

FOUNDED 1925
INCORPORATED BY
ROYAL CHARTER 1961

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

THE RADIO AND ELECTRONIC ENGINEER

The Journal of the Institution of Electronic and Radio Engineers

VOLUME 39 No. 3

MARCH 1970

Television Measuring Techniques

THERE have been many exciting developments in recent years in the field of television engineering. For example, in the United Kingdom these have included the advent of colour television using the PAL system, the introduction of transmissions at u.h.f. with the engineering of three colour networks in this band, and the increased use of insertion test signals, automatic monitoring and unattended transmitters. These innovations have meant that equipment and circuits are required to operate within very much tighter performance limits than before.

Engineers working in the field of television measurements have succeeded not only in developing the more advanced techniques necessary to ensure that the performance of the equipment and circuits can be measured to these higher standards but also in devising new techniques for the measurement of parameters which have recently acquired considerable importance. These developments have reached the stage at which it seemed appropriate for there to be held a Conference devoted to Television Measuring Techniques.

The Institution of Electronic and Radio Engineers with the association of the Electronics Division of the Institution of Electrical Engineers, the United Kingdom and Republic of Ireland Section of the Institute of Electrical and Electronics Engineers and the Royal Television Society has therefore organized a three-day Conference which will be held at the Middlesex Hospital Medical School, London, W.1, from Monday 11th to Wednesday 13th May inclusive: the outline programme of the Conference is given on pages 137 and 138 of this issue. The Conference covers measuring techniques used at all stages in the progress of the television signal from the camera tube to the domestic receiver and the papers have been arranged in four main Sessions, each of which will include a survey paper.

The Studio Origination Session will include papers on the techniques adopted in the line up of studios, measurement of colour sub-carrier phase, camera tube lag, small order non-linear distortions and video noise in the presence of synchronizing signals, also colour monitor calibration. In the Networks Session a number of papers are devoted to insertion test signal techniques and automatic measurement and there will be a paper on the additive effects of degradation. The Receivers Session covers various aspects of measurements on television monitors and receivers. The Transmitters and Aerials Session includes papers on u.h.f. klystron transmitters, u.h.f. television transposers, and aerial, feeder and field strength measurements.

In addition to the papers there will be a combined sherry party and visit to the I.T.A. Television Gallery and a visit, which unfortunately must be restricted in numbers, to the B.B.C. Television Centre.

The Middlesex Hospital Medical School is an ideal location for a Conference of this type being situated in the West End of London and adjacent to the Post Office Tower. A small exhibition of measuring equipment will be presented in an area adjoining the lecture theatre and the Conference programme has been arranged so that there will be ample time for examining the exhibits.

The Conference will provide an all too rare opportunity for engineers working in comparable fields to meet and exchange ideas. Great interest has been shown both in the United Kingdom and overseas—papers have been received from the U.S.A., Germany, Italy and Denmark—and there is every indication that they will be supported by lively and stimulating discussions.

R. LARRY

INSTITUTION NOTICES

Institution Dinner, 1970

The Guest of Honour at the Institution Dinner in London on Tuesday, 12th May next, will be His Royal Highness The Prince Philip, Duke of Edinburgh. The Dinner will take place in the Egyptian Hall of the Mansion House in the City of London in the presence of the Right Honorable the Lord Mayor of London, Sir Ian Frank Bowater, D.S.O., T.D. Other official guests will include representatives of government, the universities and industry.

Members of the Institution intending to take part in what promises a memorable evening are urged to complete, and return to the Institution without delay, the application form which is enclosed in this issue. Tickets for members' personal guests will have to be strictly limited in number, and attention is particularly drawn to the arrangements in this connexion stated on the application form.

Conference on Automatic Test Systems

Present indications are that the Conference on 'Automatic Test Systems' to be held at the University of Birmingham from 14th to 17th April, will attract a capacity attendance; a provisional programme of the Conference was published in the January issue of the *Journal*. Applications for registration forms should therefore be made without delay either by telephone to the Institution or using the tear-out forms at the back of this issue.

4th Annual Solid State Devices Conference

The Institute of Physics and The Physical Society in collaboration with the Institution of Electrical Engineers, The Institution of Electronic and Radio Engineers and the Institute of Electrical and Electronic Engineers, United Kingdom and Republic of Ireland Section, is arranging a fourth Conference on Solid State Devices from 15th–18th September 1970 to be held at the University of Exeter.

The Conference will follow the established general pattern, with the aim of providing a forum for the presentation of new work in the areas of applied solid state physics, device production and characterization, together with their associated technologies. Leading workers are being invited to present keynote and invited papers in a wide range of topics.

The programme committee also invites offers of contributed papers of either 20 minutes or 10 minutes presentation time, including discussion, in this general subject area, and would particularly welcome contributions on the following.

Electroluminescence—Light sources and displays
Photon and particle detectors—including arrays
Semiconductor transducers

Electron emission from semiconductors
Device physics
Device modelling
Effects of material defects on device yield and performance.
Microwave amplifiers and oscillators
Semiconductor memories—conventional and unconventional

Papers on materials in the specific context of device applications will also be welcomed.

The Chairman of the organizing committee is Professor D. H. Roberts (Plessey Company) and the Chairman of the programme committee is Professor E. H. Roderick (UMIST). The I.E.R.E. is represented on the organizing committee by Professor J. C. Anderson (Member) (Imperial College).

Abstracts, of about 400 words, for contributed papers, typed in double spacing on one side of A4 paper should be submitted in duplicate to the Conference Secretary, Dr. D. C. Northrop, University of Manchester Institute of Science and Technology, Sackville Street, Manchester M60 1QD, to arrive on or before Friday, 19th June 1970. Abstracts should not include figures and should be suitable for photocopying.

Residential accommodation will be available in the University Halls of Residence. Further details and application forms will be available from the Meetings Officer, The Institute of Physics and the Physical Society, 47 Belgrave Square, London, S.W.1, in May/June 1970.

Nuclear Technology in the U.S.S.R.

The Annual General Meeting of the British Nuclear Society will take place on Thursday, 30th April 1970, at 5.30 p.m. at the Institution of Civil Engineers, 1–7 Great George Street, London, S.W.1. It will be followed at 5.45 p.m. by a lecture on 'Some aspects of nuclear development in the U.S.S.R.', by Victor Rojkov (Technical College of Physics and Engineering, Moscow). Admission is by ticket only (for both members and non-members), and applications should be made to: The Secretary, B.N.E.S., 1–7 Great George Street, Westminster, London, S.W.1. (Telephone: 01-839 3611)

Postcode for the Institution's Headquarters

The postcode **WC1B 3RG** should now be used when writing to the Institution at 9 Bedford Square, London. The postcode should be typed or written, with the spacing shown, after 'London', either on the next line or placed well away, and not underlined. The object of postcodes is to enable the British Post Office to implement mechanized postal sorting techniques.

An Experimental Laser-Photochromic Display System

By

GEOFFREY G. FULLER†

B.Sc.

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

A laser display system is described which can produce real-time television standard pictures, or alternatively reversible, stored images using photochromic film. The development of components for the modulation of laser beams is also described. The spatial modulators are acousto-optic diffraction cells, and the amplitude modulator uses the electro-optic effect in ADP.

1. Introduction

In the field of real-time, store or non-store display, most requirements are currently met by the ubiquitous cathode-ray tube in its many forms. However it is recognized that besides being a fragile device because of its large vacuum envelope, the c.r.t. has certain characteristics which limit its performance, particularly in applications requiring bright large-screen displays, a reliable reversible storage capability or very high resolution. There is a need, therefore, for new display technologies which remove these limitations and it is felt that the laser could find potential application in this field. However, in order to realize this application, and indeed other suggested applications of data storage and communication, it is necessary first to develop the laser beam control technologies of high-speed temporal and spatial modulation. This laboratory has therefore developed such devices and applied them in an experimental display system which, although not purporting to reach high performance limits, does serve to demonstrate the principles of a laser display and enable a realistic assessment to be made of the potential usefulness.

For spatial modulation, there are three basic technique classes; mechanical, electro- or magneto-optical interaction, and acousto-optic interaction. Mechanical deflectors have well-defined performance limits and are mainly applicable to cyclic scan. Electrostatic or magnetic fields can induce refraction changes in an optically-active material which can be used directly to give analogue deflexion, usually with very low resolution (a materials limitation), or indirectly by inducing polarization changes and converting these into digital deflexion through a polarization analyser. The latter effect can exhibit geometrically precise positional accuracy, but since each element only produces a binary output, a deflector having a practical resolution requires many series elements, resulting in a severe compromise between speed, resolution and

optical efficiency. Acousto-optic deflexion has been chosen for the system since although there is a speed/resolution compromise, high efficiency operation is possible, and the current materials situation less restrictive.

The same three basic class divisions apply equally in the case of extra-cavity amplitude modulators but the performance criteria are now different, the emphasis being on speed of operation. Mechanical devices again are limited in speed unless used cyclically. Electro- or magneto-optic devices can have high speed and efficiency, while acousto-optic modulators are generally limited in speed by transit time effects. The field of electro- and magneto-optic modulators is very broad and well documented, the main selection parameters being optical efficiency, contrast ratio and ease of driving circuitry. A good solution from the point of view of the experimental display system is provided by an electro-optic modulator first described by Dore,¹ which has been engineered to interface with the experimental display system.

2. Acousto-optic Deflexion

2.1. Principles

The basis of beam deflexion using acoustically generated refractive index changes is well documented,^{2,3} the principle being shown in Fig. 1. A plane acoustic wave is launched through a medium, orthogonally to a parallel beam of light. The acoustic wave generates periodic refractive index variations in the medium, which are 'seen' by the light beam as a phase grating, and result in diffraction of the output beam. If the angles between a first-order diffracted beam, the input (zero-order) beam and the acoustic beam are adjusted so that the wave vectors have zero sum, only one first-order diffracted beam results, a condition analogous to that of Bragg diffraction of X-rays in crystals. For a fixed input beam angle, the output (first-order diffracted) beam may then be steered by suitably changing the acoustic vector, i.e.

† Space and Weapons Research Laboratory, Elliott Brothers (London) Ltd., Frimley, Camberley, Surrey.

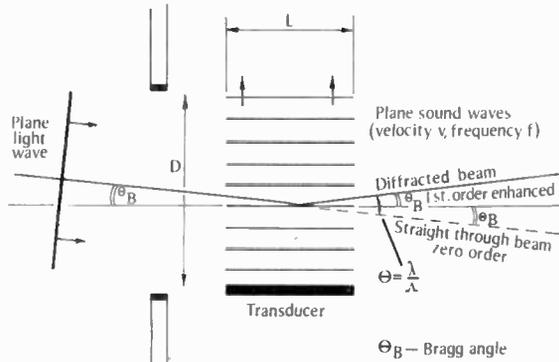


Fig. 1. Bragg diffraction.

the acoustic frequency and its angle of propagation. In practice it is possible to achieve beam steering over reasonable resolutions by simply changing the acoustic frequency whilst retaining fixed angular conditions. By using crystal anisotropy⁴ it is also possible to achieve 'automatic' angle tracking with frequency.

2.2. Design Criteria

The usefulness of a diffraction cell as a deflector is expressed in terms of its number of 'resolvable positions' *N*, that is, the angular sweep of the output beam divided by its angular spread, the latter being determined by the diffraction limit of the optical aperture, width *D*. This can be expressed as

$$N = \tau \Delta f \dots\dots(1)$$

where τ is the transit time of an acoustic wave across the cell aperture, and Δf the acoustic frequency sweep.

For the case of a cell not having a steered acoustic beam, there is a limit to the acousto-optic interaction length, *L*, for a given frequency sweep which can be defined by not allowing the light intensity to fall by more than 1 dB at the scan edges. This condition occurs when the deflexion angle varies from the true Bragg angle by greater than $\delta\theta$ where

$$\delta\theta = \frac{V}{4Lf}$$

where *V* is the acoustic wave velocity. Expressing this in terms of Δf and f_0 we have

$$L = \frac{4V^2f_0}{\lambda\Delta f(4f_0^2 - \Delta f^2)} \dots\dots(2)$$

A minimum limit restraint is imposed upon *L* by consideration of the acoustic power density *P_a* needed for complete diffraction; this is given,³ using standard symbology, by

$$P_a = \frac{9\rho V^3\lambda^2}{2L^2(n^3 + n - 2/n)^2} \dots\dots(3)$$

Since the power density available from practical transducers is limited by conditions of electrical breakdown, heating or mechanical damage, *L* has a defined minimum value for a given transducer system.

The aperture of the cell *D* is defined by eqn. (1) (since $\tau = D/V$) and by the fact that the acoustic power is attenuated as it travels across the aperture. If we consider the case of diffracting media which exhibit a loss coefficient α proportional to f^2 (typical of most liquids and of some solids over certain frequency ranges) and permit a 3 dB loss across the aperture at maximum frequency, it can be shown that

$$D_{min} = \frac{N^2V^2\alpha(2f_0 + \Delta f)}{\Delta f \log_e 2} \dots\dots(4)$$

Equations (1) to (4) form the basic design criteria for diffractor cells, and their application will be outlined in Section 2.3.

2.3. Practical Cell Development

The use of the design criteria in the previous section can be demonstrated by reference to two specific cell designs.

The first of these was designed for a system requiring a resolution of not less than 100 at an optical wavelength of 1.06 μ m, and not less than 3 dB intensity fall off at the scan edges. The diffracting medium was to be deuterium oxide (heavy water) since ordinary water is less transparent at this wavelength. For this system the cell transit time was not significant so that the aperture was made large, thus avoiding the use of a steered acoustic beam for such low resolution. With the aperture calculated at 50 mm (a figure less than the theoretically permitted maximum but limited by the manufacturing difficulties of cylinder lenses), the frequency sweep was then given as 2.8 MHz. For the practical reasons of drive circuit design and efficiency the acoustic power was generated using lead zirconate titanate ferroelectric ceramic transducers bonded on to a coupling prism of fused silica; this acoustic system has an acoustic bandwidth of 25% to 1.5 dB, thus fixing the cell centre frequency at 16 MHz. The bandwidth and centre frequency then fix the maximum sound/light interaction length if the Bragg condition is to be adequately met.

The ultrasonic power can thus be calculated and in order to reduce this to a minimum, the optical aperture was reduced to a thin slot shape in a direction normal to the diffraction plane by the use of cylindrical optics.

The completed cell was duly evaluated and with the exception of some degradation of resolution caused by imperfect cylindrical lenses, the device functions in accord with the design predictions and exhibits an optical efficiency (input light to diffracted output light) of about 70%.

The deflector is one of a pair which are incorporated into a practical field equipment which can produce a digitally-controlled raster scan of 100×100 resolution over a $5^\circ \times 5^\circ$ sweep at an overall optical efficiency of about 18%.

The second cell to be described was for the experimental display system and its specification was therefore determined by television 405-line broadcast standards. This includes a maximum transit time (for flyback) of $16 \mu\text{s}$ and a resolution of 405.

For the particular case of a linear time-varying (sawtooth) frequency sweep, the effective time of scan from one spot position to the next can be made very small compared to the cell transit time, the only effect upon the diffracted output beam being to produce an effective astigmatism which can be corrected by means of a fixed optic. However, since a sawtooth is a regenerative waveform, time must be allowed during the flyback period to allow the acoustic swept wave to leave the cell aperture and be replaced by a new acoustic swept wave; this requirement modifies the basic resolution expression to:

$$N = \tau \Delta f (1 - t/P)$$

where P is the sweep period. For the television case, t is the flyback period of $16 \mu\text{s}$ and P is the line scan time $98.8 \mu\text{s}$, fixing a frequency sweep of 31 MHz.

Again for reasons of practicality, a lead zirconate titanate/fused silica transducer combination was chosen with a 3 dB mechanical bandwidth of 42%, i.e. a minimum centre frequency of 41 MHz.

From eqn. (2) a maximum interaction length of 3.8 mm is given but this results in an unacceptably high ultrasonic power density (eqn. (3)). In order to overcome the maximum limitation, the cell was designed so that the acoustic beam was steered with changing frequency. A convenient method of achieving this is to divide a wide acoustic wave front into a number of discrete sections, and produce a predetermined phase relationship between the sections, either electronically or mechanically. The precision

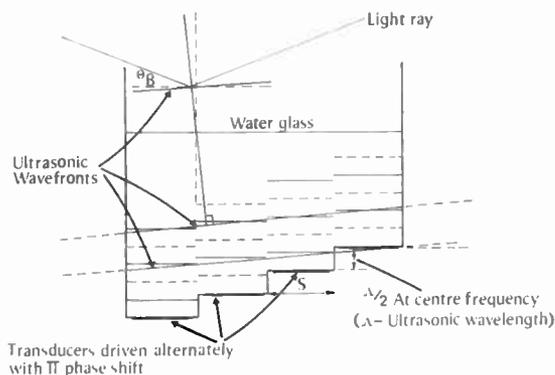


Fig. 2. Phased array steps.

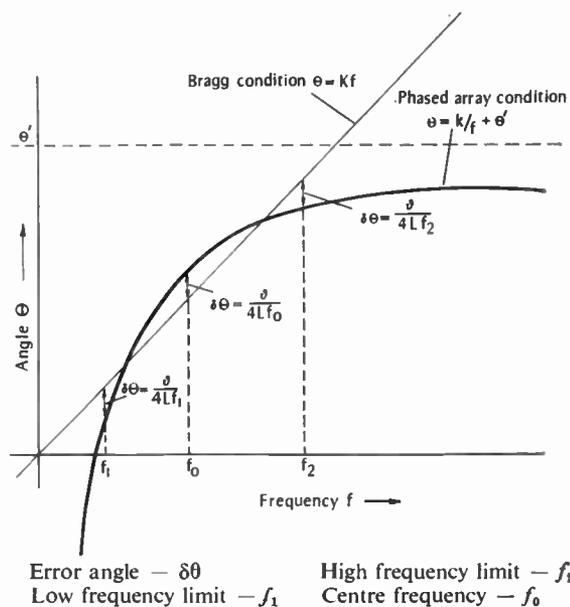


Fig. 3. Phased array beam steering conditions.

of phasing makes the implementation of electronic systems very difficult, so that a mechanical system was chosen where the transducer sections are mounted in a staggered fashion, shown in Fig. 2. This results in a characteristic where the acoustic beam angle varies as $1/f$, the required Bragg condition being a linear f dependence as shown in Fig. 3. The two characteristics can be matched over a frequency range which is defined by the points where they differ by the error angle 'criterion' given in Section 2.2 (see again Fig. 3).

Since the operating bandwidth and frequency sweep are already determined, for minimum frequency operation the frequency limits also become fixed (points f_1 and f_2 in Fig. 3). The shape of the $1/f$ characteristic is determined by the aperture (transducer) width, S , of each section of the phased array; when the error condition is related to this shape, equations result relating the total array length, and the section width, and having a unique solution.

For the particular case of this cell the resultant solutions were

- step width 3.8 mm
- array length 40 mm

As in the previous cell, the power requirement was reduced by the use of a cylindrical optical arrangement resulting in a transducer height of 2 mm, and approximate power input of 350 mW.

The cell constructed in accordance with the foregoing design procedure is shown in Fig. 4. The optical efficiency at the centre of the scan is 12% and the scan intensity falls off by approximately 3 dB at the lower frequency end, but by some 6 dB at the high frequency.

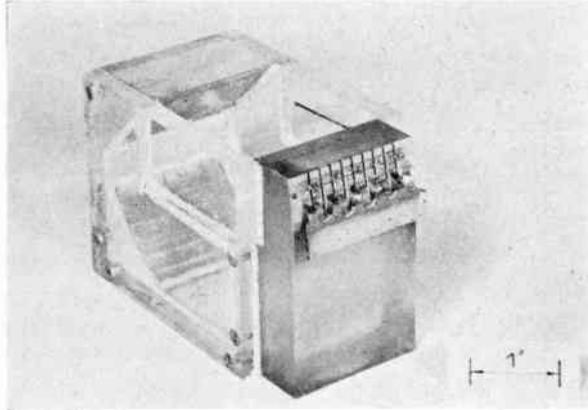


Fig. 4. Phased array cell.

The increased loss is in part due to the rapidly increasing ultrasonic attenuation at these frequencies and to the greater significance of geometrical imperfections in the manufacture of the array at short wavelengths. A resolution of between 300 and 350 was recorded, but again some degradation was due to imperfections in the ophthalmic quality cylinder lenses.

3. Amplitude Modulator

The outline specification for the amplitude modulator for use in the experimental system was that it should have a bandwidth from d.c. to 3.5 MHz, be capable of being driven continuously at 100% modulation depth with relatively uncomplicated electronic circuits, have a good extinction ratio (at least 25 to 1) and high optical efficiency.

An electro-optic modulator developed by Dore¹ is shown to meet the outline specification and additionally can be made from relatively cheap and readily available material. An engineered model has now been produced by this laboratory.

The modulator uses the transverse r_{41} electro-optic coefficient in ADP (ammonium dihydrogen phosphate) and is arranged as shown in Fig. 5. Two crystals of ADP are used, the first 45° X cut and the second 45° Y cut, separated by a natural quartz half-wave plate. This configuration has the advantages of a net addition of electrically induced phase changes for a common applied field direction in both crystals, but cancellation of natural birefringence and of phase shifts produced by temperature induced path length changes, provided that both crystals are maintained at an equal temperature. Since the electric field direction is common to the two crystals, the temperature condition can be easily met by mounting both crystals on a thermally massive metallic plane, which also serves as the ground electrode.

The effect of a changing voltage in the unit on an input beam I_0 , plane polarized at 45° to the input face is to change its polarization through elliptical to plane polarized orthogonally to the original plane. If this beam is analysed by a polarizer, the intensity of the output beam I is given by

$$I = I_0 \sin^2 \frac{\pi}{2} \left(\frac{V}{V_{\pi}} \right)$$

where V_{π} is the units' half-wave voltage.

The engineered version of the modulator uses crystals 2 mm thick, 3 mm wide and 50 mm long, these dimensions representing a compromise between maximizing the length to thickness ratio (thus minimizing the half-wave voltage) and the restrictions imposed by mechanical handling problems. The analyser is a Nicol prism, mounted integrally, and with a bloomed fused silica window to protect the input face, the whole assembly is cemented together to form a single optical unit. The measured half-wave voltage is 190 V at 0.488 μm , the transmission approximately 70% and extinction ratio greater than 25 to 1.

4. Frame Scan Unit

The frame (vertical) scan unit for the experimental system could have been another acoustic-optic cell, but it was felt that the additional optical and electronic complexity was not justified for a unit that was required to produce a sawtooth scan at only 50 Hz. Of the various mechanical methods of deflexion, those based on rotating mirrors or prisms were found to be optically incompatible (in a practical implementation sense) with the acousto-optic scanner, which produces a fundamental angular scan of 16 minutes of arc; recourse was therefore made to a piezo-electric system.

The unit consists of a pair of parallel-connexion bimorph bender strips carrying flat aluminized

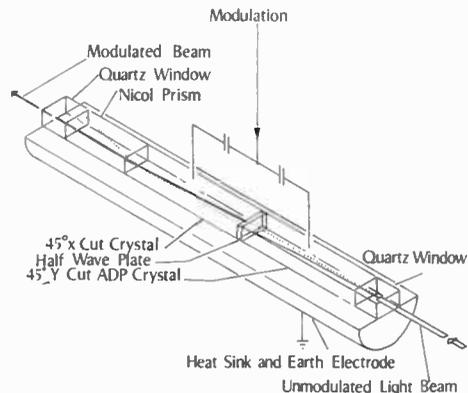


Fig. 5. Amplitude modulator schematic.

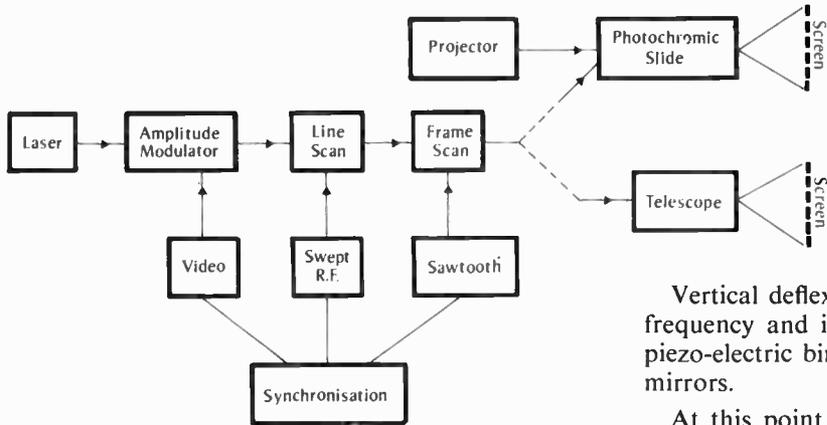


Fig. 6. System block diagram.

mirrors, and driven in phase. This configuration produces additive angular deflexion but cancellation of the positional deflexion normally resulting from single bender deflectors. The high electro-mechanical coupling coefficient for lead zirconate in conjunction with a low electrical output drive impedance tends to damp out ringing but it was found, in practice, that additional mechanical (viscous) damping was required.

Although the fundamental resonance of the benders was chosen to be compatible with the maximum frequencies in the sawtooth spectrum, the unit was deficient in high frequency response and resulted therefore in a 'folding over' of the top part of the television raster.

5. The Laser Source

The light source has to produce a uniphase wave-front of high continuous power which has an output wavelength suitable for the activation of available photochromic materials. At the present time only the argon ion laser meets this requirement and a tube has therefore been developed for the system. The tube has a 'long' cavity configuration resulting in a maximum single transverse mode output of 250 mW, single line selection being by an intra-cavity Littrow prism. The unsuitability of this type of high current plasma tube in any practical system is appreciated but the portents for the development of high efficiency frequency doubling methods for application to solid-state lasers operating at 1.06 μm are good.

6. The Display System

A block diagram of the experimental display system is shown in Fig. 6. The laser source is the argon laser described in Section 5.

Amplitude modulation is impressed on the beam by means of the electro-optic modulator, horizontal deflexion being added by means of the acousto-optic cell.

Vertical deflexion is only required at relatively low frequency and is performed by the pair of matched piezo-electric bimorph benders carrying front surface mirrors.

At this point in the system the information to be displayed appears as a modulated raster scan approximately 2 x 3 mm square. There is then the option to project this image directly on to a screen at any required magnification to be viewed as a real-time picture, or to allow it to store on a photochromic slide, the stored image being projected as in a normal slide projector with condensed, filtered white light and a projection lens.

The three modulator units have been driven directly from a television broadcast receiver chassis, from a television camera, and from a c.r.t. graphic display joystick controller.

6.1. The Optical Interface

The system optical assembly is shown in Fig. 7. The beam leaving the laser is approximately 3 mm in diameter. This is collimated into a 0.5 mm beam, passed through the amplitude modulator and then re-collimated by a reverse telescope to a diameter of 40 mm equivalent to the line scanner aperture. A pair of confocal cylindrical lenses causes the beam to come to a line focus in the line scan cell and be re-collimated at the output. The first order diffracted beam is brought to a focus, the frame scan unit being positioned in the convergent part of this beam, so that a full raster scan is produced at the focus, with

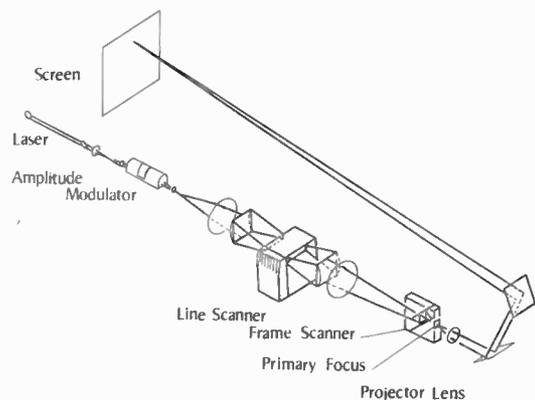


Fig. 7. System optical assembly.

a size approximately 3.0×2.5 mm. The effect of astigmatism produced by having a linear frequency sweep in the line scan cell is removed by readjustment of the cylindrical lens positioned near the frame scan unit. In the direct display mode a short focus lens projects the raster scan on to a back projection screen to any desired magnification by interchanging lenses or altering the projection throw.

All the optical components are anti-reflexion coated at $0.488 \mu\text{m}$. The short focus lenses are microscope objectives. The larger lenses are cemented achromats and the cylindrical lenses plano-convex, both working at approximately $f/12$.

6.2. The Photochromic System

When operating in the storage mode a photochromic film is placed at the primary system focus so that the display is recorded in an analogue fashion on the film. Simultaneously with the recording, the image on the film is projected on to a screen as is a colour slide in a conventional slide projector. The arrangement of the photochromic slide and projection optics is shown in Fig. 8. Just before the primary focus, the scanning laser beam is deflected through 90° by a small surface mirror. The photochromic slide is also placed near the focus of the quartz-iodine lamp condenser system, the mirror being small enough not to interfere seriously with the evenness of the projector illumination. Together with the condenser lens is the filter which matches the photochromic unactivated transmission spectrum, in this case a Schott OG2 orange. The photochromic film currently being used is Cyanamid type 63-071 which has absorption characteristics shown in Fig. 9. When projected with OG2 filtered incandescent tungsten light, an energy input of approximately 1 joule/cm^2 is required to produce an effective image density of unity (contrast ratio 10), the image appearing to the viewer as a black trace on a bright yellow/orange background.

6.3. System Performance

In the direct projection mode the system can be made to produce real-time broadcast television pictures, as shown in Fig. 10. Although notable picture degradation occurs due to the frame scan deficiencies

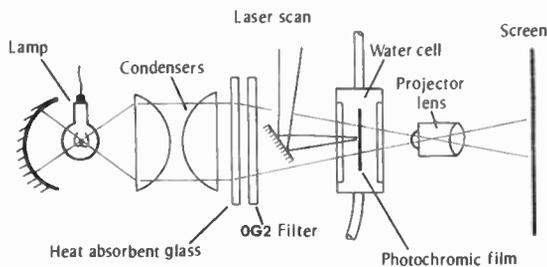


Fig. 8. Photochromic projector optics.

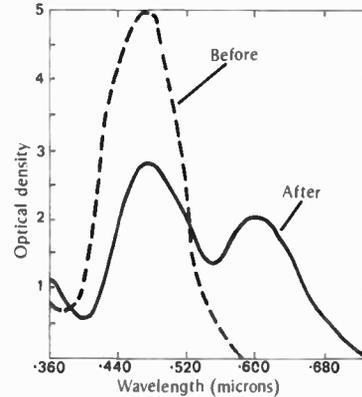


Fig. 9. Absorption spectra of photochromic film.

(cf. Sect. 4) and to severe radio frequency interference between the line cell driver and the broadcast receiver, the picture approaches the required standard, as evidenced by the 1.5 MHz bar resolution on test card D.

Most notable however is practical demonstration of the inadequacy of display brightness when attempting to use the laser light directly as the illumination. This is for two reasons, firstly the limited power levels that are currently available from c.w. visible light lasers, and secondly from the nett low transmission of any optical complex containing a number of laser modulators, typically between 1% and 10%.

However, the use of the lasers' unique characteristic of being able to be focused to high power densities has enabled the photochromic system to be realized. With the relatively simple projection system using a 100 W lamp and $f/2$ condensers, a 75×90 cm ($2\frac{1}{2} \times 3$ ft) picture has been obtained on a back projection screen at a brightness of 50 ft lamberts (538 lumens/m^2). The present laser system transmission loss limits the writing rate of a complete television frame to several seconds, but when the display is used in a line or

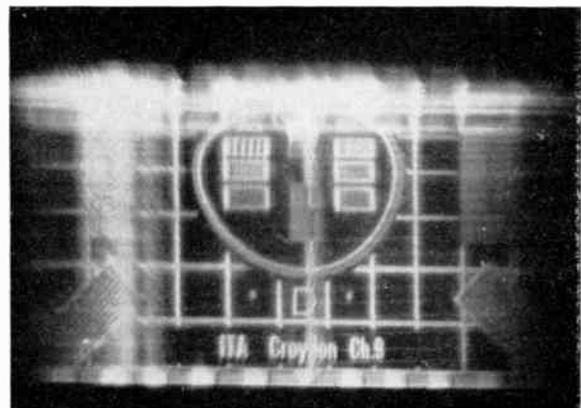


Fig. 10. Photograph of direct television picture.

character drawing mode where the effective screen information cover is reduced by a factor of 10^2 , full frame writing can be achieved in real time.

7. Conclusion

The construction and operation of this experimental display has been helpful in enabling a real assessment to be made of the current and near future potential of acousto-optical deflectors and the role of lasers in display technology.

It is felt that the potential of direct laser light displays for large screen viewing is not good but that coupled with a light amplifier system, e.g. a photochromic or other imaging light valve technique, a powerful display technology could be developed. In addition to display applications, the demonstrated ability to deflect and modulate laser beams at high random speed and efficiency should aid the development of other proposed laser system applications, for

example holographic and magneto-optic stores, servo beam tracking and etc.

8. References

1. Dore, M., 'A low drive power light modulator using a readily available material ADP', *I.E.E.E. J. Quantum Electronics*, QE-3, No. 11, pp. 555-60, November 1967.
2. Willard, G. W., 'Criteria for normal and abnormal ultrasonic light diffraction effects', *J. Acoust. Soc. Amer.*, 21, No. 2, pp. 101-8, March 1949.
3. Korpel, A., *et al.*, 'A television display using acoustic deflection and modulation of coherent light', *Proc. Inst. Elect. Electronics Engrs*, 54, No. 10, pp. 1429-37, October 1966.
4. Lean, E. G. H., *et al.*, 'Continuous deflection of laser beams', *Applied Physics Letters*, 10, No. 2, pp. 48-57, 15th January 1967.

Manuscript received by the Institution on 4th March, 1969. (Paper No. 1305/CC65).

© The Institution of Electronic and Radio Engineers, 1970

The Author



Geoffrey G. Fuller obtained his B.Sc. degree in physics from the University of Bristol. He then spent three years with Standard Telephones and Cables Ltd., where he was concerned with design and development of travelling wave tubes. He joined Elliott Automation Ltd. at Rochester to lead design and development projects on miniature radar systems. After four years he

transferred to the research laboratory of Elliott Space & Weapons Systems Ltd. at Frimley to work on optoelectronics and lasers and he is now research projects manager with responsibilities which have extended to include infra-red systems and electric propulsion for satellites.

An Optical Fixed Data Store

By

R. J. BOTFIELD†

and

A. N. HILL, B.Sc.(Eng.), A.K.C.†

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

An optical fixed data storage system is described which is designed to hold 4 M bits of information on each exchangeable photographic plate. The store, intended as an integral part of a computing system, enables any of 65,536 69-bit words to be retrieved with a random access time of a few microseconds. Writing time for each data plate is approximately 1 hour. To record and retrieve the data, a spot is positioned on the face of a high resolution electrostatic cathode ray tube, and multi-imaged on to a photographic plate by means of a lens/square cross-section mirror tunnel/lens system. Word readout is achieved using one 11-stage photomultiplier as detector for each bit of the word.

1. Introduction

In many data processing systems there is frequent reference to information which is rarely changed. This information is typically held in a storage subsystem having a large random access time but a high data transfer rate, from which blocks are transferred as required to a rapid random access store.

It was decided to develop a store providing rapid random access to large quantities of fixed information, and having an appreciable cost advantage over core stores. The store to be described is designed for this purpose with the following goals:¹

- (i) 64k words of storage;
- (ii) each word at least 50 bits;
- (iii) random word access time a few microseconds;
- (iv) information recorded in one such store to be retrievable in any other similar store.

It was further decided that the storage medium should be a photographic emulsion since their limiting resolution is upwards of 200 lines per millimetre. A photosensitive emulsion-coated glass plate of quite reasonable dimensions can be used to store the required number of bits of information, the data being retrieved by focusing on the plate spots of light whose positions are accurately controlled.

2. Evolution of Store Design

We shall now briefly discuss the progress of the optical store to its present form, and then examine more fully the performance required of each major component in order to achieve the design goal.

Clearly data can be recorded in a photographic emulsion either by incoherent or coherent radiation, and may, for example, be in the form of discrete spots or interference patterns. In view of the target speed,

† Research and Advanced Development Organisation, International Computers Limited, Stevenage, Hertfordshire.

cost and capacity for the store, a cathode-ray tube (c.r.t.) light source was chosen with its correspondingly well-proven deflexion techniques. Discrete spots on the processed photographic plate represent stored bits and light passing through the plate is detected by means of photomultiplier tubes (p.m.t.).

To achieve the desired word length an optical system is required which produces at least 50 images of the c.r.t. face on the photographic information plate. This can be done using many separate lens systems², but a single lens system with a square cross-section mirror tunnel³ was chosen to avoid the problems associated with mounting and adjusting a large number of components. Multiple reflexions within this tunnel are imaged by one lens (the projection lens) on to the information plate (see Fig. 1). The used semi-angle of the projection lens and the number of reflexions needed in the tunnel together fix the length to cross-section ratio (aspect ratio) of this component. A solid glass tunnel could be used, but would of course need to be longer and more absorptive than a hollow one.

Having chosen the projection lens semi-angle and fixed the tunnel aspect ratio, the scale of the photographic plate is determined by availability of suitably sensitive p.m.t.s, the minimum diameter for these tubes fixing the pitch from one image area on the information plate to the next. Further an examination of the probably achievable c.r.t. spot size makes it clear that 64k spots in a matrix of 256×256 could not be provided by an electrostatic c.r.t. within a sufficiently small area to result in a tunnel of reasonable length. Thus it becomes necessary to form a reduced image of the c.r.t. face within the end of the tunnel by means of a second lens (the minifying lens). The resulting physical configuration of the store is then as shown in Fig. 2, which also illustrates the addressing or spot positioning optics, yet to be described.

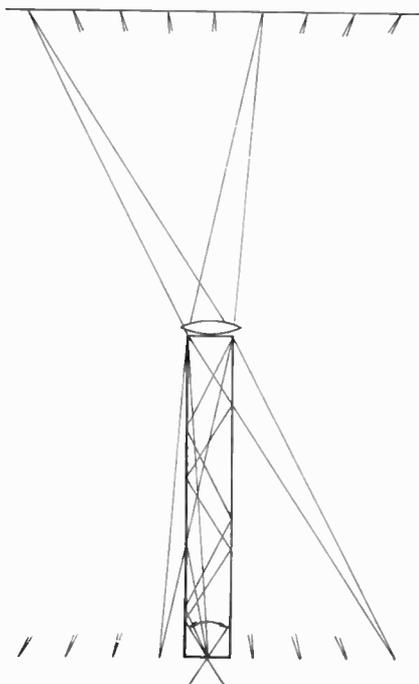


Fig. 1. Mirror tunnel.

Each of the 64k defined positions of the light spot on the c.r.t. face represents one computer address: by multiple reflexions in the mirror tunnel a single spot is imaged as more than 50 precisely positioned spots, one within each bit area on the photographic plate as shown in Fig. 1. During the writing of information only one bit area at a time is unmasked to allow exposure of the emulsion, this being achieved by a multi-position mask. After processing the plate all 50 focused images of the c.r.t. light spot are allowed to impinge on the plate during reading cycles: the many p.m.t. outputs provide the parallel read-out of 50 bits.

Since the light spot on the c.r.t. face must be positioned within limits small compared with the pitch between adjacent spots, it is clear that servo positioning of the c.r.t. spot is necessary. The positional feedback is obtained by imaging the c.r.t. light spot on a series of opaque/transparent coded plates. By use of a separate such arrangement for each c.r.t. deflexion axis the spot address can be deduced for comparison with the address requirement: necessary corrections are calculated and applied. A large proportion of the light emitted by the c.r.t. phosphor is not collected by the minifying lens. The address servo optics use some of this: the servo p.m.t.s receive more light than the data reading p.m.t.s and can therefore operate at a higher signal to noise ratio, or at a higher speed.

The data reading and address servo mechanical components must be held stably and precisely in a framework with adjustments provided as necessary.

Individual system components will now be more fully described.

3. Performance of Major Components

3.1. Photomultiplier Tubes

These components are required to operate as data bit detectors and as address servo light detectors: the former application is examined as it is the more critical due to lower light levels. These p.m.t.s are required to produce a signal of adequate signal/noise ratio after integration of their outputs for $1 \mu\text{s}$. This requires that the uncertainty in photocathode electron emission should be small compared with the charge accumulated in this time: other factors of course influence the noise level, but a current of 30 pA from the photocathode is the minimum acceptable due to the stated consideration. As high as possible a photocathode sensitivity is clearly essential and it was decided to select for high sensitivity from available

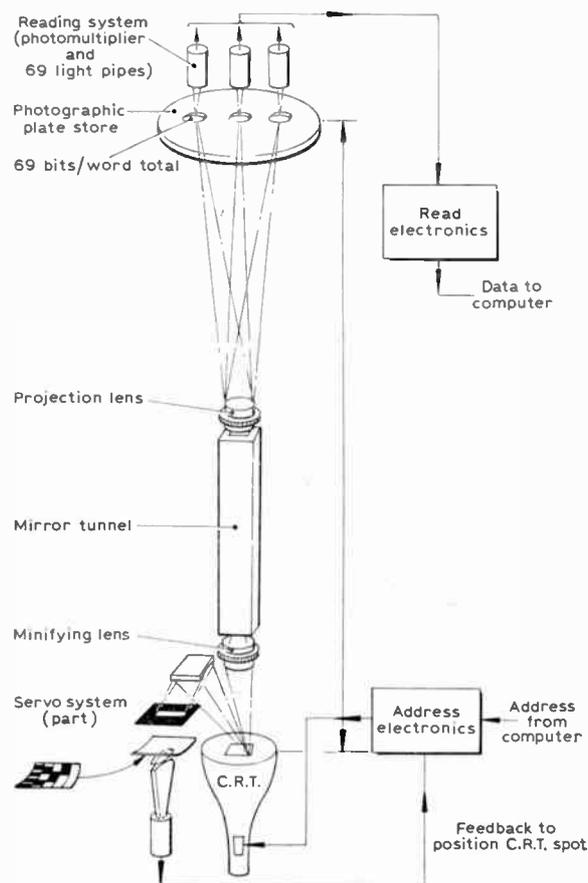


Fig. 2. Physical configuration.

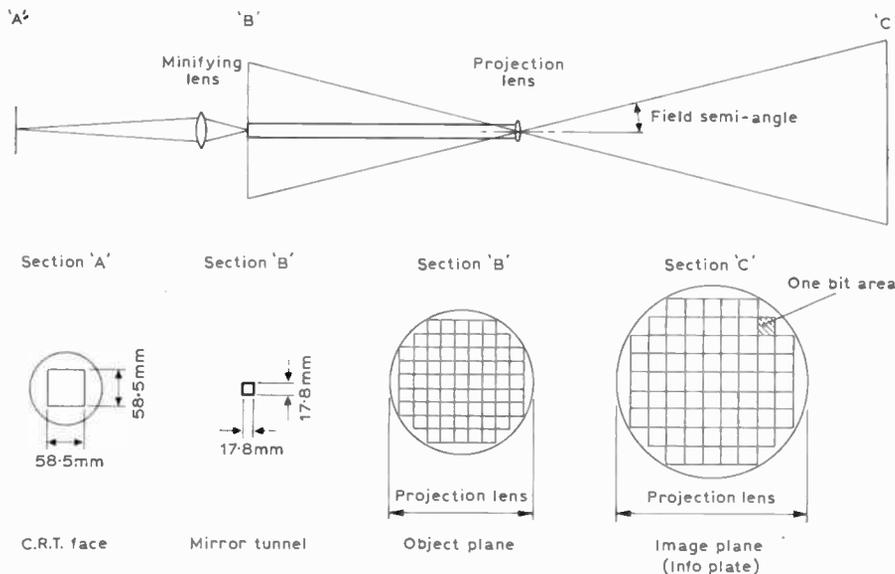


Fig. 3. Optical layout.

p.m.t.s with S11 photocathodes, this type of cathode being chosen due to its good spectral match to several very short persistence phosphors. The minimum illumination having been determined at 0.4 microlumens and the output circuit bandwidth setting a maximum for the anode load, the required p.m.t. current gain was deduced to be approximately 10^6 . A survey of available p.m.t.s having the required mechanical and electrical characteristics resulted in the choice of the Mullard XP1110 for both data and address light detectors:

Cathode	S11
Cathode sensitivity	$\geq 72 \mu\text{A/lumen}$
Overall sensitivity	150 A/lumen
Overall diameter	19 mm
Overall length	88 mm
Useful cathode diameter	14 mm

All the p.m.t.s in the store (a total of 89) are operated from a single low impedance power supply, consisting of eleven identical voltage stabilizers in series. Each sub-unit has adjustable voltage output and current trip level, being the source for one of the inter-dynode voltages. The tubes are used with their anodes near to ground potential and with the cathode ends of the glassware supported in insulating rings.

3.2. Projection Lens and Mirror Tunnel

The choice of p.m.t. fixes the minimum dimensions of the bit areas at the photographic plate as 25 mm x 25 mm.

The projection lens is required to resolve 10 spots per mm to the extremes of its image field: this implies a resolution much greater than 10 lines per mm, since

if possible it is required that 90% of any spot's luminous energy be focused within the nominal spot area. A maximum field semi-angle of 15° was chosen to ease the problems of field curvature and a focal length of 200 to 225 mm devolved from tunnel length considerations.

A tunnel length of approximately 500 mm was the largest considered since the flatness required of the glass blocks involved polishing on machinery which could conveniently cater for lengths of this order: following the decision to minify the c.r.t. face into the tunnel end, a 17.8 mm square section tunnel was calculated to be large enough. At 15° maximum semi-angle for the projection lens the tunnel operates at 75° minimum angle of incidence. In order to obtain at least 50 bits in the word (50 images of the c.r.t. on the photographic plate) it can be seen from Fig. 3 that a maximum of eight reflexions are needed, four in each axis, to give a square field of 9×9 images. A circular approximation to this removes the three images from each corner of the field leaving $9 \times 9 - 4(3) = 69$, with a maximum of six reflexions. The aspect ratio of the tunnel must be 20 : 1 for this configuration, requiring a tunnel length of 355 mm. The use of only three reflexions in each axis would give an inadequate word length of 45 or 37 bits. The tunnel is constructed of four blocks of silvered glass for the inward facing mirrors, with two end-windows to position them accurately. Calculations of the flatness and other physical requirements gave for the tunnel:

Opposite pairs of surfaces to be parallel within 3.5 seconds of arc.

Average plane of one pair of opposite faces to be perpendicular to the average plane of the other pair, within 3.5 seconds of arc.

Difference between two 'equal' widths of the tunnel at one end, within 0.0025 mm.

Difference between 'equal' widths of the tunnel at opposite ends, within 0.025 mm.

Mirror faces to be flat to $\frac{1}{4}$ of a wavelength over any 100 mm.

The projection lens chosen had:

$$F = 225 \text{ mm}$$

$$\text{aperture} = f/7.7$$

This lens is operated at $4/3$ magnification, the minimum aperture being determined by the need to avoid excessive vignetting with the tunnel end.

3.3. Minifying Lens

In order to take full advantage of the projection lens semi-angle and to allow at least 50 images to be produced with a minimum variation in brightness, the minifying lens must have an exit cone semi-angle of at least 15° over its whole image field. A minification of 1 : 4 was chosen for this lens and it became clear from discussions with manufacturers that a lens would have to be specially designed and built to provide the required performance. The specification included the following:

$$F = 51 \text{ mm} + 1.5\%$$

Object size	58.5 mm × 58.5 mm
Image size	14.6 mm × 14.6 mm
Wavelength range	370–500 μm

(peak at approximately 440 μm)

Resolution such that a 0.20 mm diameter object is imaged to give 95% of the light within a circle of 0.057 mm diameter.

An implication of the required exit angle is that an aperture of $f/1.5$ or larger is required, even though for a single spot on the c.r.t. face the lens operates at approximately $f/13$: in fact $f/1.0$ was attained, the lens being designed for an image plane within the end window of the mirror tunnel.

3.4. Cathode Ray Tube

In order to enable rapid movement of the spot across the c.r.t. face an electrostatically deflected and focused tube was specified with a phosphor decay time of a few hundred nanoseconds. Suitable phosphors with peak emissions at 410 μm , 450 μm and 500 μm are available. Since maximum luminous energy emission and maximum life are both needed, a robust efficient phosphor was chosen. In order to allow tubes to be changed, an event which could be made prematurely necessary by very frequent use of a single position on the tube face, the faceplate thickness and

tolerance are closely defined. This is necessary to avoid changes in apparent phosphor plane position causing loss of focus at the photographic plate and parallax errors in the address servo optics.

With a tunnel cross-section of 17.8 mm × 17.8 mm, containing a $\frac{1}{4}$ size image of the 256 × 256 matrix of c.r.t. spots, the used area on the tube would be a maximum of 71 mm × 71 mm. However, a guard space is left at the tunnel end around the image matrix referred to the c.r.t. This reduces the required area to 58.5 mm × 58.5 mm implying spots at approximately 0.23 mm pitch. A target of 0.15 mm was set for the spot diameter being defined as below in the tube specification.

Spot size: 90% of the light output of a spot to originate from inside a square of 0.15 mm maximum.

Phosphor: Imperfections in the phosphor shall not cause the light output to be less than 50% of the nominal value at any point.

Faceplate tolerance: Thickness, parallelism and wedge tolerance to total not more than ± 0.038 mm.

Repeatable measurements of c.r.t. resolution can be obtained by use of spatial frequency measuring techniques. This method was used to compare the performance of several tubes, but the specification quotes the resolution in a manner corresponding as closely as possible to the way in which the c.r.t. is used.

3.5. Address Optics

The optical system used to deduce the spot position for one c.r.t. axis consists of a pair of cylindrical mirrors forming line images of the c.r.t. spot as shown in Fig. 4. Only one mirror per axis is illustrated for reasons of simplicity. As the spot moves across the tube face, so the focused line follows a curve of radius related to that of the cylindrical mirrors, the focus being improved by an aperture restriction. Sections of opaque/transparent Gray-coded film are fixed along the curve of focus, 8 digital channels being used to give a resolution of 1 in 256. By means of perspex light collection optics, the light passing through each channel of Gray code is directed on to a separate photomultiplier: comparison of the digitized and decoded outputs of these detectors with the required address enables corrections to the spot position to be calculated and effected.

In addition to the digital servo system described above, two other optical/electronic channels are used to produce an analogue correction signal, thus accurately centring the light spot within the desired digital address area. These two channels consist of one clear section (reference section) and the centring

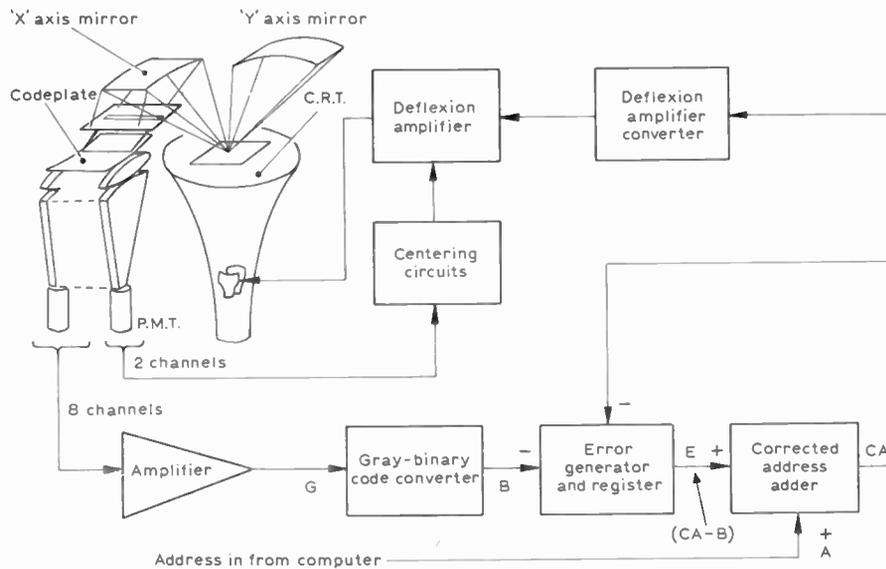


Fig. 4. Operation of address servo.

section: the latter is divided to form the Gray-coded channel next lower in significance to the finest of the digital section. A comparison of the reference and the centring p.m.t. signals enables accurate fine positioning of the c.r.t. spot, due regard being paid to the gain and bandwidth of the circuits to ensure stability of action.

The ten sections of optics needed for each c.r.t. deflexion axis are mounted five on each side of the tube face, because of the limited space available. Clearly, the tolerance required of the coded films and their positioning must be small compared with the 0.23 mm spot pitch. (Further details of the address servo will be found elsewhere.⁴)

3.6. Photography

Since the total store capacity is $256^2 \times 69$ bits (maximum), i.e. approximately 4×10^6 bits, and each bit to be written requires a separate photographic exposure, a reasonable writing time for a full plate necessitates the use of a fast emulsion: 'reasonable' writing time may be taken as a time which is very short compared with the life of the information written on to the plate.

The total energy of approximately 0.4 microlumens is contained within a focused spot of 0.08 mm diameter on the information plate and, for a one-hour plate exposure time, no more than 1 ms exposure time per bit can be permitted. Several blue sensitive, medium resolution emulsions can provide this performance, the faster emulsions recording images with the poorest definition. Writing of a full plate of information must clearly be performed under computer control, a further indication of the need for fast emulsions. By

off-line contact printing of the exposed and processed plate on to a slower emulsion, positives (i.e. clear bits on a black field) can be produced with a corresponding improvement in resolution.

Reliable reading of the written information requires:

- (i) The maximum possible plate transmittance in a nominally clear zone.
- (ii) The maximum contrast in transmittance from a nominally clear zone to a nominally opaque zone.

Reversal onto a high gamma emulsion can improve both (i) and (ii).

3.7. Mechanical Construction

In order to permit the adjustment of the optical components and to maintain their relative positions, great attention must be paid to the manner of mounting each sub-assembly. The frame to which all the components are fixed is shown in Fig. 5.

The c.r.t. is fixed accurately in the vertical, tilt and yaw modes, approximately in X and Y position, and is rotatable. The accurate location is achieved by one of two means; either the tube is fitted with a mounting ring precisely related to the apparent phosphor plane and this ring supported on three ground pins, or the tube front face may be held against three such pins. In either case the mounting pin faces must be set to 0.012 mm relative to the address servo or spot positioning optics. The address servo itself corrects for X, Y and rotational errors in c.r.t. position which are within its designed operating range.

The polished steel cylindrical mirrors of the servo are assembled on a precision plate with their reference

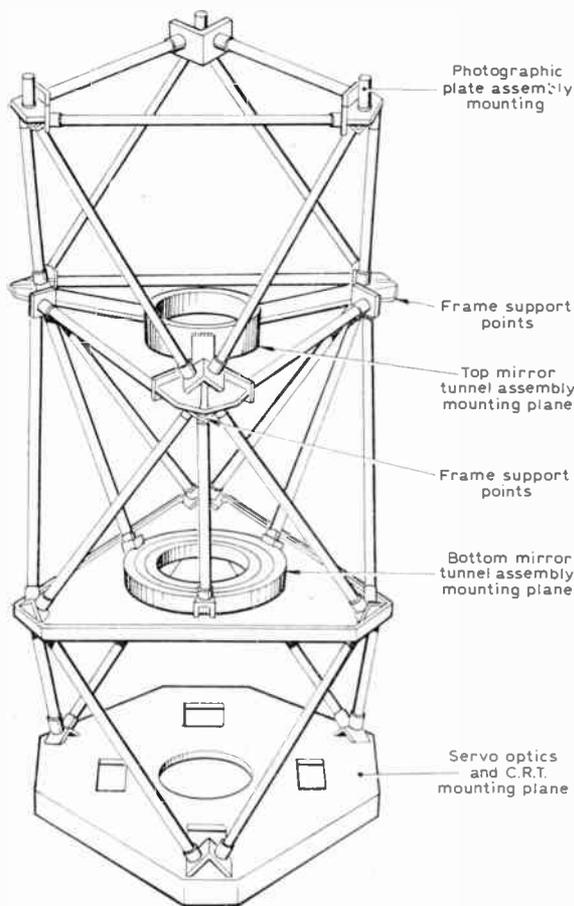


Fig. 5. Mechanical structure.

edges exactly aligned and square as needed. The plate is kinematically attached to the store framework; X, Y, tilt, yaw, rotational and vertical adjustment are all provided. The four sections of servo code plate mechanics and optics are semi-kinematically mounted on the store framework.

The mirror tunnel is held within a steel tube by fixed pins and springs, to produce no twisting or conflicting strains. This tube, with the minifying and projection lenses screwed in at either end, is mounted kinematically at its lower end in gimbals with X and Y adjustment. Three adjustable spherical ended feet provide vertical support and vertical adjustment; rotation is also possible. To avoid applying conflicting strains the upper end of the tunnel is fixed between three horizontal adjustable pins, spaced mutually at 120° .

The photographic plate is mounted kinematically in the plate holder, whose three feet rest in three grooves; tilt, yaw and vertical movements are possible by means of adjustment of the feet. It is possible to remove the plate and/or plateholder and replace the

same to a tolerance small compared with the spot pitch at the plate. The data light collection optics and photomultiplier assembly are not adjustable.

Clearly no amount of care in the kinematics of the mounting and in the adjustment of all the optical sub-assemblies will suffice, if the framework to which they are fixed is dimensionally unstable or susceptible to excessive vibration. As shown in Fig. 5 the structure is a space-frame which approaches a theoretically non-redundant design. The final step in the jiggered assembly of this framework is to join the prefabricated corner pieces by steel tubes: this is a brazing process, obviating built-in stresses and the need for stress relieving heat treatment which would result from welding. In order to isolate the framework from floor-borne vibrations it is supported within an outer frame on three damped spring assemblies.

4. Performance

The storage system described above is in the process of final commissioning and evaluation. The philosophy behind the construction of the store was that:

- (i) The store's average random access time must be of similar order to that for magnetic core stores.
- (ii) The cost per stored bit must be considerably less than that for core stores.
- (iii) It must be possible to read data which have been previously written on the same or any other optical store made to the same specification.

In (i), it follows from the method of positioning the spot that a change of address involving a large spot movement may require cycles around the address servo loop which are not needed by a small movement. Corresponding variations in access time will be detected. However, analysis of the frequencies of various computer program jump lengths shows that the average access time will approach the minimum. To date experiment shows that a random access time of $10 \mu\text{s}$ will be achieved as an average of many typical accesses. However, a long-term random accessing test has yet to be performed. Redesign of the servo loop and the use of faster logic could reduce the access time but even so it is unlikely that it will fall below $1 \mu\text{s}$, due to the non-zero phosphor persistence, maximum attainable electron beam deflexion speeds etc.

Considering point (ii) above, the estimated production cost for stores as described is a third of a core store. An enlarged development of the store having (say) four times the present capacity could be built for a small increase in cost, resulting in a cost advantage over core store of up to 10 to 1.

The reading of previously written information (iii) depends on several factors. It has been made possible to interchange c.r.t.s by providing each with a precise glass face-plate and mounting the tube within the store with reference to this face. Calculation shows that for the present steel structure a temperature stability of ± 2 deg C will be required: however the use of a constructional alloy having a temperature coefficient of expansion closer to that of glass can increase this range to ± 15 deg C. Every store will be optically adjusted to the same standard using index marks on a glass master setting plate.

Experimental evaluation of the store's reliability has not yet been undertaken. Apart from component failures, the major control of reading reliability is the signal/noise ratio of the data photomultiplier outputs after integration. This is dependent on the absolute light level at the data plate and the integration time, the former being set by the c.r.t. beam current (and thus related to the tube life), the latter being adjusted to the minimum consistent with the required signal/noise ratio.

5. Acknowledgments

The authors gratefully acknowledge the assistance given by their colleagues at I.C.L. and thank the

Directors of I.C.L. and the Ministry of Technology for permission to publish this information which results from work which was supported by a normal cost-sharing contract under the Ministry's Advanced Computer Technology Project.

6. References

1. Bowden, K. R. R., 'Random access photographic data storage', Proc. Scientific Instrument Research Assoc. Conference, Eastbourne, pp. 6-1—6-8 (Chislehurst, Kent, B.S.I.R.A., 1968).
2. Hoover, C. W., Haugk, G. and Herriott, D. R., 'System design of the flying spot store', *Bell Syst. Tech. J.*, **38**, pp. 365-401, March 1959.
3. Krolak, L. J. and Parker, D. J., 'The Optical Tunnel, a Versatile Electro-Optical Tool', Applied Research D.E.P., Camden, N. J.
4. Hill, A. N. and Steptoe, B. J., 'Servo system for c.r.t. light-spot positioning', *Control*, **12**, pp. 952-3, November 1968.
5. Botfield, R. J., Hill, A. N. and Steptoe, B. J., 'A random access optical fixed store', *J. Sci. Instrum. (J. Phys. E)*, (Series 2), **1**, pp. 810-11, August 1968.

Manuscript first received by the Institution on 13th March 1969 and in final form on 3rd December 1969. (Paper No. 1306 Comp 127)

© The Institution of Electronic and Radio Engineers 1970.

The Authors.



Ron Botfield joined the aircraft industry in 1954 as an instrument maker, became a draughtsman and obtained his H.N.C. in mechanical engineering. After working with Hawker Siddeley Dynamics for 3 years as an engineer with a small design team, he moved to I.C.T. Ltd. (later I.C.L.) in 1965. He is now a principal engineer in the Research Division.



Adrian Hill joined Ferranti Ltd. (later I.C.T. Ltd.) in 1962 on graduating in electrical engineering from King's College, London. After three years as a commissioning and site engineer with *Orion* computers, he transferred to development and later joined the research laboratories of I.C.L. as a principal engineer.

Joint Conference on 'Television Measuring Techniques'

Organized by the I.E.R.E. with the association of The Institution of Electrical Engineers, The Institute of Electrical and Electronics Engineers and The Royal Television Society

Middlesex Hospital Medical School, London, W.1, 11th to 13th May, 1970

OUTLINE PROGRAMME

Monday, 11th May

STUDIO ORIGNATION

- Survey Paper, J. S. SANSOM, *Thames Television*.
'Television Studio Equipment Performance Measurements,' P. J. DARBY and M. S. TOOMS, *ITA*.
'Measuring Techniques used in Colour Studio Line-up,' I. H. TEEAR and J. P. MEADOWS, *BBC*.
'The Monitoring and Measuring Systems adopted for Colour Camera Set-up at Teddington Studios,' B. G. SCOTT and A. FOWLER, *Thames Television*.
'A Camera Tube Lag Meter,' J. R. SANDERS, *BBC*.
'The Digital Measurement of Colour Sub-Carrier Phase,' L. E. WEAVER and J. LEWIS, *BBC*.
'Techniques for Measuring Small Order Non-Linear Distortions,' N. W. WHITE, *Marconi Instruments*.
'A New Equipment for the Measurement of Video Noise in the Presence of Synchronizing Signals,' F. H. WISE and D. R. BRIAN, *ITA*.
'Television Test Cards,' S. LENT, H. A. S. PHILLIPART, *BBC*, and J. ROYLE, *Jeremy Royle*.
'Instruments and Techniques in the Correct Setting up of Colour Television Monitors and Receivers ("Color Trak"),' M. T. DAWSON, *Fisher Controls*.
'A Colour Monitor Calibrator,' R. E. KNIGHT, *Thames Television*.

Tuesday, 12th May

NETWORKS

- Survey Paper, J. B. SEWTER, *ITA*.
'Knowing the Enemy—Degradation,' F. J. PADDOCK, *BBC*.
'Application of Insertion Test Techniques to Television Transmission Chain Operation,' D. C. SAVAGE and D. A. CARTER, *BBC*.
'Automatic Measurement of Insertion Test Signals,' I. J. SHELLEY and G. WILLIAMSON-NOBLE, *BBC*.
'Insertion Test Signal Equipment,' J. SCHAEFFER and I. LEVER, *ITA*.
'Signal to Noise Ratio Automatic Measurement in the Blanking Interval of the Video Signal,' P. D'AMATO and G. BARBIERI, *Radiotelevisione Italiana*.
'Measuring Differential Phase Distortions by using only One Line per Frame,' K. VOIGT, *Institut fur Rundfunk-technik*.
'A Set for Annulling Gain and Delay Inequalities,' D. A. COLES, *Post Office*.

RECEIVERS

- Survey Paper, B. J. ROGERS, *Rank Bush Murphy*.
'Colorimetric Error Measurements with the Vectorscope as a Measure of Picture Impairment,' C. W. RHODES and A. KASTELEIN, *Tektronix*.
'Electro-optical Measurements of Television Monitors and Receivers,' Dr. H. GROSSKOPF, *Institut für Rundfunk-technik*.
'Evaluation of Convergence Requirements of Colour Television Receivers,' A. CIUCIURA, *Mullard Mitcham*.
'Behaviour of E.H.T. Triplers under Picture Tube Conditions,' A. CIUCIURA, *Mullard Mitcham*.

Wednesday, 13th May

TRANSMITTERS AND AERIALS

- Survey Paper, W. WHARTON, *BBC*.
'U.H.F. Klystron Transmitters,' J. F. H. BINNS, *Marconi*.
'The Relationships and Measurements of Cross-modulation and Intermodulation on a Combined Signal Working C.C.I.R. System A V.H.F. Band III Relay System,' J. W. MORRIS, *ITA*.
'Measurement of U.H.F. Television Transposer Characteristics,' E. J. WILLS, *Pye TVT*.
'New Test Signal Generator for Transmitter and Studio Measurements,' P. HEJBERG, *Philips Radio, Copenhagen*.
'The Measurement of Television Service Areas,' G. A. VICKERS, *Marconi Research Laboratories*.
'The Measurement of Field Strength and Signal Quality at the Site of Proposed U.H.F. Television Relay Stations,' R. J. BYRNE, *ITA*.
'Automatic Supervision of Remote Television Stations,' E. CASTELLI, *Radiotelevisione Italiana*.

'Amplitude Shaping and Group Delay Equalisation of Television Measuring Demodulators for System I,' B. D. WATSON, *ITA*.

'Aerial Measuring Techniques,' M. B. ANDERS, *E.M.I. Electronics*.

'U.H.F. Aerial and Feeder Measurements,' R. G. WILLS, *Marconi*.

Practical demonstrations will be staged at the Medical School and there will be visits to the ITA Television Gallery on 11th May and to the BBC Television Centre on 13th May (attendance limited).

Registration Forms which include details of charges etc. are now available from The Conference Registrar, Institution of Electronic and Radio Engineers, 8-9 Bedford Square, London, W.C.1.

FORTHCOMING CONFERENCES

Instruments in Working Environments

It is easy for instrument manufacturers to concentrate on *what* rather than *where* their brain-children measure. That is unfair to instrument users. It is just as unfair to manufacturers if users expect a standard instrument to work perfectly under abuse or attack.

Users who think that manufacturers are complacent or ignorant will be able to tell them so, in instructive detail, at the 1970 Conference on instrument in working environments, organized by the British Scientific Instrument Research Association at Eastbourne on 5th and 6th May, 1970. The manufacturers will learn a lot from this, but they will no doubt have a thing or two to say about users' practice. They will probably point out that users must balance the extra capital cost of foolproof, fireproof, waterproof and anything-else-proof instruments against the higher running and maintenance cost of simpler devices.

Not only the measured variables—temperature, voltage, velocity, pressure, flow, pH, weight and the rest—but also the facts of environment are the common (and sometimes sad) lot of instrument users. So although there may be much for instrument makers to learn about users' problems, users can also learn something from each other's solutions.

The conference is being organized as six discussion sessions with named opening speakers, representative of manufacturers and users, on the following subjects: Fluid Environments; Electro-magnetic and Ionizing Radiations; Human Environment; Varying Thermal Conditions; Vibration and Shock; The Problem of Specification. Registration and hotel reservation forms can be obtained from Industrial Communications Group, S.I.R.A., South Hill, Chislehurst, Kent BR75EH. (Telephone: 01-467 2636.)

Computer Graphics Conference at Brunel University

A three-day Conference on Computer Graphics—CG.70—is to be held at Brunel University, Uxbridge, Middlesex, from April 14th to 16th. The Conference will comprise a major exhibition and a symposium, which has been organized into 20 sessions, four running parallel at any one time with video-recorded repeats. Of the 100 speakers at the conference about 50 are from the U.S.A., 35 from the U.K., and others are coming from U.S.S.R., Canada, Australia, France and Hungary.

In addition to basic sessions on the techniques of computer graphics, there are sessions on specialist applications such as in the design of electrical circuits, in the

design and construction of civil engineering projects, in architecture, hospital administration management information systems and text processing for printing.

The registration fee of £40 covers preprints of all the technical papers, admission to all sessions and the exhibition, as well as lunches and an invitation to the Conference banquet. Further information may be obtained from CG.70, Department of Computer Science, Brunel University, Kingston Lane, Uxbridge, Middlesex.

Recent Advances in Bio-Medical Engineering

A Conference and Exhibition is being held by the Biological Engineering Society at Oxford from 7th to 10th April 1970, to mark the occasion of the 10th anniversary of its formation. The Society was founded in 1960 to further co-operation between workers in the fields spanning Engineering and the biomedical sciences, and its members represent some 30 or more professions.

The field of Bio-Medical Engineering is now too wide for it to be covered adequately at one meeting, so this conference will be concerned mainly with those subjects in Bio-Medical Engineering in which progress has been most marked. Sessions include the following:

- Education and Organization of Biochemical Engineering
- Computer Applications and Science
- Clinical Application of Biomedical Engineering
- Automation in Pathology
- Biomedical Engineering in Surgery and Intensive Care
- Biomedical Engineering in Research
- External Prostheses Implants
- New Materials in Medicine and Biology
- Aids to the Aged and Infirm

The Conference is being held in the Examination Schools at Oxford, and there is residential accommodation at Christ Church.

The Organizing Committee for the Conference is under the Chairmanship of Mr. K. Copeland (Member), who is the immediate past chairman of the I.E.R.E. Medical and Biological Electronics Group Committee; the Programme Chairman is Mr. W. J. Perkins (Fellow), who was chairman of the institution's Group Committee on its formation. Applications to attend the Conference should be sent to the Secretary, Mr. J. D. Gasking, C.Eng., M.I.E.R.E., Department of Pharmacology, The Medical College of St. Bartholomew's Hospital, London, E.C.1.

Six New Delay Functions and their Realization using Active RC Networks

By

T. DELIYANNIS,

B.Sc., Ph.D.†

Six new delay functions with three right-half s -plane zeros and four left-half s -plane poles are obtained from the application of Budak's technique to the Bessel and Chebyshev all-pole delay functions. These are compared with three well-known similar delay functions and three of them found to have some advantages over the existing ones. Active networks, which are useful in the realization of the functions, are examined and experimental responses are given.

1. Introduction

King and Rideout¹ have studied the possibility of using low-pass non-minimum phase functions of order 3, 4 (3 r.h.s. zeros and 4 l.h.s. poles), in order to avoid the initial spike or precursor which characterizes the step response of all-pass delay functions. The two functions $K_{3,4}$ and $R_{3,4}$ which they introduced have a better transient response than the $P_{3,4}$, the corresponding Padè delay function.²

Budak³ has shown how to modify the all-pole Bessel delay functions so as to achieve an adjustable magnitude response. The resulting functions are of order m, n , where $m < n$, and consequently need more complex networks for their realization than the corresponding all-pole ones. A fourth-order all-pole Bessel delay can thus be modified to obtain three r.h.s. zeros and its magnitude can be properly adjusted in order to have the same half-power frequencies with each one of $P_{3,4}$, $K_{3,4}$, $R_{3,4}$. A comparison of the phase and transient responses will then reveal whether Budak's technique can produce delay functions of order 3, 4 which are more useful than the $P_{3,4}$, $K_{3,4}$, $R_{3,4}$. This is done in the first part of this paper. Budak's technique is also applied in order to modify a fourth order all-pole Chebyshev delay. Thus, three more $F_{3,4}$ delay functions are obtained each having the same half-power frequency with one of $P_{3,4}$, $K_{3,4}$, $R_{3,4}$.

The second part of the paper is concerned with the realization of these functions using RC active networks. Some recently developed networks employing operational amplifiers are used and experimental transient responses are given.

2. Application of Budak's Technique

A Bessel delay function is an approximation‡ to e^{-s} . In order to obtain delays with controllable

† West Ham College of Technology, Electrical Engineering Department, London, E.15.

‡ Delays normalized to 1 s are used throughout this paper.

magnitude response from the Bessel all-pole delays Budak has applied the following technique: Introducing a constant k and splitting e^{-s} into two terms gives

$$e^{-s} = \frac{e^{-ks}}{e^{-(k-1)s}} \quad 0 < k \leq 1 \quad \dots\dots(1)$$

e^{-ks} and $e^{-(k-1)s}$ can then be approximated independently using Bessel all-pole delay functions. If n poles are used to approximate e^{-ks} , and m poles to approximate $e^{-(k-1)s}$, the Budak-Bessel function, which results after substitution in (1), will have m r.h.s. zeros and n l.h.s. poles. This function can then be multiplied by a suitable constant in order to obtain a magnitude equal to unity at zero frequency. As k is reduced from unity the useful bandwidth increases, overshoots and transient ringing in the step response increase and the rise-time decreases.

In order to examine the usefulness of Budak's technique $m = n - 1 = 3$ was selected. Using a digital computer, k was adjusted until the Budak-Bessel 3, 4 function had approximately the same half-power frequency as $P_{3,4}$. This function will be subsequently referred to as the $B-P_{3,4}$ (Bessel-Padè) delay function. This was also done in order to obtain two more functions, the $B-K_{3,4}$ and the $B-R_{3,4}$, which have the same half-power frequencies with $K_{3,4}$ and $R_{3,4}$ respectively. The pole-zero locations of these functions are given in Table 1 together with those of $P_{3,4}$, $K_{3,4}$, $R_{3,4}$.

Next, Budak's technique was applied to an all-pole Chebyshev delay. Macnee⁴ has studied a number of all-pole Chebyshev delays with various ripple constants ϵ in the phase response. The 3- and 4-pole functions with $\epsilon = 0.05$ were selected to approximate to $e^{-(k-1)s}$ and e^{-ks} respectively in eqn. (1). Next k was adjusted as before and three more 3, 4 Chebyshev-Budak functions were obtained. These are the $C-P_{3,4}$, $C-K_{3,4}$ and $C-R_{3,4}$, each having the same half-power frequency as each of $P_{3,4}$, $K_{3,4}$ and $R_{3,4}$

respectively. Their pole-zero locations are also given in Table 1.

Table 1. Poles and zeros of the $F_{3,4}$ delay functions

Group	$F_{3,4}$ (s)	Zeros	Poles
1	$P_{3,4}$	$5.6485 \pm j0.0$	$-4.7872 \pm j1.5675$
		$4.6758 \pm j3.9135$	$-3.2128 \pm j4.7731$
	$B-P_{3,4}$	$5.4336 \pm j0.0$	$-5.0578 \pm j1.5145$
		$4.3028 \pm j4.1050$	$-3.6740 \pm j4.6408$
	$C-P_{3,4}$	$3.5711 \pm j0.0$	$-2.6244 \pm j2.1176$
		$3.0999 \pm j5.5844$	$-2.2009 \pm j6.1764$
2	$K_{3,4}$	$5.4694 \pm j0.0$	$-2.8464 \pm j1.6516$
		$4.6991 \pm j6.5529$	$-3.6436 \pm j6.0258$
	$B-K_{3,4}$	$8.1952 \pm j0.0$	$-4.0414 \pm j1.2101$
		$6.4896 \pm j6.1913$	$-2.9356 \pm j3.7082$
	$C-K_{3,4}$	$6.7925 \pm j0.0$	$-1.9730 \pm j1.5921$
		$5.8962 \pm j10.622$	$-1.6547 \pm j4.6435$
3	$R_{3,4}$	$6.2942 \pm j0.0$	$-2.6076 \pm j1.9189$
		$4.5556 \pm j5.5086$	$-3.4424 \pm j5.7933$
	$B-R_{3,4}$	$7.4549 \pm j0.0$	$-4.2066 \pm j1.2596$
		$5.9034 \pm j5.6320$	$-3.0556 \pm j3.8597$
	$C-R_{3,4}$	$5.9020 \pm j0.0$	$-2.0585 \pm j1.6610$
		$5.1232 \pm j9.2292$	$-1.7264 \pm j4.8446$

3. Comparison of the $F_{3,4}$ Functions

In order to compare the functions in Table 1 the following quantities were calculated:

- (a) The frequency at which the phase error $\Delta\phi = (\omega - \phi(\omega))$ is 0.1 rad.
- (b) The ratio of the delay time to the rise-time in the step response (T_D/T_r).
- (c) The peak value A_r of the transient ringing or precursors.
- (d) The maximum overshoot.

Quantities T_D/T_r , A_r and maximum overshoot were obtained from the actual plot of the unit step response of each delay function. This response was calculated using the conventional Laplace transform inversion method at 80 points in the range 0-4 seconds.

The corresponding all-pole Bessel and Chebyshev delay functions were also considered in the comparison.

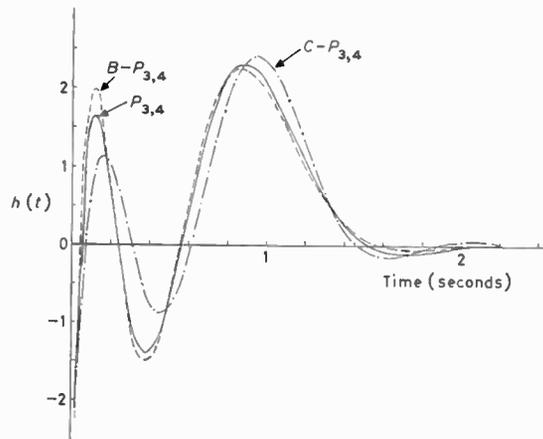
Results are summarized in Table 2.

It can be seen that the Bessel-Budak functions do not differ significantly from the corresponding $P_{3,4}$,

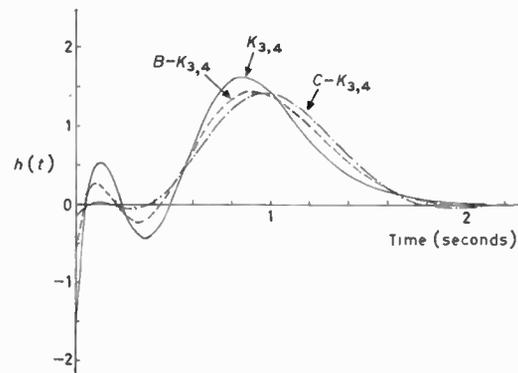
$K_{3,4}$ and $R_{3,4}$ functions. However, the Chebyshev-Budak functions show a considerable improvement

Table 2

	k	$\omega_{\Delta\phi=0.1 \text{ rad}}$ (rad/s)	$\omega_{3 \text{ dB}}$ (rad/s)	Over-shoot %	A_r %	T_D/T_R
$P_{3,4}$	—	5.25	6.99	3.11	-16.8	2.298
$B-P_{3,4}$	0.572621	4.95		1.64	-16.5	2.157
$C-P_{3,4}$	0.589602	7.47		2.19	-10.9	2.590
$K_{3,4}$	—	3.25	3.28	0.62	-4.16	1.421
$B-K_{3,4}$	0.71664	4.50		1.05	-1.95	1.441
$C-K_{3,4}$	0.784235	5.83		2.88	-1.05	1.474
$R_{3,4}$	—	4.50	3.54	1.82	+4.33	1.608
$B-R_{3,4}$	0.688502	4.65		1.12	-3.20	1.511
$C-R_{3,4}$	0.751678	6.05		2.99	-1.06	1.608
$B_{0,4}$	1	3.25	2.10	0.83	0	0.940
$Ch_{0,4}$	1	4.60	2.36	2.63	0	1.088



(a) Delay functions of group 1.



(b) Delay functions of group 2.

Fig. 1. Response to a unit impulse of the delay functions in Table 1.

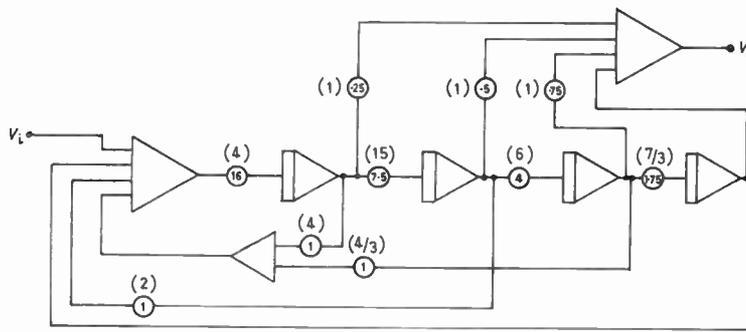


Fig. 2. Analogue computer circuit for the implementation of $P_{3,4}$. Numbers in parentheses give an alternative set of potentiometer setting values times amplifier gain values.

over the rest except for the overshoots of $C-K_{3,4}$ and $C-R_{3,4}$, which are higher than the overshoots of $K_{3,4}$ and $R_{3,4}$ respectively. The $C-P_{3,4}$ has the largest useful bandwidth and the largest T_D/T_r . It can also be seen from Table 2 that in each group the better the phase response the smaller the A_r .

The responses of the delays of the first two groups in Table 1 to a unit impulse are shown in Figs. 1(a) and 1(b). The corresponding responses of the delays of group 3 are very similar to those of group 2.

4. Realization of the $F_{3,4}$ Delay Functions

The realization of the delay functions, which have been examined in the previous sections, will now be considered.

King and Rideout have given analogue computer circuits realizing $P_{3,4}$ and $K_{3,4}$. These configurations can also be used to realize all 3,4 functions in Table 1. Some alternative circuits using operational amplifiers are given below.

4.1. Analogue Computer Circuits

The well-known nested form has been used by King and Rideout in order to implement $P_{3,4}$ and $K_{3,4}$. An alternative circuit based on the use of the dummy variable form and realizing $P_{3,4}$ (1 s delay) is shown in Fig. 2. A second set of potentiometer setting values times amplifier gain is given in parentheses on the same figure. Both have been tested on a Solartron SCD.10 analogue computer, and found to work satisfactory.

4.2. Circuits using Fewer Operational Amplifiers

For sensitivity reasons we will exclude networks whose transfer functions have the denominators formed from the difference of two quantities. Since the functions under consideration are not highly selective their realization can be achieved in one or two stages in cascade.

4.2.1. Realization in a single stage

A suitable active network employing the method of single inversion is given in Ref. 5. Its transfer voltage ratio is

$$\frac{V_o}{V_i} = \frac{Y_1 - Y_2}{-y_{12}}$$

where Y_1 and Y_2 are RC driving-point admittances associated with the forward path network and $-y_{12}$ is the transfer admittance of an RC 3-terminal network connected in the feedback of an operational amplifier. A second operational amplifier is used as a sign inverter feeding Y_1 . The feedback network can be properly designed⁵ to realize the two pairs of complex conjugate poles of each $F_{3,4}$ using only eight passive components. The normalized network realizing $P_{3,4}$ is shown in Fig. 3.

Although this network realizes a fourth-order function in one stage each component in the feedback network affects only one pair of poles. The pole sensitivity is not expected to be unduly high.

This network denormalized to an impedance level of $10^7 \Omega$ was built using Solartron AA1054 operational amplifiers†; the results agreed with theory.

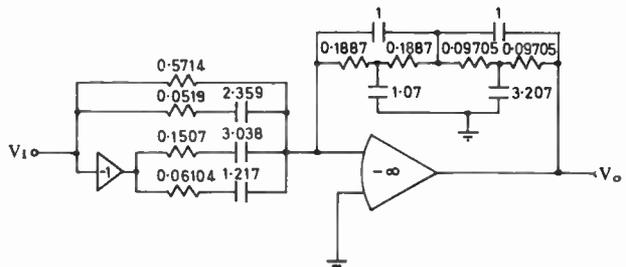


Fig. 3. Realization of $P_{3,4}$ in one stage using two computer amplifiers. Component values in ohms and farads.

† All circuits described here were built using resistors and capacitors of various types selected to within 1% of their nominal values.

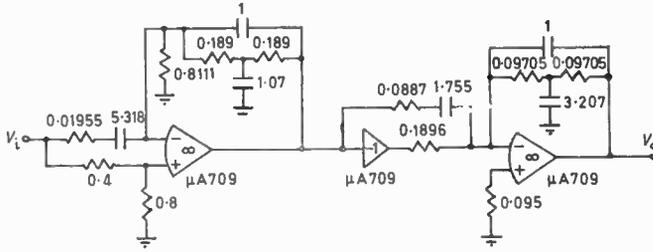
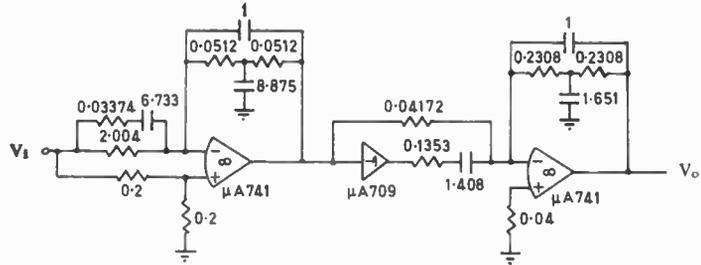


Fig. 4. Realization of $P_{3,4}$ in two cascaded stages using three integrated circuit operational amplifiers.

Fig. 5. Normalized network realizing $C-P_{3,4}$.



4.2.2. Realization by two stages in cascade

$F_{3,4}$ is written as the product of two functions f_1 and f_2 . One of these, say the f_1 , is a non-minimum phase biquadratic of the form

$$f_1 = k_1 \frac{s^2 - \alpha_1 s + \alpha_0}{s^2 + \beta_1 s + \beta_0} \quad \dots\dots(2)$$

and the other, the f_2 , a quadratic with a zero on the positive real axis, i.e.

$$f_2 = k_2 \frac{\gamma - s}{s^2 + \delta_1 s + \delta_0} \quad \dots\dots(3)$$

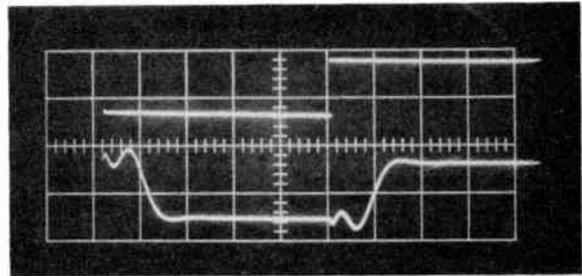
f_1 can be realized by network 1 in Ref. 6. Theoretically f_2 can also be realized by the same network but this is not achieved accurately in practice, since it means a complete suppression of the s^2 term in the numerator of the network function and this requires the operational amplifier to have infinite gain at all frequencies. However, we can use the single inversion circuit on which the network in Fig. 3 is based.

Networks realizing $P_{3,4}$ and $C-P_{3,4}$ in two cascaded stages are shown in Figs. 4 and 5 respectively. The responses of the denormalized networks ($10^6\Omega$ and 0.1 s delay) to a step and a triangular wave are shown in Figs. 6 and 7 respectively.

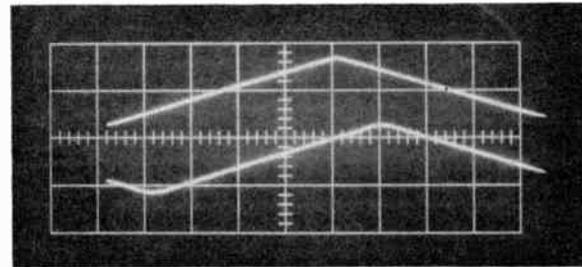
Function f_2 can also be realized by the network shown in Fig. 8. Its transfer voltage ratio V_o/V_i is given by†

$$\frac{V_o}{V_i} = \frac{K(g_2 + g_3)}{C_3} \cdot \frac{s + g_4/C_2}{s + g_5/C_3} \cdot \frac{g_1 g_3 (1 - K) - K g_2 g_3}{K C_1 (g_2 + g_3)} \cdot \frac{-s}{s^2 + \frac{g_2 + g_3 + g_4}{C_2} s + \frac{g_2 g_3}{C_1 C_2}} \quad \dots\dots(4)$$

† Conductances g_4 and g_5 can be dispensed with if a chopper stabilized operational amplifier be used as the second amplifier.



(a)



(b)

Fig. 6. Responses of the network in Fig. 4 denormalized to $R = 10^6\Omega$ and $T = 0.1$ s. (a) Step response. (b) Response to a triangular wave (time-base 0.1 s/cm).

where

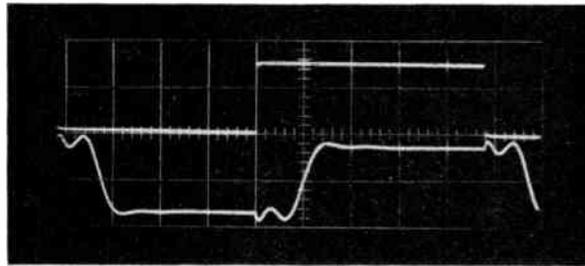
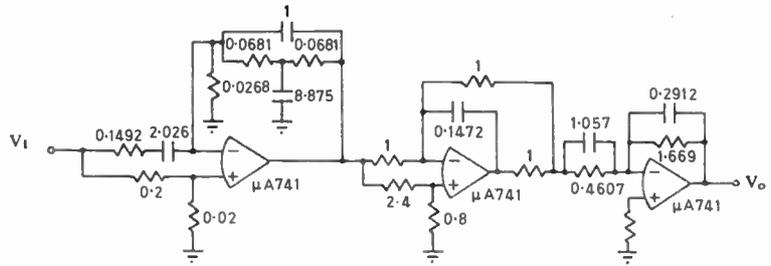
$$K = \frac{g_a}{g_a + g_b}$$

This can be simplified under the following conditions

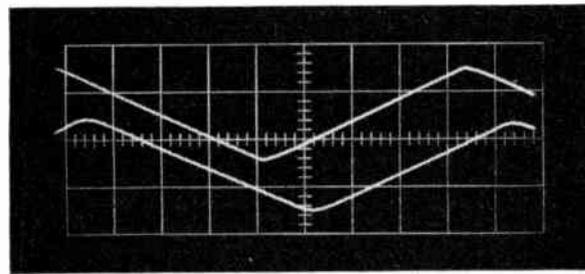
$$\frac{g_4}{C_2} = \frac{g_5}{C_3} \quad \dots\dots(5)$$

$$g_2 = g_3 = g \quad \dots\dots(6)$$

Fig. 9. Normalized network realizing $C-K_{3,4}$.



(a)



(b)

Fig. 7. Responses of the network in Fig. 5 after being denormalized to $10^6\Omega$ and 0.1 s delay (time-base 0.1 s/cm).

Then

$$\frac{V_o}{V_i} = \frac{2Kg}{C_3} \frac{g_1(1-K) - Kg}{s^2 + \frac{2g+g_4}{C_2}s + \frac{g^2}{C_1C_2}} - s \quad \dots\dots(7)$$

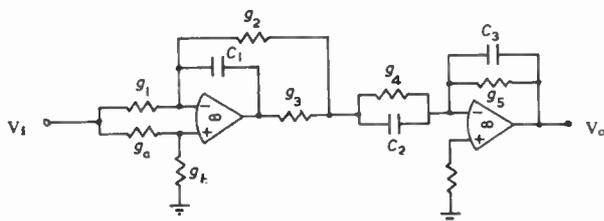


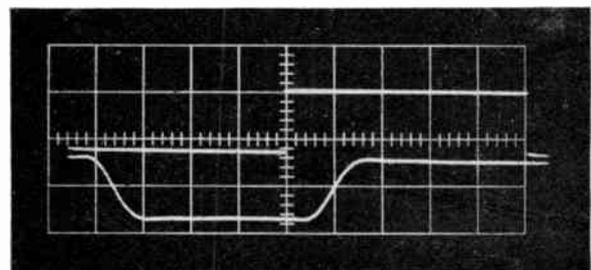
Fig. 8. General network realizing a quadratic function with a positive real zero.

It can be shown that the sensitivity of the zero on the positive real axis with respect to K can be minimized if K takes the value

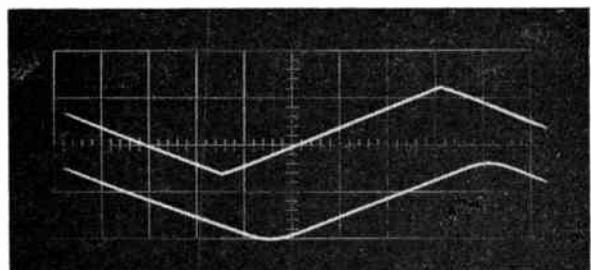
$$K = \frac{g_1}{2(g_1 + g)} \quad \dots\dots(8)$$

Component values can be found by equating coefficients of equal powers of s in eqns. (3) and (4) and making use of eqns. (5) to (8). However, since resulting equations are fewer than the components, some component values will have to be selected 'arbitrarily'.

This network was used to realize the f_2 factor of $C-K_{3,4}$. The overall network realizing $C-K_{3,4}$ is shown in Fig. 9. The responses of the denormalized network ($10^6\Omega$, 0.1 s delay) to a step and a triangular wave are shown in Fig. 10.



(a)



(b)

Fig. 10. Responses of the network in Fig. 9 after being denormalized to $10^6\Omega$ and 0.1 s delay (time-base 0.1 s/cm).

5. Conclusions

The all-pole Bessel and Chebyshev fourth-order delay functions have been modified using Budak's technique to give six new delay functions of order 3,4. Each modified Bessel and Chebyshev function has the same half-power frequency as one of $P_{3,4}$, $K_{3,4}$, $R_{3,4}$. The modified Chebyshev delays show a considerable improvement when compared to the rest.

The functions have been realized using RC networks combined with operational amplifiers. The networks which have been considered are expected to have low sensitivity to variations in the component values.

6. References

1. King, W. J. and Rideout, V. C., 'Improved transport delay circuits for analogue computer use', Proc. of the Third International Meeting of the Association pour le Calcul Analogique, Opatija, Yugoslavia, Sept. 5, 1961.

2. Morill, C. D., 'A sub-audio time delay circuit', *Trans. Inst. Radio Engrs on Electronic Computers*, EC-3, No. 2, June 1954.
 3. Budak, A., 'A maximally flat phase and controllable magnitude approximation', *Trans. I.E.E.E. on Circuit Theory*, CT-12, No. 2, p. 279, 1965.
 4. Macnee, A. B., 'Chebyshev approximation of a constant group delay', *Trans. I.E.E.E. on Circuit Theory*, CT-10, No. 2, pp. 284-285, 1963.
 5. Deliyannis, T. and Ream, N., 'An RC-active synthesis of a bi-quartic function with complex poles', *Internatl J. Electronics*, 23, No. 5, pp. 413-415, 1967.
 6. Deliyannis, T., 'RC active allpass sections', *Electronics Letters*, 5, No. 3, pp. 59-60, 6th February 1969.

Manuscript received by the Institution on 4th July 1969. (Paper No. 1307/CC66.)

© The Institution of Electronic and Radio Engineers, 1970

STANDARD FREQUENCY TRANSMISSIONS—February 1970

(Communication from the National Physical Laboratory)

Feb. 1970	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)		Feb. 1970	Deviation from nominal frequency in parts in 10 ¹⁰ (24-hour mean centred on 0300 UT)			Relative phase readings in microseconds N.P.L.—Station (Readings at 1500 UT)	
	GBR 16 kHz	MSF 60 kHz	Droitwich 200 kHz	*GBR 16 kHz	†MSF 60 kHz		GBR 16 kHz	MSF 60 kHz	Droitwich 200 kHz	*GBR 16 kHz	†MSF 60 kHz
1	-300.0	+0.1	+0.1	649	564.3	15	-299.9	+0.1	+0.1	644	562.6
2	-300.0	+0.1	0	649	563.8	16	-299.8	-0.1	+0.1	642	563.4
3	-300.2	0	0	651	563.6	17	-299.9	+0.1	0	641	562.8
4	-299.8	0	0	649	563.4	18	-300.1	0	+0.1	642	563.0
5	-300.0	-0.1	0	649	564.0	19	-300.1	-0.1	+0.1	643	564.4
6	-299.9	0	0	648	564.2	20	-300.2	+0.1	+0.1	645	563.5
7	-299.9	0	0	647	564.2	21	-299.9	0	0	644	563.9
8	-300.0	0	0	647	564.2	22	-299.8	+0.1	0	642	564.5
9	-299.8	0	0	645	564.2	23	-300.0	0	+0.1	642	564.3
10	-300.1	0	0	646	563.8	24	-300.2	+0.1	+0.1	644	563.7
11	-300.0	-	0	646	-	25	-299.8	+0.1	+0.1	642	563.1
12	-299.9	+0.1	0	645	563.8	26	-300.1	-0.1	+0.1	643	563.7
13	-300.0	-0.1	0	645	564.6	27	-300.1	+0.1	+0.1	644	562.7
14	-300.0	+0.1	+0.1	645	563.4	28	-299.9	+0.1	+0.1	643	561.9

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.

Instantaneous and Time-Varying Spectra— An Introduction

By

M. H. ACKROYD, B.Sc.†

Several definitions of 'time-varying', 'short-time' and 'instantaneous' spectra exist. The paper relates these to the time-frequency energy distribution of a signal and to the time-varying power spectrum of a non-stationary random process. The treatment emphasizes physical interpretation, rather than mathematical rigour.

List of Principal Symbols

$e(t, f)$	time-frequency energy density distribution of a signal
$e_s(t, f)$	version of $e(t, f)$ smoothed in the t -direction.
E	energy
E_T	total energy
f	frequency
Δf	frequency interval
$G_s(f)$	'short-time' spectrum according to Fano.
$h(t)$	impulse response
$H(f)$	transfer function corresponding to $h(t)$
$i(t)$	current
$p(t, f)$	'instantaneous spectrum' according to Page
$P(t, f)$	time dependent power spectrum of a random process
$q(t)$	power
$R(t, \tau)$	time dependent autocorrelation function
$\text{Re}(\cdot)$	'real part of (\cdot) '
$s(t)$	signal as a time function
$S(f)$	Fourier transform of $s(t)$
t	time
Δt	time interval
T	effective duration of a signal or an impulse response
$u(t)$	unit step function
W	effective bandwidth of a signal or filter
$X(f)$	reactance
$Y(f)$	admittance
$\overline{(\cdot)}$	'ensemble average of (\cdot) '

1. Introduction

For the usual purposes of time-invariant system analysis it is sufficient to consider a signal as a function of time alone or as a function of frequency alone. However, in certain situations (which are usually associated with the study of time-varying linear systems) one is concerned with both the time and the frequency characteristics of a signal at once. One such

situation is in the study of speech and the vocal mechanism.^{1,2} Another is in radar system theory where signals suffer both a time delay and a frequency shift between being transmitted and subsequently being received after reflection by a moving target.^{3,4}

Intuitively, it is evident that the energy of a signal does have a distribution in time and frequency in some sense. For example, the acoustic energy of a short blast on a whistle is 'obviously' located both at the frequency of the note which is blown and at the epoch in time when the whistle is sounded.

Many definitions of 'short-time', 'instantaneous' and time-varying power and energy spectra can be found in the literature.⁵⁻⁹ In addition various instruments have been constructed to measure short-term spectra.^{10,11} Thus further implicit definitions of short-term spectra couched in the mechanisms and parameters of these instruments have been introduced. The questions arise: How are these definitions related? Can they be regarded as approximations to some 'exact' or 'true' instantaneous spectrum? How are they to be interpreted in physical terms?

In what follows it is shown that an exact definition of $e(t, f)$, the energy density distribution in time and frequency of a signal, can be made. This definition applies whether the signal is deterministic, such as a pulse of specified shape, or whether it is random. If signals originate in a random process, then each individual signal from the process has its own energy distribution in time and frequency. The time-varying power spectrum of the random process, $P(t, f)$, is found by averaging all the possible energy density functions in accordance with their probability of occurrence.¹²⁻¹⁶ In other words, $P(t, f)$ is the average of $e(t, f)$ over the ensemble. $P(t, f)$ itself, for a well defined random process is not a random function any more than is, for example, the mean of the process.

The short-term spectrum measured according to a particular definition can be interpreted in two ways: it provides an approximation to the time-frequency energy density distribution of the signal and it can also provide an estimate of the time varying power spectrum if the signal is a random one.

† Department of Electronic and Electrical Engineering, University of Technology, Loughborough, Leicestershire.

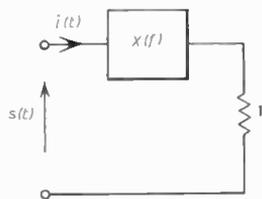
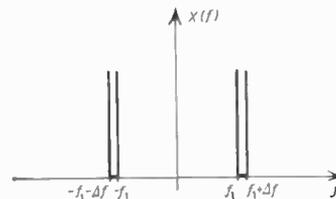


Fig. 1. Idealized circuit.

Fig. 2. Reactance of idealized circuit element.



2. Energy Density of a Signal in Time and Frequency

This section derives an expression for the distribution of the energy of a signal in the time-frequency plane.¹⁸

Replacing the real signal by the analytic signal† would have allowed the mathematics to have been made slightly more elegant and would have given substantially similar results. However, the conceptual difficulties involved in assigning physical meanings to the various results would have been increased (although this can be done in a satisfactory way). To minimize the difficulties of visualizing the results, the analysis is presented in terms of the real signal.

The definition of the energy density function can be understood by considering the circuit shown in Fig. 1. $X(f)$ is supposed to be a purely reactive circuit element having infinite reactance at all frequencies except over a narrow band where it has zero reactance. The reactance of this element is shown diagrammatically in Fig. 2. The admittance between the terminals of the circuit is

$$Y(f) = 1, \quad f_1 \leq |f| \leq f_1 + \Delta f$$

$$= 0, \quad \text{otherwise.} \quad \dots\dots(1)$$

Such an idealized admittance function is not physically realizable. However, it is nevertheless possible to calculate the current that would flow in the circuit in response to a given voltage waveform.

From the waveform of a signal, $s(t)$, the distribution of its voltage as a function of frequency is given by the Fourier transform

$$S(f) = \int_{-\infty}^{\infty} s(t) \exp(-j2\pi ft) dt \quad \dots\dots(2)$$

Conversely, the waveform is given by the inverse transform

$$s(t) = \int_{-\infty}^{\infty} S(f) \exp(j2\pi ft) df \quad \dots\dots(3)$$

Applying the signal $s(t)$ as a voltage to the terminals of the circuit, the current that flows only has frequency

components lying in the frequency range f_1 to $f_1 + \Delta f$. The waveform of the current is given by the inverse Fourier transform

$$i(t) = \int_{f_1}^{f_1 + \Delta f} S(f) \exp(j2\pi ft) df +$$

$$+ \int_{-f_1 - \Delta f}^{-f_1} S(f) \exp(j2\pi ft) df$$

$$= 2 \operatorname{Re} \left[\int_{f_1}^{f_1 + \Delta f} S(f) \exp(j2\pi ft) df \right] \quad \dots\dots(4)$$

The power entering the terminals of the circuit is given by the product of the applied voltage and the current which flows:

$$q(t) = s(t)i(t) \quad \dots\dots(5)$$

Thus the energy which enters the circuit between a time t_1 and $t_1 + \Delta t$ is given by

$$E = \int_{t_1}^{t_1 + \Delta t} s(t)i(t) dt \quad \dots\dots(6)$$

This quantity is defined as twice the part of the energy of the signal contained in the frequency range f_1 to $f_1 + \Delta f$ and in the time interval t_1 to $t_1 + \Delta t$ in the t - f plane. Thus the energy density at t_1 and f_1 is given by

$$e(t_1, f_1) = \lim_{\substack{\Delta t \rightarrow 0 \\ \Delta f \rightarrow 0}} \frac{1}{2} E / \Delta t \Delta f \quad \dots\dots(7)$$

Substituting equations (4) and (6) into (7) and writing t and f in place of t_1 and f_1 gives

$$e(t, f) = s(t) \operatorname{Re} [S(f) \exp(j2\pi ft)] \quad \dots\dots(8)$$

This is the required expression for the energy density distribution in time and frequency.†

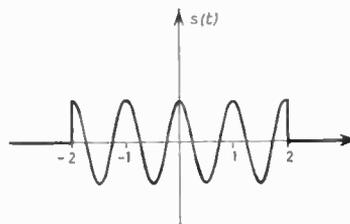


Fig. 3. Pulsed carrier signal of example. (Shown for $T = 4.0$, $f_0 = 1.0$.)

† The analytic signal^{3,5,19} is a complex function of time. Its relation to the real signal is simple; the analytic signal is derived by doubling the amplitudes of the positive frequency components and eliminating the negative frequency components of the real signal. As an example, the analytic signal corresponding to the real signal $A \cos \Omega t = A/2(\exp j\Omega t + \exp -j\Omega t)$ is $A \exp j\Omega t$.

† Reference 18 derives the complex energy density function of the analytic signal. The present definition appears in reference 8.

Example

The pulsed-carrier signal shown in Fig. 3 can be written (using Woodward's *rect* function^{3†})

$$s(t) = A \text{ rect}(t/T) \cos 2\pi f_0 t.$$

The corresponding energy density function is

$$e(t, f) = \frac{A^2}{4} \text{ rect}(t/T) \times [\cos 2\pi(f+f_0)t + \cos 2\pi(f-f_0)t] \times \left[\frac{\sin \pi(f+f_0)T}{\pi(f+f_0)} + \frac{\sin \pi(f-f_0)T}{\pi(f-f_0)} \right]$$

which is shown graphically in Fig. 4. Note that as T , the duration of the pulse, is increased the spread of the energy density on each side of the carrier is reduced. Eventually, as T becomes infinite and the signal becomes a pure cosine wave, the whole of the energy density becomes concentrated at the frequency of the carrier.

The example illustrates that, for values of t where the signal is zero, $e(t, f)$ is also zero. If the waveform of the example were applied to a bank of bandpass filters and square-law envelope detectors in the arrangement traditionally used to measure 'short-term' spectra, the output voltage of each filter would not immediately fall to zero at the end of the input pulse. Instead, it would die away gradually due to the response time of the filter. In fact, the function of time and frequency measured by a filter bank is an approximation to $e(t, f)$, as explained in Section 4 of this paper.

The function $e(t, f)$ is symmetrical about the time axis; the negative frequencies merely mirror the positive frequencies:

$$e(t, f) = e(t, -f).$$

This results directly from the symmetry properties of the Fourier transforms of real signals. Thus in plotting $e(t, f)$ it is only necessary to consider positive values of f .

Integrating $e(t, f)$ over all values of f gives the energy density (that is, the power) of the signal at time t . This is readily verified by integrating equation (8):

$$\int_{-\infty}^{\infty} e(t, f) df = s(t)^2 \quad \dots\dots(9)$$

Similarly, the energy density spectrum of the signal is got by integrating $e(t, f)$ over all time;

$$\int_{-\infty}^{\infty} e(t, f) dt = |S(f)|^2 \quad \dots\dots(10)$$

† The *rect* function is defined by

$$\text{rect}(t/T) = \begin{cases} 1, & |t| < T/2 \\ 0, & |t| > T/2. \end{cases}$$

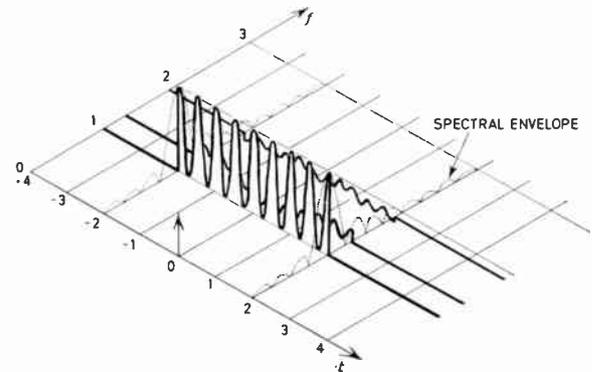


Fig. 4. Sketch of $e(t, f)$ of example. Note: Cross-sections at $f = 1.0, 1.3, 1.9$ are shown.

The total energy of the signal is thus found by integrating $e(t, f)$ over the whole time-frequency plane;

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e(t, f) dt df = E_T \quad \dots\dots(11)$$

There exists a misconception that it is not possible to measure exactly the t - f energy density function of a given waveform and that this is a consequence of Gabor's uncertainty relation.¹⁹ However, the uncertainty principle of waveform analysis is not concerned with the measurement of t - f energy density distributions; instead, it states that if the effective bandwidth of a signal is W then its effective duration cannot be less than about $1/W$ (and conversely, the bandwidth of a signal of effective duration T cannot be less than about $1/T$). In fact, given the waveform of a signal, its t - f energy density distribution can in principle be computed *exactly* using the defining relation (8). However, if $e(t, f)$ is measured *approximately* by the use of a bank of bandpass filters each having a bandwidth W , then it is evident that variations of $e(t, f)$ in the f -direction which are finer than W will be obscured. If the effective duration of the impulse response of each filter is T , then details of $e(t, f)$ in the t -direction which are finer than T will be obscured. The only sense in which the uncertainty principle applies to the measurement of t - f energy distributions is that it prohibits filters from having both short impulse responses *and* narrow bandwidths (as the impulse response of a filter, like any other signal, is subject to the $TW \geq 1$ uncertainty relation). Thus a bank of fixed bandwidth filters cannot provide both good spectral and good temporal resolution (although the resolution may be entirely adequate for practical purposes, of course).

Exactly parallel limitations on resolution apply when short term spectral analysis is performed by the process of multiplying the time waveform by contiguous 'time window' functions to produce a succession of short waveforms which are each subjected to

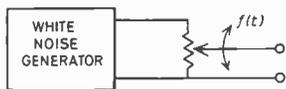


Fig. 5. Non-stationary (time varying) noise source.

Fourier analysis to produce a succession of cross-sections of a short-term spectrum. (The use of this procedure is described in detail in reference 20.)

3. Power Spectra of Non-stationary Processes¹²⁻¹⁷

A non-stationary process is one whose statistics vary with time. Here, one is interested in processes which are non-stationary in the wide sense, that is, with processes whose mean or whose autocorrelation vary with time. Figure 5 shows a model of a particular non-stationary process. Stationary white noise (from an ideal noise diode held at constant temperature, for example) is applied to a time-varying attenuator. Intuitively, one feels that the output signal has a time-varying spectral density. Many non-stationary processes can be represented as a time-varying network excited by a stationary random process.

The autocorrelation of a real process is defined

$$R(t, \tau) = \overline{s(t)s(t+\tau)} \quad \dots\dots(12)$$

The line above the product indicates that the ensemble average is to be taken, that is, the product is to be averaged over all possible pairs of values of $s(t)$ and $s(t+\tau)$ in accordance with their probability of occurring jointly. $R(t, \tau)$ exists for signals whose total energy is finite and also for signals whose total energy is infinite but whose average power is finite.

Taking the Fourier transform of equation (12) with respect to τ and denoting it by $P(t, f)$ gives

$$\begin{aligned} P(t, f) &= \int_{-\infty}^{\infty} R(t, \tau) \exp(-j2\pi f\tau) d\tau \\ &= \int_{-\infty}^{\infty} \overline{s(t)s(t+\tau)} \exp(-j2\pi f\tau) d\tau \quad \dots\dots(13) \end{aligned}$$

Assuming that the order of averaging and integrating can be interchanged, and that the F -transform of $s(t)$ exists,† equation (13) becomes

$$\begin{aligned} P(t, f) &= s(t) \int_{-\infty}^{\infty} s(t+\tau) \exp(-j2\pi f\tau) d\tau \\ &= \overline{s(t)S(f)} \exp(j2\pi ft) \quad \dots\dots(14) \end{aligned}$$

Equation (13) is the time-varying analogy of the Wiener-Khinchine relation, which states that the power spectrum of a stationary random process is the Fourier transform of its autocorrelation function.

† These assumptions are always satisfied for physical signals. The requirement for the Fourier transform of $s(t)$ to exist can be relaxed by the use of the generalized harmonic analysis.

$P(t, f)$ is generally a complex function, in contrast to the purely real power spectra of stationary processes. This happens as $R(t, \tau)$, unlike the stationary case, is not generally an even function in τ .

The real part of $P(t, f)$ is the ensemble average value of $e(t, f)$ as given by equation (8). The physical interpretation of this quantity is that $\text{Re} [P(t_1, f_1)]\Delta f$ is expected value of the power entering the circuit of Fig. 1 at time t_1 when $s(t)$ is applied as a voltage across its terminals. The imaginary part of $P(t, f)$ is of no direct interest here, but the following brief discussion explains its significance in a different context.

Had the analytic signal been used in the foregoing arguments, then $P(t, f)$ would have played an analogous part to the 'complex power' of a.c. phasor theory.²¹ Its real part would have represented the power entering the circuit of Fig. 1 and its magnitude the volt-amperes. As in the case of complex power, the imaginary part has little physical significance but is useful in analysis in that it provides the discrepancy between the real part and the magnitude of $P(t, f)$ in the correct way.

Relations similar to (9), (10) and (11) apply to $P(t, f)$.^{16, 17} The ensemble mean power of the signal is given by

$$\int_{-\infty}^{\infty} P(t, f) df = \overline{s(t)^2} \quad \dots\dots(15)$$

The expected energy of the signal, if it is finite, is given by

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} P(t, f) df dt = E_T \quad \dots\dots(16)$$

By integrating $P(t, f)$ with respect to time in the case of a signal of finite energy, or averaging over time in the case of a signal of finite power, it is possible to define a mean energy (or power) spectrum of the process. This is discussed in detail in references 16 and 17.

4. Short-term Spectra—Relations between Definitions

Many definitions of short-term spectra exist, as was mentioned in the Introduction. Some were introduced to provide Wiener-Khinchine-like relationships with corresponding short-term autocorrelation functions^{7, 9} and others to accord with the results of physical measurement, being cast in terms of the past and present values of the signal.^{6, 9} This Section shows that the various definitions can be regarded as approximations to $e(t, f)$.

The block diagram of the system which produces $e(t, f)$ from $s(t)$ is shown in Fig. 6. The relationship between this block diagram and equation (8) becomes

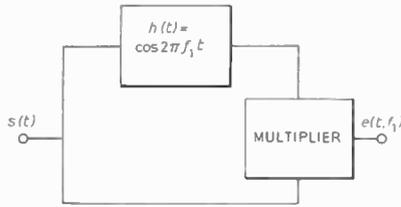


Fig. 6. Scheme for measuring $e(t, f)$.

clear when it is realized that $\text{Re} [S(f_1) \exp(j2\pi f_1 t)]$ is simply the response to $s(t)$ of the system whose impulse response is

$$h(t) = \cos 2\pi f_1 t \quad \dots\dots(17)$$

This is not a physically realizable system as its impulse response is non-zero for negative t . Nevertheless, its output can be calculated or it can be simulated.²⁴ The transfer function corresponding to the impulse response $h(t)$ is

$$H(f) = \frac{1}{2}[\delta(f+f_1) + \delta(f-f_1)] \quad \dots\dots(18)$$

which represents a filter having an infinitely narrow bandwidth (Fig. 7).

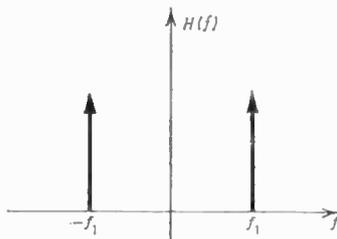


Fig. 7. Transfer function of 'ideal' filter.

4.1. Page's Definition

Page⁶ defined an 'instantaneous power spectrum' in terms of a 'running transform';

$$S_i(f) = \int_{-\infty}^t s(\tau) \exp(-j2\pi f\tau) d\tau \quad \dots\dots(19)$$

The running transform depends only on the past and present values of the signal and not on its future values. His instantaneous spectrum was defined as the rate of change with time of the squared magnitude of $S_i(f)$:

$$p(t, f) = \frac{\partial}{\partial t} |S_i(f)|^2 \quad \dots\dots(20)$$

Equations (19) and (20) can be manipulated into the alternative form

$$p(t, f) = 2s(t) \text{Re} [S_i(f) \exp(j2\pi f t)] \quad \dots\dots(21)$$

Now $\text{Re} [S_i(f_1) \exp(j2\pi f_1 t)]$ can be shown to be the response to $s(t)$ of a filter whose impulse response is

$$h(t) = u(t) \cos 2\pi f_1 t \quad \dots\dots(22)$$

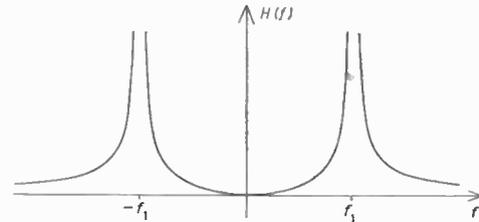


Fig. 8. Transfer function of physically realizable filter of equation (23).

where $u(t)$ is the unit step function. This impulse response is physically realizable in the sense that it is zero for negative t . The transfer function corresponding to this impulse response is

$$H(f) = jf/2\pi(f_1^2 - f^2) \quad \dots\dots(23)$$

whose magnitude is shown plotted in Fig. 8.

$p(t, f_1)$ is thus produced (except for a scale factor) by a system having the block diagram of Fig. 6 in which the 'ideal' filter of equation (18) is replaced by the less perfect (but realizable) filter of equation (23). $p(t, f)$ can therefore be thought of as an approximation to $e(t, f)$ with poorer spectral resolution.

4.2. Fano's Definition

Fano^{7,9} defined a 'short-time' power spectrum by the expression

$$G_t(f) = \left| \int_{-\infty}^t s(\tau) \exp [(-\alpha + j2\pi f)(t - \tau)] d\tau \right|^2 \quad \dots\dots(24)$$

(A scale factor which is not of interest here is omitted.) This definition was chosen by Fano as being the Fourier transform of a 'short-time' autocorrelation function. Equation (24) can be rewritten

$$G_t(f) = \left| \int_{-\infty}^t s(\tau) \exp [-\alpha(t - \tau)] \cos 2\pi f(t - \tau) d\tau \right|^2 + \left| \int_{-\infty}^t s(\tau) \exp [-\alpha(t - \tau)] \sin 2\pi f(t - \tau) d\tau \right|^2 \dots\dots(25)$$

Thus $G_t(f_1)$ is produced by the system whose block diagram is shown in Fig. 9. The signal is applied to two filters whose impulse responses are $u(t) \exp(-\alpha t) \cos 2\pi f_1 t$ and $u(t) \exp(-\alpha t) \sin 2\pi f_1 t$ and the outputs of these filters are squared and added to yield $G_t(f_1)$.

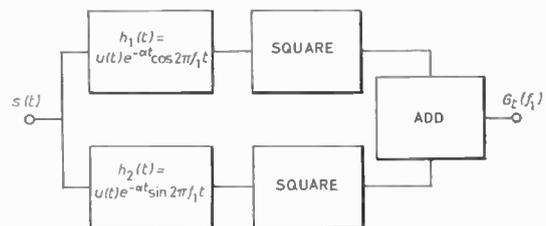


Fig. 9. Scheme for measuring $G_t(f)$.

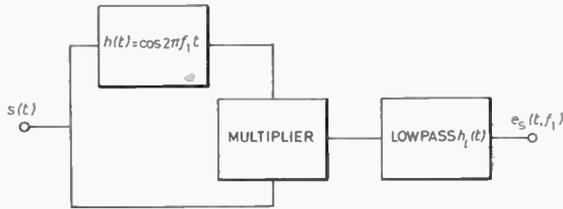


Fig. 10. Scheme for measuring $e_s(t, f)$ ($e(t, f)$ smoothed in the t -direction).

Suppose that in measuring $e(t, f)$ the fine structure in the t -direction is of no interest. This irrelevant detail can be removed by smoothing $e(t, f)$ in the t -direction with a low-pass filter. Figure 10 shows the block diagram of a system to produce $e_s(t, f_1)$, the smoothed version of $e(t, f_1)$. This is simply the system of Fig. 6 which produces $e(t, f_1)$ with the addition of a low-pass filter at its output. The smoothed version of $e(t, f)$ can be expressed by the convolution integral

$$e_s(t, f) = \int_{-\infty}^{\infty} e(\tau, f)h(t-\tau)d\tau \quad \dots\dots(26)$$

where $h(t)$ is the impulse response of the low-pass filter. Equation (26) becomes, on substituting the defining relation for $e(t, f)$,

$$\begin{aligned} e_s(t, f) &= \text{Re} \left[\int_{-\infty}^{\infty} S(f) \exp(j2\pi f\tau)s(\tau)h(t-\tau)d\tau \right] \\ &= \text{Re} \left[S(f) \exp(j2\pi ft) \times \int_{-\infty}^{\infty} s(\tau) \exp[-j2\pi f(t-\tau)]h(t-\tau)d\tau \right] \quad \dots\dots(27) \end{aligned}$$

$e_s(t, f)$ can thus also be calculated by the scheme shown in Fig. 11, which is the block diagram representation of equation (27). The similarity between Fig. 9 and Fig. 11 becomes exact if, in Fig. 11,

- (i) the filters having impulse responses $\cos 2\pi ft$ and $\sin 2\pi ft$ are replaced by filters having impulse responses $u(t) \exp(-\alpha t) \cos 2\pi ft$ and $u(t) \exp(-\alpha t) \sin 2\pi ft$, respectively; and
- (ii) the impulse response of the low-pass filter, $h(t)$, is given by

$$h(t) = u(t) \exp(-\alpha t)$$

From the similarity of these block diagrams it is evident that Fano's definition of the short-term spectrum, $G_t(f)$, can be thought of as an approximation to $e(t, f)$.

4.3. Other Definitions

The definition of $e_s(t, f)$ given above (in which $h(t)$ can be any low-pass filter impulse response) was termed by Schroeder and Atal⁹ a 'short-time spectrum

of the second kind'. This definition was used by them as it is the Fourier transform of a corresponding short-time autocorrelation function.

Their 'short-time spectrum of the first kind' represents a generalization of the short-term spectrum of Fano in which the simple band-pass filters of Fig. 9 are replaced by band-pass filters of a more general form. Their 'short-time spectrum of the third kind' is measured by a system similar to that used in measuring $e_s(t, f)$ in which the ideal band-pass filter is replaced by a physically realizable one. All these definitions thus provide approximations to $e(t, f)$.

The effect of using non-ideal band-pass filters in the various short-term spectrum measuring systems is to give reduced spectral resolution which is equivalent to a smoothing of $e(t, f)$ in the f -direction. The various definitions of short-term spectrum thus correspond to modifications of $e(t, f)$ made by smoothing in the t - and f -directions with various weighting functions.

5. Measurement of Short-term Spectra

Roughly speaking, situations in which a short-term spectrum is to be measured can be placed in two categories, although the division between categories is by no means clear. In one category, the signal itself is the prime object of interest, for example in sonar signal design. In the other category, the principal object of study is not the signal as such, but rather the process from which it originates. Examples here are the spectrographic study of the speech forming process and seismology.

In the first category, the t - f energy distribution is required, perhaps in the finest detail possible. The finest possible detail is provided by computing $e(t, f)$ directly from its defining relation.† However,

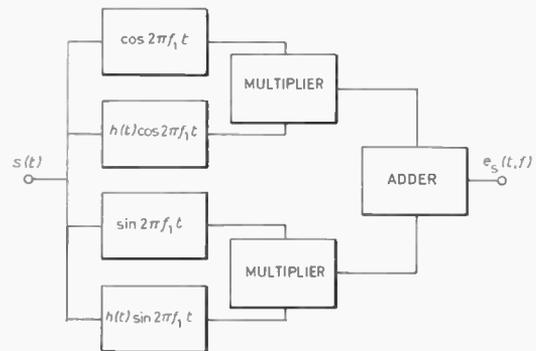


Fig. 11. Alternative scheme for measuring $e_s(t, f)$.

† This can be done economically by using the fast Fourier transform algorithm. Reference 22 is an entire issue devoted to this subject; reference 23 gives a FORTRAN sub-routine.

the fine structure of $e(t, f)$ (such as its fluctuations over a period of one or two cycles of the frequency concerned) may be of no interest. Then a degree of smoothing can be provided in the t -direction by an appropriate choice of $h(t)$ in computing $e_s(t, f)$ as given by equation (26).

In the second category of situations, one often wishes to estimate the power spectrum of a time-varying random process. If $P(t, f)$ changes with time at a rate which is of the same order as the rate at which the signals fluctuate, there is little that can be done with a single signal. With a set of signals from the process, however, $e(t, f)$ can be computed for each member of the set and the average taken to yield an estimate of $P(t, f)$.

If the time variation of $P(t, f)$ is slow compared with the rate of fluctuation of the signal then $P(t, f)$ can be estimated by smoothing $e(t, f)$ calculated for a single signal from the random process. However, at the present time the question of what the best smoothing operation is in a particular situation is not well understood. Ideally, procedures would be available by which the optimum smoothing operation could be chosen on the basis of the available *a priori* knowledge of the process being studied. (For example, in measuring the spectrum of the noise from a vehicle moving past a fixed microphone at a particular speed certain features of the non-stationary noise process such as the rate at which the level changes, the total Doppler shift, etc., are known beforehand.) Developing such procedures is a matter for further research.

In the absence of systematic methods for choosing the smoothing operation to be performed on $e(t, f)$ it is necessary to rely on the intuition of the investigator. In short-term spectrum analysis as it is usually done (using a filter bank or by Fourier analysis of weighted sections of the signal), the smoothing operation is implicitly chosen by the investigator when he chooses filter bandwidths (and so on) so that the measured short-term spectrum appears 'best' according to his subjective judgement.

6. Conclusion

The intention of this paper has been to serve as an introduction to the concepts of the t - f energy density distribution and the time-varying power spectrum and to clarify the physical meaning of the results of practical short-term spectrum measurement.

The idea of the t - f energy density distribution of a signal has been used to relate the various definitions of 'short-time' and 'instantaneous' spectra and to interpret them in physical terms. The t - f energy density distribution of a signal can in principle be calculated exactly if its waveform is known. The uncertainty

principle of waveform analysis makes no restriction on the accuracy with which $e(t, f)$ can be computed. The only sense in which it does apply is that it restricts the t - f plane resolution that can be obtained when short-term spectrum measurement is made by a bank of bandpass filters or by an analogous method.

The time-varying power spectrum, $P(t, f)$, of a random process can be regarded as being the average of the t - f energy density functions of the individual signals from the process. A smoothed version of a measured $e(t, f)$ function can provide an estimate of $P(t, f)$, but as yet rules which specify the 'best' smoothing operation for a given application do not exist. Thus the design of a short-term spectrum analysis remains a matter in which heavy reliance must be placed on intuition.

7. References

- Potter, R. K., Kopp, G. A. and Green, H. C., 'Visible Speech'. (Van Nostrand, New York, 1947)
- Flanagan, J. L., 'Speech Analysis, Synthesis and Perception'. (Springer-Verlag, New York, 1965.)
- Woodward, P. M., 'Probability and Information Theory, with Applications to Radar' (2nd edition). (Pergamon, London, 1964.)
- Benjamin, R., 'Modulation, Resolution and Signal Processing in Radar, Sonar and Related Systems'. (Pergamon, London, 1966.)
- Ville, J., 'Théorie et applications de la notion de signal analytique', *Câbles et Transmission*, 2, pp. 61-74, 1948.
Translation available by I. Selin, 'Theory and applications of the notion of complex signal', Report T-92, RAND Corporation, Santa Monica, California, USA.
- Page, C. H., 'Instantaneous power spectra', *J. Appl. Phys.*, 23, pp. 103-106, 1952.
- Fano, R. M., 'Short-time autocorrelation functions and power spectra', *J. Acoust. Soc. Amer.*, 22, pp. 546-550, 1950.
- Levin, M. J., 'Instantaneous spectra and ambiguity functions', *Trans. I.E.E.E. on Information Theory*, IT-10, pp. 95-97, 1964.
- Schroeder, M. R. and Atal, B. S., 'Generalised short-time power spectra and autocorrelation functions', *J. Acoust. Soc. Amer.*, 34, pp. 1679-1683, 1962.
- Ehrich, W. G. and Steinberg, B. D., 'Panoramic spectrum analysis in real time', *I.R.E. Nat. Conv. Rec.*, 7, Part 2, pp. 62-69, 1959.
- Gill, J. S., 'A versatile method for short-term spectrum analysis in "real-time"', *Nature*, 189, pp. 117-119, 1961.
- Lampard, D. G., 'A generalisation of the Wiener-Khinchine theorem to non-stationary processes', *J. Appl. Phys.*, 25, pp. 802-803, 1954.
- Urkowitz, H., 'Pre-envelopes of non-stationary bandpass processes', *J. Franklin Inst.*, 277, pp. 31-36, 1964.
- Bendat, J. S. and Piersol, A. G., 'Measurement and Analysis of Random Data'. (Wiley, New York, 1966.)
- Silverman, R. A., 'Locally stationary random processes', *Trans. I.R.E., on Information Theory*, IT-3, pp. 182-187, 1957.

16. Karkevich, A. A., 'Spectra and Analysis' (Translated from the Russian). (Consultants Bureau, New York, 1960.)
17. Papoulis, A., 'Probability, Random Variables and Random Processes'. (McGraw-Hill, New York, 1965.)
18. Rihaczek, A. W., 'Signal energy distribution in time and frequency', *Trans. I.E.E.E. on Information Theory*, IT-14, pp. 369-374, 1968.
19. Gabor, D., 'Theory of communication', *J. Instn Elect. Engrs*, 93, Part 3, pp. 429-457, 1946.
20. Singleton, R. C. and Poulter, T. C., 'Spectral analysis of the call of the male killer whale', *Trans. I.E.E.E. on Audio*, AU-15, pp. 104-113, 1967.
21. Thorn, D., 'An Introduction to Generalised Circuits', p. 114 *et seq.* (Wadsworth, Belmont, Calif., 1963.)
22. Special Issue on Fast Fourier Transform, *I.E.E.E. Trans. on Audio and Electroacoustics*, AU-15, pp. 43-117, June 1967. (This issue contains reference 20.)
23. Robinson, E. A., 'Multichannel Time Series Analysis with Digital Computer Programs'. (Holden-Day, San Francisco, 1967.)
24. Storey, L. R. O. and Grierson, J. K., 'Time-symmetric filters' (part 1), *Electronic Engng*, 30, pp. 586-592, 1958.

Manuscript first received by the Institution on 26th June 1969 and in final form on 29th October 1969. (Paper No.1308/CC67.)

© The Institution of Electronic and Radio Engineers, 1970

LETTERS

Advertisements in the Journal

Whilst agreeing with Mr. Shepherd's observations on the ethical reasons for banishing advertisements from the *Journal* (see his letter, page 112, February 1970), I feel he severely underrates advertisements as the best means of: (1) concisely disseminating useful information, and (2) breaking up the format to give the *Journal* greater eye-appeal.

As a consultant in the world of banking, I am only too painfully aware of the harmful effects of brash advertising. Not uncommonly the pages of some periodicals as distinguished as our *Journal* prove, on careful analysis, to be carrying advertisements that are not all they seem—which proves how difficult the task must be for those whose responsibility it is to enforce minimum standards. But there is one good rule: any advertisement manager who feels he is compelled to use lavish superlatives or claims to novelty (except in an obviously humorous context), or depends heavily for appeal on sex portrayal, is effectively advertising a weakness in the product or service he is paid to promote!

Typographically imaginative and factual advertisements enhance appearance by resting the eye from pages of uniform type and—often with advantage—jog one's memory. Even the Russians discovered that a journal without advertisements had very low appeal!

All that I would beg is that advertisement promoters should be given far more encouragement to book a series of spaces in which the text of each advertisement is different. Britain has set a standard for Europe in the

appeal of light-hearted and informative advertising, and it will be a commendable advantage if British electronic exports are promoted by this type of British media.

F. P. THOMSON

39 Church Road,
Watford WD1 3PY, Herts.

26th February 1970.

Although I am clearly an interested party, I suspect that the cloistered and superior attitude of Mr. Shepherd is shared by few members of our Institution. Whatever our personal feelings about advertising—and that much-abused word certainly covers a multitude of sins, and virtues, in the electronics field—it is hard to argue against the many readership surveys that have examined the noting and recall of material appearing in engineering journals. They show conclusively that engineers remember editorial and advertising content about equally; this would hardly be possible if we all abhorred advertising with Mr. Shepherd's passion. It appears, in fact, that we rely on good (i.e. clear and informative) advertising for much of our knowledge about what is actually happening in our own industry.

Only if Mr. Shepherd believes that the industrial side of electronics (as opposed to research and education) is irrelevant or sordid do his strictures against advertising in the *Journal* make sense.

ALFRED L. COTCHER

Howard Panton Limited,
Panton House, Howard Street,
London, W.C.2.

27th February 1970.

A Function Generator using Hybrid Techniques

By

R. N. BARNES,

B.Sc.(Eng.)†

The principle of operation of a novel function generator is outlined and some applications are reviewed. The basic function produced by the generator is a 'box-car', or discrete, type function having fixed thresholds with respect to the input voltage, and output voltage levels that may be independently set in any segment. By using a sampling technique, linear interpolation may be applied across any of the discontinuities in the 'box-car' function to produce an 'analogue', or continuous type function. This arrangement allows for the independent adjustment of the value of the function for input voltages midway between the threshold voltages. In the discrete mode, good accuracy is maintained for full amplitude input swing up to 80 kHz and in the continuous mode up to 5 kHz.

1. Introduction

Variable function generators have found considerable application in general-purpose analogue computers for the simulation of non-linear functional relationships, but their application in other fields, such as complex waveform generation for test purposes, has been somewhat limited in the past. This was because of the complexity involved in setting up the desired function, the limitation of the number of segments that may be employed, and the fact that only continuous functions may be generated. These limitations arise directly from the conventional method of function generation using diode resistor networks, and a significant improvement in versatility and simplicity of operation can only be achieved by utilizing a fundamentally different technique of function generation. The generator described in this paper radically departs from conventional techniques and overcomes the above limitations.

The normal techniques of function generation employing diode resistor networks in either input or feedback paths of an operational amplifier are well known,^{1,2} and lead to a piecewise linear approximation of the desired function as shown in Fig. 1(a). The most common method of synthesizing this function involves the addition of the individual functions $F_1 \dots F_4$ shown in Fig. 1(b). The breakpoints, $V_1 \dots V_4$, and the slopes of these individual functions are usually variable. It is clear that with this technique, the slope of any one segment of the total function is the sum of all of the slopes of the individual functions to the left. It also follows that adjustment of slope in any one segment alters the slopes of all segments to the right of that particular segment. In short considerable interaction occurs.

It should also be noted that the absolute value of the output of the function is only defined at the most negative breakpoint voltage. The above effects lead to complex setting-up procedures, and limit the number

of segments that may be employed since any drifts that occur cause large errors at the extreme right of the function due to cumulative action.

Miura *et al.*³ have described a method where the slopes of the segments in the total function may be set independently. This is achieved by summing individual functions of the form shown in Fig. 1(c), and leads to a simpler setting-up procedure, and is less susceptible to the effects of drift. However the mechanization of the individual functions $G_1 \dots G_4$ requires matched dual-ganged potentiometers, which is a distinct disadvantage.

The function generator described in this paper employs techniques that allow the absolute value of the synthesized function to be set independently (i.e. without interaction) at particular input voltages. This approach eliminates the effects of drift and results in an extremely simple setting-up procedure. If, for example, calibrated potentiometers are used, then functions may be set up directly without the need of applying test voltages, and if the potentiometers are rectilinear types and are mounted adjacently and in correct sequence, then the shape of the function is graphically displayed by the potentiometer indicators. There is no limit to the number of segments that may be employed, and the method allows for discontinuous as well as continuous functions to be simulated.

2. Principle of Operation

The generator basically produces the discrete type functional relationship between input and output as shown by the broken line on Fig. 2 and the amplitudes of the equal width segments may be set independently. Selective interpolation may be applied to the discontinuities of the basic function to produce a continuous functional relationship as shown by the continuous line on Fig. 2. It may therefore be appreciated that discontinuities in otherwise continuous functions can be arranged by omitting interpolation at the required input level. (An example of this is shown in Fig. 8.)

† West Ham College of Technology, Electrical Engineering Department, London, E.15.

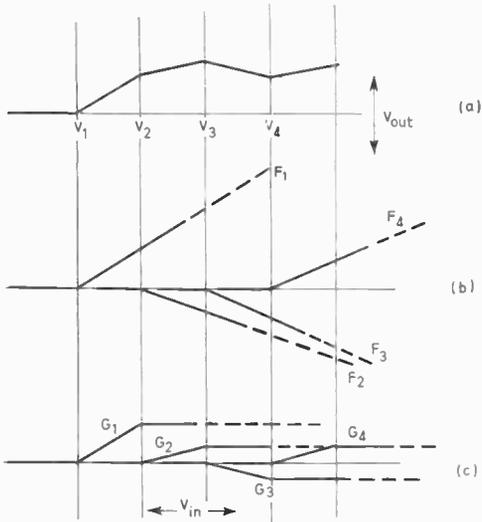


Fig. 1. Action of conventional diode resistor function generators.

It will be noticed on Fig. 2 that the breakpoints of the continuous function occur at the mid-point of the discrete segments, and that the amplitude, or ordinate, of the function at these points equals the amplitude of the basic discrete segment. The setting-up of the generator is therefore much more simple than with conventional types, since not only are the ordinates of continuous functions set independently but they are directly related to the potentiometers that set the discrete function.

The prototype generator is capable of producing 13 discontinuities corresponding to the input levels $V_1 \dots V_{13}$ of Fig. 2, and thus 12-segment discrete functions or 13-segment continuous functions can be generated. A partial block diagram for three discontinuities is shown in Fig. 3.

The input voltage is applied to 13 high-speed comparators, $C_1 \dots C_{13}$ which are biased to switch when the input voltage is $V_1 \dots V_{13}$ respectively. Let

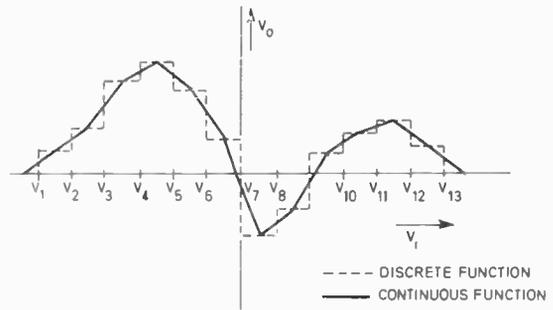


Fig. 2. Action of new function generator.

the logic outputs of these comparators be denoted $A_1 \dots A_{13}$.

In general therefore the output of the n th comparator is

$$A_n = 1 \text{ if } V_i > V_n.$$

The comparator outputs are connected to the logic elements which generate the function given by

$$B_n = A_n \cdot \overline{A_{n+1}} \text{ (when } I_1 = 1, \text{ and } I_2 = 1)$$

Thus the output of the logic following the n th comparator is at logic 1 when the input voltage is given by

$$B_n = 1, \quad V_n < V_i < V_{n+1} \text{ (} I_1 = 1, I_2 = 1)$$

The voltage corresponding to the B_n signal is then adjusted by the potentiometer K_n whose output D_n is the amplitude of the discrete segment corresponding to an input voltage described by

$$V_n < V_i < V_{n+1}$$

The voltages $D_1 \dots D_{13}$ are then summed, with inversion if necessary, to produce the discrete characteristic of the form shown by the broken line in Fig. 2.

The generator has a number of important applications in the discrete mode of operation and some applications are reviewed in Section 5.2.

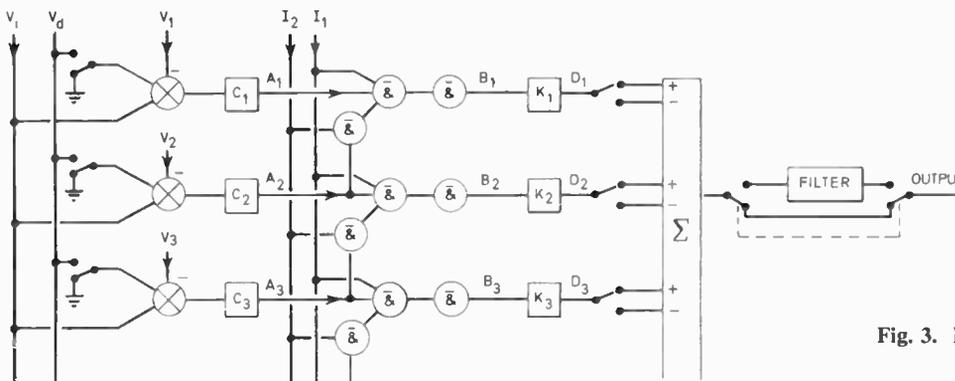


Fig. 3. Partial system block diagram.

2.1. Interpolation

To obtain interpolation across a discontinuity, a high-frequency triangular wave of peak-to-peak amplitude equal to the discrete segment width is added to the input of the comparator corresponding to that discontinuity. The action of the interpolation voltage (V_d , see Fig. 4) is to sample the magnitudes of the discrete segments alternately on either side of the discontinuity. This produces a rectangular wave at the output with high and low voltage levels equal to the respective discrete segment magnitudes, and with a mark/space ratio dependent on the instantaneous value of the input voltage relative to the comparator bias voltage. The average value of the rectangular wave so produced lies on a straight line drawn between the mid-points of the two discrete segments.

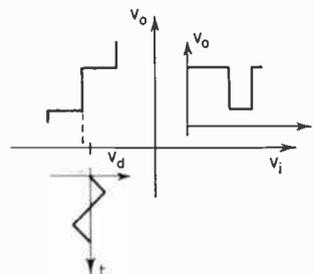


Fig. 4. Development of the rectangular waveform from the interpolation voltage.

2.2. Logic Inhibiting

The logic circuits associated with each comparator have additional inhibiting inputs. In the normal mode of operation previously outlined, the inhibit lines are set to

$$I_1 = 1 \quad \text{and} \quad I_2 = 1$$

Reference to Fig. 3 shows that if

$$I_1 = 1 \quad \text{and} \quad I_2 = 0$$

the output of the n th logic element is given by

$$B_n = A_n$$

Under these circumstances, the output of the generator, for a given input voltage, is the cumulative sum of the settings of all segments in the region V_1 up to V_i . In general therefore the output voltage is given by

$$V_0 = \sum_{r=1}^n GLK_r$$

where G is the operational gain of the summing amplifier,

L is the logic 1 voltage,

K_r is the potentiometer setting of the r th segment,

n is the number of the segment corresponding to the input voltage.

This expression for the output voltage should be compared with that for the normal independent mode, namely

$$V_0 = GLK_n$$

In the cumulative mode the potentiometers set the change in level between segments rather than the absolute level as in the independent mode, the setting-up procedure is therefore more complicated. However, the cumulative mode does have application for discrete functions as less switching is involved, and consequently higher speeds are possible.

When the I_1 inhibit input is set to logic 0 the outputs of all the logic elements $B_1 \dots B_{13}$ are at logic 0 irrespective of the input voltage to the generator and of the state of the I_2 inhibit line. Under these circumstances the output of the generator will be zero. This is a very valuable feature and finds considerable application in the generation of pulse shapes as outlined in Section 5.1.

3. Bandwidth

The question of bandwidth raises some interesting points in this form of function generator. It is obvious that completely different situations exist for the continuous and discrete forms of operation. As far as the discrete mode is concerned, the performance can be related to the rise-time of the discontinuities in the function. However, since the switching actions that occur as the input voltage varies do not depend on positive feedback (i.e. on a bistable action) but rather on the gain (including the logic elements) between input and output, the rise-time depends on the rate of change of the input signal. In the limiting case, namely when the rate of change of input signal is very high, the summing amplifier will limit the rise-time. In the present generator this is better than 100 ns and thus 'good' pulses of 0.5 μ s width may be obtained in the discrete mode with independent setting, and slightly better pulses of the same duration with cumulative setting. Thus for a 12-segment generator, the maximum input frequency corresponds to a half-period time of about 6 μ s, i.e. approximately 80 kHz bandwidth.

In the continuous mode of operation, bandwidth considerations are not so straightforward as sampling action is involved in the interpolation process. The ultimate bandwidth obtainable will depend on the rise-time of the comparators and associated logic and summation, since the rise-time dictates the maximum interpolation frequency that may be used for a given linearity of interpolation. The maximum bandwidth possible in a sampling process (without ambiguities occurring) is $f/2$; however, as filtering is required to

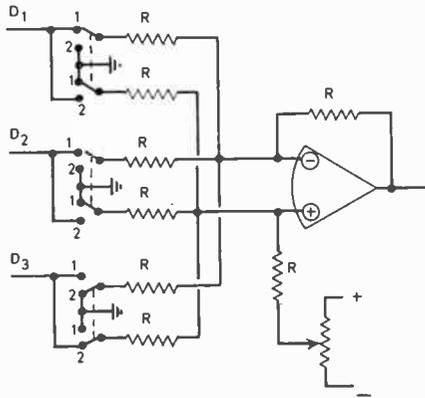


Fig. 5. Bipolar summing amplifier.

extract an average, the maximum bandwidth is likely to be limited to $f/20$ or even less. As the rise-time of the comparator-logic-summer chain is better than 100 ns, then reasonable pulses (for averaging purposes) of 200 ns duration can be achieved. If the periodic time of the interpolating triangular wave is made 10 times greater than the minimum pulse length of 200 ns, it is possible to obtain good linearity up to 10 : 1 mark/space ratio in the sampling process. This then yields linear interpolation over at least 90% of the segment width.

In the prototype generator the interpolation frequency is about 450 kHz, and thus after filtering, the maximum practical bandwidth is about 20 kHz. However the actual bandwidth has been set to approximately 5.0 kHz.

4. Circuitry

Integrated circuits are used exclusively in the instrument, and the comparator and logic arrangements are completely standard and do not require further description. The summing, filtering and high frequency triangular wave frequency generator elements are functionally more complex and further details are given below.

4.1. Summation

Figure 5 shows the arrangement adopted for summing (with or without inversion) the signals $D_1 \dots D_{12}$. Inversion and summation of an input is obtained when the switch for that signal is in position 1. The gain in the channel in this case is -1 .

Non-inversion is obtained when the switch is in position 2. In this case there is an attenuation to the positive input terminal of the amplifier such that,

$$e_+ = \frac{1}{1+n} D_r$$

by virtue of the potential divider action.

However, the gain of the amplifier for signals applied to the positive input is

$$V_o = (n+1)e_+$$

Thus the overall gain for non-inversion of a signal is again unity.

This is a convenient way of achieving 'bi-polar' summation as no great demands are placed on the operational amplifier for the following reasons:

- (i) The common mode range does not have to be high since the maximum voltage that can occur at the positive input terminal is $1/n$ when the generator is in the independent mode.
- (ii) Although the gain in any one channel is unity, the feedback fraction is $1/(1+n)$ and stabilization presents few difficulties. The large signal bandwidth of most operational amplifiers depends on the output slew-rate limit and this invariably depends directly on the degree of stabilization (either lag or lead) that is required. Thus since the degree of stabilization required is small, a relatively high large signal bandwidth is achieved.
- (iii) The source resistance of both input terminals of the amplifier is $R/(1+n)$ and thus the effects of bias current on output offset are negligible and temperature effects are also small since the source resistances are relatively small if n is large.
- (iv) Since the source resistances are small the input resistance of the operational amplifier does not have to be high.

4.2. Filter

The gain of the summing amplifier configuration is unity and since the maximum input signal to any channel is L volts, i.e. about 4 V, the maximum output of the summer will also be about 4 V, and this is the maximum output when the generator is used in the discrete mode.

For generation of continuous functions, the output of the summer must be filtered, and is amplified to give a peak output of approximately 12 V. The filter and final amplifier are shown in Fig. 6. The

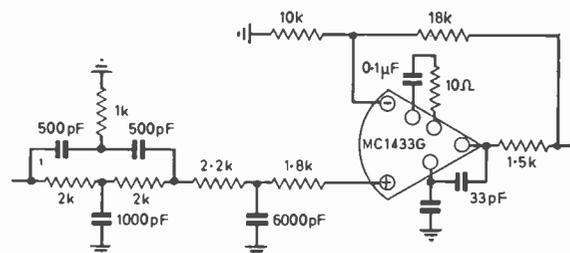


Fig. 6. Filter and final amplifier.

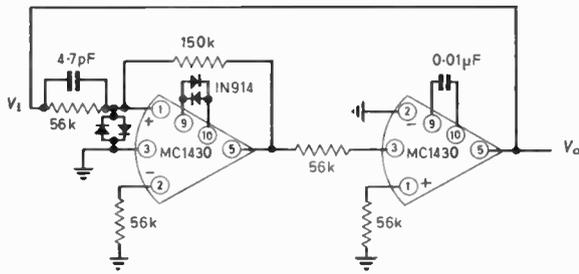


Fig. 7. Triangular wave generator.
 R_1 56 kΩ input resistor; R_2 150 kΩ feedback resistor.

twin-T network provides a notch at the sampling frequency, and this is followed by a first-order lag filter and a non-inverting gain of $\times 3$ is provided by the operational amplifier. The output of the amplifier is buffered by a 1.5 kΩ output resistor to render the output proof against short circuits.

4.3. *Triangular Wave Generator*

The standard method of using an integrator and a defined hysteresis switch in a closed loop has been adopted in the oscillator (see Fig. 7).

The hysteresis switch operates when the voltage at the input terminal of the amplifier is zero, i.e. when

$$\frac{V_i}{R_1} = \frac{V_o'}{R_2}$$

where V_o' is the positive or negative saturated output voltage of the amplifier, and is ± 4.9 V.

The threshold levels are therefore ± 1.8 V and the hysteresis width is therefore 3.6 V.

The integrator in the circuit is in fact an open-loop operational amplifier with the output slew-rate defined by 0.01 μF capacitor across terminals 9 and 10 which are the collectors of the first stage differential amplifier. For the MC 1430 device the output slew-rate is characterized for the capacitance between terminals 9 and 10, and is 3.0 V/μs for 0.01 μF.

The half-period time of the oscillator is the time required for the output of the integrator to slew across

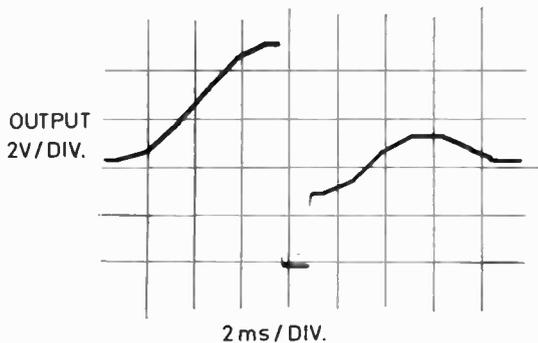


Fig. 8. An arbitrary function with two discontinuities.

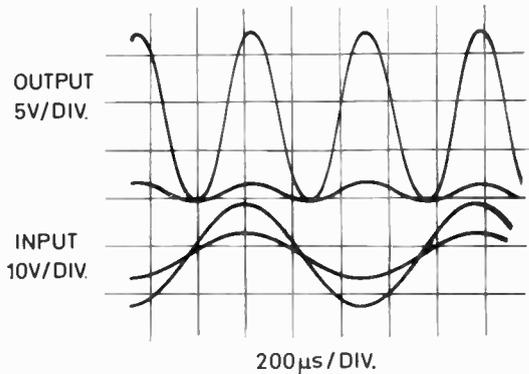


Fig. 9. Showing the frequency doubling action of a square-law function.

the hysteresis width thus

$$\frac{T}{2} = 1.2 \mu s.$$

The frequency of the oscillator is therefore calculated to be approximately 425 kHz, and was found to be 450 kHz.

5. Applications of the Generator

5.1. *Continuous Mode*

5.1.1. *Non-linear function*

The generator may be used to simulate non-linear functions with discontinuities if required, as described in Section 2. A typical response is shown in Fig. 8. Figure 9 shows the response of the generator to a sinusoidal input with the generator set up to a square-law function, the frequency doubling action being clearly apparent.

5.1.2. *Waveform generation*

If the input to the generator is a triangular waveform, then the waveform at the output will be the same shape as the function set up on the generator for the positive half-cycle of the input waveform. The negative-going half-cycle of the input also generates the

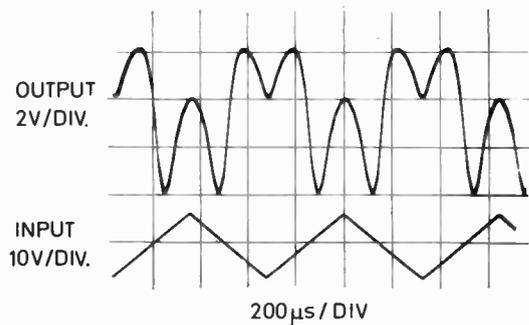


Fig. 10. Generation of an arbitrary waveform.

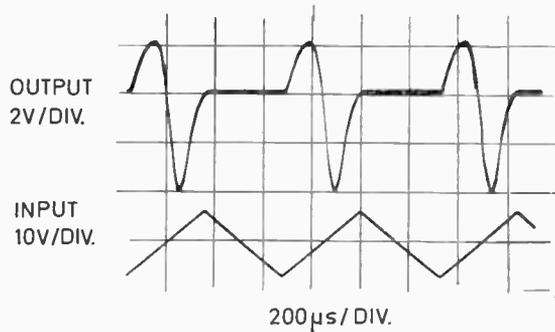


Fig. 11. Generation of an arbitrary pulse. This is the same setting as for Fig. 10, but with the output suppressed for negative-going input signals.

shape of the set function but in a reverse direction. The generated waveform will therefore be symmetrical about the point in time corresponding to the peak of the input. Figure 10 shows an example of arbitrary waveform generation from a triangular wave input.

5.1.3. Pulse generation

Repetitive shaped pulses may be generated using similar techniques to those in the previous application but with the output of the generator suppressed during either the positive or negative-going half-cycle of the drive waveform. The majority of triangular waveform generators simultaneously produce a synchronous square wave, and this may be connected either directly or through an inverter to the I_1 inhibit input of the function generator to suppress the output during the required half-cycle.

The pulse shown in Fig. 11 was obtained with the generator set as for the response shown in Fig. 10 but with the output suppressed during the negative-going half-cycle of the input.

Figure 12 shows another example of pulse generation where the pulse has the form

$$\frac{\sin x}{x}$$

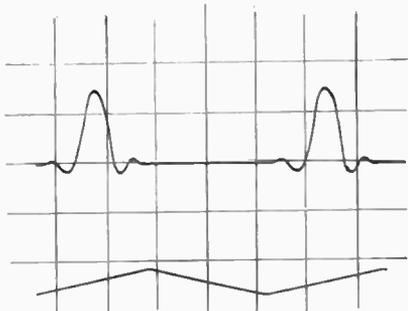


Fig. 12. $\sin x/x$ pulse.

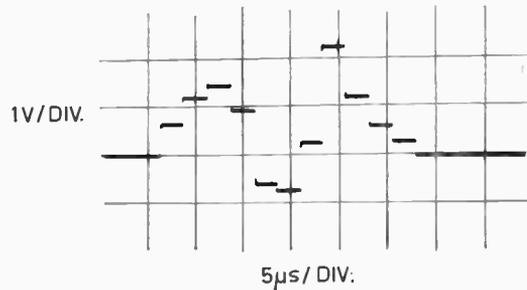


Fig. 13. A pulse amplitude sequence.

On examining the response shown on Fig. 11, it is clear that if the input to the generator was a saw-tooth waveform with fast flyback, then the pulse would be expanded out into a continuous waveform. This technique of waveform generation is not restricted to 'symmetrical' type waves as is the technique previously discussed.

5.2. Discrete Mode

Figure 13 shows an example of a pulse amplitude sequence of approximately $25 \mu s$ duration at a repetitive rate of 1 kHz. This was achieved by driving the input of the generator with a truncated triangular wave so that the input window of the generator was scanned in a time very much shorter than the half-period of the input, the reverse sequence being suppressed as described before. By generating this sequence with the instrument in the independent mode, the amplitude of any individual pulse may be adjusted without affecting the amplitudes of the other pulses.

5.2.1. Quantization

If the outputs of successive segments increase by equal increments, the generator may be used as a quantizer, and an example of this application is shown in Fig. 14.

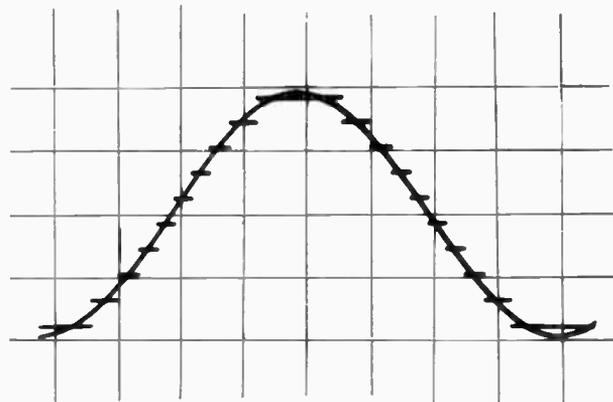


Fig. 14. The quantization of a 1 kHz sine wave.

6. Some Possible Extensions

All of the applications reviewed in the previous section are capable of being achieved on the prototype instrument described in Section 2. However, certain not very extensive additions to the basic generator can yield some useful additional functions.

If the logic outputs $B_1 \dots B_{12}$ are made externally available, the following applications become possible.

6.1. Pulse Height Analyser

If a pulse is applied to the input of the generator, one of the $B_1 \dots B_{12}$ logic signals will be at logic 1 for the duration of the pulse, indicating that the magnitude is between the voltage levels corresponding to that channel. The I_1 inhibit input may be used to strobe the input pulse if necessary.

6.2. Amplitude Density Function Measurement

If a random signal is connected to input of the generator, the smoothed output of each of the B logic signals gives the amplitude density of the input signal corresponding to the mean bias level of that input channel. The collective results of all of the outputs yield the amplitude density function. The cumulative density function will be given if the generator is used in the cumulative mode.

6.3. Two Functions of a Variable

By duplicating the potentiometers, summer and filter, it is obviously possible to generate two functions of the input variable. One possible application of this is the sequential positioning of an electron beam in

beam welding processes. The generator would allow the independent adjustment of the X and Y coordinates of the beam at any point in the sequence.

6.4. External Control of the Function

The ordinates of the generated function are set by adjusting the potentiometers which effectively control the amplitude of the B_1 to B_{13} signals. Now the level of these logic signals could alternatively be set by external voltages controlling simple clamps. Thus the generated function, whether it be continuous or discrete, could be controlled by external voltages. This might find application in adaptive control schemes.

7. Acknowledgment

The author wishes to express his thanks to Messrs. Feedback Ltd for supporting the development of the Function Generator.

8. References

1. Jackson, A. S., 'Analogue Computation' (McGraw-Hill, New York, 1960).
2. Tomovic, R. and Karplus, W. J., 'High Speed Analogue Computers' (Wiley, New York, 1962).
3. Miura, T., Amemiya, H. and Numabura, T., 'A new diode function generator', *Trans. Inst. Radio Engrs on Electronic Computers*, EC-1, pp. 95-100, June 1957.

Manuscript first received by the Institution on 11th July 1969 and in final form on 19th December 1969. (Paper No. 1309/IC21.)

© The Institution of Electronic and Radio Engineers, 1970

The Author



he was appointed a senior lecturer at the College. Mr. Barnes is at present a consultant to the Minerva Fire Detector Co. Ltd. and he has held consultancies with Feedback Ltd. and with Cossor Electronics Ltd.

R. N. Barnes graduated from Battersea College of Technology in 1959, and after a period at Mullard Research Laboratories, returned to Battersea to follow a post-graduate course in Automatic Control. He spent a short time as a lecturer in the University of the West Indies, and in 1963 was appointed a lecturer at West Ham College of Technology. Two years later

Contributors to the Journal



Dr. E. J. S. Becklake graduated in 1964 from the University of Exeter. He then carried out post-graduate research on upper atmospheric scattering using an optical radar employing a Q-switched ruby laser, for which he was awarded his Ph.D. degree in 1968. In 1967, he joined E.M.I. Electronics Ltd., Wells, Somerset, where he worked on sub-millimetric wave research

until 1969. He is currently engaged in research work on upper atmospheric physics at the University of Victoria, British Columbia, Canada.



C. D. Payne graduated in physics at University College London in 1964 and received his M.Sc. degree from the same University for his work on microwaves and the study of the impact ionization characteristics of silicon at 4·2°K. In 1965 he joined the Mullard Research Laboratories, where he carried out research on the liquid-helium-cooled parametric amplifier for the Goon-

hilly satellite communications station. He moved to E.M.I. Electronics Ltd., Wells, Somerset, in 1967 as a research engineer in the sub-millimetric laser group.



M. A. Smith graduated in 1960 from the University of London and joined the Atomic Weapons Research Establishment, Aldermaston, to work on solid-state lasers. In 1963, he took up a post with Thermal Syndicate Ltd., where he carried out research on discharge tubes and gas and ruby lasers. Some of this work was submitted to the University of Durham in 1967 and gained

him the M.Sc. degree. In 1968 he took up his present appointment as a senior research engineer with E.M.I. Electronics Ltd., Wells, Somerset.



B. E. Prewer worked on radar and radio systems in the Royal Air Force. After a short period as a test engineer with E. K. Cole Ltd., he joined E.M.I. Electronics Ltd., Wells, Somerset, in 1958. He has worked on various projects involving microwave systems and ferrite harmonic generation. He is now a research engineer in the group working on detection and the appli-

cation of radiation at sub-millimetric wavelengths.



Martin H. Ackroyd was recently appointed lecturer in industrial instrumentation in the Department of Electronic and Electrical Engineering at Loughborough University of Technology. In 1966 he graduated at Birmingham University with first class honours in electronic and electrical engineering. Before his staff appointment at Loughborough University he held a research

studentship in the Department for a period. His research interests include Hilbert transform networks and radar signal design.



Dr. T. Deliyannis received his B.Sc. degree in physics from the University of Athens in 1958, and he was then employed for a year as a research scientist with the Greek Atomic Energy Commission. Following a period at Battersea College of Technology at the conclusion of which he obtained the College's post-graduate diploma in control engineering, he worked for a

doctorate of the University of London, his thesis on active network synthesis being accepted in 1966. Since September 1967, Dr. Deliyannis has been a lecturer in the Department of Electrical Engineering at West Ham College of Technology where he is continuing his research interests.

Cyanide Gas Lasers for Sub-millimetric Wavelengths

By

E. J. BECKLAKE,

Ph.D., A.Inst.P. †

and

M. A. SMITH,

M.Sc., A.Inst.P. ‡

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

Progress is outlined in the design of c.w. and pulsed lasers, using methyl cyanide, dimethylamine and a mixture of methane and nitrogen, for use at wavelengths in the 337 μm region. Output stability is stated to be a function of cathode material, gas type and pressure and discharge current for any given laser design. Brass appears to be the best cathode material. With input power at 1 kW, outputs of 7 mW (c.w. laser) and 1 W (pulsed laser) were obtained, the gas used making no significant difference to these figures. An optimum working pressure of 0.55 torr was established using a mixture of 60% methane and 40% nitrogen. Tube life with all three gases is typically 100 hours. Initial investigations into transmission properties of various materials at sub-millimetric wavelengths show that p.t.f.e. and t.p.x. (polyolefin thermoplastic) make the best lenses and windows, and that black polythene makes the best attenuator. Investigations into atmospheric absorption at this wavelength indicate that, dependent on temperature and relative humidity, attenuations of between 36 dB/km and 204 dB/km can be expected.

1. Introduction

Progress made in the design of cyanide gas lasers operating in the sub-millimetric region of the electromagnetic spectrum is reported in this paper. The laser emissions are produced by gaseous discharges in the vapours of various cyanides and other organic compounds, and in gas mixtures containing H, C and N. Laser action at a wavelength of 337 μm , using methyl cyanide in a pulsed discharge, was first reported by Gebbie *et al.*¹ in 1964. Continuous operation, with emission at 337 μm and 311 μm , was obtained by the same authors in 1966.² Cyanide laser lines have also been found at a number of other wavelengths between 72 μm and 373 μm by Mathias *et al.*³ using pulsed operation, and many of these lines have been observed with continuous operation in the experiments of Müller *et al.*⁴ Outputs in the region of 10 W with pulsed discharges and a few milliwatts continuous have been obtained for the strongest lines at 337 μm and 311 μm .

A mechanism involving the vibration/rotational transitions of the CN radical, suggested by Chantry *et al.*,⁵ did not explain all of the above laser lines and Lide and Maki⁶ proposed a mechanism based on HCN as the emitting species. This explanation is supported by the experimental findings of Hocker

*et al.*⁷ However, Steffen *et al.*⁸ have reported laser emission in BrCN and ICN at 310 μm , 337 μm , 538 μm and 774 μm which cannot be explained by the HCN theories because of the absence of hydrogen atoms.

2. Design Considerations

The long wavelength of the emitted radiation, relative to the optical laser, necessitates a careful analysis of the resonant cavity of the laser in terms of its Fresnel number ($F = a^2/L\lambda$, where a = mirror radius, L = mirror separation, λ = wavelength) to avoid excessive diffraction losses. Unlike their optical counterparts, resonators for lasers at sub-millimetric wavelengths usually have Fresnel numbers around unity. These lasers are therefore more susceptible to diffraction loss, which is usually the major factor governing oscillation, especially in the higher-order modes. Furthermore, with optical lasers, there are frequently many cavity modes possible within the Doppler-broadened linewidth of the amplifying medium. The gain-width of cyanide lasers at 337 μm is in the region of 3 MHz and the separation of the axial modes or higher-order transverse modes of the cavity is at least an order of magnitude larger.

The cyanide laser will thus usually operate in a single mode, tunable by adjustment of the Fabry-Perot resonator length. The presence of more than one transverse mode from a c.w. laser with a high cavity Fresnel number (where diffraction losses for high and low-order modes are comparable) or from a high-gain pulsed laser has, however, been indicated by

† Formerly with E.M.I. Electronics Ltd.; now in Physics Department, University of Victoria, British Columbia, Canada.

‡ Systems Research Division, E.M.I. Electronics Ltd., Wells, Somerset.

a beat frequency in the 1 MHz region in the detected laser signal.

Theoretical calculations of the diffraction loss for various modes of a symmetric confocal resonator, with and without coupling aperture, have been made by McCumber⁹ for Fresnel numbers between 0.6 and 2.0, and by Li¹⁰ for confocal and non-confocal resonators.

Increase in laser power output can be obtained by an increase in resonator length, with a suitable increase in diameter to keep the diffraction loss down to a practical figure of about 1% per pass. For a 5 cm bore tube, the maximum possible length between confocal resonators is about 180 cm. However, a 660 cm long × 5 cm diameter laser has operated with a pulsed discharge, the theoretical loss per pass with a hole-coupled output being about 70%. This should lead to an unrealizable threshold condition and tends to confirm the hypothesis of Steffen and Kneubühl¹¹ that the effect of the walls of the tube cannot be neglected in sub-millimetre lasers.

3. Wavelength Measurement and Radiation Detection

A simple method of measuring the wavelength of the emitted radiation is to monitor the laser output as a function of resonator length. The distance moved between successive resonance peaks corresponds to $\lambda/2$. This method suffers from the disadvantage of low resolution and the difficulty in distinguishing between different transverse modes of the same wavelength or peaks corresponding to two or more laser emissions at different wavelengths. However, in most cases it can be used to separate the high-gain 311 μm and 337 μm lines.

Greater accuracy and resolution is provided by a reflection-grating spectrometer, whose layout follows the style of Czerny and Turner (M. F. Kimmitt¹²). Several blazed gratings are available to cover a wavelength range from 28 μm to 2300 μm . The resolution of this instrument is normally limited (to about 1% at 300 μm) by the accuracy with which the vernier on the grating-angle scale can be read, i.e. ± 6 minutes of arc. Interpolation between the vernier scale divisions is obtained by mechanically driving the grating turntable and using a revolution indicator on the drive spindle. In this way a resolution of better than 0.2% can be attained, sufficient to permit the separation of the close laser lines around 310 μm and 311 μm . The physical size of this spectrometer is a disadvantage and, due to its path length of 250 cm, evacuation of the instrument is necessary when making comparative measurements at the shorter wavelengths where attenuation by atmospheric water vapour is appreciable.

The radiation detectors used in this work are the commercially-available Golay cell, a specially-developed point-contact gallium arsenide diode¹³ and a pyroelectric detector.¹⁴ The output power of the lasers is usually monitored by a Golay cell. A c.w. laser beam is mechanically chopped at 11 Hz and the manufacturer's calibration figure is used to convert the Golay output voltage, measured on an oscilloscope, to output power. Allowance is made for the difference between the detector aperture and the laser beam diameter with the detector set at a fixed distance from the laser. Black polythene attenuators are used to avoid overloading.

Point-contact gallium arsenide diodes, mounted in waveguide, are also used as video monitors.¹³ Due to their nanosecond response time, these diodes are also suitable for observation of the pulsed laser output (0.5 μs rise-time, several microseconds duration) without distortion.

A pyroelectric detector with frequency-compensated amplifier is also used to monitor pulsed laser outputs, but microphony has been a problem with c.w. lasers due to acoustic breakthrough from the mechanical chopper. One advantage over the point-contact diode is the wider acceptance angle.

Indium antimonide (liquid helium cooled) detectors¹⁵ have the best sensitivity in the sub-millimetre region, but the ancillary cryogenic equipment required is a disadvantage.

Table 1. Comparison of detector parameters

Detector	Wave-length range (μm)	Typical respon-sivity (V/W)	N.e.p. for 1 Hz band-width (watts)	Response time (seconds)	Operat-ing temperature
Golay cell	50-5000 (quartz window)	10^5	10^{-10}	10^{-2}	room
Point-contact GaAs diode	down to about 100	0.1	10^{-7} at 337 μm 10^{-6} at 100 μm	$< 10^{-8}$	room
Pyro-electric ¹⁴	sub-mm down to 1 μm	100	5×10^{-8}	5×10^{-6}	room
InSb (with magnetic field) ¹⁵	200 500 1000	200 500 1000	2×10^{-11} 1×10^{-11} 5×10^{-12}	2×10^{-7}	1.5°K
InSb (without magnetic field)	300 8000	300	5×10^{-13}	10^{-4} (trans-former dependent)	1.8°K

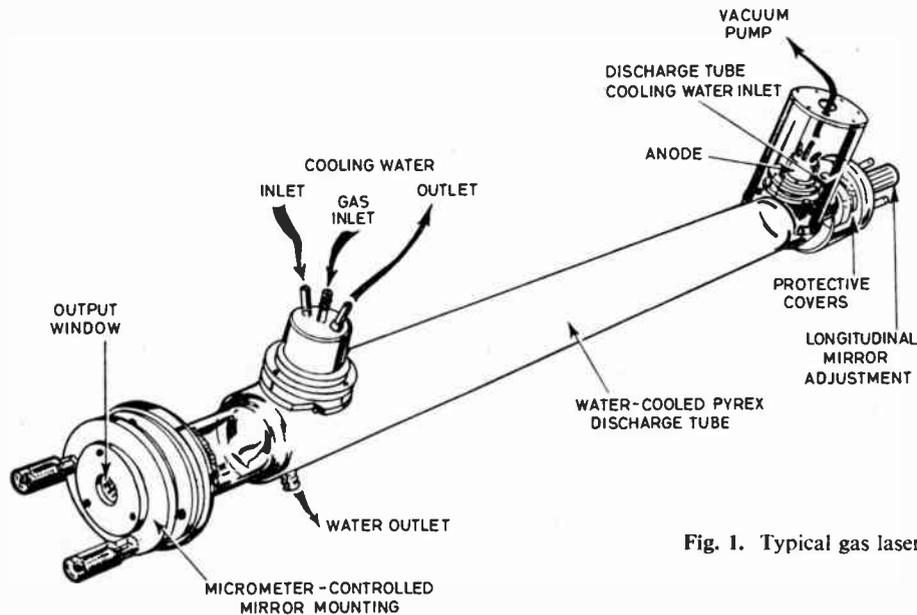


Fig. 1. Typical gas laser.

The detector parameters are compared in Table I. A comprehensive review of sub-millimetric radiation detection techniques has been made by Putley and Martin.¹⁶

4. Experimental Details and Results

4.1. C.W. Laser System

A schematic diagram of a typical c.w. laser is shown in Fig. 1. The Pyrex discharge tube is 150 cm long, 5 cm internal diameter, with an anode-to-cathode separation of 115 cm. The Fabry-Perot type resonator consists of two polished aluminium mirrors in micrometer-controlled mounts, one of which can be moved along the axis of the tube without tilting. The aperture in the output mirror is sealed with a quartz or p.t.f.e. window. An alternative coupling system has also been used, with one mirror external to the laser tube and a p.t.f.e. beam-splitter within the resonant cavity. The laser tube window is 0.4 mm thick p.t.f.e. and is normal to the cavity axis. This system allows the extraction of two beams from the cavity.

The anode consists of a water-cooled brass tube which also serves as the connexion to the vacuum system. The concave shape of the water-cooled cathode, which was determined experimentally, minimizes cathode spot wander.

The gas under investigation, methyl cyanide, dimethylamine, or the mixture of methane and nitrogen suggested by Stafsudd *et al.*,¹⁷ is fed into the discharge region via the central hole in the cathode, and a rotary vacuum pump maintains the gas pressure in the tube between 0.2 torr and 0.8 torr, dependent on the flow rate.

The 2.5 kV d.c. power supply unit is smoothed to minimize current fluctuation of the arc, and to reduce amplitude and frequency modulation of the laser output. Ballast resistors are used to limit the discharge current, once the arc has struck. Under operating conditions the voltage drop across the 115 cm discharge path at a current of 1.5 A is in the region of 800 V.

4.2. Pulsed Lasers

The construction of a pulsed HCN laser is simpler than that of the c.w. laser because, since the power dissipated is much less, i.e. 10^{-2} W for the pulsed laser against 1 kW for the c.w. laser, water cooling is not necessary. Mechanically, therefore, long pulsed lasers, with their greater gain per pass, are more easily produced than long c.w. lasers.

A typical pulsed laser 2.5 m long, 5 cm internal diameter, can produce pulses with peak powers of 1 W at p.r.f. rates up to 10 per second. The power required for such a laser is of the order of 10 kV at 100 A. This is supplied by a simple modulator fed from a 20 kV, 20 mA mean d.c. supply. Using a recently-constructed laser 6.6 m long and 7.5 cm internal diameter, peak powers of up to 20 W have been obtained at 337 μ m.

Although the mean power obtainable from a pulsed laser is considerably less than that obtainable from a typical c.w. laser, the high gain per pass produces strong outputs at 373 μ m, 310 μ m, 211 μ m and 128 μ m as well as at 337 μ m. These outputs are useful when investigating the properties of various materials in the sub-millimetric region.

4.3. Stability of Output

The laser output stability is dependent to a large extent on the stability of the discharge arc. Slow frequency drifts, due to thermal variation in cavity dimensions, have been reduced by water-cooling the discharge tube. (Under operating conditions, about 1 kW is dissipated in the discharge.) The arc stability depends on the cathode material and design, gas type and discharge current.

Aluminium, brass, carbon, molybdenum, stainless steel, tungsten and zinc have been used as cathode materials. Brass has been found to be the best cathode material with the least cathode-spot wander. A concave electrode shape also assists in limiting the extent of cathode spot movement. The current density in the cathode spot is in the region of 150 A/cm^2 , and Joule heating plays an important part in actual evaporation of the cathode material. This is explosive in nature and gives rise to localized melting of the cathode. Transitions between thermal and cold electron emission take place as the cathode spot shifts. Kesaev¹⁸ has related the arc stability of a cold cathode in a vacuum arc to the nature of the cathode material, and his results show that the most stable arcs are obtained with solid cathodes with low thermal and electrical conductivity and low boiling-point. In this work, it has been found that high boiling-point materials, e.g. tungsten, cause a marked reduction in arc stability and it is probable that the arc stability with a brass cathode is a direct result of its high zinc content (melting point 420°C). The arc spot on a pure zinc cathode is more stable, but rapid erosion limits the electrode life to a few hours.

The i.f. output signal from a point-contact mixer, obtained by mixing the output of two $337 \mu\text{m}$ lasers, can be adjusted by cavity-tuning between 100 kHz and 1 MHz and the stability of this signal is directly related to the mean arc duration of the discharge. Discharges in methane/nitrogen provide the most stable beat-frequency for periods up to several seconds. These gases are fed from cylinders at a constant pressure. The variations of vapour pressure of the volatile methyl cyanide and dimethylamine cause fluctuation in discharge gas pressure and flow rate, and reduce the stability of the discharge.

4.4. Output Power

The output power of a cyanide laser of given dimensions is dependent on gas type, pressure and flow rate, discharge current, output coupling and reflector configuration.

The optimized outputs of c.w. lasers of similar dimensions and mirror configurations were compared, using methyl cyanide or dimethylamine vapour and a mixture of methane and nitrogen gas. The outputs obtained with all three gases were the same within the

limits of experimental error. The maximum output of 7 mW at $337 \mu\text{m}$ was obtained using CH_4/N_2 with the 160 cm c.w. laser, at a power input to the discharge tube of about 1 kW. This system is the most convenient to use, since the constant-pressure feed from the methane and nitrogen cylinders requires no adjustment once the optimum gas mixture (in the region of 60% CH_4 , 40% N_2) has been set. Furthermore, the mixture is far less toxic than methyl cyanide and dimethylamine.

The gas pressures, flow rates, pulse-repetition frequency and power input were also optimized with each gas in a 250 cm long, 5 cm diameter, pulsed laser. The peak power output at $337 \mu\text{m}$ was in each case in the region of 1 W.

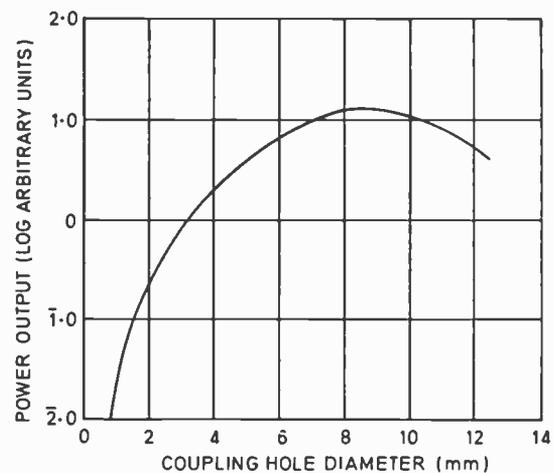


Fig. 2. Total power output from a 160 cm c.w. laser as a function of coupling hole size.

To investigate the variation in output with gas pressure, the discharge current was altered at a series of fixed gas pressures and the optimum output recorded. The gas used was a mixture of CH_4/N_2 and maximum output was obtained at a pressure of 0.55 torr. Further increase in pressure results in an increase in plasma resistance, and measurements at higher pressures are at present limited by the maximum voltage obtainable from the power supply. Below the optimum pressure, gain saturation of the medium occurs and no increase in output is obtained with increased current. Further increase in excitation then causes a reduction in output.

The variation in output power of the 160 cm c.w. laser with coupling hole diameter is shown in Fig. 2. The maximum output is obtained with an 8 mm coupling-hole in the 50 mm diameter mirrors and the power is within 3 dB of this peak for hole sizes between 6 mm and 11 mm. The laser gain is a function of the tube parameters, and variation in these alters the position of the peak on the power/coupling curve.

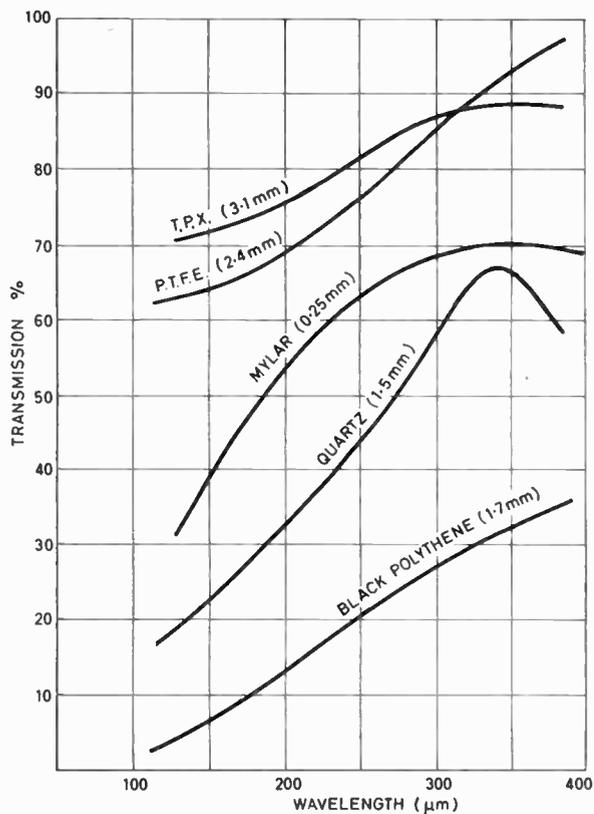


Fig. 3. Variation of attenuation with wavelength.

A high-gain 660 cm long pulsed laser, 7.5 cm internal diameter gives maximum output with a 2.5 cm diameter coupling aperture.

Recent work with a 300 cm long, 7.5 cm internal diameter c.w. laser has produced a significant increase in output power over the 160 cm long, 5 cm internal diameter c.w. lasers. The output from these lasers has not yet been optimized, but powers of the order of tens of milliwatts have been obtained at 337 μm.

4.5. Tube Life

The typical polymerization products of C, H and N discharges coat the wall of the discharge tube in the form of an orange film with all of the gases investigated. For safety reasons, samples of this film have been analysed and show less than 1 portion in 10⁶ of cyanide. The life of a discharge tube is typically in the region of 100 hours before polymer, flaking from the tube wall, interrupts the radiation in the laser cavity. The CH₄/N₂ discharge produces less polymer but has not given, in the lasers used, the clean operation described by Stafsudd *et al.*¹⁷

5. Components

Work in this new region of the electromagnetic spectrum has necessitated investigations into the

transmission properties of various materials for use as lenses, windows and attenuators. The materials investigated included t.p.x. (polyolefin thermoplastic, dielectric constant = 2.12), p.t.f.e., translucent polythene, mylar, fused quartz and black polythene. Graphs of the transmission properties at different wavelengths are shown in Fig. 3.

T.p.x. and p.t.f.e. have the lowest transmission losses and lenses are normally made from these materials; t.p.x. has the added attraction of being transparent. Commercially-available black polythene is the best attenuator, 1.5 mm thick material realizing attenuations of 5 dB at 337 μm and 13 dB at 128 μm (the reflexion loss amounts to only 4% of the attenuation at 337 μm.)

6. Atmospheric Attenuation Measurements

The absorption of radiation in the sub-millimetre range of the spectrum due to water vapour is appreciable. The radiation at 337 μm is near the centre of an atmospheric window, but the fact that atmospheric water content can vary considerably from day to day necessitates a knowledge of the attenuation under various humidity conditions. A beam of 337 μm radiation from a c.w. laser was transmitted over a known atmospheric path length and received by a parabolic aerial. Goly cells monitored both transmitted and received signals. The air temperature and humidity were recorded along the signal path. The

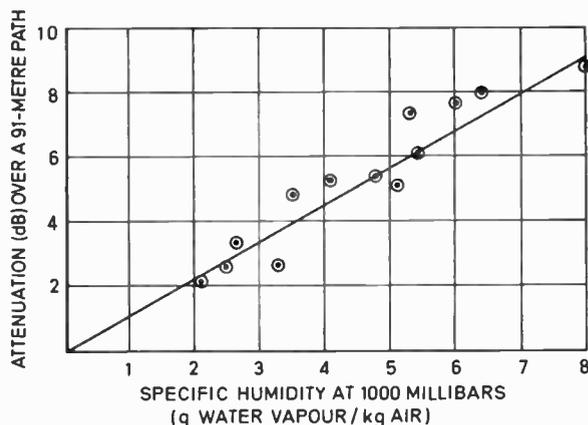


Fig. 4. Variation in atmospheric attenuation at 337 μm with atmospheric water content.

relationship between attenuation and atmospheric water vapour content is shown in Fig. 4. For saturated air at 0°C, the attenuation of 337 μm radiation is 52 ± 5 dB/km. This figure is close to that obtained by Burroughs *et al.*¹⁹ For relative humidity and room temperature changes between 50% and 100% and 4°C and 20°C respectively, losses of between 36 dB/km

and 204 dB/km may be expected at this wavelength. This atmospheric attenuation will limit any communications application at this wavelength and variations of attenuation with humidity have to be considered in applications involving metrology.

7. Acknowledgments

The authors wish to thank the Royal Radar Establishment, Malvern and the Ministry of Technology for their support and for permission to publish this work.

8. References

1. Gebbie, H. A., Stone, N. W. B. and Findlay, F. D., 'A stimulated emission source at 0.34 millimetre wavelength', *Nature*, **202**, p. 685, 16th May 1964.
2. Gebbie, H. A., Stone, N. W. B., Slough, W., Chamberlain, J. E. and Sheraton, W. A., 'Sub-millimetre maser amplification and continuous wave emission', *Nature*, **211**, p. 62, 2nd July 1966.
3. Mathias, L. E. S., Crocker, A. and Wills, M. S., 'Laser oscillations at sub-millimetre wavelengths from pulsed gas discharges in compounds of hydrogen, carbon and nitrogen', *Electronics Letters*, **1**, No. 2, pp. 45-6, April 1965 and 'Spectroscopic measurements on the laser emission from discharges in compounds of hydrogen, carbon and nitrogen', *I.E.E.E. J. Quantum Electronics*, **QE-4**, No. 4, pp. 205-8, April 1968.
4. Müller, W. M. and Flesher, G. T., 'Continuous wave sub-millimetre oscillation in discharges containing C, N, and H or D', *Applied Physics Letters*, **10**, No. 3, pp. 93-4, 1st February 1967.
5. Chantry, G. W., Gebbie, H. A. and Chamberlain, J. E., 'A suggested mechanism for the 337 μm CN maser', *Nature*, **205**, p. 377, 23rd January 1965.
6. Lide, D. R. and Maki, A. G., 'On the explanation of the so-called CN laser', *Applied Physics Letters*, **11**, pp. 62-4, 15th July 1967.
7. Hocker, L. O. and Javan, A., 'Absolute frequency measurements on new c.w. HCN sub-millimetre laser lines', *Physics Letters*, **25A**, pp. 489-90, 9th October 1967.
8. Steffen, H., Steffen, J., Moser, J. F. and Kneubühl, F. K., 'Comments on a new laser emission at 0.774 mm wavelength from ICN', *Physics Letters*, **21**, No. 4, pp. 425-6, 1st June 1966.
9. McCumber, D. E., 'Confocal laser resonator modes and their perturbation by output coupling apertures', *Bell Syst. Tech. J.*, **44**, pp. 333-63, February 1965.
10. Li, T., 'Circular mirror maser resonators', *Bell Syst. Tech. J.*, **44**, pp. 917-32, May 1965.
11. Steffen, H. and Kneubühl, F. K., 'Dielectric tube resonators for infrared and sub-millimetre lasers', *Physics Letters*, **27A**, No. 9, pp. 612-3, 23rd September 1968.
12. Kimmitt, M. F., 'A far infrared spectrometer', Royal Radar Establishment Technical Note No. 716. December 1965.
13. Payne, C. D. and Prewer, B. E., 'Semiconductor diodes as detectors and mixers at sub-millimetre wavelengths', *The Radio and Electronic Engineer*, **39**, No. 3, pp. 167-71, March 1970.
14. Ludlow, J. H., Mitchell, W. H., Putley, E. H. and Shaw, N., 'Infra-red radiation detection by the pyroelectric effect', *J. Sci. Instrum.* **44**, pp. 694-6, 1967.
15. Putley, E. H., 'The detection of sub-millimetre radiation', *Proc. Inst. Electrical Electronics Engrs*, **51**, pp. 1412-22, November 1963.
16. Putley, E. H. and Martin, D. H., Chapter 4, p. 113, in 'Spectroscopic Techniques for Far Infra-red Sub-millimetre and mm Waves', Ed. Martin D. H. (North Holland, Amsterdam, 1967).
17. Stafstudd, O. M., Haak, F. A. and Radisavljevic, K., 'Laser action in selected compounds containing C, N, and H or D', *I.E.E.E. J. Quantum Electronics*, **QE-3**, No. 11, pp. 618-20, November 1967.
18. Kesaev, I. G., 'Stability of metallic arcs in vacuum', *Soviet Physics, Technical Physics*, **8**, No. 5, pp. 447-62, November 1963.
19. Burroughs, H., Pyatt, E. and Gebbie, H. A., 'Transmission of sub-millimetre waves in fog', *Nature*, **212**, p. 387, 22nd October 1966.

Manuscript first received by the Institution on 21st February 1969 and in final form on 8th December 1969. (Paper No. 1310/CC68.)

© The Institution of Electronic and Radio Engineers, 1970

Semiconductor Diodes as Detectors and Mixers at Sub-millimetric Wavelengths

By

C. D. PAYNE, M.Sc.†

and

B. E. PREWER †

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

The application of germanium, silicon and gallium-arsenide point-contact diodes and gallium-arsenide Schottky-barrier diodes as detectors and mixers at sub-millimetric wavelengths is discussed. Figures of merit for these diodes, when used as video detectors or mixers, are also discussed and prove greatly dependent on junction radius and carrier mobility. The fabrication of point-contact diodes in a waveguide mount is described and comparisons are made between various combinations of whisker and semiconductor materials. Measurements of conversion loss, n.e.p. and voltage responsivity have been made, using both c.w. and pulsed HCN lasers. The n.e.p. of the germanium point-contact diode, used as a superheterodyne mixer, is superior to that of the well-known Golay cell; it also has a much shorter response time. Schottky-barrier diodes, whilst not being so sensitive, are proving useful at much shorter wavelengths than hitherto reported.

1. Introduction

The development of a coherent source giving useful power at around 0.3 mm wavelength (the cyanide gas laser) has established a requirement for a detector, fast enough in response time to observe microsecond output pulses or the beat frequencies (100 kHz to 1 MHz) which result from the non-linear mixing of two cyanide lasers.

Detectors may be classified into three groups according to the physical process responsible for their operation:

- (i) thermal,
- (ii) photoconductive,
- (iii) semiconductor junction.

The relative merits of all these have been discussed in a review article by Putley.¹ Thermal detectors in general have a slow response time with good sensitivity, but do require cooling to low temperatures to improve their sensitivity. The photoconductive detector (for example, cooled bulk indium antimonide) requires cooling to at least liquid nitrogen temperature, is very sensitive and has a response time which can be as fast as a few tenths of a microsecond. The metal-semiconductor junction in the form of the point-contact diode has a response time which is only limited by the time taken for the current carriers to cross the junction. Hence these detectors may be used for observing nanosecond pulses or as superheterodyne mixers with microwave intermediate frequencies.

There are other detectors for this region of the spectrum, for example the Josephson superconducting junction,² which are in an early stage of development, but at present the point-contact diode is the only detector combining very fast response time with reasonable sensitivity at room temperature.

Point-contact diodes have been used since the early days of radio, but they have always suffered from instability, and uncertain lifetime (at least for the smallest junctions). For these reasons their use at the lower microwave frequencies has now been largely superseded by the Schottky-barrier diode, which offers stability and long lifetime, lower noise and more efficient mixing. Even at the higher microwave frequencies Schottky-barrier diodes are competing with point contacts,³ but where the very smallest junction areas are required the point-contact diode must be chosen. Nevertheless, what is believed to be the first mention of the operation of a Schottky-barrier diode at 0.337 mm wavelength is reported here.

The performance of point-contact diodes at sub-millimetric wavelengths depends on the mobility, doping and hardness of the semiconductor material, and on the whisker metal.

Because of its superior electrical properties, point-contact diodes were initially made with n-gallium arsenide. However, because of the extra hardness of p-silicon and n-germanium, it has recently been found possible to produce smaller contact areas on these materials. The strong dependence of diode performance on contact area more than compensates for the lower mobility of silicon and germanium and results in a better performance at sub-millimetric

† Systems Research Division, E.M.I. Electronics Ltd., Wells, Somerset.

wavelengths. For diodes whose contact area does not depend on pressure, e.g. the Schottky-barrier diode, n-gallium arsenide is expected to be the superior material.

2. Theory

The criteria by which the performance of a diode is assessed depend on the use to which the diode is to be put.

When used as a video detector, two criteria commonly used are the voltage responsivity and noise equivalent power (n.e.p.). The noise equivalent power is the level of signal power into the diode which produces a detected signal equal to the r.m.s. noise level for a specified bandwidth. As a mixer, the performance of a diode may be specified by its noise figure and the conversion loss from r.f. (in this case 890 GHz) to i.f.

The conversion loss is made up of three separate parts:

- (i) mismatch losses due to incomplete coupling of the incident signal power into the diode junction;
- (ii) inefficiency of conversion by the junction from r.f. to i.f.;
- (iii) loss due to power dissipated in the ohmic resistance of the diode.

Loss due to (iii) may be written in terms of the diode's ohmic resistance R_s and junction capacitance C_j and the mean non-linear resistance R_j . In this way a figure of merit may be derived in terms of mobility of the semiconductor and the contact radius of the junction.

The same expression may also be shown to be a figure of merit for the diode used as a video detector. This figure of merit is strongly dependent on the junction radius.

Since, of the three most commonly used materials for diodes (n-germanium, p-silicon, n-gallium arsenide), n-gallium arsenide has the highest mobility, it would also be expected to have the highest figure of merit. However, this does not take into account the technological difficulty of producing small contacts with a pointed whisker pressing on the surface of a semiconductor.

Taking into consideration the relative hardness coefficients of the three semiconductor materials, germanium has the highest figure of merit with silicon and gallium arsenide next, in that order. However, for diodes with evaporated metal-junctions, gallium arsenide should be the best.

The noise power generated within a metal semiconductor junction limits the n.e.p. and, since the noise spectrum is not 'white', the n.e.p. will vary with

the video or modulation frequency of the r.f. signal. Measurements on a gallium arsenide point-contact diode show that above about 300 kHz, with only a few microamperes of current flowing, the noise power is approximately thermal in nature and thus is proportional to temperature and bandwidth only. Below 300 kHz measurements indicate an inverse frequency dependence of the noise power.

3. Fabrication of Diode Mount Suitable for Mixing and Detection

3.1. Description of Detector and Mixer Mount

A conventional microwave mixer or detector consists primarily of a diode mounted in fundamental-mode waveguide. Fundamental-mode waveguide at 0.337 mm wavelength has dimensions 0.254 mm × 0.127 mm; the difficulty of fabricating waveguide of this size, together with problems associated with mounting the diode in such a small cavity, has led, initially, to the use of circular guide approximately 10 wavelengths in diameter.

The power is coupled into the diode structure through a waveguide horn feed and a phase-correcting lens at the mouth of the horn. This method of coupling, whilst not significantly improving the match, does make the alignment of the mount with respect to the laser beam less critical. The lens and horn are designed to accept a collimated beam.

The mount consists of two parts:

- (i) a removable insert containing the semiconductor and whisker assembly and i.f. output connector;
- (ii) the main body with its horn feed, waveguide structure and micrometer short-circuiting plunger.

This latter feature is analogous to the short-circuit used in a conventional microwave detector mount.

The semiconductor chip is attached to a removable post held in an insulating bush so that the surface of the chip is level with the waveguide wall. A second post with the pointed whisker attached is placed diametrically opposite the first. The whisker post is arranged to be a press fit in the insert, enabling the whisker point to meet the surface of the semiconductor at the required pressure.

This method of construction greatly facilitates interchange of semiconductor material and whisker metal for rapid diode assessment.

The mount may be used either for detection or mixing. In the latter case both the local oscillator and the signal are coupled into the diode using a beam splitter to combine them prior to entry into the single horn. A simple modification to the detector mount includes a second feed horn, orthogonal to the first, allowing mixing to take place without the use of a beam splitter and its associated losses.

3.2. Diode Fabrication

Several different semiconductor-whisker combinations have been assessed both as detectors and mixers. The results are summarized in Tables 2 and 3.

Certain whisker-semiconductor combinations produce good sub-millimetric diodes; these are indicated by a tick in Table 1, and those combinations that do not make good sub-millimetric diodes are indicated by a cross.

Table 1. Whisker-semiconductor combinations

Whisker material	Semiconductor		
	n-Ge	p-Si	n-GaAs
phosphor bronze	—	—	✓
platinum	—	—	✓
platinum iridium (85 : 15)	—	—	✓
gold/copper (75 : 25)	×	×	✓
tungsten	✓	✓	×
titanium	×	—	—

The whiskers are 0.025 mm in diameter and are electrolytically-pointed to an included angle of approximately 30°. The table shows that while n-type gallium arsenide will form a useful diode at sub-millimetric wavelengths with a variety of whisker metals, it does not, however, form a reliably-stable diode, and deterioration of its characteristics may occur. The gallium arsenide diodes require the use of forward bias which produces a large inverse frequency noise component at low video frequency. Somewhat improved stability may be obtained with a combination of p-type silicon and tungsten. Such a diode can be operated without bias and the noise output is then reduced.

The best sub-millimetric performance has recently been obtained with a combination of n-type germanium and tungsten. This diode has similar characteristics to the silicon diode with the same whisker and also operates without bias both as a detector and a mixer.

The factors governing diode lifetime are not well known, and diodes are easily destroyed by stray fields or circulating earth currents. However, these particular diodes do have the outstanding advantage that unless the diode is destroyed by excessive energy input to the junction, it may be reformed by re-pointing and repositioning the whisker.

4. Measurements

A description of the point-contact diode mount has been given in Section 3 and measurements taken with

this mount will be described here. The HCN lasers used for this series of measurements are described in a companion paper⁴ and typically have 1 to 5 mW output. The radiation is coupled out through a small hole in one end-mirror and the beam divergence is approximately 3.5°. For this reason collimation of the beam is necessary for operation at a distance from the laser.

4.1. Detection

The detected output voltage for most point-contact diodes in this wavelength region is less than 0.5 mV so some amplification is required. This is generally accomplished with a low-noise transistor amplifier with an input impedance high enough to prevent current flow and loss of output due to voltage drop across the internal impedance of the diode. Bias, where necessary, is applied via the amplifier. To facilitate the measurements, the laser radiation is mechanically modulated at 2 kHz and the bandwidth reduced by passing the detected voltage through a tuned amplifier at 2 kHz.

The voltage responsivity is deduced from the laser output power, measured with a Golay cell, and the detected diode voltage. Because of the mismatch losses due to the use of oversize waveguide, the actual power arriving at the junction will be one or two orders less than the power radiated by the laser. Consequently, even with the unattenuated laser power incident on the diode, it will still operate as a square-law device.

The voltage responsivity obtained with point-contact diodes of silicon, germanium and gallium arsenide and with a gallium arsenide Schottky-barrier diode are given in Table 2, together with the n.e.p. measured at 2 kHz video frequency.

Table 2. Performance of diodes as detectors

Diode		N.e.p. (at $f_v = 2$ kHz) in 1 Hz	Voltage responsivity (β)
p-Si	PC	3.8×10^{-6} W	20 mV/W
n-Ge	PC	9.5×10^{-7} W	60 mV/W
n-GaAs	PC	3.8×10^{-5} W	47 mV/W
n-GaAs	SB	—	65 mV/W

PC point contact. SB Schottky-barrier

4.2. Mixing

Mixing is a useful way of increasing the sensitivity of a system. By arranging that signal and local oscillator sources are about 1 MHz apart in frequency, the intermediate frequency falls in that region of the noise spectrum of the diode where inverse frequency noise is negligible (this is only the case at the low bias currents flowing in these sub-millimetric detectors; at higher bias currents, inverse frequency noise may

exceed thermal and shot noise even into the microwave region).

Although the linewidth of the transition involved in the 0.337 mm cyanide laser is very small, Doppler and pressure broadening, together with plasma instabilities, cause the output power to have a spectrum of frequencies extending from 1 to 2 MHz on either side of the centre frequency. Moreover, the centre frequency of oscillation may be changed within this broadened line width by slightly altering the length of the cavity. This enables two such lasers to be mixed in a diode to produce an i.f. in the required region (100 kHz to 1 MHz).

It has already been pointed out that the parameter which limits the sensitivity is the diode conversion loss. The noise generated by the diode is not far from the thermal level when inverse frequency noise is insignificant and low bias currents are used, so that the limiting sensitivity is worse than the thermal level by about the magnitude of the conversion loss.

Conversion loss measurements have been made to show the dependence on diode material and on local oscillator power. To extend the range of measurement of the latter, a pulsed laser was used to supply up to 0.5 W of local oscillator power.

Table 3 gives the results of the conversion loss measurements for different materials together with the n.e.p. measured in a bandwidth from 100 kHz to 1 MHz. Comparison with the results of Table 2 clearly demonstrates the improvement in sensitivity as a result of mixing.

Table 3. Performance of diodes as mixers

Diode	Conversion loss (<i>L</i>) (at 3 mW local oscillator)	N.e.p. in 1 Hz (from 100 kHz to 1 MHz)
p-Si	89 dB	3.8×10^{-11} W
n-Ge	77 dB	2.4×10^{-12} W
n-GaAs	85 dB	1.2×10^{-11} W

The variation of conversion loss with local oscillator power has been measured for one particular diode and approached an optimum at about 140 mW of laser power. At this point the improvement was 13 dB and represents a conversion loss of 64 dB for the best diode shown in Table 3. Comparing this figure of 64 dB loss with a commonly-encountered value for a mixer operating at 80 GHz with optimum local oscillator power, the degradation is considerably greater than the expected f^2 dependence would suggest.

The major part of the discrepancy is thought to be due to the mismatch loss between input waveguide and the junction. In addition, some loss will accrue since the junction itself is not expected to be as efficient a mixer at the higher frequencies.

As a result of this, careful attention will be given in future designs to methods of reducing the mismatch loss which will also have the added advantage of reducing the required local oscillator power.

5. Discussion

The parameter $1/R_S C_J^2$ may be shown to be a figure of merit, expressed in terms of the diode parameters, for sub-millimetric diodes used as detectors and mixers. This figure of merit may be expressed in terms of the material parameters and shown to be largely independent of the doping concentration of the semiconductor. In our work to date no correlation of sub-millimetric performance with doping concentration has yet been observed.

Table 4 shows figure-of-merit values for four types of diode evaluated from both conversion loss and voltage responsivity measurements. In both cases, due allowance has been made for mismatch loss (a value of 20 dB was assumed) and 10 dB junction loss to ensure that the figure of merit is calculated from the effect of the parasitic elements of the diode only.

In addition, a further 13 dB is subtracted to reduce the results of Table 4 on conversion loss to a value consistent with optimization with respect to local oscillator power.

Table 4. Figures of merit for diodes

Diode	Figure of merit $\frac{1}{R_S C_J^2}$ from loss	Figure of merit derived from voltage responsivity (β)	Junction radius derived from β
n-GaAs (pc)	2×10^{24}	1×10^{25}	5×10^{-6} m
n-Ge (pc)	6×10^{24}	1×10^{25}	3×10^{-6} m
p-Si (pc)	4×10^{23}	5×10^{24}	2×10^{-6} m
n-GaAs (sb)	—	2×10^{25}	5×10^{-6} m

The figures of merit for all three diodes are seen to be lower when evaluated from the conversion loss; this may be due to the junction loss being greater than the 10 dB assumed. n-type germanium has the highest figure of merit with n-type gallium arsenide next.

The voltage responsivity results have been used to evaluate the contact radius by expressing the product $R_S C_J^2$ in terms of the mobility and dielectric constant of the semiconductor, and contact radius, a . The results are shown in Table 4 (column 4). Not too much significance should be attached to the absolute values obtained because of the uncertainty of the mismatch loss, but the relative values are such that silicon has the smallest contact radius with germanium next. This is in the same order as their hardness coefficients, as indicated in Section 2.

The smallest contact areas may therefore be achieved on silicon. Unfortunately, the low mobility of p-type silicon reduces the figure of merit to below that for either gallium arsenide or germanium, but the lower barrier voltage of the silicon diode compared with gallium arsenide means that it can operate without bias. However, since no bias is required with the silicon diode, it therefore produces less noise than gallium arsenide, and the result is that the sub-millimetric performance lies between that of germanium and gallium arsenide.

The Schottky-barrier diode has a figure of merit approximately equal to that of the germanium or gallium arsenide point-contact diodes, but, because of the higher barrier voltage of gallium arsenide, it does require biasing. However, the lower inverse frequency noise properties of Schottky-barrier diodes should enable a lower n.e.p. to be obtained at lower video frequencies. These diodes are in an early stage of development and few measurements have yet been taken, but it is interesting to note that the effective junction radius achieved is similar to that of the point-contact diode.

Table 5. Comparison of diodes for detection and mixing

Diode	N.e.p. detection	N.e.p. mixing (3 mW)
Ge	9.5×10^{-7} W	2.4×10^{-12} W
Si	3.8×10^{-6} W	3.8×10^{-11} W
GaAs	3.8×10^{-5} W	1.2×10^{-11} W

Table 5 compares the n.e.p. on detection and mixing for different diodes. The results agree with the predictions of Section 2, germanium and silicon making the best detectors and mixers. The improvement brought about by mixing is seen to be considerable—this is due in part to the adoption of a high intermediate frequency where inverse frequency noise is much reduced. Measurements show that the reduction in inverse frequency noise in going from 2 kHz to 100 kHz is approximately 20 dB for a gallium arsenide diode. The remaining 30 dB and the major part of the improvement for the germanium and silicon diodes is due to the mixing action.

The results of Table 5 were taken with a local oscillator power of 3 mW. The improvement in n.e.p. to be obtained by increasing the local oscillator power to over 100 mW could be up to 13 dB (due to the improvement in conversion loss). An n.e.p. of 10^{-13} W should then be possible with the best diode of Table 5 in a 1 Hz bandwidth.

6. Conclusions

In agreement with the predictions of Section 2, germanium and silicon form the best sub-millimetric detectors and mixers. As a detector, a germanium diode is able to detect about 10^{-6} W at a chopping frequency of 2 kHz and, as a mixer, a power of 10^{-12} W, the bandwidth being 1 Hz in both cases. The latter result implies that point-contact diodes may be used in a mixer system with a measurement capability better than that of the Golyay cell and with appreciably wider bandwidth. (Alternatively, the diode may be used for observing very short pulses, in which case the sensitivity is correspondingly reduced.)

Further improvements to the structure, designed to improve the coupling of the sub-millimetric power into the junction, should lead to even lower n.e.p. of about 10^{-8} W as detectors and 10^{-14} W to 10^{-15} W as mixers. This may provide a serious competitor to any known sub-millimetric detector or mixer in terms of response time and noise equivalent power.

The Schottky-barrier diode is beginning to compete with the point-contact diode even at sub-millimetric wavelengths and promises long lifetime and robustness.

7. Acknowledgments

The authors wish to thank the Royal Radar Establishment, Malvern and the Ministry of Technology for their support and for permission to publish this work.

8. References

- Putley, E. H., 'The detection of sub-mm radiation', *Proc. Inst. Elect. Electronics Engrs*, 51, No. 11, p. 1412, November 1963.
- Grimes, C. C., Richards, P. L. and Shapiro, S., 'Josephson-effect far-infrared detector', *J. Appl. Phys.*, 39, No. 3, p. 3905, July 1968.
- Landkammer, F. J., 'A millimetre mixer with Schottky-barrier diodes', Northeast Electronics Research and Engineering Meeting, Boston 1967, pp. 58-9, Boston I.E.E.E.
- Becklake, E. J. and Smith, M. A., 'Cyanide gas lasers for sub-millimetric wavelengths', *The Radio and Electronic Engineer*, 39, No. 3, p. 161-6, March 1970.

Manuscript first received by the Institution on 21st February 1969 and in final form on 8th December 1969. (Report No. 1311/CC69.)

The Two-Frequency M.T.I. System

By

J. KROSZCZYŃSKI

Dr.habil.Eng.†

Clutter motion relative to the radar may seriously degrade the efficiency of moving target indication (m.t.i.) systems used in radio location. This problem arises in shipborne radar, airborne radar, or when the clutter is being carried by the wind. Although various methods have been devised to overcome the effect, the problem still exists because these have their inherent drawbacks. As another approach, the two-frequency m.t.i. system has been proposed occasionally in the past few years, but it seems that a more detailed consideration of this method is as yet not available in the literature. A basic theory of the two-frequency m.t.i. system, together with a discussion of its significant characteristics, is presented in this paper.

1. Introduction

In practical applications of the m.t.i. systems used in radio location it is often required to cancel clutter in situations where there is relative motion between the radar and the clutter. This can occur either when the interfering reflecting objects are moving, for example, the water droplets in a drifting cloud, or the radar itself is in motion, as in the case of a shipborne or airborne radar. The clutter spectrum will be then Doppler-shifted, which degrades the efficiency of clutter attenuation,¹ unless some method of compensation is used.

The compensation for the relative motion by adding a correcting frequency to the coherent oscillator (coho) output may be used, but the difficulty of this method lies in deriving an accurate control signal for the shift frequency.² For example, if a programmed variation of the coho frequency with azimuth angle is used, it may happen that the frequency shift will be not optimum for some range intervals. Various 'clutter-locking' schemes may be more effective, but have also their drawbacks in confrontation with intricate clutter configurations.³ The same can be said about the non-coherent m.t.i.; its chief limitation is that the target must be in the presence of relatively large and uninterrupted clutter signals if moving target detection is to take place.²

Another method has been proposed for overwhelming the difficulties mentioned above, in which the target is illuminated simultaneously at two radio frequencies and two coherent receivers are interconnected so as to produce an output dependent on

the phase difference between the received signals at the two frequencies.^{3,4} Although this system shows many interesting features, it seems that an analysis of its theory and characteristics has not yet appeared in the available literature. It is the purpose of this paper to present an elementary theory of the two-frequency m.t.i. system, and to discuss its basic characteristics.

2. Theory and Characteristics of the System

Consider a radar transmitter, emitting coherent signals at two different carrier frequencies f_1 and f_2 . Signal appearing at the input of the first receiving channel may be written as:

$$V_1 = E_1 \cos(\alpha_c \omega_1 t + \varphi_{1t} + \varphi_{1e}) + C_1 \cos(\alpha_c \omega_1 t + \varphi_{1t} + \varphi_{1c}) \quad \dots\dots(1)$$

where E_1 = amplitude of the target echo;

$\alpha_c = 1 + v_e \simeq 1 + 2v_e/c$, where v_e is the radial target velocity and c is the velocity of propagation;

φ_{1t} = initial transmitter phase in channel 1;

φ_{1e} = effective phase of echo reflected from the target;

C_1 = clutter amplitude; $\alpha_c = 1 + v_c \simeq 1 + 2v_c/c$,

where v_c is the radial velocity of passive reflectors;

φ_{1c} = effective clutter reflexion phase.

Similarly, signal received by the second channel is:

$$V_2 = E_2 \cos(\alpha_c \omega_2 t + \varphi_{2t} + \varphi_{2e}) + C_2 \cos(\alpha_c \omega_2 t + \varphi_{2t} + \varphi_{2c}) \quad \dots\dots(2)$$

The beat signal, obtained at the output of a narrow-band multiplicative mixer having V_1 and V_2 at the input, is given by:

$$V_b = \frac{1}{2} \{ E_1 E_2 \cos[\alpha_c(\omega_1 - \omega_2)t + (\varphi_{1t} - \varphi_{2t}) + (\varphi_{1e} - \varphi_{2e})] + C_1 C_2 \cos[\alpha_c(\omega_1 - \omega_2)t + (\varphi_{1t} - \varphi_{2t}) + (\varphi_{1c} - \varphi_{2c})] + E_1 C_2 \cos[(\alpha_c \omega_1 - \alpha_c \omega_2)t + (\varphi_{1t} - \varphi_{2t}) + (\varphi_{1e} - \varphi_{2c})] + E_2 C_1 \cos[(\alpha_c \omega_1 - \alpha_c \omega_2)t + (\varphi_{1t} - \varphi_{2t}) + (\varphi_{1c} - \varphi_{2e})] \} \quad \dots\dots(3)$$

In the single-subtraction m.t.i. canceller, direct and delayed (by T , time interval between pulses) signals

† Przemyslowy Instytut Telekomunikacji, Warsaw, Poland.

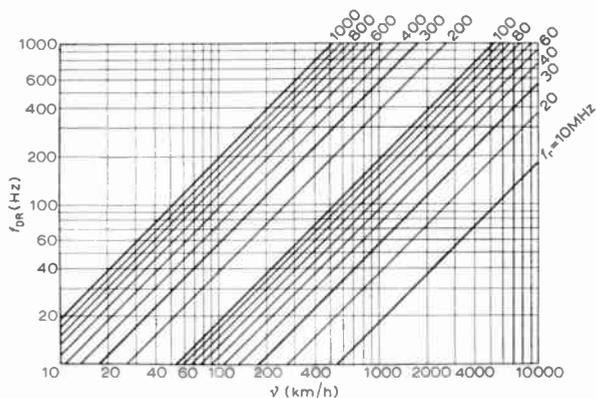


Fig. 1. The differential Doppler frequency f_{Dr} as the function of the carrier separation f_r and the radial velocity v of a reflecting object.

are subtracted from each other. Using known trigonometric relations and considering the necessary condition for i.f. cancellation, which in this case may be written in the form $(\omega_1 - \omega_2)T = n.2\pi$, where n is an integer, the canceller output is:

$$\begin{aligned}
 V_b(t) - V_b(t-T) = & E_1 E_2 \sin(\Delta\omega_c^D T/2) \sin[\Delta\omega t + \Delta\omega_c^D(t-T/2) + \Delta\phi_t + \Delta\phi_c] + \\
 & + C_1 C_2 \sin(\Delta\omega_c^D T/2) \sin[\Delta\omega t + \Delta\omega_c^D(t-T/2) + \Delta\phi_t + \Delta\phi_c] + \\
 & + E_1 C_2 \sin[(\omega_{1e}^D - \omega_{2c}^D)T/2] \sin[\Delta\omega t + \omega_{1e}^D(t-T/2) - \omega_{1c}^D(t-T/2) + \Delta\phi_t + (\phi_{1e} - \phi_{2c})] + \\
 & + E_2 C_1 \sin[(\omega_{1c}^D - \omega_{2e}^D)T/2] \sin[\Delta\omega t + \omega_{1c}^D(t-T/2) - \omega_{2e}^D(t-T/2) + \Delta\phi_t + (\phi_{2e} - \phi_{1c})]. \dots(4)
 \end{aligned}$$

where $\Delta\omega = \omega_1 - \omega_2$;

$\omega_{1e}^D = v_e \omega_1$ is the signal Doppler angular frequency for the first channel;

$\omega_{1c}^D = v_c \omega_1$ is the clutter Doppler angular frequency for the first channel;

ω_{2e}^D and ω_{2c}^D are the similar Doppler frequencies for the second channel;

$\Delta\omega_D = \omega_{1e}^D - \omega_{2e}^D = v_e \Delta\omega$ is the differential Doppler angular frequency for the signal;

$\Delta\omega_c^D =$ similarly for the clutter;

$\Delta\phi_t = \phi_{1t} - \phi_{2t}$;

$\Delta\phi_e = \phi_{1e} - \phi_{2e}$;

$\Delta\phi_c = \phi_{1c} - \phi_{2c}$.

Figure 1 shows the differential Doppler frequency f_{Dr} as the function of the carrier frequency difference $f_r = f_1 - f_2$ and the radial velocity v of a reflecting object.

The above considerations were formulated for c.w. signals, but it is easy to see that for coherent carrier waves, similar results are obtained for a train of short pulses having the recurrence frequency equal to $1/T$.

From eqn. (4), the following conclusions can be drawn:

- (1) For target in the clear ($C_1 = C_2 = 0$), the average output signal amplitude depends on the

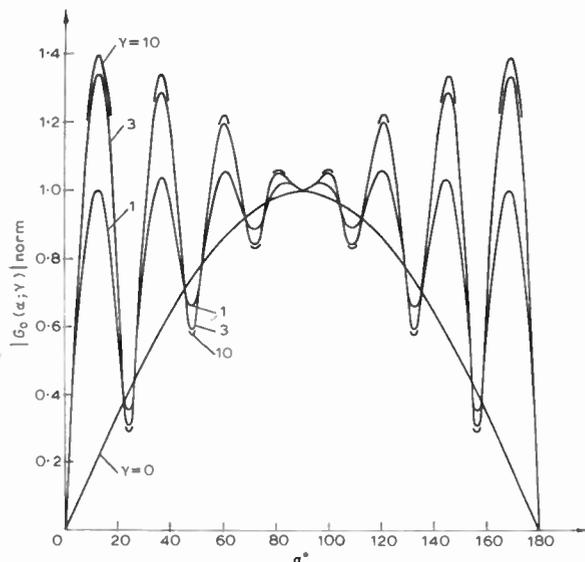


Fig. 2. The normalized two-frequency m.t.i. velocity response (single subtraction canceller, stationary clutter).

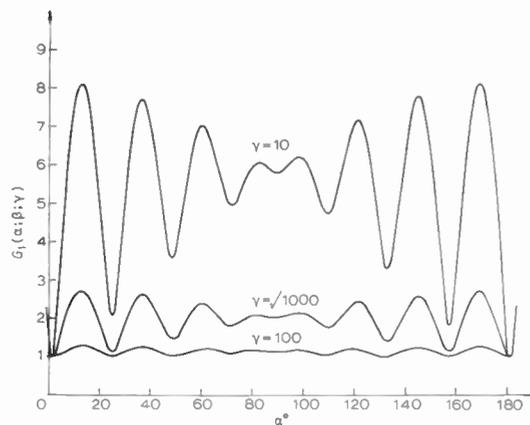


Fig. 3. The characteristics of single-subtraction two-frequency m.t.i. system (moving clutter, $\beta = 1^\circ$).

radial target velocity as in the case of a conventional single-frequency m.t.i. system having carrier frequency equal to $(f_1 - f_2)$.

- (2) When clutter only is present ($E_1 = E_2 = 0$), the amplitude of clutter residue at the canceller output depends on the radial drift velocity v_c in the same way as the target echo amplitude in case (1).

(3) For a target in clutter, the relations are somewhat complex, and will be discussed in detail below.

Assume that the target echo amplitudes in both channels are equal, $E_1 = E_2 = E$; a similar assumption, $C_1 = C_2 = C$, is quite reasonable for the clutter also. Consider at first the case of stationary clutter ($v_c = 0$). As the phase angles $\Delta\phi_i, \phi_e, \phi_c$ are random, the average amplitude of the output pulse train may be, according to eqn. (4), written as the r.m.s. value:

$$G_0 = \sqrt{E^4 \sin^2 (\Delta\omega_c^D T/2) + E^2 C^2 [\sin^2 (\omega_{1c}^D T/2) + \sin^2 (\omega_{2c}^D T/2)]}. \quad \dots\dots(5)$$

Denoting $\gamma = C/E$; $\alpha = \Delta\omega_c^D T/2$; $p = \omega_1/\Delta\omega$; $q = \omega_2/\Delta\omega$, eqn. (5) takes the form:

$$G_0(\alpha; \gamma) = E^2 \sqrt{\sin^2 \alpha + \gamma^2 (\sin^2 p\alpha + \sin^2 q\alpha)} \quad \dots\dots(6)$$

Figure 2 shows the normalized† $G_0(\alpha; \gamma)$ function for $\gamma = 0; 1; 3; 10$ and $p = 7, q = 8$. The curve for $\gamma = 10$ differs not much already from the result for $\gamma \rightarrow \infty$. It may be noted that for large values of γ the shape of G_0 is similar to the characteristics of a staggered p.r.f. single-frequency m.t.i. system³ for the stagger ratio T_1/T_2 corresponding to $p/q = \omega_1/\omega_2$ in eqn. (6).

$$G_1 = \frac{\sqrt{E^4 \sin^2 (\Delta\omega_c^D T/2) + E^2 C^2 [\sin^2 (\omega_{1c}^D - \omega_{2c}^D) T/2 + \sin^2 (\omega_{1c}^D - \omega_{2c}^D) T/2] + C^4 \sin^2 (\Delta\omega_c^D T/2)}}{C^2 \sin (\Delta\omega_c^D T/2)} \quad \dots\dots(8)$$

Using notation similar to that in eqn. (6), and $\beta = \Delta\omega_c^D T/2$, we obtain

$$G_1(\alpha; \beta; \gamma) = \frac{\sqrt{\sin^2 \alpha + \gamma^2 [\sin^2 (p\alpha - q\beta) + \sin^2 (p\beta - q\alpha)] + \gamma^4 \sin^2 \beta}}{\gamma^2 \sin \beta} \quad \dots\dots(9)$$

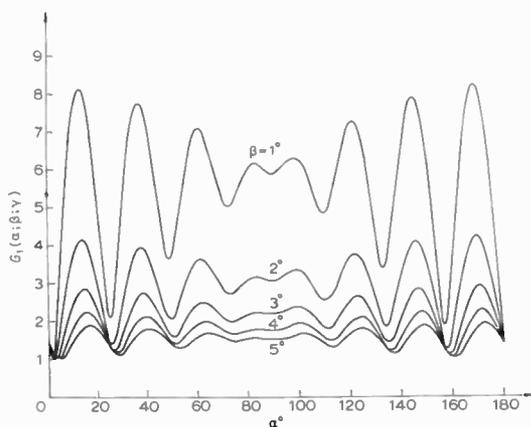


Fig. 4. The characteristics of single-subtraction two-frequency m.t.i. system (moving clutter, $\gamma = 10$).

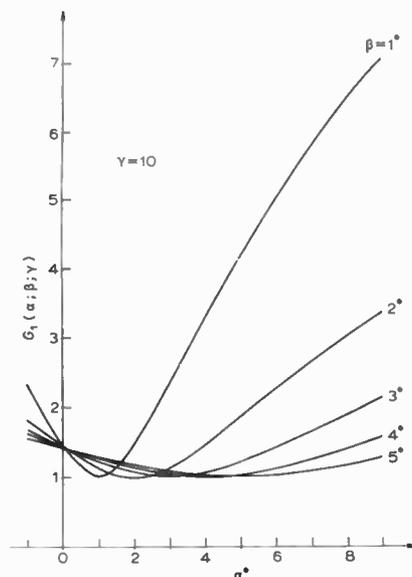


Fig. 5. Expanded portion of Fig. 4.

Figure 2 indicates that the velocity characteristics of the two-frequency m.t.i. system are fairly good in the case of targets in strong clutter, but for targets in the clear there is some signal loss for slower objects. Some method of gating may be therefore advised, allowing the m.t.i. system to operate mainly in clutter regions, outside of which a simpler receiving system could be used.

For the case of drifting clutter objects, it is convenient to define the ratio of r.m.s. amplitudes of

(target+clutter)/(clutter alone) at the output of the canceller. This can be expressed as

$$G = \frac{\text{r.m.s. output amplitude in presence of the target}}{\text{r.m.s. output amplitude in absence of the target}} \quad \dots\dots(7)$$

where the r.m.s. value of eqn. (4) for a given set of parameters should be inserted in the numerator, while the r.m.s. value of eqn. (4) for the same parameter values, except that $E_1 = E_2 = 0$, should be used in the denominator. It follows, that for a single canceller,

† $|G_0(90^\circ; \gamma)|_{\text{norm}} = 1$.

It may be noted that the radial clutter drift velocity corresponding to a given β value is:

$$v_c = \beta c / \Delta\omega T \quad \dots\dots(10)$$

Several diagrams of the $G_1(\alpha; \beta; \gamma)$ function are shown in Figs. 3-5. Figure 3 contains curves for $\beta = 1^\circ$ and $\gamma = 10; \sqrt{1000}; 100; p = 7; q = 8$.

Figure 4 shows the results for $\gamma = 10$ and $\beta = 1^\circ, 2^\circ, \dots 5^\circ$. A shift—equal to β in value—in the position of the minimum can be observed (this may be seen better on the expanded scale in Fig. 5) as well as the deterioration of signal/interference ratio resulting from increased β values.

It may be seen also from eqn. (9) and Figs. 3 and 4 that to obtain a good signal detection the approximate condition $\gamma \sin \beta \ll 1$ must be fulfilled.

Similarly, taking eqn. (4) and forming a second difference equation, the expression for a double-subtraction canceller may be obtained:

$$G_2(\alpha; \beta; \gamma) = \frac{\sqrt{\sin^4 \alpha + \gamma^2 [\sin^4 (p\alpha - q\beta) + \sin^4 (p\beta - q\alpha)] + \gamma^4 \sin^4 \beta}}{\gamma^2 \sin^2 \beta} \quad \dots\dots(11)$$

In this case, the approximate signal-detection condition $\gamma \sin^2 \beta \ll 1$ is valid. It follows, that the double-subtraction circuit gives approximately $(1/\sin \beta)$ -times better results than the single canceller. Figure 6 shows diagrams of the $G_2(\alpha; \beta; \gamma)$ function for $\beta = 1^\circ$ and $\gamma = 100; \sqrt{10^5}; 1000; p = 7; q = 8$. Figure 7 contains the curves for $\gamma = 100$ and $\beta = 1^\circ, 2^\circ \dots 5^\circ$. As it may be seen from the above, the double canceller should be possibly preferred for use in the two-frequency m.t.i. system.

3. Clutter Attenuation in the Two-frequency M.T.I. System

The diagrams in Figs. 3-7, although instructive by themselves, should of course not be taken as the actual standards of performance for the two-frequency m.t.i. system. In real situations, the influence of antenna scanning modulation and stochastic clutter fluctuations must be taken into account; these (besides of equipment instabilities) are the main factors

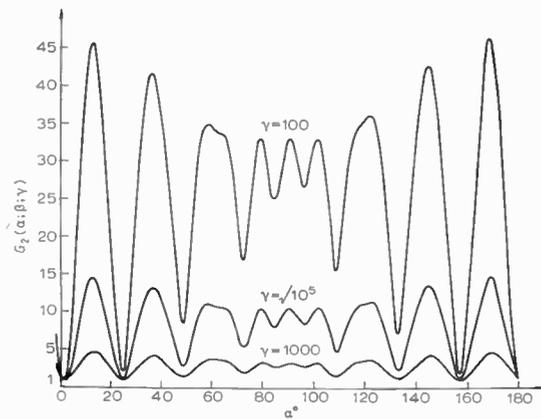


Fig. 6. The characteristics of double-subtraction two-frequency m.t.i. system (moving clutter, $\beta = 1^\circ$).

degrading the effectiveness of signal detection in clutter.

For most practical applications the clutter spectra in both receiving channels of the two-frequency m.t.i. system may be assumed to have equal and Gaussian shapes. The spectrum at the canceller input is also Gaussian in shape, but has a $\sqrt{2}$ -times broader bandwidth.† This effect brings about a slight decrease in clutter attenuation. On the other side, the two-frequency m.t.i. system is more immune to the Doppler shift of the interfering clutter signals. The clutter attenuation of moving clutter in a conventional system has been derived¹; for the system described here, in place of the Doppler frequency shift ω_c^D , the differential Doppler frequency shift $\Delta\omega_c^D$ must be taken. Taking together the two effects mentioned above, the clutter attenuation in the case of the two-frequency m.t.i. system may be written for the single canceller:

$$CA_1 = \frac{0.5}{1 - e^{-7 \cdot 12k^2} \cdot \cos(\Delta\omega_c^D T)} \quad \dots\dots(12)$$

and for the double canceller:

$$CA_2 = \frac{0.5}{3 - 4e^{-7 \cdot 12k^2} \cdot \cos(\Delta\omega_c^D T) + e^{-28 \cdot 48k^2} \cdot \cos(2\Delta\omega_c^D T)} \quad \dots\dots(13)$$

† This may be derived by taking into account that the spectrum of the product of two independent stochastic processes is equal to the convolution of the spectra of these processes.

where k is the ratio of the half-power clutter bandwidth to the pulse repetition frequency.¹

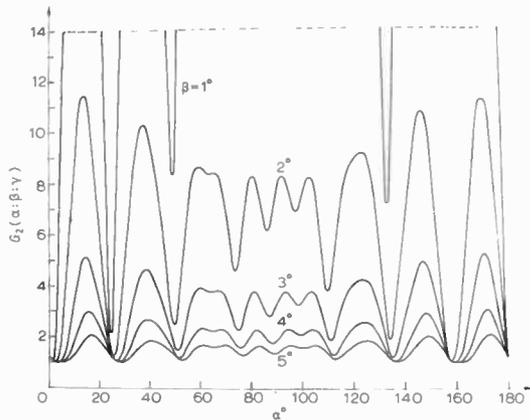


Fig. 7. The characteristics of double-subtraction two-frequency m.t.i. system (moving clutter, $\gamma = 100$).

As an example, the curves in Fig. 8 show the clutter attenuation CA_1 as the function of v_c for $k = 0.02$; curve 1 is valid for the conventional single-frequency system, while curve 2 applies for the two-frequency system with $\Delta\omega = 0.01(\omega_1 + \omega_2)$. It may be noted that already for a comparatively small clutter drift speed the effectiveness of a single-frequency system falls off rapidly, while for a two-frequency m.t.i. system the clutter attenuation remains fairly good for much higher drift speeds.

4. Acknowledgments

Acknowledgments are due to the Management of the P.I.T. for permission to publish this paper, and to Mrs. B. Mazurek, who programmed the computer.

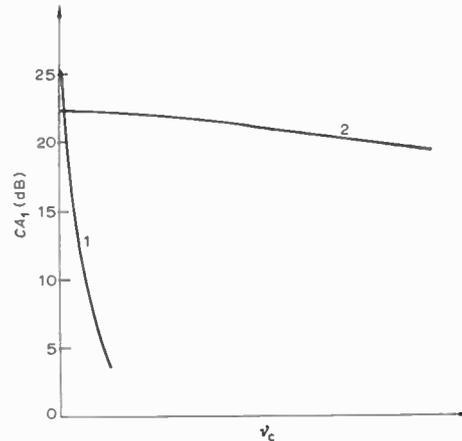


Fig. 8. Clutter attenuation in conventional (1) and two-frequency (2) m.t.i. system as the function of clutter drift speed.

5. References

1. Kroszczyński, J., 'Efficiency of attenuation of moving clutter', *The Radio and Electronic Engineer*, **34**, pp. 157-9, September 1967.
2. Skolnik, M. I., 'Introduction to Radar Systems' (McGraw-Hill, New York, 1962).
3. Bradsell, P. 'Moving Target Indication—A Survey of Developments since 1948', 8th AGARD Avionics Panel Symposium, September 1964.
4. Drushynin, W. W., *et al.* 'Spravotchnik po Osnovam Radiolokatsyonnoy Tekhniki' (Handbook on the Fundamentals of Radiolocation Technique), (Moscow, 1967).

Manuscript first received by the Institution on 7th May 1969 and in final form on 7th October 1969. (Paper No. 1312/AMMS28).

© The Institution of Electronic and Radio Engineers, 1970

The Author



Dr. Jan Kroszczyński received his first degree in electronic engineering from Warsaw Technical University. He was awarded the D.Sc. degree in engineering from the Institute of Fundamental Technical Problems of the Polish Academy of Sciences in 1964 and the D.Sc. habil. degree from the Warsaw Technical University in 1967. Since 1948 he has been with the Institute of Telecommunications (P.I.T.) in Warsaw. He is author of more than forty papers, mainly on optimum signal processing and signal design, author of a book on m.t.i. theory, and co-author of several other books of electronics. He has recently become a member of the Electronics and Telecommunications Committee of the Polish Academy of Sciences.

FOUNDED 1925
INCORPORATED BY
ROYAL CHARTER 1961

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

THE RADIO AND ELECTRONIC ENGINEER

The Journal of the Institution of Electronic and Radio Engineers

VOLUME 39 No. 4

APRIL 1970

Innovation and Industrial Expansion

To the economist the electronics industry in Great Britain and in other major industrial countries presents a paradox. It is, relatively speaking, a labour-intensive sector of manufacturing industry, despite the extension of automated production methods, and while most labour-intensive industries should display low overhead costs per unit of output, this is not the case in the electronics industry: its problem lies in its high rate of technological advance and innovation which call for far greater capital and current expenditure on research and development. In discussing the implications of this state of affairs the recently published Economic Assessment Report of the Electronics EDC* points out that the industry is hampered by the exclusion of R & D expenditure from the Government's investment grant scheme. The significance of this may be seen from the EDC's calculation that R & D costs in electronics are approximately five times as important in relation to conventional capital investment as the average for manufacturing industry.

The key to prosperity of the electronics industry is seen in the Report to be the encouragement of innovation but, in looking forward to the 1970s, concentration of effort over a more limited range of activities is regarded as essential. This calls for identification of areas and formulation of policies to assist industry in their development, processes which must take place both within firms and between industry and government, and the Report recalls some of the ways in which the Ministry of Technology has acted in recent years.

Computers have been the object of considerable industrial reorganization and this seems to have been successful in promoting central processor production and sales; there is, however, a serious need to expand production and sales of peripheral equipment for which the industry still depends heavily on imports. By contrast, while British companies have not been laggard in innovations in industrial automation systems, yet the Report is forced to point out that the market is at present only a quarter of the size of the computer market and has hardly grown in the last three years. Clearly this calls for co-operation between manufacturers, users, trade unions and government to build up a home base for what will be a world-wide growth area in the coming decade.

The EDC's Report finds that in general the telecommunications industry is improving its competitive strength, although it is suggested that the well-developed telecommunications network in the U.K. has tended to reduce the incentive to the industry to produce the most modern equipment required by those countries now building up their own networks. The association with the computer industry through the expansion of data transmission services is an important factor in promoting the spread of technological innovation and developments.

Because electronics is essentially in the role of a service industry to marine technology and depends on decisions to innovate and exploit resources that must be taken in other industries, the potentialities of electronic techniques have been realized rather slowly. The Report pays tribute to the stimulation which has been given in this direction by professional society activities such as the I.E.R.E. Conference on Electronic Engineering in Oceanography in 1966; a second Conference, on Electronic Engineering in Ocean Technology in September of this year, will give further support to Governmental efforts such as are provided by the Interdepartmental Advisory Committee on Marine Technology and by the National Electronics Council.

Several other areas for concentration of effort are identified in the Report—for instance integrated circuits and medical electronics. The policies for finding financial resources to maintain internationally competitive capability in all these sectors are seen to fall into the categories of tax concessions, direct grants and non-commercial loans, as well as market enlargement based on a large rapidly growing home market. These points have recently been made with some feeling by the Electronic Engineering Association. Innovation clearly cannot flourish without adequate financial backing—nor, we may add, without adequate manpower resources, another thorny subject which the Electronics EDC is investigating and will report on later this year.

F.W.S.

* 'Economic Assessment to 1972. Industrial Report by the Electronics EDC.' Published by the National Economic Development Office, Millbank Tower, London, S.W.1. (Free).

INSTITUTION NOTICES

Appointments to the Indian Council

The following appointments were made by the Indian Council at its meeting in February 1970:

Professor K. S. Hegde, M.A., B.E. (Fellow), of the College of Engineering, Madras, succeeds Professor J. N. Bhar as Chairman. Dr. A. N. Daw (Member) of the Institute of Radio Physics and Electronics, Calcutta, succeeds Mr. C. P. Joshi as Honorary Treasurer. Air Vice-Marshal K. Narasimhan, B.A. (Member), Director of Signals at Air Headquarters, Ministry of Defence, joins the Council.

Postponement of London Meeting

The joint meeting of the I.E.R.E. and I.E.E. Computer Groups announced for Wednesday, 27th May has been postponed. This Colloquium on Economics of Computer Maintenance and Operation will, it is hoped, be included in the Programme of Meetings for the 1970-71 Session with a slight change of emphasis in its theme. Offers to contribute to this Colloquium will be welcomed by the Joint Committee and should be sent to the Secretary of the Computer Group Committee, I.E.R.E., 9 Bedford Square, London, WC1B 3RG.

Cancellation of London Engineering Congress

The Council of Engineering Institutions has announced with regret that LECO 70 (which was to have been held from 4th to 7th May next) has been cancelled. Whilst interest had seemingly been high with a very brisk demand for information, the number of registrations received by the end of March was judged unlikely to have led to sufficient delegates to make the Congress viable as planned.

Reprints of Journal Papers

Reprints are prepared of all papers published in the *Journal* and copies may be obtained from the Institution, price 5s. 0d. each (post free). Requests for reprints may be made using the form which is included in the end pages of most issues of the *Journal*. It is particularly asked that remittances be sent with orders to avoid book-keeping entries and thus reduce handling costs.

Institution Giro Account

Members are advised that for the convenience of those who wish to remit their annual subscriptions and other payments through the National Giro, the Institution now has a Giro account. The number is 578 0101 and it may be used in precisely the same manner as any other Giro transaction.

Conference on Laboratory Automation

In recent years, automatic techniques have been introduced into many laboratories to speed up experimental and analytical procedures, and to reduce time spent by staff on repetitive work. On-line computers have been installed, facilitating the handling of large quantities of data, with immediate processing and presentation to the experimenter. In some cases, the computers have been used to control the actual experiments, for example the movement of a set of detectors to new, accurately controlled positions when sufficient data have been accumulated. Automatic equipment has been designed for carrying out chemical and biochemical analyses on a large number of samples simultaneously, such as blood and urine in a pathological laboratory. Similar applications have occurred in other research laboratories.

A Conference on Laboratory Automation has therefore been arranged to bring together workers who are already applying automatic techniques in their laboratories, or who may be interested in so doing, and designers and manufacturers of such equipment.

Organized by the Institution of Electronic and Radio Engineers with the association of the Institution of Electrical Engineers, the Institution of Chemical Engineers, the Royal Institute of Chemistry, the Institute of Physics and the Physical Society and the Institute of Measurement and Control, it will be held at the Middlesex Hospital Medical School, Cleveland Street, London, W.1, from Tuesday, 10th November to Thursday, 12th November, 1970.

The main themes of the Conference will be Automatic Analysis and Computer Controlled Experiments, and will include:

- Sensors and measurement techniques
- Automatic tests and calibration
- Data acquisition
- On-line data reduction, presentation and analysis
- Control techniques and instrumentation
- Automatic sample preparation and handling
- Automatic sample processing and dispensing of reagents

The term 'Laboratory' is intended to cover scientific observatories and routine testing laboratories as well as research and development laboratories. It is hoped to receive papers from workers in many fields, including geophysical, agricultural, chemical, biochemical, nuclear, electrical, rubber and building laboratories.

Synopses of proposed contributions are invited and should be sent to the I.E.R.E. as soon as possible. Further information and registration forms for the Conference will be available in due course from the Conference Registrar at 9 Bedford Square, London WC1B 3RG.

The Synthesis of Asynchronous Digital Pattern Generators

By

M. C. WATERS. B.Tech.†

and

D. P. BURTON, Ph.D., M.Sc.,‡

The paper describes a method for designing circuits which generate predetermined digital waveforms when triggered by an input. To do this an asynchronous sequential machine is caused to move through a number of states under its own control. Individual portions of the waveform can be adjusted without interfering with the rest of the output sequence. The circuit can also be made to give different output sequences for different inputs.

1. Introduction

Many digital processes require a circuit which generates predetermined waveforms when triggered by an input signal. This paper considers the flow table structure of such devices and gives a general design method. The principle used is to introduce a series of multiple transitions into an asynchronous sequential machine by using delayed feedback from output to input. Lewin¹ has described a method for using the transition time of an asynchronous machine to produce a delay which is a function of the gate propagation time. However, gate delays tend to drift and the present system overcomes this problem by controlling the transition time with delays external to the machine. Figure 1 shows a schematic diagram of the pattern generator which is composed of a sequential machine and one or more feedback delays. The delay can be a simple CR network, a self-controlled shift register or any other device which can effectively delay the rise and fall of a single pulse.

2. Basic Principles

Suppose two pulses of known width are to be generated from a single poorly defined input pulse as shown in Fig. 2(a). Figure 2(b) gives the flow table of an asynchronous machine which will perform this function; the single output Z is fed back through a delay τ to the input.

Let the machine be in stable state 1 with x and Z equal to '0' and let x now change to '1'. The machine will move to state 2 as defined by input column 10. For state 2 the output Z is '1' and after passing through the delay it will arrive at the input. The machine now has inputs $Z = '1'$ and $x = '0'$ or '1' and this causes the machine to move to state 3. In state 3 the output goes to '0' and after a delay will change the machine into state 4. Here again the output alters and the device goes to state 5. At this stage the machine remains in state 5 until x returns to '0'. If x has gone to '0' during the transitions the machine transits to state 1 after the output for state 5 has

† Department of Electronics, University of Southampton.

‡ Formerly at the University of Birmingham; now with Husband & Co., 388 Glossop Road, Sheffield, S10 2JB.

reached the input. Thus each input pulse generates two output pulses whose width is primarily governed by the feedback delay.

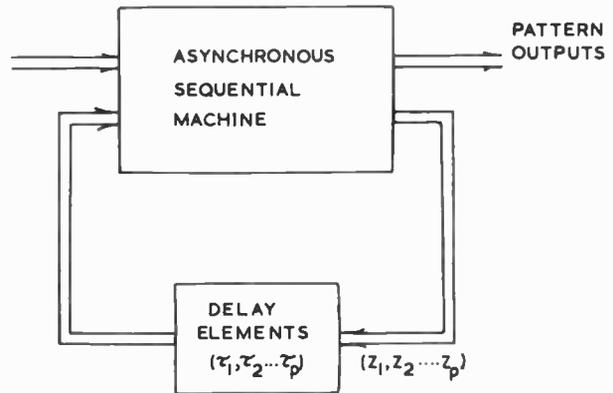
In effect the machine oscillates through a number of states and the oscillating variable, in this case, performs the dual purpose of being the output and providing the 'clock' pulse for initiating changes of state. By varying the delay the pulse width can be adjusted. This type of pulse multiplier flow table can be generalized for any number of output pulses. Rows 1 and 5 represent the starting and finishing rows respectively and appear at the top and bottom of every flow table of this type. The remainder of the flow table is of conventional binary counter form with the sole exception of the 'don't care' state in row 2. It is well to consider how this 'don't care' state can be used. If the entry is assigned as a transition to state 1 then the input x must be '1' at least until the machine has entered state 3; this implies a minimum input pulse width of $(\tau + \delta)$, where δ is the transition time of the sequential machine. But if the entry is filled in as a stable state 2 then the next input pulse need only be long enough to take the machine into state 2, which can be as little as three gate delays. Figure 3 shows the general form of a pulse multiplier flow table. The number of states in the flow table is equal to $(2n + 1)$ where n is the number of output pulses required from the machine.

3. Extensions of the Method

So far only tables with one input and one feedback delay have been considered. In this Section examples are used to show how the procedure can be extended to the design of pattern generators which give different outputs for different inputs and those which use more than one delay.

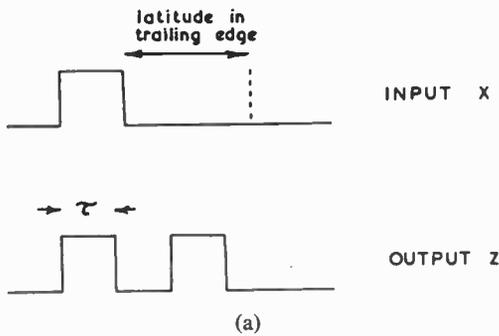
Figure 4 shows the flow table for a circuit which gives three output pulses if x_1 is put to '1' and two output pulses if x_2 is triggered. It is assumed that x_1 and x_2 are mutually exclusive and that a second trigger pulse does not occur until the machine has completed the transitions in hand. If x_1 is put to '1' the machine successively moves through states 1 to 7 generating three output pulses in the process, but if

Fig. 1. System model. (x_1, x_2, \dots, x_q)



x_2 is put to '1' the machine 'jumps' to state 4 before beginning its series of transitions.

More flexible behaviour can be obtained by using a number of feedback delays and Fig. 5(a) shows the flow table for a pattern generator which uses two delays τ_1 and τ_2 ; Z_1 is fed back via τ_1 and Z_2 via τ_2 , Z_3 is the pattern output and its theoretical shape is shown in Fig. 5(b). If τ_1 is altered only those parts of the waveform that depend upon τ_1 are changed and similarly for τ_2 . The circuit is primarily intended



(a)

	X Z				Z
	00	01	11	10	
1	①	-	-	2	0
2	-	3	3	②	1
3	4	③	③	4	0
4	④	5	5	④	1
5	1	⑤	⑤	⑤	0

(b)

Fig. 2. Pulse multiplier.

	X Z				Z
	00	01	11	10	
1	①	-	-	2	0
2	-	3	3	②	1
3	4	③	③	4	0
2n	②n	2n+1	2n+1	②n	1
2n+1	1	②n+1	②n+1	②n+1	0

Fig. 3. Generalized pulse multiplier flow table.

to function for the case when τ_2 is greater than τ_1 and a series of possible transitions made by the machine under these circumstances is shown by the arrows in the right-hand half of the flow table.

Suppose that the machine is stable in state 1 and all inputs are at '0'. Now let x go to '1' and remain so for the rest of the transitions. In this way the discussion of the circuit's behaviour is restricted to the right-hand half of the flow table although in fact it is only necessary for x to be at '1' long enough for the machine to get into state 2. To continue: the x change causes the machine to enter input column 100 where it makes a transition to state 2. In state 2 outputs Z_1 and Z_3 go to '1'. The machine remains in state 2 for time τ_1 until the Z_1 change reaches the input, whereupon it moves to input column 110 and goes to state 3 causing Z_1 to return to '0' and Z_2 to change to '1'. For a time neither of these changes will reach the machine's input and the circuit will be stable for input 110. Eventually the Z_1 change will arrive (assuming $\tau_1 < \tau_2$) and the machine will move to column 100 which is also stable for state 3. When the Z_2 signal gets to the input the machine moves to

	$x_1 x_2 Z$						Z
	000	001	011	010	100	101	
1	①	-	-	4	2	-	0
2	-	3	-	-	②	3	1
3	4	③	-	-	4	③	0
4	④	5	5	④	④	5	1
5	6	⑤	⑤	6	6	⑤	0
6	⑥	7	7	⑥	⑥	7	1
7	1	⑦	⑦	⑦	⑦	⑦	0

Fig. 4. Pulse multiplier with two inputs.