided by a camera tube test bench, which is essentially a normal television camera channel with facilities particularly suitable for camera tube measurement.

The window generator uses field drive and mixed synchronizing pulses to produce the 'window' pulses which are used to select the two areas of the television

raster to be sampled. One set of pulses is added in the mixer unit to the video signals for display on a television monitor to allow adjustment of the relative positions of the light source and the timing of the 'window' pulses, as shown in Fig. 1. Each window pulse is also fed to an AND gate to select the two areas of picture to be sampled, one window pulse corresponding to an unilluminated area of the camera tube target and the other to the illuminated area.

3.2 The Logic Unit

For the measurement of decay lag the light source must be switched off at time t_1 (see Fig. 2), i.e. just before the commencement of a window pulse. For the measurement of build-up lag, however, the light must switch on at time t_2 which occurs just after a window pulse. The light source timing is therefore partly controlled by either the leading edge or the trailing edge of the field component of the 'window' pulses depending on whether decay lag or build-up lag measurements are to be made. This information is obtained from the leading or trailing edges of the field frequency pulses fed to the light source timing unit from the 'window' generator.

Field frequency pulses are also fed to a counter which counts the number of fields during which the light is on. At a preselected count, determined by the duration of 'light on' selected, the counter resets and the light is switched off. The counter then counts the number of field pulses during which the light is off. Again, at a preselected count, determined by the duration of 'light off' selected, the counter is reset and the light switched on so starting again the cycle of light on/off. The precise time at which the light is switched is determined by both the counter and the field frequency pulses from the window generator. In order to select the field on which a measurement is to be made, the outputs of the counter stages are also taken through an array of switches to an AND gate, this performing the function of 'field selector switching' shown in Fig. 3. The array of switches is arranged so that any particular field in the light on/off cycle can be selected by setting on the switches the binary number of the field required. If decay lag is to be measured, the

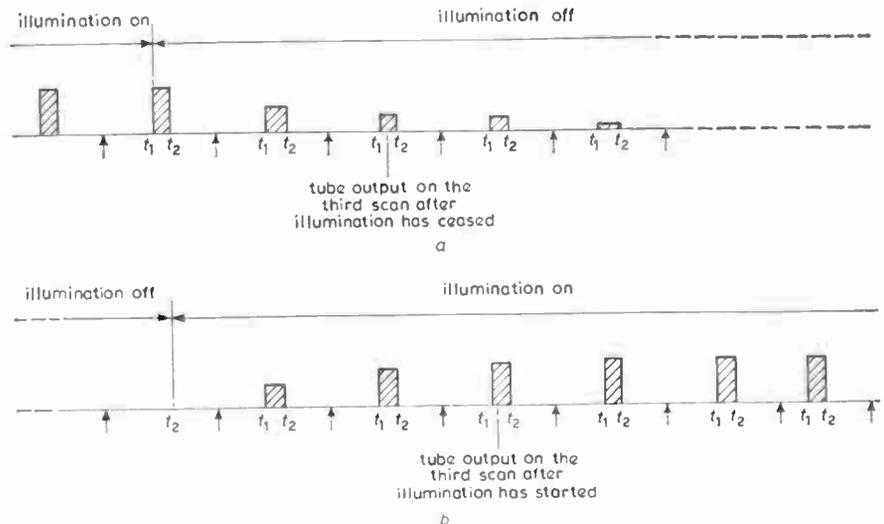


Fig. 2. Switching times for decay lag and build-up lag measurements.

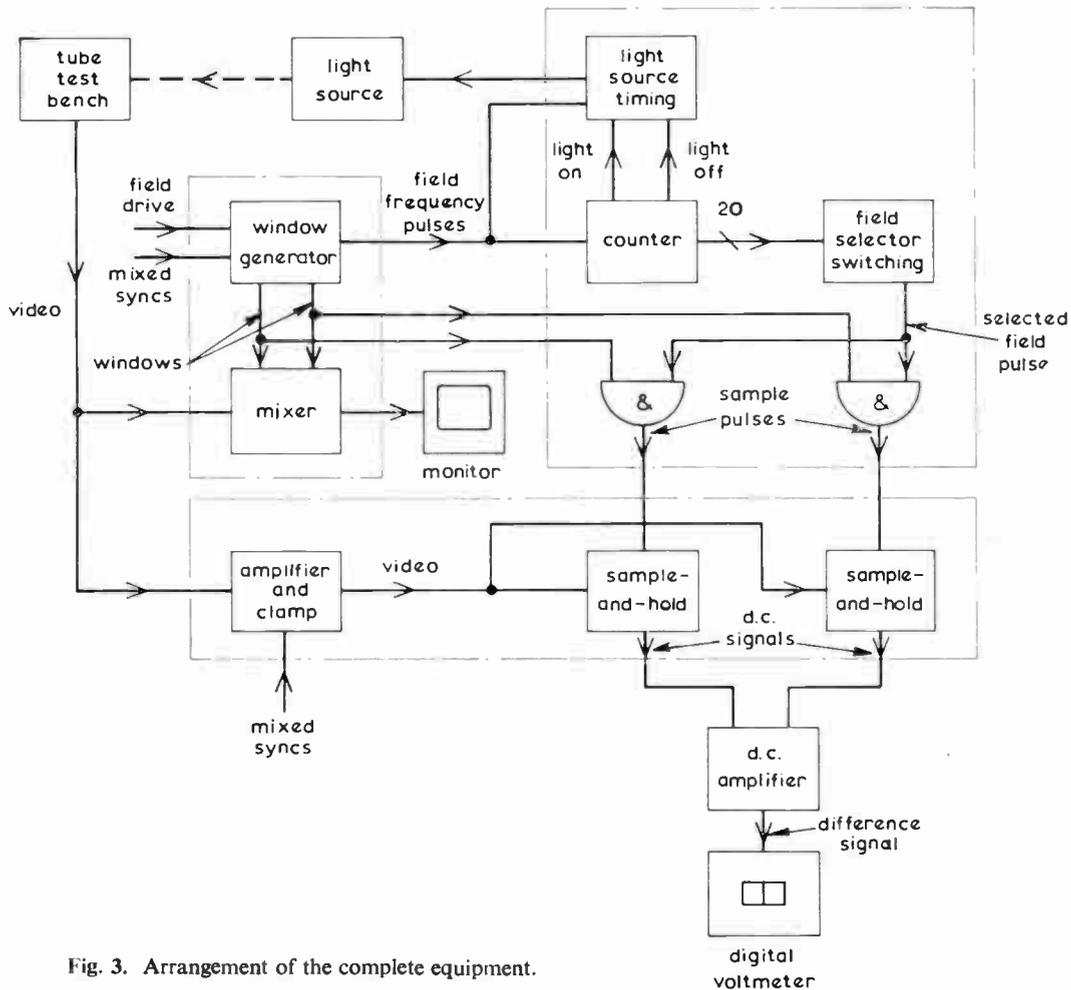


Fig. 3. Arrangement of the complete equipment.

number of the field selected refers to the number of fields which have elapsed since the light source turned off. If build-up lag is to be measured, the number refers to the number of fields which have elapsed since the light source turned on. The output of the field selector switch is a pulse on the selected field and is fed to two AND gates. One AND gate is also fed with the window pulse corresponding to the unilluminated area of the camera tube target and the other AND gate with the window pulse corresponding to the illuminated area. The pulse outputs of these gates are thus timed to occur when the magnitude of the video signal is to be examined and form the sample pulses used in the sample-and-hold stages.

3.3 Sample-and-Hold and D.C. Amplifier

The video signal from the test bench is amplified and clamped to reduce variations in the d.c. value of black level before it is passed to two identical sample-and-hold stages which sample the signal during each of the 'window pulses' and hold this value until the next sample pulse. The difference in potential of the d.c. signals from the sample-and-hold stages represents the difference in video signal magnitude from the two sampling areas. This difference is the value of the lag signal to be measured and is displayed on a digital voltmeter after amplification. In order to make the measurements substantially im-

mune to random noise added to the signals, a long time-constant is included in the charging circuits of the 'hold' capacitors. Since the measurement is repeated once every light on/off cycle, the noise measurement will tend to average to zero but the signal magnitude will be correctly measured since it will have the same value at each measurement. The time-constant used is such that the correct measurement is available after approximately five samples. This time-constant is a compromise between noise immunity and speed of measurement.

4. Performance

The accuracy of the meter is such that readings between 1% and 10% of white level are accurate to $\pm 0.1\%$ of white level. The addition of white noise at -20 dB r.m.s. relative to the signal level corresponding to white, causes variations in the output signal of not more than 0.1% of white level. D.c. drift resulting from changes in the voltage offsets across the field effect transistors used in the sample-and-hold stages is not more than 0.1% of white signal after 15 minutes warming up time.

5. Conclusions

The meter enables lag in camera tubes to be measured to an accuracy of $\pm 0.1\%$ of white signal in the presence

of noise. The main advantage of the technique is the unambiguous indication of lag, without the uncertainty that is introduced when small signals are examined on an oscilloscope in the presence of noise.

6. Acknowledgments

The author wishes to thank the Director of Engineering of the BBC for permission to present this paper, and for the assistance in design and construction of the equipment by his colleagues I. D. Chisholm, D. J. Meares and M. J. Mortimore.

7. References

1. de Hann, E. F., van der Drift, A. and Schampers, P. P. M., 'The "Plumbicon", a new television camera tube', *Philips Tech. Rev.*, **25**, No. 6/7, pp. 133-51, 1963-64.
2. Beurle, R. L., 'Low-light-level television', *Proc. Instn Elect. Engrs*, **110**, No. 10, pp. 1735-46, October 1963.
3. Van de Polder, L. J., 'Target-stabilization effects in television pick-up tubes', *Philips Res. Reports*, **22**, pp. 178-207, April 1967.

Manuscript received by the Institution on 7th January 1970. (Paper No. 1379/IC43.)

© The Institution of Electronic and Radio Engineers, 1971

STANDARD FREQUENCY TRANSMISSIONS—March 1971

(Communication from the National Physical Laboratory)

March 1971	Deviation from nominal frequency in parts in 10 ¹⁰ (24-hour mean centred on 0300 UT)			Relative phase readings in microseconds N.P.L.—Station (Readings at 1500 UT)		March 1971	Deviation from nominal frequency in parts in 10 ¹⁰ (24-hour mean centred on 0300 UT)			Relative phase readings in microseconds N.P.L.—Station (Readings at 1500 UT)	
	GBR 16 kHz	MSF 60 kHz	Droitwich 200 kHz	*GBR 16 kHz	†MSF 60kHz		GBR 16 kHz	MSF 60 kHz	Droitwich 200 kHz	*GBR 16 kHz	†MSF 60 kHz
1	-299.9	+0.1	+0.1	647	637.4	17	-300.0	+0.1	0	637	634.9
2	-299.6	+0.1	+0.1	643	636.5	18	-299.9	0	-0.1	636	635.0
3	-300.3	+0.1	+0.1	646	635.6	19	-300.0	+0.1	-0.1	636	634.2
4	-299.8	+0.1	+0.1	644	634.9	20	-299.9	+0.1	0	635	633.6
5	-300.0	+0.1	+0.1	644	634.1	21	-300.0	+0.1	0	635	632.5
6	-300.0	+0.1	+0.1	644	632.7	22	-300.1	0	+0.1	636	632.2
7	-299.9	0	+0.1	643	632.4	23	-299.9	0	+0.1	635	631.5
8	-299.9	+0.1	+0.1	642	631.7	24	-299.9	+0.1	+0.1	634	631.0
9	-299.9	0	+0.1	641	630.3	25	-299.9	+0.2	+0.1	633	629.0
10	-299.9	+0.1	+0.1	640	629.3	26	-299.8	+0.2	+0.1	631	627.4
11	-299.9	+0.2	+0.1	639	627.3	27	-299.8	+0.1	+0.1	629	626.4
12	-299.9	+0.1	+0.1	638	626.0	28	-299.9	+0.1	+0.1	628	625.2
13	-299.9	+0.2	+0.1	637	624.5	29	-300.0	+0.2	+0.1	628	623.4
14	-299.8	+0.1	+0.1	635	623.5	30	-300.1	0	+0.1	629	622.2
15	-299.9	+0.2	0	634	621.9	31	-299.9	+0.1	+0.1	628	621.0
16	-300.3	+0.1	0	637	635.6						

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 *Mediator* Air Traffic Control System Comes into Operation



Operations Room at the London Air Traffic Control Centre, West Drayton, Middlesex.

In the early 1960s, *en-route* air traffic control in the United Kingdom was conducted from many separate units by civil and military controllers in two separate organizations. It was not unsafe, but two things were needed to prepare for the challenge of the 1970s.

The first was an organizational structure within which a.t.c. could grow together into a single organism, and the National Air Traffic Control Services form this structure.

The second was a plan and programme for evolution and re-equipment, to enable the air traffic service to condense progressively into a smaller number of units and to provide these centres of condensation with progressively better facilities. *Mediator* was the name given to this plan and programme and a new Air Traffic Control Centre, which went into operation on 1st February this year at West Drayton and will remain the biggest of these concentrations.

The National Air Traffic Control Service

Established in 1962 to achieve, progressively, a common national civil/military air traffic service within a unified organization, the N.A.T.C.S. has a joint staff of civil and military personnel. It provides and operates air traffic control, a.t.c. support services such as radar and communications, navigation facilities and certain information services, for all users of the airspace.

For the purposes of *en-route* air traffic service, the United Kingdom airspace is divided into four component airspaces. Firstly, there are the airways and associated control areas extending downwards at least to 8000 feet and in some areas to ground level, but with a ceiling of 25 000 feet, together constituting Controlled Airspace (CAS). Secondly, there is the airspace outside the lateral confines of Controlled Airspace but still below 25 000 feet and known as the Middle Airspace (MAS). Thirdly, there is the airspace below CAS and MAS, the Lower Airspace (LAS). Fourthly, there is the airspace above 25 000 feet, the Upper Airspace (UAS). Responsibility for the control of air traffic flying in accordance with international civil procedures in the UAS over much of Europe, including the U.K., is formally vested in the Eurocontrol Agency but that agency delegates to N.A.T.C.S. the a.t.c. service in the UAS over the U.K.

Evolution and improvement of the a.t.c. system must of necessity keep pace with increases and forecast increases in the volume and speed of air traffic. Civil air traffic alone could double by the end of the 1970s. The planned introduction of supersonic transport aircraft, the expected rapid growth in air cargo operations and an increase in general aviation flying will ensure that the problems to be faced will be more than a simple increase in numbers alone.

Outline of *Mediator*

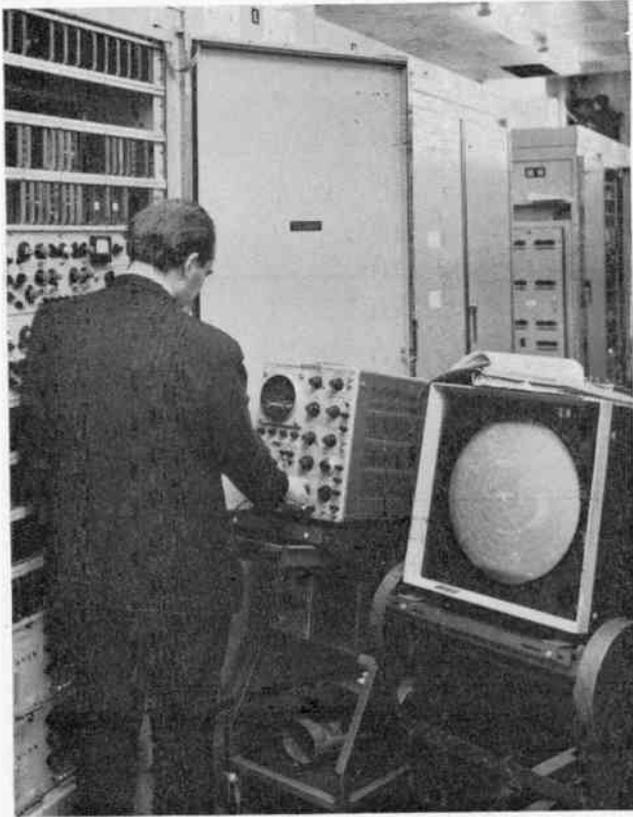
The *Mediator* plan will transform the present manual system progressively into a largely automated system. It will be introduced in a series of logical stages and will steadily extend the area of responsibility of the London Air Traffic Control Centre, West Drayton.

The main objective is to increase the capacity of the a.t.c. system by modernization, integration and improvement of all aspects of the a.t.c. service. Advantage will continue to be taken of all available means of improving the efficiency and safety of the *en-route* air traffic service by the introduction of new techniques (for example, the automatic processing of a.t.c. data), the better exploitation of existing or emerging systems (for example, the secondary surveillance radar system) and better system integration and data interchange between air traffic controllers carrying out related functions or responsible for aircraft in adjacent parts of the total airspace.

The facilities that go to make up Stage I of the *Mediator* plan may be classified under four basic headings.

Firstly, there are the peripheral radar and communications stations located at a number of strategic points throughout the South of the country. These stations are essentially the nerve ends of the *Mediator* system in that they sense the location and identity of aircraft flying in the U.K. airspace and provide the communications with them.

Secondly, there are communication links connecting the peripheral stations with the West Drayton centre and others connecting West Drayton with adjacent air traffic control centres in Europe and with airports and other air traffic service units within the U.K.



Setting-up one of the 400 cabinets of electronic equipment forming part of *Mediator*.

Thirdly, there are the data processing and distribution systems by means of which collected information is converted into a form suitable for use by the air traffic controllers and then routed to the required operating positions.

Fourthly, having collected, processed and distributed the information required, it is necessary to display this information at the various a.t.c. operating positions by means of radar display systems, closed circuit television and runway visual range systems, and on operating consoles which represent the man/machine interface.

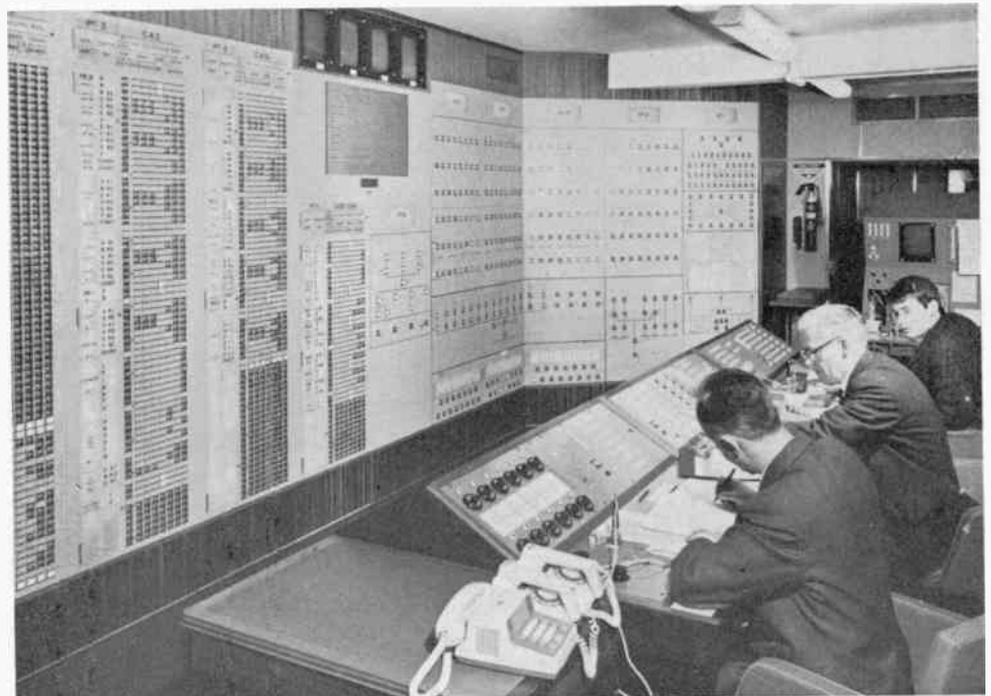
Integration of all these facilities into a viable a.t.c. system counts for nothing unless the whole system can be operated effectively and maintained to a high order of reliability. A self-contained Maintenance Communication Network has therefore been provided for the engineers responsible for running the whole *Mediator* complex.

The initiation of Stage 1 of *Mediator* before the Flight Plan Processing System was in its envisaged fully computerized form has been regarded by the N.A.T.C.S. as not entirely a disadvantage. The whole concept of regarding radar as the primary source of control information is certainly revolutionary and understandably two virtually completely new systems simultaneously introduced could well have been an added cause of what the Deputy Chief of N.A.T.C.S., Air Vice-Marshal E. D. Crew, referred to as 'stage fright'. Currently the Flight Plan Processing System is a Ferranti *Hermes* computer which is an expanded store version of the Interim Flight Plan Processor (*Minicap*) in use since 1967.

Although the central (*Hermes*) processor has proved to be highly reliable, having on one occasion run for one year without an equipment fault, the total system was not designed to have the equipment redundancy required for the continuous operation of this type of 'real-time' system. When out of service for programme testing, planned maintenance or fault conditions, preparation of flight plan strips at West Drayton reverts to being a manual operation. F.p.p.s. systems with this redundancy (i.e. having a number of computers in-system) will enter service in later stages or sub-stages of *Mediator* and a *Myriad II* computer is scheduled to come into operation next year.

A further article describing in fuller detail some of the elements of this notable achievement by the British electronics industry will be published in an early issue of the *Journal*.

System Control for monitoring the *Mediator* complex. A 'tote' presentation of remote indications from the various equipment rooms enables system controllers to assimilate the mass of equipment serviceability state information. At present the remote indications are set up by maintenance personnel in the equipment rooms but fully automatic indication is a goal to be effected progressively.



Application of Insertion Test Signal Techniques to Television Transmission Chain Operation

By

D. C. SAVAGE, C.Eng., M.I.E.E.,†

and

D. A. CARTER (Graduate)†

Reprinted from the Proceedings of the Conference on Television Measuring Techniques held in London from 12th to 14th May 1970.

This paper discusses the advantages to be obtained from insertion test signal techniques. Consideration is given to the types of distortion which need to be measured and the consequent requirements of the test signal waveform. Methods of measurement are discussed and typical results obtained in practice are shown. Probable accuracy of measurement is considered and the paper concludes with a suggestion of some possible future developments.

1. Introduction

Insertion test signal techniques are playing an increasing part in the operation of television transmission chains. These techniques enable optimum transmission performance to be maintained in an economic manner, since test signals inserted during the field-blanking period can be monitored whilst the chain is carrying normal traffic. Manual monitoring, measurement and control are possible, but already automatic methods are supplementing or replacing the manual ones. A further description of automatic measuring techniques is given in a companion paper by Shelley and Williamson-Noble.¹

2. Parameters to be Measured

Complete specification of the performance of a television transmission chain requires the measurement of a large number of different parameters. Reference 2 lists some thirty different impairments which can afflict a colour television signal in transit and further types of impairment are still being discovered. In practice however it appears that a reasonable assessment of performance can be made from a much smaller number of measurements.

The authors consider measurement of the following distortions and impairments to be adequate for normal routine purposes:

- Luminance gain
- Luminance *K*-rating
- Chrominance–luminance gain inequality
- Chrominance–luminance delay inequality
- Luminance non-linearity
- Differential gain
- Differential phase
- Chrominance–luminance crosstalk
- Signal/unweighted random noise ratio

Some organizations, however, consider other impairments equally important, or more important, than those listed above. Some, for instance, favour checking the amplitude frequency response at a number of spot frequencies, and others regard it as important that the linearity of the chrominance channel with varying amplitudes of sub-carrier should be assessed. Whilst agreement concerning the checking of international

transmission networks is in sight, it seems likely that detail differences between domestic network practices will persist for some time.

3. Test Signal Requirements

The requirements for a test signal depend upon the nature of the parameters which it is desired to measure. There are also other requirements, e.g. the signal must produce minimum annoyance when viewed on a domestic receiver and it must occupy the minimum of time in the field-blanking interval, since there are other customers for this valuable commodity. The signal must also be suitable for measurement by the simplest and most robust equipment, both manual and automatic.

Most test signals so far proposed occupy two adjacent lines in each field blanking interval. It has thus been possible to agree on an international recommendation³ that, for 625-line systems, lines 17 and 18 (330 and 331) should be reserved for international test signals, whilst lines 19 and 20 (332 and 333) may be used for national test signals.

One fundamental difference between the proposals concerns the choice between originating similar test waveforms in each field, or alternatively sending waveforms which differ between odd and even fields. The latter proposal allows a greater number of impairments to be investigated, but it suffers from the disadvantages of a possibility of a 25 Hz flicker component at the top of the received picture, a rather dimmer oscilloscope trace, poorer signal/noise ratio and more complex instrumentation.

4. The U.K. National Test Signal

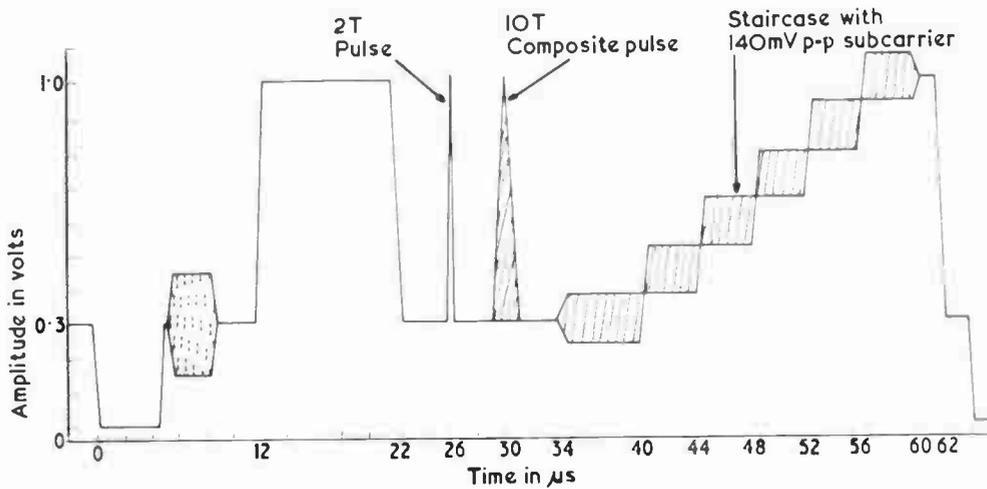
A national test signal has been evolved in the U.K., based upon the considerations outlined in the previous two sections. This signal, which has now been agreed between the B.B.C. and the I.T.A., is shown in Fig. 1. The signal is identical on both fields since, when all factors had been considered, it was felt that this arrangement was preferable to the other.

The individual components of this waveform are described in the following sub-sections.

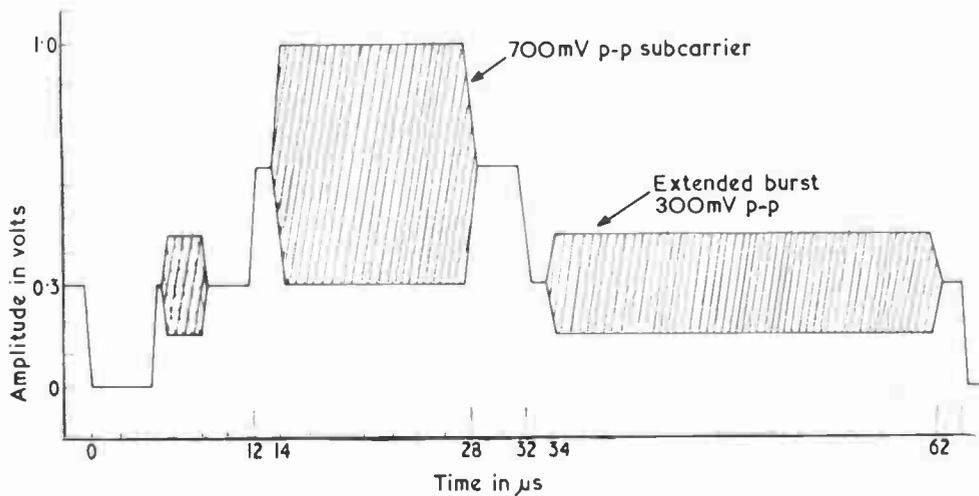
4.1 Sub-carrier Reference Burst

A sub-carrier reference burst is shown immediately after each synchronizing pulse in Fig. 1. This burst may be required in order to establish a reference-phase

† Designs Department, British Broadcasting Corporation, London, W1A 1AA.



(a) Line 19 (and 332)



(b) Line 20 (and 333)

Fig. 1. U.K. national test signal

for differential phase measuring equipment. Alternative equipment has however been devised to determine the value of differential phase impairment in the absence of a burst, since the reference burst would not normally be present during monochrome transmission periods.

4.2 Peak White Bar

A 10 μs white bar follows the burst on line 19 (and 332). This bar defines the luminance level at the point of measurement. It also enables an approximation to the bar *K*-rating impairment to be made. (The approximation here, as in the case of pulse/bar ratio below, arises from the use of a bar having a duration of only 40% of that specified for conventional *K*-rating.⁴)

4.3 2T Sine Squared Pulse

The 2*T* pulse enables the pulse *K*-rating to be measured. In conjunction with the white bar it also allows an approximation to the 2*T* pulse/bar ratio to be obtained. The value of *T* for the U.K. (System I) is of course 100 ns.

4.4 10T Composite Pulse

The 10*T* composite pulse is formed from two signals. The first of these is a 10*T* sine-squared pulse, having an amplitude equal to 50% of that of the 10 μs bar. The second is the envelope resulting from 100% amplitude modulation of colour sub-carrier by the first signal. Addition of these two signals results in a composite pulse envelope, sine-squared in shape, with a half-amplitude duration of 10*T*, see Fig. 2(a).⁵ Since this resultant pulse contains both luminance and chrominance information it can be used to assess chrominance-luminance gain and delay at a remote point. The phase of the sub-carrier in this signal, and in the other signals described below, is locked to that of the reference burst when the latter is present.

4.5 Five-riser Staircase

The 5-riser staircase which occurs at the end of line 19 (and 332) has a luminance amplitude equal to that of the 10 μs white bar. Superimposed upon this staircase is a

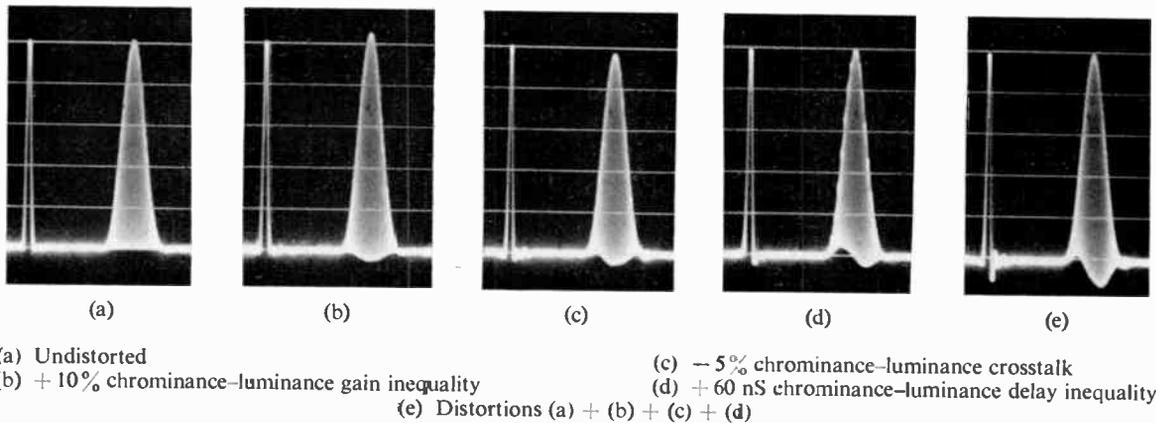


Fig. 2. $2T$ pulse and composite $10T$ pulse sections of U.K. insertion test signal.

colour sub-carrier envelope having a peak-to-peak amplitude equal to 20% of that of the bar. Information on luminance non-linearity can be obtained from the luminance component of the composite staircase signal although the accuracy of this information may be reduced by any large amount of chrominance-luminance crosstalk (see Sect. 4.6) which is present at the same time. Differential gain and phase impairment can be assessed from the chrominance component of the staircase waveform.

4.6 Chrominance Bar

A $14\ \mu\text{s}$ chrominance bar follows the reference burst on line 20 (and 333). This bar has a peak-to-peak amplitude equal to that of the $10\ \mu\text{s}$ white bar. It is superimposed on a 50% luminance pedestal. Comparison of the amplitude of the chrominance bar with that of the luminance bar gives information on gain inequality. Comparison of the mean value of the chrominance bar with that of the pedestal upon which it is superimposed, reveals the magnitude of any chrominance-luminance crosstalk impairment.

4.7 Extended Chrominance Burst

The final element of the U.K. test signal is a $28\ \mu\text{s}$ burst of sub-carrier having a peak-to-peak amplitude equal to that of the reference burst. This element may be used as a reference in one method of determining the amount of differential phase present in the sub-carrier superimposed upon the staircase.

5. Other Test Signals

Two-line insertion test signals proposed by administrations outside the U.K. have many similarities with that described in the previous section. It is not possible to describe all these signals in this paper, but a few of the suggested variations are noted below.

The $10\ \mu\text{s}$ white bar appears to be common to all proposals and a $2T$ pulse is included in most, but it should be remembered that the value of T is dependent upon the television standard with which the signal is associated. The standard also affects the duration of the composite pulse where proposed. In the case of many European countries, for instance, the equivalent composite pulse has a half-amplitude duration of $20T$.

It has been proposed that the composite staircase should be used for chrominance linearity measurements only and that a staircase, without superimposed sub-carrier, should be alternated with it on sequential fields, specifically for luminance non-linearity measurements. In this case the sub-carrier superimposed upon the staircase would have an amplitude equal to 40% of that of the white bar.

Some administrations intend using a $14\ \mu\text{s}$ bar containing three levels of sub-carrier, instead of the one level shown at the beginning of line 20 (and 333) (Fig. 1(b)), and it has been suggested that the extended burst could be dispensed with. It has also been proposed that a multi-burst should either replace the signal shown on line 20 (and 333), or else that these two signals should be alternated on a field-by-field basis.

The composition and timing of the U.K. signal has been tailored to be as near compatible as possible with many of the above proposals, and any re-instrumentation required to deal with international chains should not be too difficult.

In passing it may be noted that experiments with yet another type of insertion test signal, based upon a colour bar, are currently taking place. This type of signal may well have applications inside a television studio complex where coders and decoders, etc., are involved.

6. Methods of Measurement

Methods of measurement can be divided into two classes—manual and automatic. The latter class is described in the companion paper,¹ but in general it is true to say that what can be measured manually can also be measured automatically, and that the basic measurement techniques involved in both classes have a number of similarities.

Conventions in television transmission measurement were set before the advent of insertion test signals,⁴ and because the waveforms in the latter differ in some respects from those previously specified, it follows that some of the results may also differ. To be accurate, therefore, one should perhaps prefix some of the i.t.s. results with a qualifying adjective, e.g. 'quasi' or 'pseudo'. On the other hand, it may be that as i.t.s. techniques

become more widely adopted the existing conventions will be brought into line with them.

Various methods of manual measurement are described below and details of current B.B.C. practice are given in the Appendix.

A first essential in any system of measurement is to locate and isolate the relevant test waveform. This involves some form of strobing technique, each sequence being initiated by the field synchronizing pulses. For manual measurement it is also essential to use an oscilloscope having adequate trace brightness when operated in a field strobe mode at a time-base velocity of the order of 2 cm/ μ s.

6.1 Luminance Gain

The luminance gain of a transmission chain is conventionally specified in terms of the amplitude of the 25 μ s bar in a pulse and bar waveform. In the absence of waveform distortion the apparent gain of a system will be independent of bar duration and the 10 μ s white bar in the i.t.s. will be equally satisfactory. Waveform distortion which tilts the top of the bar presents a problem, however, and an arbitrary decision has to be made as to which points on the i.t.s. waveform should be measured to determine the system gain (see Appendix).

6.2 Luminance *K*-Rating

Line frequency bar *K*-rating is also normally specified in terms of a 25 μ s bar, and an arbitrary decision has therefore to be taken as to how best to relate measurements on a 10 μ s bar to those which would have been obtained on the longer bar. The authors feel that a reasonable solution is to treat the 10 μ s bar for *K*-rating purposes as if it was the first part of a notional 25 μ s bar, utilizing a slightly modified conventional graticule (see Appendix).

Pulse/bar ratio is conventionally expressed in terms of a 25 μ s bar, but for i.t.s. measurements it seems reasonable to relate the *2T* pulse to the 10 μ s bar amplitude as discussed in Section 6.1.

K-rating the *2T* pulse itself places stringent requirements upon the display oscilloscope, but it can be done, in conjunction with the normal *K*-rating graticule.

50 Hz bar *K*-rating cannot even be roughly estimated on the i.t.s. waveform. The absence of low frequency test information appears to be a fundamental limitation to all i.t.s. techniques, but some qualitative information may be obtained from the 'sit' of the whole blanking period.

6.3 Chrominance-Luminance Gain Inequality

Chrominance-luminance gain inequality can be measured on either the *10T* composite pulse or the 14 μ s chrominance bar, the choice being conditioned by the apparatus available and also by the nature of other distortions present simultaneously on the chain under test. Because the *10T* pulse contains both luminance and chrominance components its shape is dependent upon delay as well as gain inequality (see Fig. 2). The 14 μ s bar, however, is not composite in the same sense, and its shape is unaffected by chrominance-luminance

delay inequality. In practice gain and delay inequalities are generally found to coexist so that the 14 μ s chrominance bar is preferred for gain measurement if only an oscilloscope is available.

Measuring equipment has been developed in the form of a calibrated gain and delay corrector which can be inserted between the test point and the oscilloscope. When using this equipment on a linear chain the controls are adjusted to restore the shape of the original *10T* composite pulse, judged by the flatness of its base line. The gain and delay inequalities can then be read directly from the settings of the controls.

The presence of non-linearity on a chain renders all linear impairment measurements ambiguous to some extent. This is particularly true when gain inequality has to be measured over a chain which is contributing a significant amount of chrominance-luminance crosstalk distortion. As an example consider a typical chain which introduces 5% of chrominance-luminance crosstalk. Measurement of gain inequality by comparing the peak-to-peak amplitude of the chrominance with that of the 10 μ s white bar might indicate 0% inequality, whereas measurement on a basis of endeavouring to restore the base line of the *10T* pulse might show a 10% inequality. This discrepancy could well be greater than the normal maintenance limit for the chain! Once again an arbitrary decision has to be made, and in the author's opinion the peak-to-peak chrominance to luminance bar definition is to be preferred. It is also their opinion that the 14 μ s chrominance bar is the better waveform for chrominance amplitude measurement.

6.4 Chrominance-Luminance Delay Inequality

Chrominance-luminance delay inequality can be measured on the *10T* composite pulse. The presence of this impairment is indicated by the appearance of a positive and a negative lobe on the base of the pulse, (gain inequality on its own giving rise to a positive or a negative lobe) (see Fig. 2). If only an oscilloscope is available, then the magnitude of delay inequality can be estimated from the amplitudes of the lobes, (see Appendix). The accuracy of this estimate is markedly reduced by any appreciable amounts of gain inequality or chrominance-luminance crosstalk occurring simultaneously. An alternative method is to use the calibrated gain and delay corrector mentioned in the previous sub-section. This latter method obviates errors due to gain inequality.

6.5 Luminance Non-linearity

This distortion, which is also known as line-time non-linearity or picture signal distortion factor, can be assessed from the staircase waveform. The method normally adopted is to filter off the sub-carrier, differentiate the resulting luminance staircase and then measure the five pulses whose amplitudes correspond to the height of the received staircase risers (see Appendix).

The conventional waveform for carrying out this test does not carry any superimposed sub-carrier and, as mentioned previously, the validity of the i.t.s. measurement can therefore be lessened by any chrominance-luminance crosstalk which is present. However, the

sub-carrier amplitude is low and the error is thus not significant, unless the chain produces a large amount of crosstalk. It should be noted that the conventional test procedure for this distortion, and for differential gain and phase (see below) involves measurements at average levels of 10% and 90%, and that these conditions are unlikely to be met for i.t.s. measurements made during normal programme transmission periods.

6.6 Differential Gain

This distortion can be measured from the staircase waveform by filtering off the luminance component and examining the shape of the resultant sub-carrier envelope (see Appendix). An oscilloscope display which is rather easier to interpret can be obtained by passing the sub-carrier envelope through a full-wave detector.

6.7 Differential Phase

Differential phase can also be measured from the staircase waveform but the methods tend to be rather more complicated than those previously described. Essentials are a phase-sensitive detector and some method of establishing a reference signal with which to drive it. Five possible methods of measurement are outlined below.

The first method is the one in general use within the B.B.C. In this the reference carrier is obtained from a conventional burst locked oscillator which is fed with reference bursts from the signal to be measured. The resultant reference signal is fed, through a calibrated adjustable phase-shifter to one input of a phase-sensitive detector. The sub-carrier envelope of the signal to be measured is fed into the other input. The phase-shifter is first adjusted to bring the output of the detector to zero for the black level tread. The relative phase of the sub-carrier on the other five treads can then be determined by readjusting the calibrated phase-shifter for zero detector output on each tread in turn (see Appendix).

The second method, which has been used experimentally within the B.B.C., utilizes the same phase shifter and detector as above, but derives the reference signal from the extended burst at the end of the test signal (Fig. 1). The sub-carrier envelope of the signal to be measured is delayed 64 μ s, by a standard PAL delay line, in order that both test and reference signals enter the detector coincidentally. This method has the advantage that differential phase measurements can be made during monochrome television transmission periods, but the disadvantage that the signal/noise ratio at the detector output is a little worse than with the first method.

A third method, which has been tried on occasions, is to utilize a vector display oscilloscope. The resolution accuracy of this method tends to be rather poorer than that of the previous two.

A fourth method has been proposed which involves splitting the received sub-carrier envelope into two paths. One path is connected directly to the phase-sensitive detector, whilst the other path is connected to the second input of the detector via a 4 μ s delay line (4 μ s being the duration of one tread on the staircase waveform), and a calibrated phase shifter. The result of this arrangement

is that the phase of the sub-carrier on each tread can be measured relative to the phase of the sub-carrier on the adjacent treads.

A fifth method has been developed by the I.R.T. in Munich. In this a reference sub-carrier is derived from a local oscillator which is locked to the sub-carrier on the black level tread of the received staircase waveform.

The latter two methods require the presence of neither a reference burst nor an extended burst during the i.t.s. period.

In practice the main difficulty found with the measurement of differential phase, and to a lesser extent differential gain, is that of accuracy of reading in the presence of noise levels found on typical transmission chains. Equipment to integrate out the noise from the phase-sensitive detector over a number of fields has been developed experimentally, but none is yet in service, within the B.B.C. at any rate.

6.8 Chrominance-Luminance Crosstalk

The nature of the non-linearity which gives rise to chrominance-luminance crosstalk is often dependent upon instantaneous circuit loading. This dependence can result in a difference between the distortion suffered by the 10T composite pulse and that suffered by the 14 μ s chrominance bar. There are at present no plans to measure this variation with time. Since the effects of this distortion are probably more perceptible on large areas of colour it would seem that chrominance-luminance crosstalk should preferably be measured on the 14 μ s chrominance bar section of the i.t.s. waveform. The signal to be measured is connected, via a suitable sub-carrier removing filter, to the oscilloscope. Any difference between the mean value of the sub-carrier bar and the value of the pedestal upon which it is superimposed can then be easily detected (see Appendix).

6.9 Signal/Noise Ratio

Signal/noise ratio can be assessed on any 'quiet' lines occurring during the field-blanking period, e.g. line 16. Oscilloscope measurement of noise on a peak-to-peak basis requires a certain amount of skill and experience, and even then the results are probably not very reliable for separations of more than about 30 dB peak-to-peak. In particular it is questionable whether the figure of 18 dB, conventionally used for converting peak-to-peak to r.m.s. results, is appropriate to practical i.t.s. assessments. There seems to be a real need for a gated noise measuring set suitable for this purpose and development is at present proceeding in this field.

7. Results of Measurements

An insertion test signal waveform similar to the first line in Fig. 1 was added to line 18 (and 331) of BBC-2 transmissions as far back as mid-1966. In 1967 this waveform was transferred to line 19 (and 332) and at the beginning of 1969 the second line of waveforms was added.

An experimental i.t.s. generator producing a signal similar to that shown in Fig. 1, but adapted to 525-line N.T.S.C. standards, was sent to Mexico City for use by the E.B.U. during the 1968 Olympic Games.

The B.B.C. started daily routine checks on the performance of their first colour distribution system (BBC-2) at the beginning of 1967. These checks utilized semi-conventional line-by-line test signals transmitted from the Television Centre during one hour of the mid-day break in trade transmission,⁶ and covered most of the impairments listed in Section 2.

In April 1969 the routine checks were transferred from a line-by-line to an i.t.s. basis. This transfer enabled the test period to be changed to the operationally more convenient time of 1000–1030 hours. It was arranged that a test card should be radiated between these times in order to ensure uniformity of results. The transfer also enabled the time required to take a set of measurements to be reduced from 30 minutes to around 10 minutes. In November 1969 the checks were extended to the second colour distribution network (BBC-1).

Two sets of results are presented in Sections 7.1 and 7.2 in order to give some idea of typical distortions which can be met with in practice. The first is for an international chain including one satellite. The second is for a complex national distribution chain.

7.1 Mexico—London

An insertion test signal was present on transmissions over the Mexico City—Tulancingo—ATS 3—Goonhilly—Television Centre 525-line chain during the period of the Olympic Games. This signal was found to be invaluable for manually monitoring chain performance—being of course immune to source impairments. It was also very useful for manual regulation of gain and trimming of equalization during programme transmission.

The photographs in Fig. 3 were taken during a swimming programme.

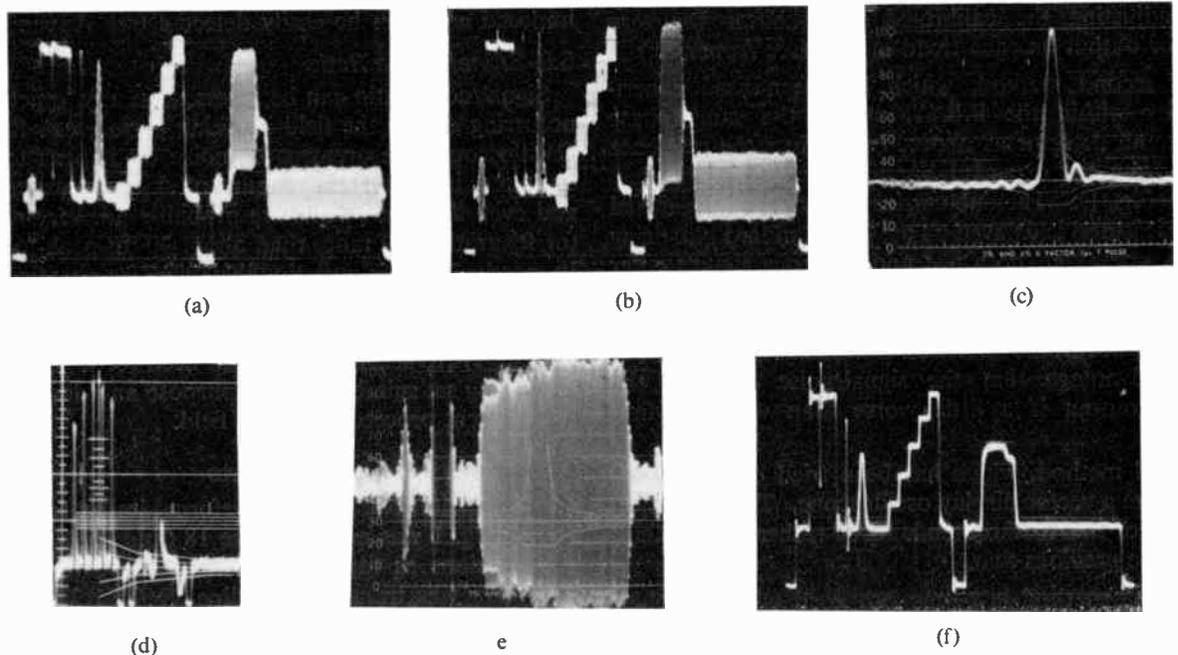
The chain performance after equalization was assessed from these waveforms as being:

luminance <i>K</i> -rating	4%
chrominance–luminance gain inequality	+ 3%
chrominance–luminance delay inequality	+ 10 ns
luminance non-linearity	23%
differential gain	25%
chrominance–luminance crosstalk	+ 10%
signal/unweighted r.m.s. noise ratio	40 dB

7.2 London—Scotland

One of the longest 625-line colour distribution chains at present in service with the B.B.C. is that between London and Angus. This chain contains eight minor local links, five main links, four B.B.C. switching centres, an unattended transmitter and a rebroadcast link receiver. The chain is terminated by the attended Angus transmitter. Manual adjustment of gain, chrominance–luminance gain inequality and chrominance–luminance delay inequality is possible at two of the intermediate switching centres, but such adjustments are normally only made at midday. An automatic gain and chrominance–luminance gain inequality regulator (operating on the i.t.s.) is fitted at the input to the unattended transmitter (Durriss).

The performance of the chain is measured, on an i.t.s. basis, at the output of the final transmitter, via a transmitter demodulator. The limits within which 90% of the results lay during the period July–October 1969 are:



(a) Whole signal unequalized
 (b) Whole signal after equalization
 (c) 2T pulse section after equalization
 (d) Luminance non-linearity display
 (e) Differential gain display
 (f) Chrominance–luminance crosstalk display

Fig. 3. Insertion test signals received from Mexico

luminance <i>K</i> -rating	5%
chrominance-luminance gain inequality	18%
chrominance-luminance delay inequality	65 ns
luminance non-linearity†	20%
differential gain†	18%
differential phase†	12°
chrominance-luminance crosstalk†	10%
signal/unweighted r.m.s. noise ratio	44 dB

† Average picture level \approx 40%

8. Relative Accuracy of Line-by-line and I.T.S. Results

In the preceding Sections attention has been drawn to differences which may occur between measurements obtained by conventional line-by-line methods, and those obtained with insertion test signals, resulting from the difference in the test waveforms. In practice other differences may occur as a result of instrument and operator measuring inaccuracies.

The probable degree of accuracy is particularly difficult to estimate where manual measurement is concerned, since the human element is involved. A carefully controlled series of subjective tests involving a number of operators would be required to arrive at reliable figures for accuracy under practical measuring conditions.

What can be done, however, is to compare the spread of results obtained on a particular chain during a period in which line-by-line test signals were used, with the spread for a similar period after conversion to i.t.s.

Parameter	Short link			Medium length chain		
	90% of spread		I.T.S. Median value	90% of spread		I.T.S. Median value
	Line-by-line	I.T.S.		Line-by-line	I.T.S.	
Luminance gain error in dB	± 0.4	± 0.3	0	± 0.8	± 1.1	-0.1
2 <i>T</i> pulse bar ratio error in %	± 1.5	± 1.0	-1.0	± 5.0	± 7.0	-3.0
Bar <i>K</i> -rating in %	± 0.5	± 0.5	0.5	± 0.8	± 1.7	3.2
C-L gain inequality in %	± 2.0	± 2.0	-1.0	± 9.0	± 13	+6.0
C-L delay inequality in ns	± 5.0	± 5.0	0	± 30	± 30	-40
Differential gain in %	± 0.6	± 0.6	1.0	± 4.0	± 5.0	6.3
Differential phase in degrees	± 0.5	± 0.4	0.7	± 3.5	± 3.2	5.6
C-L crosstalk in %	± 0.5	± 0.5	0	± 1.3	± 1.6	-5.0
Signal/unweighted noise in dB	± 3.0	± 3.5	58	± 3.0	± 2.5	44

techniques. Two sets of results from which such comparisons can be made are shown in the table below. The first set covers fifteen weeks before and fifteen weeks after the introduction of i.t.s. techniques on a minor local link. Because of the high stability of transmission performance inherent in this link, it seems reasonable to assume that, with one exception, these figures represent the accuracy which can be expected in practice for 90% of measurements made on signals suffering very small impairments. The exception is noise of which varying amounts were occurring within the Television Centre itself. The second set covers two ten-week periods on a longer chain, containing two minor local links, a main link, two B.B.C. switching centres, a transmitter and a re-broadcast link receiver. A large amount of each spread in this case is probably attributable to variations in chain performance, but differences between spreads are likely to be indicative of relative measurement accuracies in the presence of reasonable amounts of distortion, including noise.

From the above it would appear that there is very little to choose between line-by-line and i.t.s. measurements as far as accuracy on nearly undistorted signals is concerned, but the differences in spread become more marked when the signals are more distorted.

9. Future Developments

At the present time insertion test signals are added to BBC-1 and 2 625-line transmissions at the output of Television Centre only. It is likely that in future insertion test signal generators will also be installed in all studio centres and outside broadcast vehicles. This proliferation of equipment will raise questions of whether signals should be erased and new ones added *en route*, and if so where. It has been proposed, for instance, that one field should carry a test signal from the programme source, which would enable overall chain performance to be checked, whilst the signal on the alternative field would be replaced at each switching centre, so that individual links could be tested more accurately. In this context it must be remembered that a high standard of transmission performance will be required from the i.t.s. erasers and adders, in order that the test equipment itself shall not significantly distort the television signal.

The test signal itself may be modified in the light of experience, in particular the extended burst may be found to be redundant.

New measurement techniques will undoubtedly be developed. One, which has been tried experimentally, involves instrumenting the test equipment in such a way as to display simultaneously all the parameters listed in Section 6. This display can be helpful when adjusting complex equipment. Alternatively the display can be recorded rapidly on one photograph for subsequent analysis.

10. Conclusions

Insertion test signal techniques are already being used to some extent in many countries. Their use is likely to be extended until the majority of operation and maintenance of national and international television transmission chains is carried out on an i.t.s. basis. The

convenience of these techniques, and their suitability for automation of monitoring and control, should result in improved signal quality and reduced operating costs.

11. Acknowledgments

The authors wish to express their thanks to the Director of Engineering of the B.B.C. for permission to publish this paper, and to their colleagues in the Communications and Transmitter Departments who carried out the measurements.

12. References

1. Shelley, I. J. and Williamson-Noble, G., 'Automatic measurements of insertion test signals', *The Radio and Electronic Engineer*, 41, No. 3, pp. 137-43, March 1971.
2. MacDiarmid, I. F., 'Performance requirements of links for colour television', I.E.E. Conference Publication No. 46, 1968.
3. 'Insertion of Special Signals in the Field-blanking Interval of a TV Signal', Doc. CMTT/1025-E, New Delhi 1970.
4. 'Requirements for the Transmission of Television Signals over Long Distances', C.C.I.R. Recommendation No. 451, Oslo 1966.
5. Weaver, L. E., 'Sine-squared Pulse and Bar Testing in Colour Television', B.B.C. Engineering Division Monograph No. 58, 1965.
6. Savage, D. C., 'The BBC 2 colour television distribution network', I.E.E. Conference Publication No. 46, 1968.

13. Appendix: Waveform Measurement

This Appendix describes the methods at present used to measure the insertion test signals transmitted on BBC-1 and BBC-2.

13.1 Luminance Bar Amplitude ('a' on Fig. 4)

Measured between A1 at the centre of the luminance bar and A2, immediately prior to the chrominance-luminance pulse.

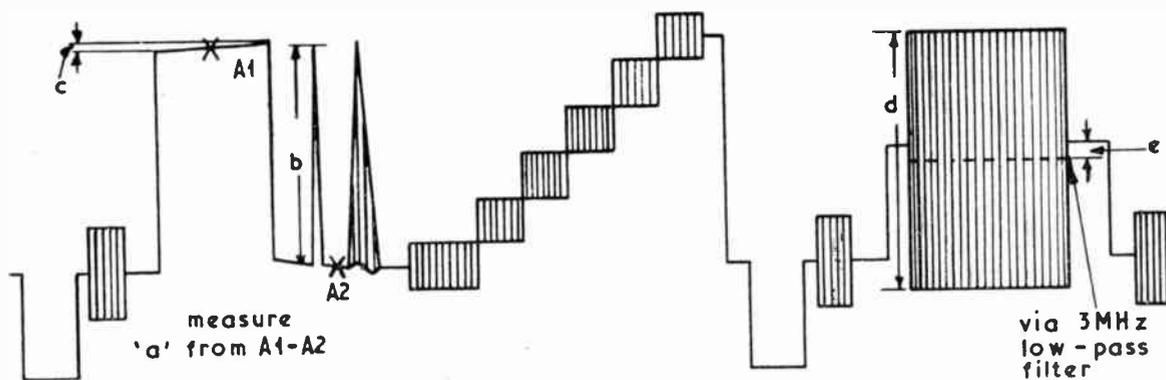


Fig. 4. Part of a distorted U.K. national signal.

13.2 Luminance K-Rating

The bar *K*-rating is measured as the overall tilt on the luminance bar top, neglecting the first and last 0.6 μs, relative to the bar amplitude. ($c/a \times 100\%$ on Fig. 4.) This measurement is assisted by a modification to the standard pulse-and-bar graticule to cater for the shorter bar time, M3 in Fig. 5.

The *2T* pulse *K*-rating is measured using a standard pulse and bar graticule (Fig. 5).

The pulse-to-bar ratio is measured directly as ($b/a \times 100\%$) on Fig. 4.

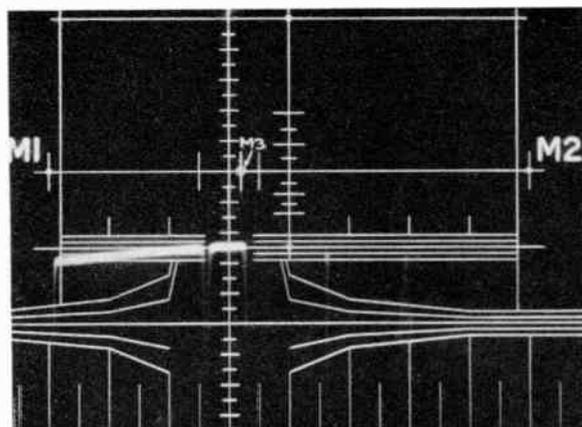


Fig. 5. Distorted 10 μs bar (containing experimental network identification pulse) positioned in modified *K*-rating graticule (*K* bar = 4%).

13.3 Chrominance-Luminance Gain Inequality

This is measured directly as the deviation of the chrominance bar amplitude from the luminance bar amplitude. ($(d/a - 1) \times 100\%$ on Fig. 4.) This is positive if the chrominance is greater than the luminance and negative for the converse.

13.4 Chrominance-Luminance Delay Inequality

This is measured on the chrominance pulse, by cancelling the distortion using a calibrated delay in the luminance path. Lagging chrominance is indicated by a positive sign, leading by a negative sign.

In the absence of the calibrated delay, the measurement may also be estimated as

$$600 \frac{\text{(p-p amplitude of ripple in pulse base)}}{\text{pulse amplitude}} \text{ ns}$$

but this method decreases in accuracy as the chrominance-luminance gain inequality increases

13.5 Luminance Non-linearity

For this measurement, a band-pass filter is used to remove the sub-carrier and low frequency components from the staircase leaving five pulses corresponding to the risers. The measurement is taken as the difference

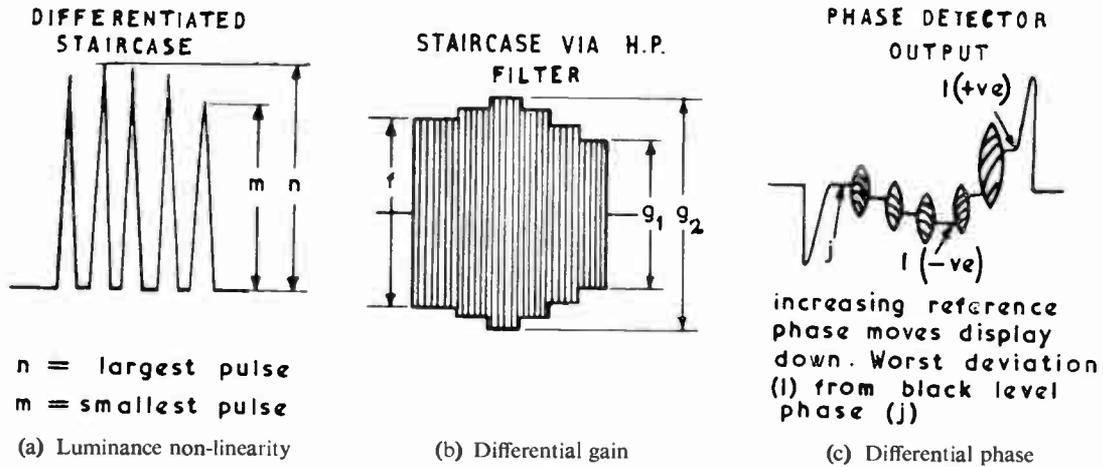


Fig. 6. Measuring equipment output waveforms

between unity and the ratio of the smallest to the largest pulse, as a percentage, i.e. $(1 - m/n)100\%$ in Fig. 6(a).

13.6 Differential Gain

The staircase portion of the waveform is passed through a 4.43 ± 1 MHz band-pass filter to give the display shown in Fig. 6(b). The poorer of the two measurements $(g_1/f - 1)100\%$ or $(g_2/f - 1)100\%$ is used with the sign negative for g_1 or positive for g_2 .

13.7 Differential Phase

The output from a phase detector is displayed as in Fig. 6(c). Using a calibrated phase shifter in the reference feed to the detector, the worst phase variation (e) from the phase corresponding to black level (l) is determined. This measurement is denoted positive for leading and negative for lagging phase relative to black level phase.

13.8 Chrominance-Luminance Crosstalk

The chrominance bar is passed through a 3 MHz low-pass filter and the difference in amplitude between the part from which the sub-carrier has been removed and the short reference pedestals at each end, is measured as a percentage of the luminance bar amplitude. This is $(e/a \times 100\%)$ on Fig. 4. The sign is positive if the sub-carrier mean is above the reference and negative if it is below.

13.9 Signal/Noise Ratio

The signal bandwidth is restricted to 5.8 MHz and the peak-to-peak noise is measured on line 16 relative to the luminance bar amplitude. The result is converted to decibels and 18 dB added to convert this to peak picture/r.m.s. noise unweighted.

Manuscript received by the Institution on 30th January 1970. (Paper No. 1380/Com. 41.)

© The Institution of Electronic and Radio Engineers, 1971

Contributors to this issue



Mr. D. C. Savage received his technical education at Regent Street and Northampton Polytechnics. He joined the B.B.C.'s Lines Department in 1944 and after two years' service in the Royal Electrical and Mechanical Engineers, returned to the Engineering Designs Department of the B.B.C. in 1949. Since then he has been engaged in the design of television transmission apparatus and the development of measurement techniques and performance standards for national and international television circuits.

ment techniques and performance standards for national and international television circuits.



Mr. D. A. Carter (Graduate 1964) obtained his technical education at Twickenham Technical College. He served an electrical engineering apprenticeship with Vickers-Armstrong (Aircraft) Ltd. and, on its completion in 1959, worked as an electrical design draughtsman with the Company. In 1962 he joined the Designs Department of the B.B.C. as a laboratory technician and since 1964 he has held the position of Engineer with the Department. During this latter period he has been associated mainly with the design of transmission measurement equipment.

position of Engineer with the Department. During this latter period he has been associated mainly with the design of transmission measurement equipment.



Dr. G. Barbieri obtained his Doctorate in Electronic Engineering in 1966 from the Turin Polytechnic. He joined the Electronic Research and Development Department of FIAT and worked on the design of digital control systems. Since December 1967 he has been with the RAI Research Laboratory where his work has been mainly concerned with television, in particular telecine, video recording, cameras and other studio techniques.

and other studio techniques.



Dr. P. D'Amato received his Doctorate in Electronic Engineering from the University of Naples in 1966, where he remained as an assistant until 1967. He then joined the RAI Research Laboratory where he is now head of the Television Measurement Department; he is mainly engaged on the development of the insertion test signal technique. Dr. D'Amato has already published five papers on various aspects of television.



Mr. J. E. Morris received the degree of B.Sc. from the University of Auckland in 1965 and the M.Sc. degree in physics in 1967. He then carried out thin films research in the Department of Electrical Engineering of the University of Saskatchewan and since 1969 has been a lecturer in electrical engineering at the University. Mr. Morris has published two papers on inductive effects in tunnel diode switching circuits

and the conduction mechanism in discontinuous thin metal films.



Dr. A. D. Booth (Fellow 1956) has been Dean of the College of Engineering at the University of Saskatchewan since 1963; he also occupies the Interdisciplinary Chair of Autometrics at the Western Reserve University, Cleveland. He graduated with an external B.Sc. degree of the University of London in 1940 and obtained his Ph.D. at the University of Birmingham in 1944; his D.Sc. was granted by the University of London in 1951. Dr. Booth is best known for his work in the design and application of computers. He is the inventor of the magnetic storage drum and he holds basic patents for magnetic structure and multi-core magnetic storage devices. Between 1950 and 1962 Dr. Booth held successively the positions of Director of the Birkbeck College Computer Project, University Reader in Computational Methods, and Head of the Department of Numerical Automation, Birkbeck College. During his professional life Dr. Booth has contributed over 250 papers to scientific and technical journals, four of them in the Institution's *Journal*, and he has written several books, mainly in computer subjects.

Dr. Booth is best known for his work in the design and application of computers. He is the inventor of the magnetic storage drum and he holds basic patents for magnetic structure and multi-core magnetic storage devices. Between 1950 and 1962 Dr. Booth held successively the positions of Director of the Birkbeck College Computer Project, University Reader in Computational Methods, and Head of the Department of Numerical Automation, Birkbeck College. During his professional life Dr. Booth has contributed over 250 papers to scientific and technical journals, four of them in the Institution's *Journal*, and he has written several books, mainly in computer subjects.

Dr. Booth served on the Council of the Institution and was Chairman of the Programme and Papers Committee from 1957 to 1962; he was appointed the first Chairman of the Computer Group Committee in 1959. While in Canada he has been closely associated with the Canadian Divisional Council of which he is Chairman and he has also served a term as one of the overseas Vice-Presidents of the Council during this period.



Mr. J. R. Sanders read for his degree of M.A. in physics at Pembroke College, Oxford. After completion of his National Service with R.E.M.E. as an instructor in electronics, he joined the B.B.C. and served a graduate apprenticeship until 1959; he was attached to the Marconi Company during part of this time. Since then he has been permanently based at the B.B.C. Research Department and has worked on a wide range of projects connected with all aspects of colour television picture origination. Mr. Sanders' present position is assistant to the head of the image scanning section of the Research Department.

wide range of projects connected with all aspects of colour television picture origination. Mr. Sanders' present position is assistant to the head of the image scanning section of the Research Department.