THE RADIO AND ELECTRONIC ENGINEER

The Journal of the Institution of Electronic and Radio Engineers

FOUNDED 1925 INCORPORATED BY ROYAL CHARTER 1961

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

VOLUME 32 JULY 1966 NUMBER 1

A COMBINED RESEARCH EFFORT

ONE of the declared purposes of the National Electronics Research Council (N.E.R.C.) is to encourage co-ordination of electronics research activity in Great Britain and fostering co-operation of effort throughout the British Commonwealth.

The June 1966 issue of *N.E.R.C. Review* includes a report of the Secretary of the Institution on the discussions he had with research administrators in many countries during his recent world tour on the development of Commonwealth collaboration in certain fields of electronics research.

The first project of N.E.R.C. on Selective Dissemination of Information (S.D.I.) has attracted world-wide attention and in several Commonwealth countries active co-operation in the project has now been secured. New Zealand, for instance, which has recently set up its own National Electronics Research Council, is arranging for research workers to take part in the S.D.I. project. The New Zealand National Research Advisory Council is anxious to avoid duplication of effort and has stated: 'New Zealand scientists do, of course, make maximum use of basic research results obtained in other countries. But . . . it is only right that we in our turn should put something back into the pool.'

The S.D.I. project has aroused considerable interest in Canada and was the subject of a lecture given in Ottawa in April last by Lord Mountbatten, Chairman of N.E.R.C. and Past President of the Institution. Lord Mountbatten's lecture was based on an Address, 'Controlling the Information Explosion', given before the Royal Institution of Great Britain in February 1966.† Lord Mountbatten also met members of the scientific staffs of the Defence Research Board and the National Research Council and discussed how Canadian scientists and engineers could best participate in the project.

Present information services within Australia operate on rather restricted lines within organizations or groups, for instance within the Weapons Research Establishment or the Australian affiliates of American and Japanese manufacturers. The National Science Foundation and other organizations are now considering how the S.D.I. project may be applied to Australian needs.

Probably the greatest problems in coping satisfactorily with research in electronics (as well as in other disciplines) exist in India, where the requirements for expansion of production and development over the next ten years should rise to £100 million per annum, according to the report prepared by the late Dr. H. J. Bhabha, F.R.S. Financial difficulties resulted in the contraction of some of the work of the Indian National Scientific Documentation Centre; the possibilities of collaboration in the S.D.I. project are, therefore, seen to be a particularly attractive proposition. The mutual advantages of Indian participation are demonstrated by the great contributions which India has made and will continue to make in ionospheric exploration, new materials and environmental studies.

The *N.E.R.C. Review* aptly epitomizes the advantages of Commonwealth co-operation through a chain of electronics research councils, with a quotation, again from New Zealand:

> 'Furthermore, there is little use in assessing priorities unless there is sufficient basic scientific information available to give an increased scientific effort a chance of achieving the economic or social aims for which it was undertaken and unless the research programme is administered in such a way as to offer our best scientists a congenial career.'

Few engineers would dispute this statement and the achievement of N.E.R.C. in initiating a valuable project and promoting Commonwealth co-operation in the whole field of electronics research is a commendable contribution to the advancement of science.

F. W. S.

† *The Radio and Electronic Engineer*, **31**, No. 4, pp. 195–208, April 1966.

INSTITUTION NOTICES

Institution Premiums and Awards

The Council of the Institution announces that the following awards are to be made for outstanding papers published in *The Radio and Electronic Engineer* during 1965:

CLERK MAXWELL PREMIUM

'Statistical Optimization of Antenna Processing Systems' by G. O. Young (February).

HEINRICH HERTZ PREMIUM

'The Mechanism and Device Applications of High Field Instabilities in Gallium Arsenide' by J. S. Heeks, A. B. Woode and C. P. Sandbank (December).

LORD RUTHERFORD AWARD

'A Correlator for Investigating Random Fluctuations in Nuclear Power Reactors' by F. D. Boardman, E. L. E. Harrington and D. J. A. Carswell (September).

A. F. BULGIN PREMIUM

'The Mechanism of Interference Pick-up in Cables and Electronic Equipment with special reference to Nuclear Power Stations' by D. Harrison (March).

MOUNTBATTEN AND BOSE PREMIUMS

'Dielectric Loaded Waveguides—a Review of Theoretical Solutions' by S. K. Chatterjee and Mrs. R. Chatterjee (September, October, November, December).

(The Mountbatten Premium is for the most outstanding paper read before a meeting in India, while the Bose Premium is for the most outstanding paper by an Indian scientist or engineer published in the *Journal*.)

MARCONI AWARD

'A Radar Receiving Array with I.F. Multiple-beam Forming Matrix' by J. Salomon, S. Pichafroy and P. Hurbin (May).

LESLIE MCMICHAEL PREMIUM

'An Analysis of Results obtained on an Aircraft Data Linkout to 1300 nautical miles (2400 km)' by W. J. Battell (August).

J. LANGHAM THOMPSON PREMIUM

'The Determination of the Parameters of a Dynamic Process' by P. C. Young (June).

BABBAGE AWARD

'Large Capacity Magnetic Film Stores—A Design Approach' by P. I. Bonyhard and W. S. Carter (March).

DR. NORMAN PARTRIDGE MEMORIAL PREMIUM

'A Groove Feed and Depth Gramophone Control System for Phonograph Disk Cutting Equipment' by H. Lindskov Hansen (December).

REDIFFUSION TELEVISION PREMIUM

'The Performance Requirements of a Television Monitor Receiver (Nyquist Demodulator) and Methods of Measurement' by F. G. Johannesen (September).

The following Premiums and Awards are withheld as papers of suitable standards have not been published during the year: the P. Perring-Thoms Premium, the Lord Brabazon Award, the Arthur Gay Premium, the Zworykin Award and the Hugh Brennan Premium.

The Premiums and Awards will be presented by the President of the Institution at the Annual General Meeting in London in December.

Institution Dinner in Southampton

The Institution's Conference on Electronic Engineering in Oceanography at Southampton University from 12th to 15th September 1966 will be marked by an Institution Dinner to be held on Wednesday, 14th September, at the Polygon Hotel, Southampton, 7 p.m. for 7.30.

Members of the Institution are cordially invited to the dinner, which will also be attended by those taking part in the Conference. The cost of tickets is £2 each, which includes wines at table; the presence of ladies and other guests will be welcomed.

It is believed that members both in the Southern Section area and in other parts of the country will find this opportunity for an Institution function to be held outside London convenient and pleasant. There will not be an Institution Dinner in London this year.

The C.E.I. Register of Chartered Engineers

The Council of Engineering Institutions is authorized under the Royal Charter of Incorporation granted to that body on the 3rd August 1965, to maintain a Register of Chartered Engineers. Admission to that Register can only be achieved through corporate membership of one of the constituent Institutions of C.E.I.

The situation regarding admission to the C.E.I. Register of Chartered Engineers of corporate members of this Institution is as follows:

(a) All those who were corporate members of the Institution on the 3rd August 1965 are automatically included in the C.E.I. Register of Chartered Engineers and are entitled to use the abbreviated designation C.Eng. (See *The Radio and Electronic Engineer* for August 1965, page 69.)

(b) Subject to final approval by the Privy Council, all those who achieve election to corporate membership of this Institution after 3rd August 1965 and not later than the 1st January 1974 will be admitted to the C.E.I. Register of Chartered Engineers by virtue of having satisfied the Institution's own requirements for election to corporate membership.

(c) All those elected to corporate membership of this Institution after the 1st January 1974 will be eligible for admission to the C.E.I. Register of Chartered Engineers, providing they have satisfied the academic requirements of the C.E.I.

U.D.C. 535: 681.39

Some Methods of Signal Processing Using Optical Techniques

By

D. C. COOPER, Ph.D., C.Eng.†

Presented at a meeting of the Institution in London on 13th October 1965.

Summary: In recent years optical techniques have been utilized in a number of different signal processing systems. Certain types of operation such as multiplication and Fourier transformation are easily implemented using optical techniques, and this paper describes a selection of signal processing systems which incorporate some optical components and have possible applications in many fields such as seismology, radar, sonar, speech recognition and vibration analysis.

Particular attention has been directed towards systems in which the interaction of acoustic waves and monochromatic light is used to provide a dynamic means for modulating the light with signal waveforms, and some adaptations of existing systems of this type are proposed.

1. Introduction

During the past decade there has been a marked increase in interest in physical optics and optical methods for processing information. The application of Fourier transform ideas has completely changed the physical concept of image formation and the utilization of the Fourier transform properties of converging lenses has produced many useful signal processing systems.[1] In particular optical techniques may well be used in the future for many applications in the field of signal processing.

In this paper we shall consider a representative selection of signal processing systems and some new adaptations will be proposed. It is convenient to classify the systems on the basis of whether or not the proper system operation depends on the use of collimated monochromatic light. Systems requiring the use of monochromatic light will be called coherent light systems and those that do not have this requirement will be called incoherent light systems.

In both coherent and incoherent light systems the information or signal may be stored and processed as a spatial pattern of varying optical density, and a photographic plate or film is an obvious choice for a recording medium. Here we immediately find an advantage and a disadvantage of optical processing systems.

The advantage is the ability to deal with a large quantity of information in a relatively small system. The number of resolvable elements on a good quality 35 mm transparency is in the region of 10^6 and there are many density levels available for each element.

† Department of Electronic and Electrical Engineering, University of Birmingham.

The disadvantage is the need to use a recording medium which requires chemical processing. In some applications this may not be a real disadvantage but in other cases the difficulty is avoided by various means. In addition the optical transmission factor of any transparency can only lie in the range from zero to unity so that a bias must be added to any bipolar signal before it is recorded.

A basic operation in optical signal processing is effected by passing light through a transparency. The output light intensity at any point in the plane of the transparency is given by the product of the incident intensity at the point and the local transmission factor. Thus the basic operation is one of *multiplication*. The presence of a bias component in the transmission factor will obviously lead to an output light intensity which has an unwanted component corresponding to the product of the bias and the input light intensity. Very often this unwanted component only produces a shift in the mean level of the output of a signal processing system but in certain circumstances special arrangements must be made to eliminate the effects of the recording bias.

So far we have implied that information is represented, in an optical system, by the intensity of the light at a point in a plane, and we have assumed that the signal functions are recorded as density variations on a photographic plate or film. In coherent light systems the phase of the light at any point can also carry information and media which modify the light phase can be used to introduce the signal functions into the system. A convenient method for producing phase modulation is the use of acoustic waves in transparent media and examples of this type of system will be described.

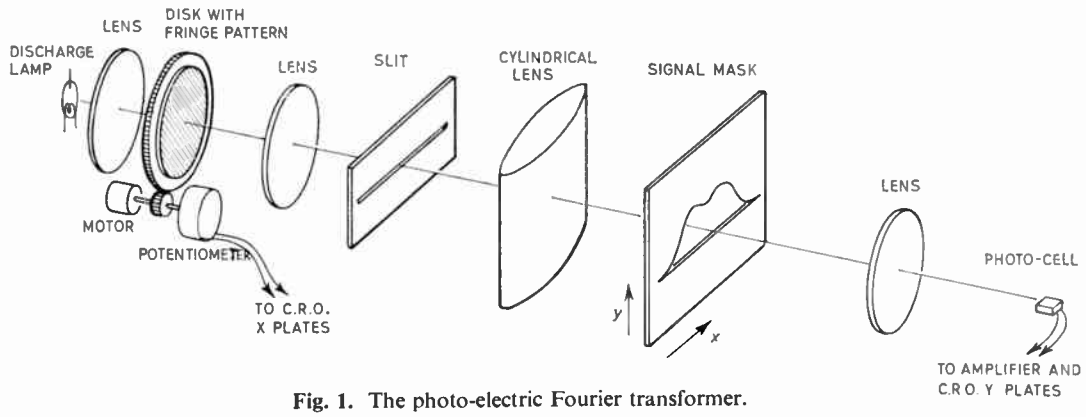


Fig. 1. The photo-electric Fourier transformer.

From the foregoing remarks we see that an optical signal processing system offers a compact method for storing and processing large quantities of data, but we shall see that the processing operations which can be performed are somewhat limited in variety.

2. Incoherent Light Systems

2.1. *The Photo-electric Fourier Transformer*

This device is an early example of an optical processing system and the original description was given by Born, Furth and Pringle.[2] Figure 1 shows the arrangement of the system.

The integral which is evaluated is of the form

$$F'(\omega) = \int_a^b f(x) \cos(\omega x + \theta)\, dx \qquad \ldots\ldots(1)$$

and it is readily seen that if $F'(\omega)$ is obtained for $\theta = 0$ and $\theta = \pi/2$ we have information which prescribes the complex Fourier transform

$$F(\omega) = \int_{-\infty}^{\infty} f(x) e^{-j\omega x}\, dx \qquad \ldots\ldots(2)$$

providing that $f(x) = 0$ for $x < a$ and $x > b$, where a and b define the limits of the optical aperture of the system.

The integral of eqn. (1) is evaluated in the system by the production of a light illumination, at the signal mask, which has an intensity that varies harmonically in the x direction. The intensity corresponds to a constant illumination or bias plus a part which varies as $\cos(\omega x + \theta)$. The aperture in the signal mask represents $f(x)$ measured in the y direction, and the total light emerging from the mask and imaged onto the photo cell is proportional to

$$\int_a^b f(x) B\, dx + \int_a^b f(x) \cos(\omega x + \theta)\, dx \qquad \ldots\ldots(3)$$

where B is the bias illumination and the integration is the result of the imaging process.

The presence of the term containing the bias factor causes no difficulty in this case since it only produces a mean output level which does not change as ω is changed.

The harmonically varying light pattern is produced by a glass disk carrying a photographically produced fringe pattern in which the transmission factor has a component which varies sinusoidally. Obviously a bias must be included in the overall transmission factor for the disk and this results in the bias term in the output which we have already discussed.

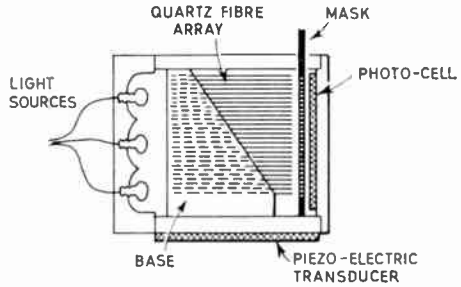
A beam of light of uniform intensity illuminates the glass disk and the transmitted light is imaged on a slit. The light intensity along the slit will have a sinusoidal component with a period which varies with the angular position of the glass disk. Thus ω can be changed by rotating the disk.

In use the motor driving the glass disk rotates continuously to change ω and the X deflection of a cathode ray oscilloscope is synchronized to the rotation. The output of the photo-multiplier is displayed as a vertical deflection of the cathode-ray tube spot and in this way a plot of $F'(\omega)$ versus ω is formed. The ω scale will be non-linear unless special steps are taken to avoid this. The value of θ is varied by changing the centre of rotation of the glass disk.

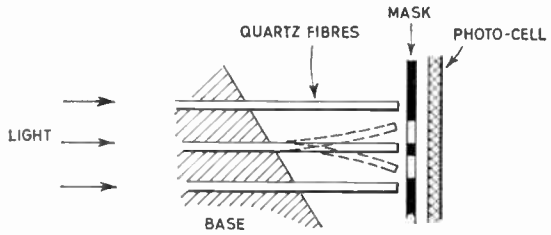
2.2. *The Sceptron*

This is a real-time filter, operating at audio frequencies, which was described recently by Hawkins.[3]

Figure 2(a) shows the arrangement of the Sceptron. The heart of the device is an array of quartz fibres which are supported by a shaped base. The base and its fibres are mounted on an electro-mechanical



(a) Arrangement.



(b) Acceptance mask detail.

Fig. 2. The Sceptron.

transducer to which is applied the electrical signal which is to be filtered.

It will be observed that each optical fibre is mounted as a cantilever beam and because of the low mechanical losses of the quartz fibre it has a mechanical resonance of high Q factor at a frequency which is determined by the fibre length. The contour of the base provides an array of fibres of various lengths so that a large frequency spectrum can be investigated at any instant of time.

The quartz fibres form part of the optical path of the Sceptron and they transmit light from the source to a mask which either passes the individual light contributions to the photocell, or blanks them off.

The most useful role of the Sceptron involves the use of what is called an acceptance mask, and a few fibres together with a section of an acceptance mask are illustrated in Fig. 2(b). The acceptance mask is formed by superimposing the negative produced by exposure to the static fibre system, and a positive (i.e. reversed negative) produced by an exposure to the fibre system excited by the signal that is to be recognized.

With this type of mask the maximum output from the photocell will be the result of the application to the system of a signal whose spectral components match those of the required signal.

The Sceptron can be programmed in this way to recognize very complex signal forms. Hawkins[3] describes the results of a successful test of the ability of the Sceptron to recognize the spoken word.

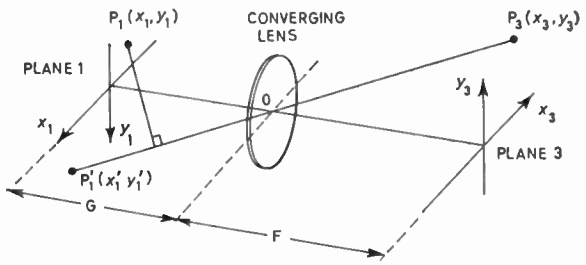


Fig. 3. Diagram relating to the transform properties of a converging lens.

3. Coherent Light Systems

3.1. *The Fourier Transform Properties of a Converging Lens*

A converging lens such as is shown in the arrangement of Fig. 3 transforms the distribution $\hat{E}_1(x_1 y_1)$ of the monochromatic light incident at plane 1 into another distribution $\hat{E}_3(x_3 y_3)$ in plane 3. The relationship between these light distributions is now well known[4] but for completeness it is formulated in Appendix 1 where it is shown that

$$\hat{E}_3(x_3 y_3) = \text{const} \times \iint \hat{E}_1(x_1 y_1)\, e^{-j\omega_x x_1}\, e^{-j\omega_y y_1}\, e^{j\beta}\, dx_1\, dy_1 \qquad \text{.....(4)}$$

where

$$\omega_x = \frac{2\pi x_3}{\lambda F}, \qquad \omega_y = \frac{2\pi y_3}{\lambda F}$$

and

$$\beta = -\left(1 - \frac{G}{F}\right)\left(\frac{x_3^2 + y_3^2}{2F}\right)$$

If the region in which $\hat{E}_1(x_1 y_1)$ is non-zero is within the angular field of view of the converging lens the limits of integration in the above expression may be considered to be infinite. Under this condition we see that the relationship is a two-dimensional Fourier transform multiplied by a phase factor $e^{j\beta}$, and it is important to note that this phase factor becomes unity if G is made equal to F.

Thus we see that there is a true two-dimensional Fourier transform relationship between the light distributions in the focal planes of the converging lens.

Many electronic engineers are used to thinking in terms of the Fourier frequency components of time varying waveforms and in the optical case there is a similar relationship. Hence if the light distribution in the first (input) focal plane of a converging lens is related to a signal waveform it is natural to say that the signal spectrum is produced by the lens at the second focal plane. In many optical systems the signal transparency provides information which is a function of only one variable, say, x, and astigmatic

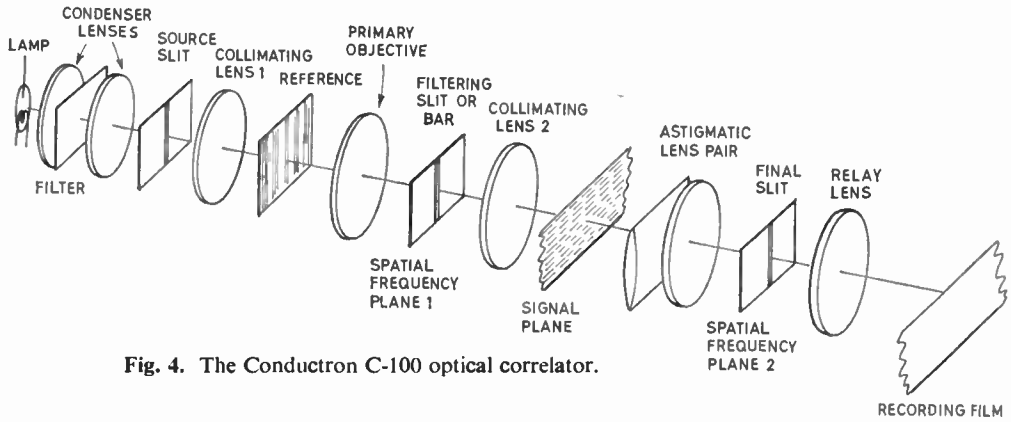


Fig. 4. The Conductron C-100 optical correlator.

lenses are used to obtain a Fourier transform relationship in the x direction while imaging in the other dimension y.

A point worth noting at this stage is the value of $\hat{E}_3$ at the point $x_3 = 0$, $y_3 = 0$, that is, on the optical axis. We see that at this point

$$\hat{E}_3(0,0) \propto \int\int \hat{E}_1(x_1 y_1)\, dx_1\, dy_1 \qquad .....(5)$$

giving a light intensity proportional to the square of the integral of $\hat{E}_1(x_1 y_1)$ over the first focal plane.

A further point of importance is worth noting. The detail structure of the spatial frequency distribution produced by a lens will have a size which is inversely proportional to the detail size in the original light distribution. The scale of the spatial frequency distribution can be increased by increasing the focal length of the transforming lens or by increasing the wavelength of the light used in the system. As an example let us assume that we have a transparency which produces light distribution which has a sinusoidally varying component in one direction only repeating at intervals of 1 mm (10^{-1} cm). If we take the wavelength of the coherent light to be 5000 Å and assume that we use a lens of focal length 50 cm to obtain the spatial frequency distribution, we find that we get two light spots representing the sinusoidal component and these are spaced from central or mean illumination (bias) component by a distance of only 0·25 mm.

Obviously a long focal length lens will be required if extremely fine detail is to be avoided in the spatial frequency distribution when the original detail is not very fine.

3.2. *The C-100 Optical Correlator*

This is a system manufactured by the Conductron Corporation of Ann Arbor, Michigan, U.S.A., and its arrangement is shown in Fig. 4. The arrangement of the light source and the use of a slit is such that a beam of monochromatic light is produced which is only collimated for the transverse x direction. A reference transparency is placed in the reference plane and the Fourier transform of the horizontal or x distribution of the light leaving the reference plane is produced in the first spatial frequency plane by the primary objective.

Filtering of the spatial frequency components of the reference function is possible in the first frequency plane. One very useful form of filtering is the removal by a stop bar of the light carrier component which lies on the optical axis. This results in a spectrum which corresponds to that for the reference transparency without bias.

The second collimating lens again transforms the light distribution in the horizontal plane so that the filtered reference function appears as the horizontal distribution of the light incident on the signal plane.

A signal transparency may contain up to one thousand separate signal tracks stacked vertically, and the filtered reference light distribution is applied to all the tracks. Each signal function and the reference function are multiplied in determining the light distribution at the output of the signal plane at each track position.

The astigmatic lens pair now provides a further Fourier transformation of the horizontal distribution of the light output from the signal plane, with focusing in the vertical plane to retain channel separation. The final slit accepts only the components on the vertical through the optical axis and, as we have already noted, this corresponds to accepting components which are proportional to the integrals of the reference-signal products.

Hence the output of the final slit at each channel position corresponds to the correlation integral of the reference function (perhaps filtered) and the signal in the channel, and the final relay lens images these outputs on the recording film. The correlation integral for each track is therefore recorded as a

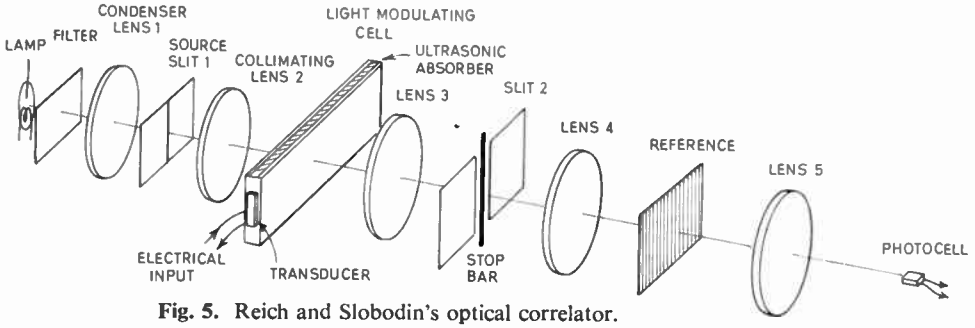


Fig. 5. Reich and Slobodin's optical correlator.

spot of varying density and by moving the signal film and the recording film synchronously a density record of the cross-correlation function with the given reference is obtained for each channel.

The unit may also be used for recording signal spectra by removing the reference transparency and opening out the final slit to allow all components of the Fourier transforms of the signal functions to pass to the recording film. Again the spectra are recorded as density variations on the recording film.

The C-100 correlator has been designed as a general purpose correlator and spectrum analyser and it can be used for the analysis of signals of many different origins, such as for example signals recorded in seismological work.

3.3. *A Correlator using Ultrasonic Light Modulation*

It is often desirable to avoid the need for recording a signal on film and the use of ultra-sound in a transparent medium is one way of achieving this end. A correlator using this technique has been described by Reich and Slobodin[5] and its arrangement is illustrated in Fig. 5.

The signal to be processed is applied in the form of an electrical waveform to the piezo-electric transducer of the ultrasonic light modulator, and an acoustic version of the waveform is propagated in a transparent liquid, such as water, across the collimated light beam. The pressure variations produced in the liquid cause local changes in its refractive index and after passing through the ultrasonic light modulator the light emerges with a spatial phase modulation pattern which is determined by the pressure distribution in the modulator at the appropriate instant.

Lens 3 is used to produce the Fourier transform of the light distribution from the light modulating cell in the plane of diaphragm 2. This Fourier transform distribution will consist of discrete lines which correspond to the unmodulated or 'carrier' light distribution and 'sideband' lines corresponding to the spatial components of the original phase modulation.

If the peak light phase deviation produced by the ultrasonic light modulator is small the intensity of each of the side band line pairs produced by different space harmonics of the phase modulation pattern will be nearly independent of the intensities of all the others. This corresponds to the well-known result that superposition is approximately applicable to phase modulation spectra when the peak phase deviations are small. However, what is more important is the fact that for the case of small peak phase deviations the suppression of the 'carrier' spectral component produces what is normally thought of as the spectrum corresponding to a suppressed-carrier amplitude-modulated waveform. The modulation in this case will be a virtually undistorted amplitude version of the original phase modulation.

At slit 2 a stop bar is positioned to block the 'carrier' component of the phase modulation spectrum but the remaining components are passed and the 'carrier' suppressed spectrum is again Fourier-transformed by lens 4 to produce the required amplitude modulated light distribution at the reference plane. The modulation of this light distribution is, under the conditions of low light phase deviation, a replica of the applied electrical signal over a time duration which corresponds to the time taken for the acoustic wave to propagate across the light beam.

The correlation process is now completed by passing the intensity modulated light pattern through a reference transparency and finally integrating by taking the central component of the light after one more Fourier transformation. As a function of time the output of the photo-multiplier is the cross-correlation function of the reference and the input signal, with the correlation integration limited to the time duration corresponding to the time of acoustic wave propagation across the light beam.

Reich and Slobodin have used this type of correlator as a pulse compression system or matched filter for 'chirp' radar signals. The intermediate frequency signal of the radar system was applied to the correlator,

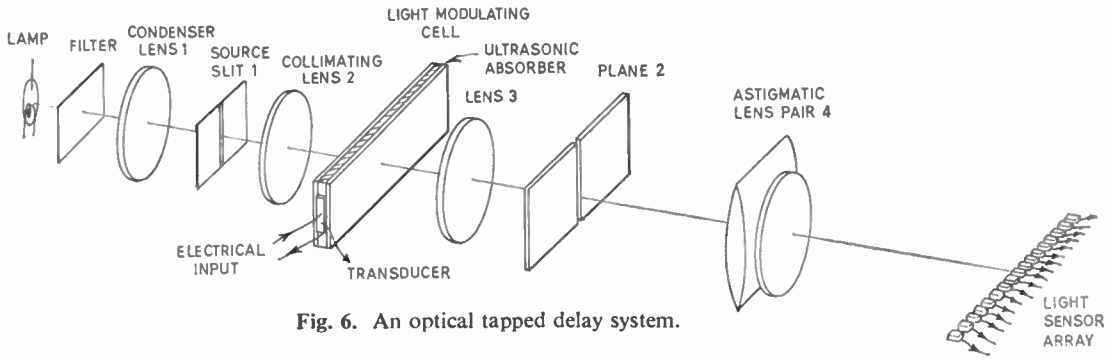


Fig. 6. An optical tapped delay system.

the frequency being chosen to be approximately 15 MHz.

The use of ultrasonic light modulation techniques in correlation systems has the obvious limitation that the integration time is restricted by the size of the light beam at the modulating cell and the relatively high velocity of propagation of acoustic waves in transparent liquids. Using water a 10 cm aperture will provide an integration time of approximately 66 μs.

3.4. *A Proposed Delay System with Adjustable Taps*[6]

It would appear that a system very similar to that described by Reich and Slobodin could be used to produce a delay system with a number of tapping points which can have easily adjusted delays.

The system which is being investigated at the University of Birmingham is illustrated in Fig. 6 and it will be seen to be the same as that described in Section 3.2 and illustrated in Fig. 5 up to the slit plane 2. Instead of the stop bar plane 2 will contain a means for changing the phase of the light 'carrier' spectral component by $\pi/2$ radians with respect to the remaining spectral components.

This relative phase change will result in the production of an envelope modulation of the light distribution at the reference plane. Again the modulation envelope of the light distribution will be a replica of the input electrical signal during a time corresponding to the time of acoustic propagation across the optical aperture, and the reference transparency can be replaced by an array of photo-detectors which sample the light distribution at various positions and can be easily adjusted. The position of each photo-detector determines the delay to each tapping point. Figure 7 shows the light pattern in the detector plane produced by a standing wave in the ultrasonic light modulating cell.

Since the overall size of the optical aperture at the ultrasonic light modulator may be small when only short delays are required some magnification in the final light pattern will be desirable. This will permit the photo-detectors to sample the light distribution at points which are closely spaced in time but reasonably spaced physically. The required magnification can be produced by using a final transforming lens with a focal length which is much greater than that of the first transforming lens.

The production of an envelope modulated light distribution is necessary in this application because a photo-detector cannot utilize the phase reversal information which is present in a 'carrier' suppressed light distribution.

It should be noted that the principle by which a phase-modulated light distribution is converted into an amplitude or intensity modulated one is similar to the technique of phase contrast microscopy introduced by Zernike in 1942.

3.5. *A Proposal for Visualizing an Acoustic Image in Water*

This system, which is also being investigated at the University of Birmingham, will utilize an optical

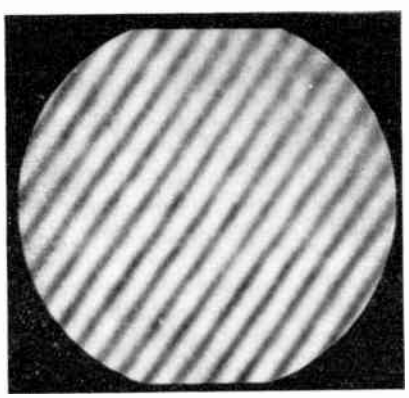


Fig. 7. Pattern produced by a standing wave in the tapped delay system.

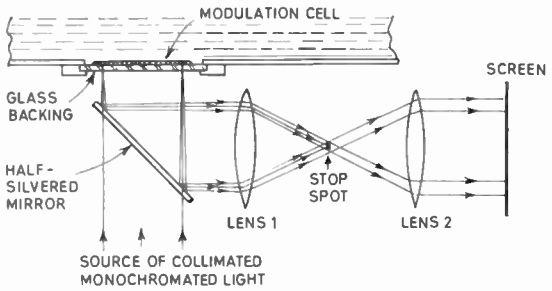


Fig. 8. Proposed system for visualizing an acoustic image.

system which will be very similar to that described for the delay system with adjustable taps. However, a two-dimensional suppressed 'carrier' final distribution is required so that a spot stop will be used on the optical axis in the spatial frequency plane and no astigmatic lenses will be used. A possible arrangement for the system is shown in Fig. 8.

In the arrangement the ultrasonic modulation cell consists of a layer of transparent silicone rubber (such as Silcoloid 201) on a glass backing. The free surface of the silicone rubber layer is arranged to be optically flat by allowing the fluid constituents to polymerize from a pool dropped on the glass backing held accurately in a horizontal position. A metallic optically reflecting film must be deposited on this surface.

The reflecting surface of the silicone rubber layer will be in contact with the water and the acoustic image to be visualized will be focused on this surface. The metallic layer will have little effect on the passage of acoustic energy into the silicone rubber layer which has a similar acoustic impedance to the water. Thus the acoustic image pressure pattern will appear in the rubber layer and will produce local variations in its refractive index.

Collimated monochromatic light is applied to the modulation cell from the normal direction to the glass backing and this light passes into and through the rubber layer, is reflected at the metallic surface and passes back through the cell. The optical phase at each point in the output light beam will be modulated by an amount which depends on the acoustic intensity at the point. Thus the cell produces a phase-modulation pattern which corresponds to the instantaneous acoustic pressure pattern in the image applied to the cell.

A half-silvered mirror is used to deflect the reflected light into the final part of the optical system in which the 'carrier' component after the first Fourier transformation is suppressed and the resulting spectral distribution is again Fourier-transformed. The resulting 'carrier' suppressed amplitude modulated version of the phase modulated pattern produced in the modulation cell is then displayed on a viewing screen.

An observer viewing the screen should not be conscious of the fluctuations in the light intensity which will be at twice the frequency of the ultrasonic image, and he should observe intensities which are proportional to the ultrasonic intensity at the appropriate point in the modulation cell.

The optical parts of this proposed arrangement can be considered to form a Schlieren system and such systems have been successfully used[7] in presenting the information contained in a phase-modulating medium such as thermoplastic sheet.

3.6. *An Optical Method for Ultrasonic Direction Finding*

This system has been described by Groginsky and Young[8] as a means for indicating the direction and the frequency of incident ultrasonic waves in water. They do not mention, however, the ambiguous information that can be displayed with multiple waves and we shall comment on this problem shortly. Figure 9 shows the optical arrangement of the system.

In this system collimated monochromatic light is passed through a modulation cell in a direction normal to the plane in which the ultrasonic wave propagates. The water in the modulation cell is bounded by parallel glass plates.

The phase modulation pattern of the light emerging from the cell in this arrangement is a two-dimensional one in which the regions of peak phase are parallel bands coincident with the peak pressure regions in the water. Thus the phase-modulation pattern is aligned and moves with the ultrasonic pressure pattern in the modulation cell. The situation is illustrated in Fig. 10(a).

The Fourier transform (two-dimensional) of the light distribution at the output of the modulation cell is all that is required to indicate the direction and frequency of the ultrasonic wave. The Fourier trans-

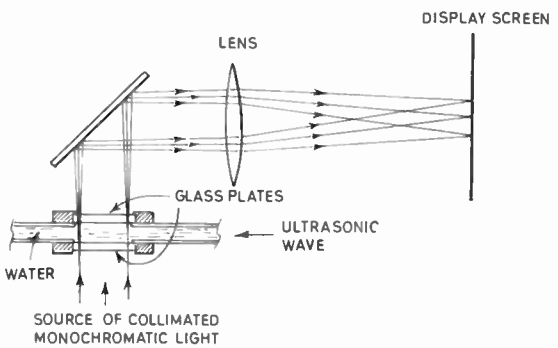


Fig. 9. Ultrasonic direction finding system.

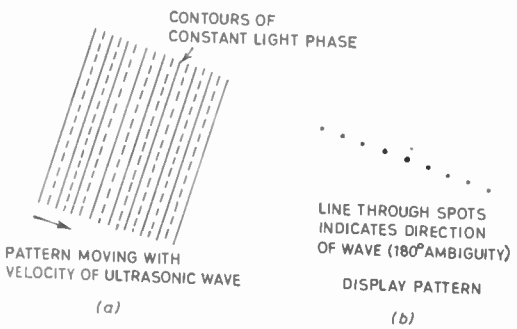


Fig. 10. Light patterns in the direction finding system.
(a) At the output of the modulating cell (phase pattern).
(b) Display intensity pattern.

form pattern is illustrated in Fig. 10(b) and we see that the direction of arrival is indicated (with a 180 deg ambiguity) by the direction of a line of bright spots and the wavelength of the ultrasonic wave is indicated by the spot spacing. The central spot is very bright in comparison with the others and the spot intensity falls off with distance from the centre.

When the peak phase deviation of the modulated light pattern is small the displayed pattern reduces to the central spot and the adjacent spots on each side. Direction and frequency are still indicated in this case but superposition of individual patterns applies when a number of ultrasonic waves of different frequencies and directions cross the modulating cell. If the peak phase deviation is not limited the presence of a number of ultrasonic waves results in the formation of a display pattern which provides very ambiguous information. The situation for the cases of one or two ultrasonic waves in the modulation cell each producing large peak phase deviations, is illustrated by the photographs of Fig. 11.

Groginsky and Young[8] proposed that this system be used to indicate ultrasonic wave direction and frequency in an ultrasonic simulation of an electromagnetic wave receiving system. This would seem to be a suitable application for the optical direction indicating system but it has been proposed that further use could be made of its properties by utilizing the frequency indicating facility to provide range information in a frequency modulated sonar or radar system.[9]

4. Conclusions

It has been possible to describe only a limited number of optical signal processing systems in this paper and the limitations of the performance of the individual systems have not been discussed in detail. However, it must be stated at this point that high quality optical equipment is necessary if precise results are to be obtained in correlation or delay applications, and the non-linear effects which appear in the conversion of phase to intensity modulation will limit the system performance when the light phase is modulated by the signal waveform.

The precise effect of this non-linear behaviour is difficult to estimate and further studies are required to determine whether or not the resulting distortion effects can be tolerated in practice.

5. Acknowledgments

The author is indebted to Professor D. G. Tucker for drawing his attention to the photo-electric Fourier transformer described in Section 2, and also for his helpful criticism of the manuscript.

Many thanks are due to Mr. R. A. K. Said for many interesting discussions and for providing the photographs of Figs. 7 and 11.

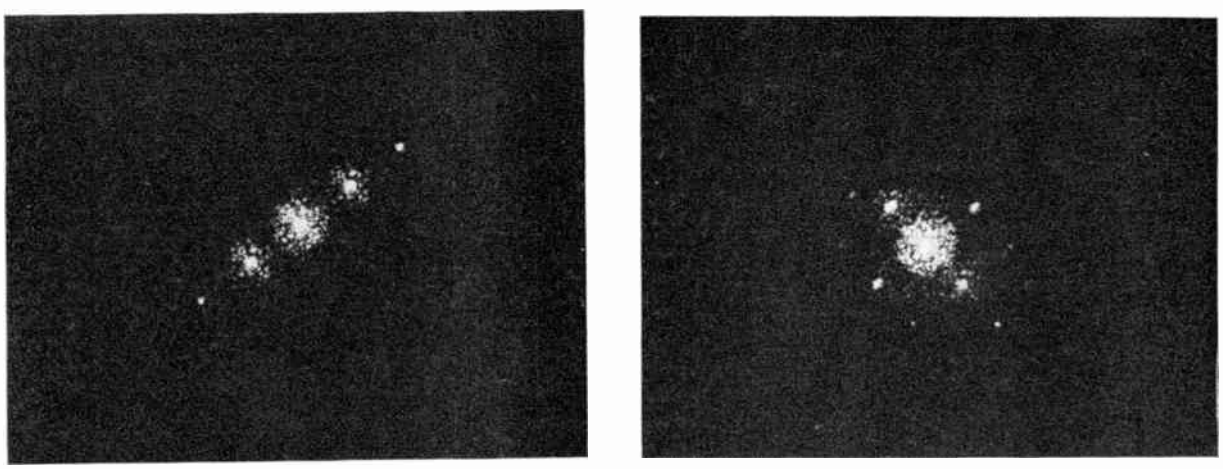


Fig. 11. Display patterns for the ultrasonic direction finding system.

6. References

1. L. J. Cutrona, *et al.*, 'Optical data processing and filtering systems', *Trans. Inst. Radio Engrs on Information Theory*, **IT-6**, No. 3, p. 386, June 1960.
2. M. Born, R. Furth and R. W. Pringle, 'A photo-electric Fourier transformer', *Nature*, **156**, No. 3973, p. 756, 22nd December 1945.
3. R. D. Hawkins, 'Vibrating Optic Fibres—A new technique for audio frequency information processing and pattern recognition', Chapter 14 in 'Optical Processing of Information', Ed. D. K. Pollock, C. J. Koester and J. T. Tippett, (Cleaver-Hume, London, 1963.)
4. M. Born and E. Wolf, 'Principles of Optics'. (Pergamon, London, 1959.)
5. A. Reich and L. Slobodin, 'Optical pulse expansion compression', *Proc. I.R.E., National Aerospace Electronics Conference*, p. 163, 1961.
6. D. C. Cooper, 'A Tapped Delay System Using Optic-Acoustic Signal Processing Techniques', University of Birmingham, Department of Electronic and Electrical Engineering, Memorandum No. 210, December 1964.
7. H. M. A. El-Sum, 'Information retrieval from phase-modulating media', Chapter 7 of 'Optical Processing of Information', Ed. D. K. Pollock, C. J. Koester and J. T. Tippett, (Cleaver-Hume, London, 1963.)
8. H. L. Groginsky and J. D. Young, 'A new technique for simultaneous radar observation of multiple targets within a broad surveillance region', *I.E.E.E. Intl Conv. Rec.*, **11**, Pt. 8, p. 82, 1963.
9. D. C. Cooper, 'An Optical Processing System for Sonar Signals', University of Birmingham, Department of Electronic and Electrical Engineering, Memorandum No. 194, July 1964.

7. Appendix

Transform properties of a converging lens using monochromatic light

We will interest ourselves in the transformation produced by a converging lens in the situation shown in Fig. 3. The lens is considered to be used for converting the incident distribution of monochromatic light at plane 1 into another distribution in plane 3, the focal plane.

The analysis of the situation is based on the use of Huygen's principle and the complex representation for the light field at any point (x_3,y_3) in plane 3 is obtained by summing the contributions from all the points in plane 1. This summation can be written in the form of a double integral as follows:

$$\hat{E}_3(x_3, y_3) = \frac{1}{j\lambda} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{(1+\cos\theta)}{2d} \times \times \hat{E}_1(x_1, y_1) e^{-j\frac{2\pi r}{\lambda}} dx_1\, dy_1 \quad \ldots\ldots(6)$$

where r = optical distance between (x_1, y_1) and (x_3, y_3),

λ = optical wavelength,

d = attenuation factor dependent on r but assumed constant in this analysis,

$\frac{(1+\cos\theta)}{2}$ = obliquity factor,

θ = angle between ray path and normal to wavefront at (x_1, y_1).

We regard θ as being small since we can regard the illumination of plane 1 to be limited to a region near the optical axis and we assume that the normal to the wavefront does not deviate appreciably from being parallel to the optical axis.

Hence $\frac{1+\cos\theta}{2}$ is approximately unity and we can write

$$\hat{E}_3(x_3, y_3) = K \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \hat{E}_1(x_1, y_1) e^{-j\frac{2\pi r}{\lambda}} dx_1\, dy_1 \quad \ldots\ldots(7)$$

where K is a constant.

To obtain an expression for r we draw a straight path from (x_3, y_3) through to the centre of the lens to intersect plane 1 at the point P_1' with coordinates (x_1', y_1') given by

$$x_1' = \frac{G}{F} x_3, \qquad y_1' = \frac{G}{F} y_3$$

Now a light ray originating at P_1 must be initially parallel to OP_1' if it is ultimately to pass through (x_3, y_3) so that we can write

$$r = OP_3 + \text{projection of } OP_1 \text{ on } OP_1' \quad \ldots\ldots(8)$$

We can now evaluate r as a function of x_1, y_1, x_3, y_3, F and G, and since our original assumption that θ is small implies that x_1, y_1, x_3, and y_3 are all small compared to F or G we can obtain a simple expression for r by neglecting terms of the type $(x_3/F)^2$, etc. In this way we find that

$$r \simeq F + G + \frac{x_1 x_3 + y_1 y_3}{F} + \frac{F-G}{2F^2}(x_3^2 + y_3^2) \quad \ldots\ldots(9)$$

Substituting for r in eqn. (7) we now obtain

$$\hat{E}_3(x_3, y_3) = K' \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \hat{E}_1(x_1, y_1) \times \times e^{-j\frac{2\pi}{\lambda F}(x_1 x_3 + y_1 y_3)} dx_1\, dy_1\, e^{-j\beta(x_3 y_3)} \quad \ldots\ldots(10)$$

where $K' = \text{constant} = K e^{-j\frac{2\pi}{\lambda}(F+G)}$

and $\beta(x_3, y_3) = \frac{2\pi}{\lambda}\left\{\frac{F-G}{2F^2}\right\}(x_3^2 + y_3^2)$

Manuscript first received by the Institution on 25th June 1965 and in final form on 23rd September 1965. (Paper No. 1052.)

Nuclear Power Developments

Formation of Exports Organization

A new organization, The British Nuclear Export Executive, (B.N.X.), has been formed to promote the export of nuclear power stations of British design. The members are the United Kingdom Atomic Energy Authority and the three industrial Consortia which design and build nuclear power stations: Atomic Power Constructions Ltd.; Nuclear Design and Construction Ltd.; and The Nuclear Power Group Ltd.

The Executive will bring to the notice of electrical power undertakings abroad the specific advantages of British reactor systems and the extent of the nuclear manufacturing resources and fuel services of this country. The Executive intend also to make the optimum use of the nuclear design ability and manufacturing capacity available in the United Kingdom and of suitable suppliers in the purchasing country.

The first Chairman of the Executive is Sir William Penney, K.B.E., F.R.S. The representatives of the member organizations are:

U.K.A.E.A.	Mr. J. C. C. Stewart
Atomic Power Constructions Ltd.	Col. G. W. Raby
Nuclear Design and Construction Ltd.	Mr. E. M. Price
The Nuclear Power Group Ltd.	Sir Edwin McAlpine

The day-to-day affairs of the Executive will be controlled by a Director-General, and Mr. B. E. Eltham has been appointed to this position. The permanent headquarters of the Executive will be in London, at Dorland House, 14–16 Regent Street, S.W.1.

The importance of the formation of B.N.X. is apparent from the fact that British industrial consortia have built or are building nuclear power stations with a total output of 6,400 megawatts for the Central Electricity Generating Board and the South of Scotland Electricity Board (the 'target' to be achieved by 1975 is twice this figure). Stations already in operation in the United Kingdom account for 55 per cent of the world's nuclear generating capacity, as may be seen from the opposite map.

Figures published in a recent issue of the *International Atomic Energy Agency Bulletin* show that in the next four years a total 105 reactors with an installed capacity of over 17000 MWe will be operating. This represents an increase of some 40 reactors having a capacity of about 10000 MWe over the number shown on the map, indicating a marked trend towards bigger reactors, such as are in use or projected in the United Kingdom.

To date two nuclear power stations have been designed and built abroad by British consortia: one at Latina in Italy and the other at Tokai Mura in Japan.

The British Nuclear Forum

The Annual Report for 1965 of the British Nuclear Forum (B.N.F.), which covers its second year of operations, records a very successful year.

The primary objective of B.N.F. has been to establish an active link with the influential and ever-growing international network of nuclear forums which have been set up in nearly every European country and in Canada, Japan and the U.S.A. At the same time it has sought to establish confidence in the ability of British industry to design and supply nuclear reactors and the wide range of plant and equipment directly associated with British reactor technology.

The past year has been marked by the success of the British Advanced Gas-cooled Reactor over the two American systems. The Report states that the award of the Dungeness 'B' contract to the A.G.R. design of a British company ended the long and unfortunate period of uncertainty in the British nuclear industry. The Central Electricity Generating Board's detailed assessment of the tenders has shown clearly that under the U.K. conditions the A.G.R. is not only the most competitive nuclear design, but that this station will produce electricity at a cost which will be at least 10 per cent lower than from equivalent conventional coal fired stations.

The fact that the A.G.R. system is capable of further advances, both by technological development, as well as by increasing the size of the reactors, shows that the efforts put into this system will undoubtedly have a continuing and long-term return. An immediate result of the lower electricity costs from the A.G.R. is reflected in the Government's decision to increase the second Nuclear Power Programme from the 5000 MW at first announced to 8000 MW over the period 1970–1975.

Another feature of the year was the success of the Foratom Congress held in Frankfurt. Particularly outstanding was the British participation organized by the B.N.F., in which the Generating Boards and the A.E.A. played a notable part with industry in an effort to sell British technology on the continent.

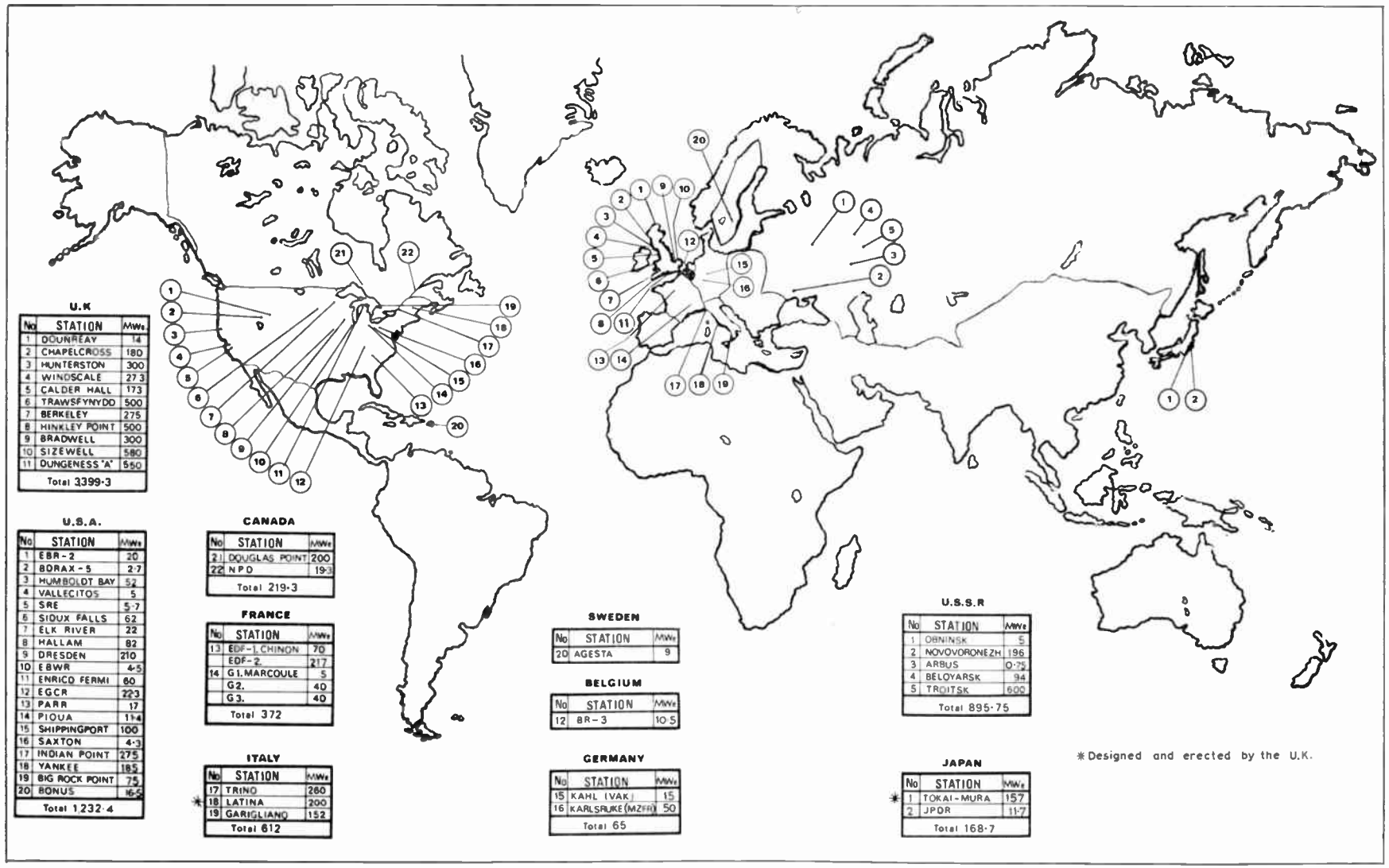
With the increasing rate of installation of nuclear power in European countries as well as in Britain, there is a growing opportunity for all European nuclear industries to work closer together than has been possible hitherto. Indeed the very fact that the British Nuclear Forum has been invited by Foratom to hold the Third Foratom Congress in London in 1967 is an indication of the importance attached by the other European countries to the participation by the U.K. in European nuclear affairs.

The affairs of the B.N.F. are guided by a Council of Management formed by senior executives from all parts of the nuclear industry. Its chairman is Colonel G. W. Raby, C.B.E., President of the I.E.R.E.

Indian Atomic Energy Commission

Dr. Vikram A. Sarabhai has been appointed to succeed the late Dr. H. J. Bhabha as Chairman of the Atomic Energy Commission and Secretary to the Department of Atomic Energy, Government of India. Dr. Sarabhai was Professor of Cosmic Ray Physics and subsequently Director of the Physical Research Laboratory, Ahmedabad. He is the first Chairman of the Indian National Committee for Space Research (INCOSPAR).

A group at the Trombay Establishment of the Commission is carrying out a feasibility study to assess the economics of the dual-purpose nuclear desalination and power plants in certain parts of India.



U.K

No	STATION	MWe
1	DOUNREAY	14
2	CHAPELCROSS	180
3	HUNTERSTON	300
4	WINDSCALE	27·3
5	CALDER HALL	173
6	TRAWSFYNYDD	500
7	BERKELEY	275
8	HINKLEY POINT	500
9	BRADWELL	300
10	SIZEWELL	580
11	DUNGENESS "A"	550
	Total 3399·3	

U.S.A.

No	STATION	MWe
1	EBR-2	20
2	BORAX-5	2·7
3	HUMBOLDT BAY	52
4	VALLECITOS	5
5	SRE	5·7
6	SIOUX FALLS	62
7	ELK RIVER	22
8	HALLAM	82
9	DRESDEN	210
10	EBWR	4·5
11	ENRICO FERMI	60
12	EGCR	22·3
13	PARR	17
14	PIQUA	11·4
15	SHIPPINGPORT	100
16	SAXTON	4·3
17	INDIAN POINT	275
18	YANKEE	185
19	BIG ROCK POINT	75
20	BONUS	16·5
	Total 1,232·4	

CANADA

No	STATION	MWe
21	DOUGLAS POINT	200
22	N P D	19·3
	Total 219·3	

FRANCE

No	STATION	MWe
13	EDF-1, CHINON	70
	EDF-2.	217
14	G1, MARCOULE	5
	G2.	40
	G3.	40
	Total 372	

ITALY

No	STATION	MWe
17	TRINO	260
18	LATINA	200
19	GARIGLIANO	152
	Total 612	

SWEDEN

No	STATION	MWe
20	AGESTA	9

BELGIUM

No	STATION	MWe
12	BR-3	10·5

GERMANY

No	STATION	MWe
15	KAHL (VAK)	15
16	KARLSRUKE (MZFR)	50
	Total 65	

U.S.S.R

No	STATION	MWe
1	OBNINSK	5
2	NOVOVORONEZH	196
3	ARBUS	0·75
4	BELOYARSK	94
5	TROITSK	600
	Total 895·75	

JAPAN

No	STATION	MWe
1	TOKAI-MURA	157
2	JPDR	11·7
	Total 168·7	

* Designed and erected by the U.K.

The output of nuclear power stations throughout the world, 1966. (Compiled by Fairey Engineering Ltd. from data published by *Nucleonics*, Euronuclear, I.A.E.A. Reactor Card Index and in papers presented at the 2nd Foratom Congress.)

U.D.C. 621.372.622:621.382.3

Some Observations on Parametric Action in Transistor Mixers

By

I. GÖK, Ph.D.†

AND

Professor F. J. HYDE, D.Sc., C.Eng. (*Member*)†

Summary: Experimental results are presented for the type 2N502 micro-alloy graded-base transistor, which show the significance of parametric amplification at the collector-base diode in addition to normal mixing at the emitter-base diode. The latter is primarily due to the resistive non-linearity at the emitter, the parametric action at the collector being due to the non-linear junction capacitance.

1. Introduction

From the inception of the transistor the non-linearity of the emitter-base diode has been used to promote mixing action and in particular down-conversion from a high frequency to a lower 'intermediate' frequency.[1,2] If the local oscillator (or pump) frequency is ω_3, then down-conversion to ω_1 can occur from ω_4 $(=\omega_3+\omega_1)$ or ω_2 $(=\omega_3-\omega_1)$. In early work it was assumed that the mixing action was due to the resistive non-linearity of the emitter-base diode. However, in consequence of the development of parametric amplifiers based on varactor diodes[3] it was suggested that parametric action could also arise in the emitter-base circuit.[4-9] Very high values of down-conversion gain together with good noise performance have been claimed, indicating that something other than resistive mixing is involved.

It is not difficult to show,[10] however, that direct parametric action at the emitter-base diode by means of the non-linear junction capacitance is insignificant in relation to resistive mixing, except at very small emitter currents and then the transistor gain at ω_1 is correspondingly low. The emitter-base diode has a low Q-factor for normal emitter currents, because the diffusion conductance greatly exceeds the shunt capacitive susceptance: on the other hand the collector-base diode has a relatively large Q-factor since the collector capacitance is virtually unshunted and significant pumping of the collector junction capacitance can be achieved over a wide frequency range, with corresponding parametric mixing.

A first-order theory of parametric mixing at the collector-base diode has been presented,[10] together with some experimental evidence in support of this theory.[11,12] The experimental system which was investigated involved a common-base-connected 10 MHz amplifier incorporating a single u.h.f. transistor. It was shown that the amplification of a 10 MHz signal injected at the emitter was considerably enhanced when a pump voltage at about 1000 MHz was simultaneously applied and that the gain was maximized when the collector-base impedance around the pump frequency was made high. High gain was not associated with instability at 10 MHz.[13]

In this paper some complementary results are presented for the same micro-alloy diffused-base transistor type 2N502, operated as a down-converter from u.h.f. to 10 MHz. The value of f_1 (the frequency at which the magnitude of the external short-circuit common-emitter current gain is unity) is less than 200 MHz for this type of transistor.

The pattern of behaviour is similar to that for the pumped 10 MHz amplifier. It was found that the magnitude of the conversion gain did not depend on whether the mixing was inverting (ω_2 the input frequency) or non-inverting (ω_4 the input frequency).

2. Outline of Mixer Operation including Parametric Action at the Collector-base Diode

The basic circuit to be considered, which incorporates a single u.h.f. transistor, is given in Fig. 1. Although all the resonant circuits are shown with lumped components, those which support voltages at ω_2, ω_3 and ω_4 will generally be distributed circuits in u.h.f. applications. The situation depicted is therefore an ideal one and it is correspondingly assumed that the effects of stray reactance elements are absorbed in the resonant circuits shown.

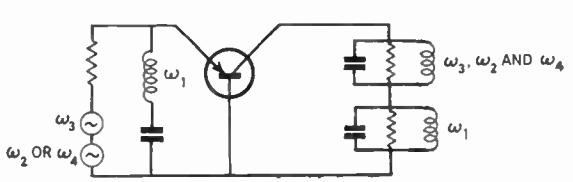


Fig. 1. Basic mixer circuit.

† Electronics Division, School of Engineering Science, University College of North Wales, Bangor.

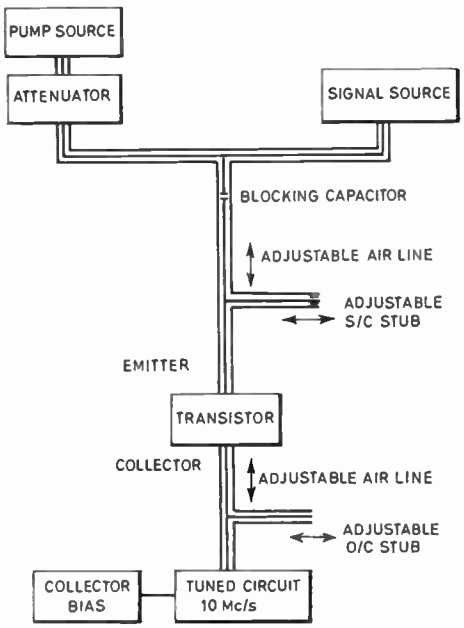


Fig. 2. Block diagram of the experimental mixer.

On the input side a pump voltage at ω_3 is connected. Because of both capacitance coupling and normal transistor action, a corresponding voltage at ω_3 is also developed across the collector-base terminals. An input signal voltage at either ω_2 or ω_4 is applied in series with the pump voltage on the emitter side. By virtue of the non-linearity of the admittance of the emitter-base diode (primarily the conductance non-linearity) a small-signal current at ω_1 is generated. Since this is at a comparatively low frequency it is readily transmitted to the collector by transistor action. At the collector parametric mixing occurs due to the pumping action of the collector capacitance by the voltage at ω_3. It is assumed that the tuned circuit supporting the voltage at ω_3 on the collector side is sufficiently wide-band to support voltages at ω_2 and ω_4 as well: these are the products of parametric mixing at the collector-base diode, any transmission at these frequencies from the emitter being ignored. Under such conditions, if it is assumed that the time-varying collector capacitance has the form

$$C = C_0 + 2C_1 \cos \omega_3 t + 2C_2 \cos 2\omega_3 t \ldots\ldots(1)$$

it follows that as far as ω_1 is concerned the parametric action at the collector can be represented by the effective generation of a resistance in parallel with C, given by[10]

$$R = \frac{G_2 G_4 - \omega_2 \omega_4 C_2^2}{\omega_1 C_1 (\omega_4 G_2 - \omega_2 G_4)} \qquad \ldots\ldots(2)$$

Here G_2 and G_4 are the conductances of the external collector-base circuit at ω_2 and ω_4 respectively. At low pump levels R has a large positive value. This decreases monotonically as the pump level is increased, passes through zero and then becomes increasingly negative. It has been shown theoretically that maximum power gain at ω_1 from emitter to collector arises when $-R$ is approximately equal to $r_{bb'}$, the high frequency base resistance of the transistor.

3. The Experimental Down-converter

A schematic diagram of the down-converter is shown in Fig. 2. On the input side the signal and local oscillator supplies both have coaxial-line output, with a characteristic impedance of 50 Ω. They are coupled at a T-junction beyond which is a coaxial blocking capacitor and a matching section consisting of an adjustable short-circuited stub and an adjustable series air line. The 'short-circuit' of the stub is in fact a capacitive termination, whose reactance is very small at pump and input signal frequencies, and is also small at ω_1. Emitter bias is fed in at the free end of the stub. The transistor mount is similar to that supplied with the General Radio Transfer Function Meter type 1607. On the output side an adjustable series air-line and open-circuited stub are followed by a lumped tuned circuit resonant at the intermediate frequency of 10 MHz. Variable tappings are provided to permit matching at 10 MHz. The voltage across the resonant load of 60 kΩ is indicated by a millivoltmeter. Collector bias is fed in via the inductor of this tuned circuit. Conversion gain was measured for various conditions. In each case the adjustable stub and series air line on the input side were used to match the signal source and pump into the transistor. The corresponding coaxial elements on the output side could be adjusted to provide a range of terminating impedances at ω_3, ω_2 and ω_4 and the adjustable tapping on the inductor to match

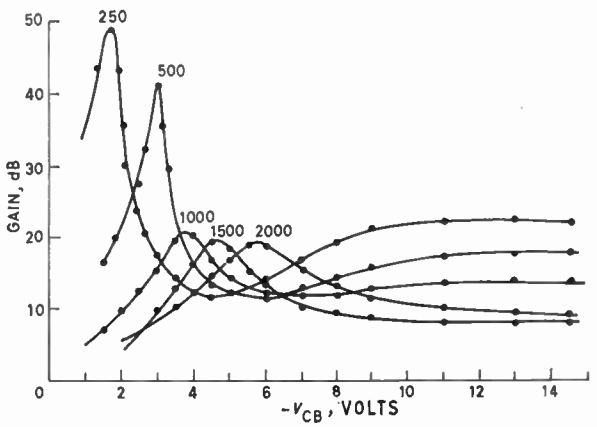


Fig. 3. Dependence of conversion gain on $-V_{CB}$. Signal frequency in MHz is the parameter $I_E = 1{\cdot}0$ mA.

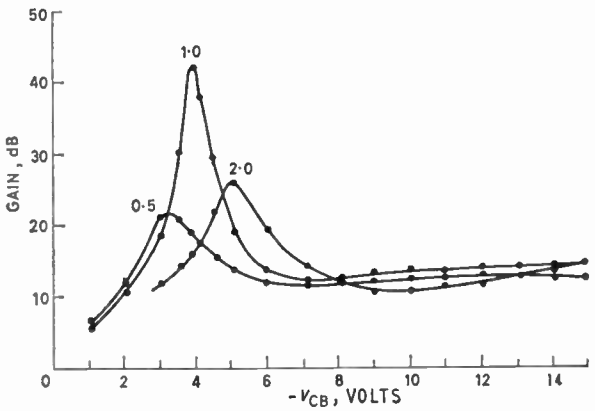


Fig. 4. Dependence of conversion gain on $-V_{CB}$. I_E in mA is the parameter. Pump power = 1·8 mW; signal frequency = 1000 MHz.

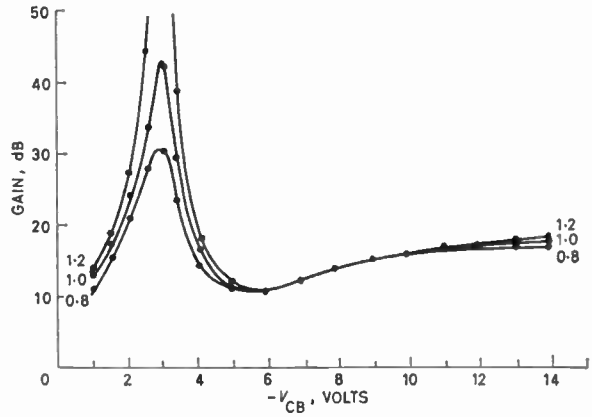


Fig. 5. Dependence of conversion gain on $-V_{CB}$. Pump power in mW is the parameter. I_E = 1 mA; signal frequency = 500 MHz.

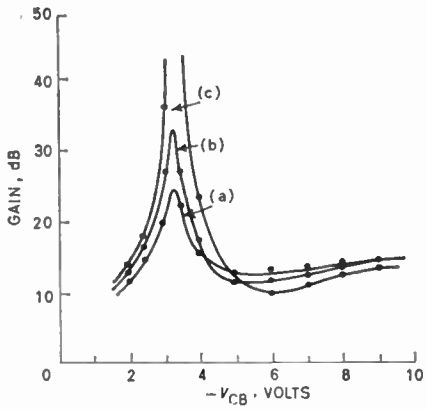


Fig. 6. Dependence of conversion gain on $-V_{CB}$. The parameter is the length of short-circuited line behind the collector-base terminals: (a) 0·5λ, (b) 0·19λ, (c) 0·22λ. Signal frequency = 1000 MHz, I_E = 1·0 mA, pump power = 1 mW.

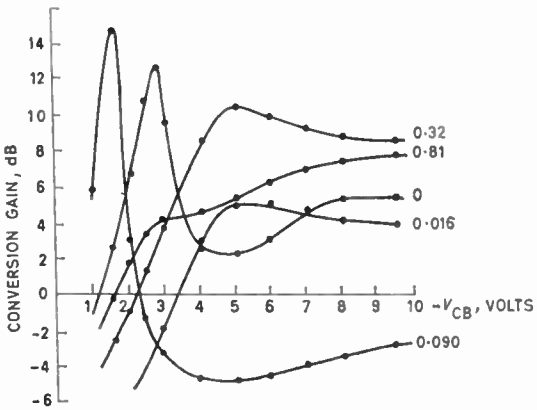


Fig. 7. Dependence of conversion gain on $-V_{CB}$. The length of short-circuited line terminating the collector-base terminals is the parameter and is specified in terms of wavelength at the pump frequency. Signal frequency = 2500 MHz, I_E = 1·0 mA, pump power = 5·0 mW.

the output tuned load into the transistor at the output frequency. Maximum conversion gain for a given available signal power was indicated by maximum reading on the millivoltmeter.

4. Experimental Results

In Fig. 3 conversion gain is shown as a function of collector-base bias voltage for different signal frequencies. In each case the emitter current was 1 mA, but the pump power was chosen to give a reasonably large gain peak in each case. Pump powers for the curves shown were as follows:

Frequency (MHz)	250	500	1000	1500	2000	2500
Pump power (mW)	0·5	1·0	1·0	3·0	3·0	5·0

The dependence of conversion gain on collector-base bias voltage is shown in Fig. 4 with emitter current in milliamperes as parameter. The pump power in this case is 1·8 mW and the signal frequency 1000 MHz. For other pump frequencies and powers the same general pattern is observed. In Fig. 5 the dependence of conversion gain on collector-base bias voltage is shown with pump power in milliwatts as parameter. The signal frequency is 500 MHz and the emitter current is 1 mA.

In the experimental work described above the length of open-circuited stub on the collector side and the adjustable series air line were set to be approximately $\lambda/4$ at the pump frequency in each case. Figure 6 shows the effect of varying the external u.h.f. impedance between the collector and base terminals on the conversion gain/$(-V_{CB})$ characteristic for a signal frequency of 1000 MHz. A similar set of curves is shown in Fig. 7 for a signal frequency of 2500 MHz. For these experiments the open-circuited

stub on the output side was made $\lambda/4$ long at the pump frequency. The distance from the near-short-circuit so created on the collector line to the collector-base terminals is the parameter in Figs. 6 and 7. It is specified in fractions of a wavelength at the pump frequency and effectively indicates the range of impedance variation. A precise indication of the actual terminating impedance cannot be given because of the effect of stray reactance associated with the transistor lead wires and the encapsulation. The plane to which measurement was made was the base of the transistor capsule.

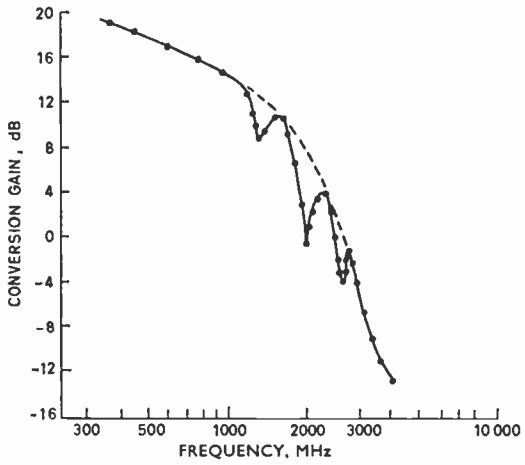


Fig. 8. Dependence of conversion gain on frequency. $I_E = 1{\cdot}0$ mA, pump power = 1·0 mW.

The dependence of conversion gain on frequency for a normal collector-base bias voltage of -10 V is shown in Fig. 8 for frequencies up to 4000 MHz. The emitter current is 1 mA and the pump power 1 mW. This curve was obtained with conjugate-matching at input and output at signal frequencies.

5. Discussion

Analytical study of conversion gain based on normal mixing action at the emitter junction is not straightforward because of the inherent frequency dependence of the transistor parameters. Assuming that these are invariant then the indications are that resistive mixing is dominant so that the 'gain' arises because of transistor action at the intermediate frequency. Because the emitter junction admittance increases with frequency, it is to be expected that conversion efficiency will be less than the value calculated assuming invariance of parameters, and will fall as the frequency is raised if only emitter mixing is significant. In the present work this is the situation depicted by Fig. 8, which relates to normal collector-base bias voltage, chosen to make the active base width of the transistor small.[14]

It is found in practice that at reduced collector voltage infinite conversion gain (oscillation at the intermediate frequency) can be achieved at certain input frequencies if sufficient pump power is applied. For lower pump powers stable gain of ~ 20 dB can be achieved. The general pattern is for the conversion gain to be higher the greater is the pump power used; for a given pump power the gain is greater the smaller is the signal frequency; the value of $-V_{CB}$ at which the peak gain occurs is greater the higher is the signal frequency; the value of $-V_{CB}$ for maximum gain at a given frequency does not depend very markedly on the level of pump power although it does increase slightly with this level.

These observations are consistent with the premise that parametric action is occurring at the collector-base diode, and are directly in accord with the experimental evidence already published for the corresponding 10 MHz amplifier with u.h.f. pump.[12]

Reference to equation (2) shows that to achieve a negative resistance either the numerator or denominator must be negative. Either condition can be achieved, but as $\omega_4 > \omega_2$ it will be assumed that a negative numerator is the more likely. It is evident that this condition can most readily arise when G_2 and G_4 are small. A key parameter is C_2, which is the second Fourier harmonic component of collector capacitance variation; this is greater the smaller is the value of $-V_{CB}$ and the larger is the pump voltage-swing developed across the collector depletion layer. It should be noted that in the experimental work pump power has been used as a parameter rather than pump voltage because, by the nature of things, it is not possible to determine the latter at u.h.f. In view of the fall in collector-base diode Q-factor at fixed V_{CB} as the frequency is raised it is evident that to obtain a given voltage-swing more pump power is required the higher is the pump frequency. If the bias voltage $-V_{CB}$ is increased the value of C falls and the Q-factor is greater at a given frequency, so that for a particular pump power a bigger voltage swing is achieved. The trend to larger values of $-V_{CB}$ for maximum conversion gain as the signal frequency is increased can be explained on this basis. The validity of the argument is substantiated by the fact that peak conversion gain is achieved when a high terminating impedance is presented to the collector-base diode.

In some papers reference has been made to 'current-tuning' in relation to the establishment of a condition of high conversion gain. Figure 4 gives data which could be used to support such a term. With a fixed collector-base bias of -4 V it is evident that variation of emitter current would result in a peak in conversion gain: this actually arose for $I_E = 1{\cdot}0$ mA in this case.

The results presented in Fig. 7 are apparently not in accord with the general pattern of results presented for lower frequencies. It is found that maximum conversion gain is obtained at a very low value of $-V_{CB}$ (cf. Fig. 3). However, it should be borne in mind that there are conflicting requirements for the establishment of negative values of R of appropriate magnitude. As ω_2 and ω_4 are increased smaller values of C_2 can be utilized. Since C_2 depends on V_{CB} and the available pump voltage swing at the collector, it is possible that $\omega_2\omega_4 C_2$ can have a comparatively large value at reduced $-V_{CB}$.

It is clearly of importance to consider the practical usefulness of the parametric mode of operation of transistors. If, as is suggested in this paper, parametric mixing at the emitter-base diode is relatively unimportant, then a very low overall noise figure cannot be expected. The particular usefulness of the parametric mode of operation of transistors appears to lie in the extension of the normal frequency range. Because the thermal noise of the base resistance is common to the emitter mixing action, the normal transistor action, and parametric action at the collector-base diode, it is possible that useful noise performance may be achieved.

6. Conclusion

Experimental results have been presented which support the theory that significant parametric mixing can occur at the collector-base diode of a transistor. In consequence useful conversion gain can be obtained well beyond the normal frequency range in which a transistor is used as a straight amplifier. In the present investigation results have been presented up to 2500 MHz for a transistor type having an f_1 value less than one-tenth of this frequency. Further understanding of parametric action in transistors depends on more detailed theoretical and experimental work.

7. Acknowledgment

The authors wish to acknowledge the support of this work by the Ministry of Defence.

8. References

1. J. Zawels, 'The transistor as a mixer', *Proc. Inst. Radio Engrs*, **42**, p. 542, 1954.
2. D. D. Holmes, T. O. Stanley and L. A. Freedman, 'A development pocket size broadcast receiver employing transistors', *Proc. I.R.E.*, **43**, p. 662, 1955.
3. L. A. Blackwell and K. L. Kotzebue, 'Semiconductor Diode Parametric Amplifiers'. (Prentice-Hall, Englewood Cliffs, N.J., 1961.)
4. R. Zuleeg and V. W. Vodicka, 'Parametric amplification properties in transistors', *Proc. I.R.E.*, **48**, p. 1785, 1960.
5. U. L. Rohde, 'Parametric amplification with transistors', *Wireless World*, **67**, p. 498, 1961.
6. R. Zuleeg and V. W. Vodicka, 'Microwave operation of drift transistors in the transit-time mode', *Trans. I.R.E. on Circuit Theory*, **PGCT-8**, No. 4, p. 426, 1961.
7. U. L. Rohde, 'Pushing transistors above their frequency limits', *Electronics*, **35**, No. 25, p. 46, 22, June 1962.
8. U. L. Rohde, 'Very low noise transistor amplifier in the u.h.f. band using the parametric conversion mode', *J. Brit. Instn Radio Engrs*, **24**, p. 223, 1962.
9. H. G. Dill and R. Zuleeg, 'A 1 Gc/s coaxial transistor for mixer and oscillator application', *Semiconductor Products*, p. 15, 1963.
10. F. J. Hyde, 'Parametric action in transistors: Theory', *Proc. I.E.E.*, **113**, p. 209, 1966.
11. F. J. Hyde and I. Gök, 'Parametric amplification in transistors', *Electronics Letters*, **1**, p. 8, 1965.
12. I. Gök and F. J. Hyde, 'Parametric action in transistors: Experiment', *Proc. Instn. Elect. Engrs.*, **113**, p. 214, 1966.
13. J. Lindmayer and C. Wrigley, 'On parametric amplification in transistors', *Proc. I.R.E.*, **49**, p. 1335, 1961.
14. H. Kröemer, 'The drift transistor', 'Transistors 1' p. 202, (R.C.A. Laboratories, 1956.)

Manuscript received by the Institution on 16th September 1965.
(Paper No. 1053.)

U.D.C. 546.431.821 : 621.375.7.029.64

Application of Barium Titanate to Microwave Parametric Amplification

By

S. N. DAS†

AND

K. W. H. FOULDS, Ph.D., C.Eng.‡

Summary: This paper investigates the usefulness of barium titanate rods as the non-linear element of an S-band resonant cavity parametric amplifier, the pump power being at X-band. An analysis is made of the variation of reactance of a small diameter loss free rod as a function of diameter (d) and relative permittivity (ε_r'). This shows that there is a large change in the resonant frequency of the cavity for a small change in the permittivity of the rod at a number of values of $\frac{\pi d}{\lambda}\sqrt{\varepsilon_r'}$. Taking the loss into account removes all except the smallest optimum value of $\frac{\pi d}{\lambda}\sqrt{\varepsilon_r}$, and this is about unity.

For solid polycrystalline barium titanate ($\varepsilon_r' \simeq 5000$) the optimum diameter becomes impractically small. This paper therefore deals particularly with the application of composition rods 0·2 cm diameter made of various percentages of powdered barium titanate and powdered polythene. Microwave measurements at X- and S-bands have shown that relative permittivity values between 10 and 65 can be obtained quite easily, but that these values change only by one or two percent with a d.c. biasing field of 7 kV/cm. A quality factor for these rods is defined analogously to that for varactor diodes, but the measured values compare very unfavourably with that for a typical diode.

Based on one of these composition rods a resonant cavity parametric amplifier was built for a signal frequency at S-band and the pump power at X-band but no evidence of gain was detected.

1. Introduction

Several non-linear effects have been used to obtain parametric amplification. Comprehensive reviews of the developments in this field are given by Mumford[1] and Rafuse,[2] and in particular the microwave properties of ceramic ferroelectrics have been discussed by Johnson.[3]

This paper deals with some aspects of the application of such a ferroelectric material, in particular barium titanate, as the non-linear element in a microwave parametric amplifier.[4] Large electric fields are required and in order to avoid frequency doubling it is necessary to use a d.c. biasing field. Cassedy[5] has treated this problem theoretically and brief results of parametric action[6] and parametric oscillation[7] have been reported. It would be expected that ferroelectric materials, like ferrites, could be used for loading a system uniformly to obtain a wide band device.

The design of a microwave parametric amplifier using a ferroelectric requires a knowledge of the variation of its permittivity as a function of the electric field at microwave frequencies. At low frequencies the corresponding information can be obtained by displaying the D–E relation directly on an oscilloscope. Unfortunately existing techniques do not enable this to be done at microwave frequencies. It is possible to obtain some information on the high frequency properties of ferroelectric materials by measuring the incremental relative permittivity at microwave frequencies as a function of an applied d.c. or low frequency biasing electric field. For a practical ferroelectric parametric amplifier it is the non-linearity of the relative permittivity which is important, and any inherent power dissipation is undesirable. For this reason it is advantageous to work above the Curie temperature, since the loss tangent is less.

2. Properties of Non-Linear Dielectrics at Microwave Frequencies

The kind of measurements suggested, i.e. the measurement of the relative permittivity of a ferroelectric material at microwave frequencies, as a

† Formerly of the Department of Physics, Battersea College of Technology, London.

‡ Department of Physics, Battersea College of Technology, London.

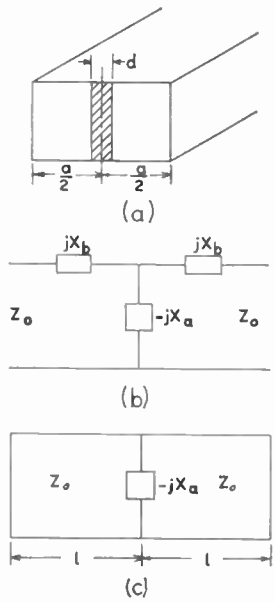


Fig. 1. Dielectric circular rod in rectangular waveguide.

(a) Position and dimensions of the rod.
(b) The equivalent circuit.
(c) Resonant circuit of short-circuited waveguide containing the dielectric rod.

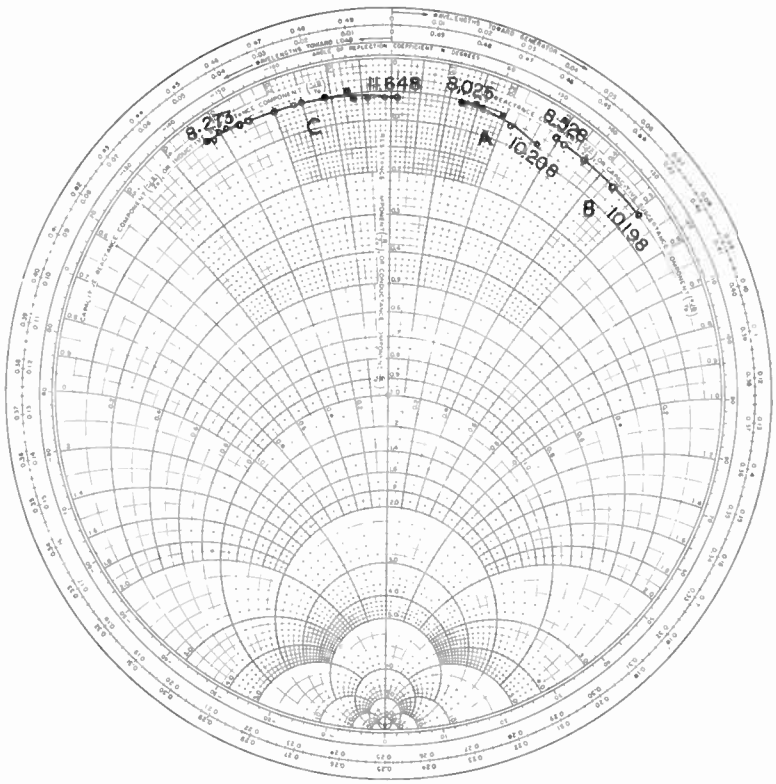


Fig. 2. Variation of normalized shunt impedance of the ferroelectric rods.

function of d.c. bias electric field, have been made by several investigators.[8-15]

In the present work circular rod shaped specimens were used. This was convenient for preparation as well as being the shape most convenient for analysis.

2.1. *Normalized Reactance of a Dielectric Rod*

2.1.1. Theoretical analysis

A thin cylindrical dielectric rod (Appendix 7.2) can be represented by the equivalent circuit shown in Fig. 1. The relative permittivity of the rod and its variation with the applied d.c. field can be calculated from the measured values of the normalized reactance of the rod. The simplified equivalent circuit of a ferroelectric rod placed in a cavity is shown in Fig. 1(c). An analysis of the reactance of a dielectric rod is made in Appendix 7.2. It is shown that to obtain a large change in the reactance for a small change in the dielectric constant, the optimum diameter of a material of high dielectric constant should be small ($\simeq$ 0·5 mm) and this was not convenient for the machining equipment available. On the other hand, if $BaTiO_3$ is mixed with a non-polar binder as suggested by Cassedy,[5] the permittivity of the composite material could be controlled and one could expect to obtain the required value of θ by trial and error.

2.1.2. Measurements

Some preliminary measurements were made using solid rods of polycrystalline $BaTiO_3$ and composition rods made of the ground-up polycrystalline $BaTiO_3$† and polythene. The diameters of the solid rods A and B were 0·199 cm and 0·054 cm respectively. The diameter of the composition rod C was 0·192 cm so that the terms α and S_0 were essentially the same as for rod A.

Figure 2 shows the variation of the normalized shunt impedance of each rod measured as a function of frequency. These results show that one must be careful in guessing the effective impedance of unfamiliar obstacles in a waveguide. Although it is common knowledge that a rod of ordinary dielectric ($\varepsilon_r \simeq 2{\cdot}5$ say) placed across the waveguide, parallel to the electric field, can be regarded as a capacitance across the waveguide, and that the corresponding metal rod behaves as an inductance, it is often not realized that a rod of high permittivity material can behave like an inductive obstacle. This behaviour is predicted by Marcuvitz,[16] and in fact, rod C changes

† $BaTiO_3$ obtained from Stanford University and stated to have the following properties at S-band, $\tan\delta = 0{\cdot}02$ and $\varepsilon_r' = 5000$, and that this changes to 50% with an applied d.c. bias field of 20 kV/cm.

from being a capacitive to an inductive obstacle as the frequency changes.

In these preliminary experiments the d.c. bias was applied with the biasing wire merely twisted around the centre of the rod. With an applied d.c. bias of 8 kV/cm the normalized reactance of rod C at 9000 MHz changed by 10%, but there was no significant change for the solid rods A and B (nor of any other solid rods having intermediate diameters). For this reason the rest of the project was concerned primarily with the composition rods.

The relative permittivity of the composition rod C has been calculated at various frequencies from the results of Fig. 2, using eqn. (9). At any particular frequency eqn. (9) gives an infinite number of discrete values of ε_r' corresponding to the infinite number of branches shown in Fig. 8. This ambiguity is resolved by taking the solutions which lead to the value of ε_r' that is sensibly independent of the frequency. Table 1 includes the values of ε_r' from the four lowest values of θ and shows that the most probable value of ε_r' is about 25. Neglect of loss will partly account for the small variation of dielectric constant with frequency. At 9000 MHz this gives the associated value of θ equal to 0·95, and this coincidentally happens to be the smallest optimum value shown in Fig. 9.

Table 1

Possible values of relative permittivity of specimen C at different frequencies

Frequency MHz	8273	8582	9090	9151	9985	11 359
ε_r'	24·9	24·8	26·5	26·6	27·4	30·5
	563	523	475	470	397	320
	1784	1681	1508	1483	1250	979
	3756	3496	3123	3081	2599	2019

2.2. *Relative Permittivity of Composition Rods*

Following these preliminary measurements, specimens of known compositions of powdered $BaTiO_3$† and polythene were made up by heating the mixture in a glass tube. Measurements at a frequency of 850 kHz showed that for zero bias the relative permittivity and loss tangent of the polycrystalline material were 6500 and 0·004. A biasing electric field of 8 kV/cm reduced the relative permittivity to 50% of the zero bias value.

During the preparation of the rods either a small copper electrode or a folded piece of wire was embedded at the centre of the specimen for biasing purposes. Glass tubes of similar diameters (nominal 2 mm) were used throughout so that θ of the specimens varied primarily because they had different permittivities.

The variation of relative permittivity of the rods was measured as a function of the percentage of barium titanate in the composition using the E_{010} mode cylindrical cavity technique developed by Horner *et al.*[17]

2.2.1. Theory of method

The theory of the method is developed in Appendix 7.3.

2.2.2. Measurements and results

It was necessary to measure the properties of the composition rods at the pump frequency (X band), the signal frequency (S band), and at the idler frequency. It was convenient to make detailed measurements at S band (3816 MHz) because the ratio of the cavity radius to the rod radius was more favourable, and then to confirm the behaviour at the other frequencies in less detailed experiments.

The cavity was made of brass and silver-plated. The top and bottom lids were secured by six bolts. The input and output coaxial sockets were mounted on the wall of the cavity diametrically opposite to each other.

The resonant frequency of the cavity was measured for (a) empty cavity, (b) cavity containing specimen but with zero bias, and (c) cavity containing specimen and with bias. The values of relative permittivity of some of the specimens, and the percentage change with bias, are given in Table 2.

Table 2

Variation of dielectric constant of specimens with applied d.c. bias

% of $BaTiO_3$	ε_r' zero bias	% decrease in ε_r'	d.c. bias (kV/cm)	1/(tan δ) (without biasing wire)
78·6	11·1	1·0	7·1	107
85·2	21·9	0·5	6·9	67
88·8	26·8	2·2	6·9	62
90·2	35·4	1·1	6·9	36

During the experiments it was noticed that the wire connected to the specimen for biasing purposes picked up microwave power, and for these reasons the results of Table 2 must be somewhat approximate. Furthermore, because of the method of preparation each rod had a shallow groove along half of its length. To avoid these difficulties the wire was pulled out of each specimen and the rod reformed by reheating in the glass tube. The relative permittivity of each specimen was re-measured in the cavity but with the

† This $BaTiO_3$ was supplied by Technical Ceramics Ltd., Northants, England, and the stated nominal Curie temperature was 20°C.

bias feed hole blanked off with a copper rod. The results are shown in Fig. 3 and they are very little different from the zero bias results of Table 2. The figure shows that a smooth variation of relative permittivity up to about 70 can be obtained simply by controlling the relative proportions of barium titanate and polythene in the composition. It was found very difficult to prepare specimens containing more than about 93% of $BaTiO_3$.

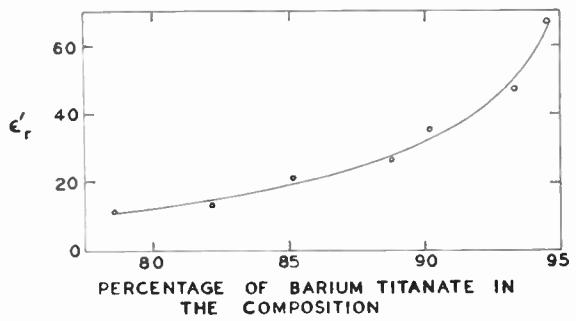


Fig. 3. Variation of ε'_r with percentage barium titanate.

The shift in the resonant frequency caused by these high permittivity rods is large, e.g. about 500 MHz for ε'_r of 50, and it is important to check that the cavity is still resonating in the E_{010} mode. Perturbation measurements were therefore made using a glass bead 2·5 mm diameter to check the radial variations of the electric field in the cavity. The higher the relative permittivity of the rod the smaller is the field outside the rod, and this makes the change in the resonant frequency of the cavity caused by the perturbing effect of the glass bead rather small. The results showed, however, that there was no higher order radial variation present.

2.3. *Quality Factor of the Ferroelectric Composition Rods*

The suitability of the composition barium titanate polythene rods as the variable reactance element in a parametric amplifier can be assessed in much the same way as can a variable capacitance diode. The ratio (the change in susceptance/conductance) is a suitable quality factor for any non-linear coupling element. Pearson and Trevena [18] have developed a very simple technique for measuring the quality factor of a variable capacitance diode and this has been adopted to measure the quality factor of the various ferroelectric composition rods. These experiments have already been described by Das.[19]

The basis of the technique is that if the rod is placed across the waveguide and this is then matched for zero bias, there is in fact some plane at which the normalized admittance remains on the unit conductance circle as bias is applied to the rod. The quality factor (A) is then simply equal to the normalized susceptance and this can be given directly in terms of the voltage standing wave ratio S

$$A = \sqrt{S} - \frac{1}{\sqrt{S}} \qquad \text{.....(1)}$$

Some typical experimental results for the normalized admittance at the appropriate plane are shown in Fig. 4. One of the best quality factors measured in this way was 2, obtained with a specimen containing 90·1% $BaTiO_3$ with a bias field of 9·6 kV/cm. Even this compares very unfavourably with the corresponding quality factor of a typical variable capacitance diode, for example a quality factor of 10 for a reverse bias of 6·0 volts.

Some measurements were made to assess the importance of self-heating within the rod caused by its conductivity since this would cause the temperature of the rod to depend upon the bias voltage. A 50 Hz bias voltage was applied to the rod and the output of the standing wave detector together with a fraction of the 50 Hz biasing voltage were displayed on the vertical and horizontal traces of an oscilloscope respectively. Having matched the rod at zero bias the v.s.w.r. at any other bias could be deduced quickly (but not very accurately) from the oscilloscope trace. The trace was somewhat complicated because of the 50 Hz hysteresis characteristic of the specimen. These results are also discussed by Das.[19] The results showed that no significant error had been introduced into the quality factor measurements by the temperature effect.

Some resonant cavity measurements were also made at X-band to confirm which specimens gave an appreciable change of resonant frequency with applied bias. The rods were placed at the centre of a

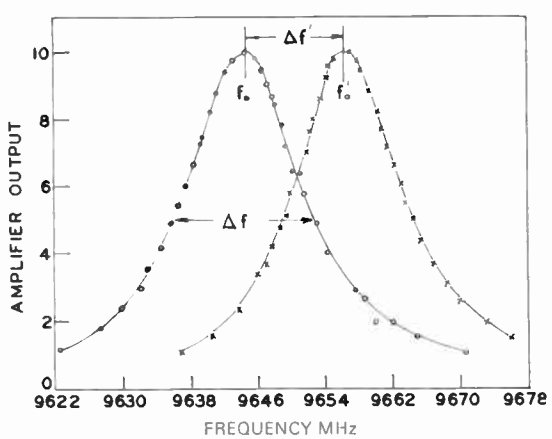


Fig. 5. Resonance curves of the waveguide cavity.

o—o—o—o without bias

×—×—×—× with bias

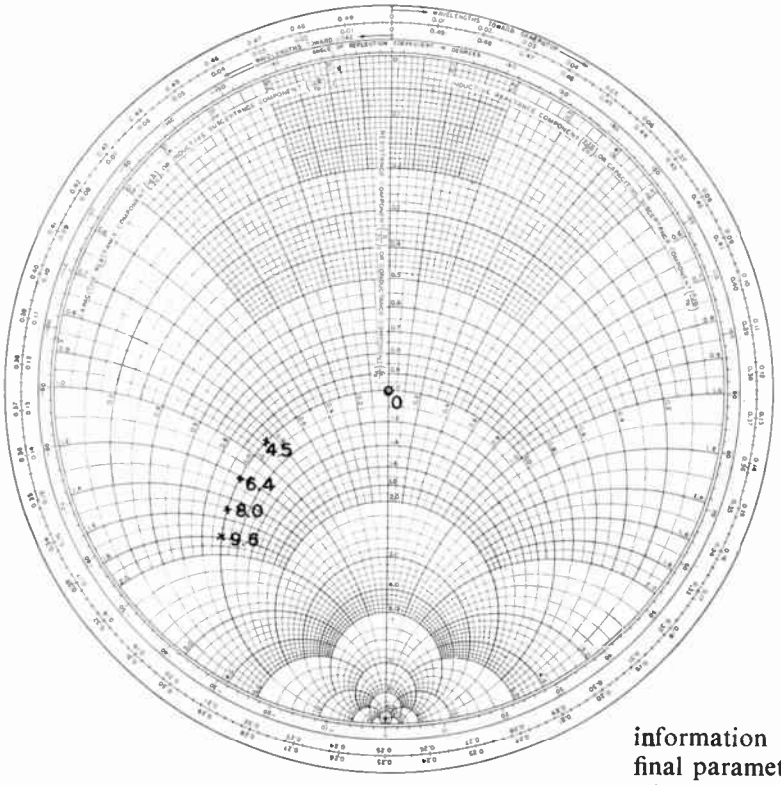


Fig. 4. Plot of change of susceptance/conductance for specimen 90·2% $BaTiO_3$ as a function of d.c. bias (kV/cm).
$f_0 = 9415$ MHz

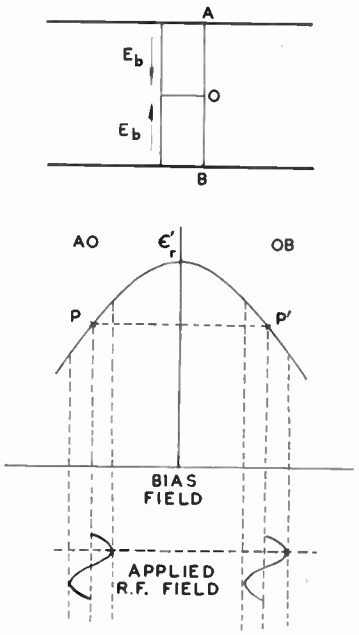


Fig. 6. Effect of central bias.

simple rectangular waveguide cavity 5 or 7 half wavelengths long. Figure 5 shows the resonance curves for a particular rod (90·2% $BaTiO_3$) with and without bias voltage. The ratio $\Delta f'/\Delta f$ is a measure of the quality factor of the composition as defined earlier in this section.

This experimental set-up is that discussed in Appendix 7.2 and it is interesting to calculate the estimated frequency shift from eqn. (11). At S-band this specimen gave a 1·1% change of relative permittivity for a bias field of 6·9 kV/cm. Assuming that this increases linearly with the bias and that the behaviour at X-band is similar, the change $\delta\theta$ in θ is 1·9%, and θ is 1·2. The predicted frequency change is then 13·5 MHz which agrees well with the measured value of 12 MHz.

3. Design of the Parametric Amplifier

It is shown in the Appendix that for the parametric amplifier to have a high gain the ratio $\Delta f'/\Delta f$ of the actual cavity should be about 20. Although the quality factor of the composition rods was only about 1/100 of this estimated value it was decided to continue with the project in order to gain some information on the problems involved. Since the final parametric amplifier device showed no evidence of gain at all this section of the work will be reported only briefly. We shall discuss (a) biasing requirements, (b) the design of the cavity, and (c) control of the pump power into the cavity.

3.1. *Biasing Requirements*

It is necessary to examine the biasing arrangements rather carefully. In the experiments described above the specimens were biased from the centre. Suppose that the working points, with central biasing, are P and P′ on Fig. 6. If at any instant the applied high frequency field increases the polarization of one half of the specimen, it simultaneously reduces the polarization of the other half. The changes in the relative permittivity in the two sections of the rod, caused by the applied signal, are of opposite sign, and if the magnitudes are equal they cancel. In low power experiments to measure the effective relative permittivity at a given electric field central biasing is correct. It has the added advantage, especially with the use of the short specimens that the large bias fields can be obtained from moderate direct voltages.

But for application to parametric amplification it is the change in the permittivity caused by the pump electric field which is important. In this case central biasing is not correct and any output would be primarily at the second harmonic frequency. To avoid this the biasing voltage must be applied across the whole length of the specimen rod.

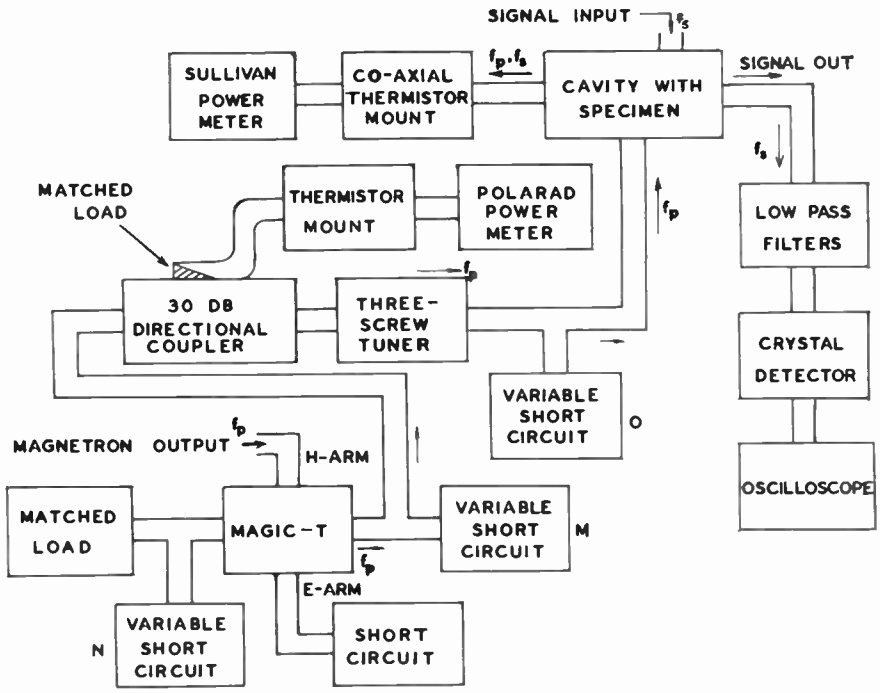


Fig. 7. Circuit arrangement for high power measurements.

3.2. *Design of the Cavity*

The essential restriction upon the cavity is that it will resonate at the signal frequency, f_1, the idler frequency, f_2, and the pump frequency, f_3. These frequencies are related by the fundamental relation:

$$f_3 = f_1 + f_2 \qquad \text{.....(2)}$$

In order to obtain the necessary coupling between the various frequency fields the three modes must have their maximum electric fields in the vicinity of the non-linear element. To reduce the likelihood of the cavity's resonating in unwanted modes the height of the cavity was kept small. This determined the choice of the modes to be the TE_{101}, TE_{301}, and the TE_{303}.

The pump power was obtained from a fixed frequency pulsed X-band magnetron. This limited the pump frequency to 9420 MHz. It was decided to have the signal frequency in S-band and this placed the idler frequency in C-band.

By writing down the resonant frequencies of the empty cavity of these modes in terms of the length l and breadth a of the cavity and assuming:

$$f_3 : f_2 : f_1 :: y : x : l$$

one obtains $x = 1{\cdot}679$ and $l = 1{\cdot}843a$. Then

$$\left(\frac{f_3}{c}\right)^2 = \frac{9}{4a^2} + \frac{1}{4 \times 3{\cdot}396a^2} \qquad \text{.....(3)}$$

where c is the velocity of light.

Substituting the required value of f_3 in eqn. (3) determines the values of a and l. An estimate of the frequency shift caused by the presence of the rod, is taken into account in deciding the value of f_3.

The pump power was fed into the cavity from the input waveguide through a $\frac{1}{4}$ in diameter coupling hole at the centre of the plate. The output was coupled from the point where the pump field should be zero. Six tuning screws were fitted in the top plate, and these were so located that the resonant frequencies of the three modes could be controlled more or less independently. These enabled the pump resonant frequency to be made 9420 MHz, the magnetron frequency, and the signal and idler resonant frequencies to satisfy eqn. (2). The final dimensions of the cavity enabling it to be tuned to the three separate resonant frequencies, with several of the specimens, were 8·56 cm by 4·57 cm by 0·68 cm.

In order to apply an adequate unidirectional biasing electric field across the specimen the length of each specimen was reduced to half the cavity height. Typical values of the ratio $\Delta f'/\Delta f$ were 0·25 at the pump frequency, 0·16 at the idler frequency and 0·15 at the signal frequency. These are lower than the value indicated in Section 2.3 and this was due mainly to the lower Q value of the cavity.

3.3. *High Power Measurements*

The experimental arrangement for controlling the large pump power fed into the cavity and the detection of the signal output, etc., is shown in Fig. 7.

The pump power incident upon the cavity could be controlled between zero and the maximum power of the magnetron by varying the position of the short circuit M. Simultaneously the v.s.w.r. was kept between 1·5 and 1·0 in the magnetron output arm by keeping the short-circuit N at predetermined positions. The actual power fed to the cavity was found by measuring the forward and reflected power by a 30-dB directional coupler. The two low-pass filters, reduced the output at the pump frequency by more than 50 dB. As a result, the output of the filter was free of pump frequency power when observation was made on the oscilloscope at the signal frequency.

There was never any real amplification of this signal during the application of the pulsed pump power. With high average pump powers, for example 3 μs pulses at a repetition frequency of 250 per second, the cavity had to be retuned to the signal frequency and the unamplified signal output increased by about 10%. This process happened gradually and was shown to be simply a temperature effect. A large proportion of the pump power is dissipated in the rod and as this heats up, the cavity is detuned but the Q of the cavity increases.

A simple calculation shows that the initial temperature rise in the rod was about 0·2 deg C per pulse and the centre of the rod could reach a temperature of about 40 deg C above room temperature. The temperature drift was not observed when the pulse repetition frequency was decreased from 250 to 50/s.

If we assume that not only is all the power dissipated in the rod but that the whole of the energy stored is stored in the rod as well we can write:

$$P = \frac{\omega W}{Q} = \omega W \tan\delta \qquad \text{.....(4)}$$

where P is the power supplied and W is the energy stored ($\frac{1}{2}\varepsilon_r'\varepsilon_0 . E_1^2 .$ volume of rod). For example, take the rod of 90·2% $BaTiO_3$ ($\varepsilon_r' \simeq 35$, $\tan\delta = 1/36$). For an input power of 3·5 kW, the peak electric field is 7·8 kV/cm, which is of the same order as the d.c. bias fields which had been applied.

4. Conclusions

Analysis shows that the reactance of a loss-free dielectric rod in a waveguide can have any value between plus and minus infinity as the parameter $\frac{\pi d}{\lambda}\sqrt{\varepsilon_r'}$ increases indefinitely from zero. For a simple waveguide cavity it is not difficult to derive the dimensions of a rod which would give the maximum change in the resonant frequency of the cavity for a given small change in the permittivity of the rod.

For other cavities it is more convenient to fix the diameter of the rod and to vary the permittivity as required. This paper has shown that it is possible to make rods to have any value of relative permittivity up to about 65 by combining suitable proportions of powdered $BaTiO_3$ and powdered polythene. It was found very difficult to make satisfactory rods containing more than about 93% of $BaTiO_3$. These rods still exhibit the property that the relative permittivity depends upon the applied d.c. electric field but to a very much reduced extent. The experiments reported here have not shown whether this non-linearity exists right up to microwave frequencies. A definite answer to this problem could be obtained by finding second harmonic frequency signals present as a result of high power microwave pump drive. The experiments showed no evidence at all of any parametric amplification using the composition rods. Even if the non-linearity behaviour is still effective at microwave frequencies the measured quality factors, or $\Delta f'/\Delta f$ were all too small for there to be much likelihood of significant amplification.

It is interesting in this connection to compare the results of this work with those mentioned by the people in the field. Goldstein[6] obtained parametric interaction with c.w. pump power of 1 watt and less. The field at the specimen in this device is not known, but the c.w. pump power is surprisingly small. Cassedy[5] discusses the pump field required in his proposed surface wave parametric amplifier. He considers a field of about 1 kV/cm, for a composition of $\varepsilon_r' = 100$, to be sufficient for obtaining parametric amplification.

Barium titanate is not the only ferroelectric available. As a ceramic, it is rugged and can be prepared easily in several convenient shapes. It has been extensively investigated because of its simple crystal structure. For application in a parametric amplifier, ferroelectric materials with large non-linearity and low loss are desirable. There are many ferroelectrics which have not yet been investigated at microwave frequencies. Investigation of microwave properties of ferroelectrics[20] like dicalcium strontium propionate might reveal a more suitable material.

5. Acknowledgment

The authors would like to thank Professor L. R. B. Elton, Head of the Physics Department of Battersea College of Technology, for the facilities made available for this work and for his encouragement, and the Decca Radar Company for the loan of the two coaxial filters. In addition, the authors also thank Dr. Euyen Gott and the computing centre of the University of Hawaii respectively for helping in

programming and numerical computations of eqns. (10) and (11). One of the authors (S.N.D.) is grateful to the late Dr. D. R. Workman and to his co-author for supervising, respectively, the beginning and later stages of his Ph.D. thesis[4] work.

6. References

1. W. W. Mumford, 'Some notes on the history of parametric transducers', *Proc. Inst. Radio Engrs*, **48**, pp. 848–53, 1960.
2. R. Rafuse, 'Parametric devices", *J. Res. Nat. Bur. Stand.* (U.S.A.), **68D**, p. 661, May 1964.
3. D. A. Johnson, 'Microwave Properties of Ceramic Nonlinear Dielectrics', Air Force Office of Scientific Research and Air Research and Development Command, Washington, D.C., Contract No. AF 49(683)–415, July 1961.
4. S. N. Das, 'Application of Ferroelectricity to Parametric Amplification', Ph.D. Thesis, London University, February 1962.
5. E. S. Cassedy, Jr., 'A surface wave parametric amplifier', *Proc. I.R.E.*, 47, pp. 1374–75, August 1959.
6. I. Goldstein, 'Interaction of two microwave signals in a ferroelectric material', *Proc. I.R.E.*, **48**, p. 1665, September 1960. (Letters.)
7. R. A. Pucel, *et al.*, 'A ferroelectric microwave parametric oscillator', *Proc. Inst. Elect. Electronics Engrs*, **51**, pp. 1660–61, November 1963. (Letters.)
8. J. G. Powles and Willis Jackson, 'The measurement of the dielectric properties of high-permittivity material at centimetre wavelength', *J. Instn Elect. Engrs*, **96**, Part III, pp. 383–89, September 1949.
9. L. Davis and L. G. Rubin, 'Some dielectric properties of barium-strontium ceramics at 3000 megacycles', *J. Appl. Phys.*, **24**, pp. 1194–97, September 1953.
10. T. S. Benedict and J. L. Durand, 'Dielectric properties of single domain crystals of $BaTiO_3$ at microwave frequencies', *Phys. Rev.*, **109**, pp. 1091–93, February 1958.
11. M. Didomenico, Jr., D. A. Johnson and R. H. Pantell, 'Ferroelectric harmonic generator and the large-signal microwave characteristics of a ferroelectric ceramic', *J. Appl. Phys.*, **33**, pp. 1697–1706, May 1962.
12. I. P. Kaminow and G. O. Harding, 'Complex dielectric constant of KH_2PO_4 at 9·2 Gc/s', *Phys. Rev.*, **129**, pp. 1562–66, February 1963.
13. C. B. Sharpe and C. G. Brockus, 'Method for measuring the dielectric constant of ferroelectric ceramics at S-band frequencies', *J. Amer. Ceram. Soc.*, **43**, pp. 302–5, June 1960.
14. E. Stern and Allen Lurio, 'Dielectric properties of $BaTiO_3$ single crystals in the paraelectric state from 1 kc/s to 2000 Mc/s', *Phys. Rev.*, **123**, p. 117, July 1961.
15. G. Rupprecht, B. D. Silverman and R. J. Bell, 'Investigation of the microwave properties of ferroelectrics', ARCRC Labs., Bedford, Mass., Contract No. AF 19(604)–4085, 16th June 1958–15th October 1960.
16. N. Marcuvitz, 'Waveguide Handbook', pp. 266–67, 1st ed. (McGraw-Hill, New York, 1951).
17. F. Horner *et al.*, 'Resonance methods of dielectric measurements at centimetre wavelengths', *J. Instn Elect. Engrs*, **93**, Part III, pp. 53–68, January 1946.
18. D. H. Trevena and J. D. Pearson, 'Evaluation of the Quality of a Variable Capacity Diode", C.V.D., England, Memo No. C.V.D. 204/14/4.
19. S. N. Das, 'Quality of a ferroelectric material', *Trans. I.E.E.E. on Microwave Theory and Techniques*, **MTT–12**, No. 4, pp. 440–45, July 1964.
20. Alan D. Franklin, 'Progress in Dielectrics', p. 210, Vol. 1, Edited by J. B. Birks and J. H. Schulman. (John Wiley, New York, 1959.)
21. S. N. Das, 'Application of barium titanate compositions to parametric amplification', *Trans. I.E.E.E. on Microwave Theory and Techniques*, **MTT-13**, No. 2, pp. 245–47, March 1965.
22. National Bureau of Standards, 'Table of Bessel Functions $J_0(Z)$ and $J_1(Z)$ for complex arguments', p. 22. New York, MTP, Columbia University Press, 1947.
23. H. Heffner and G. Wade, 'Gain, bandwidth and noise characteristics of the variable parametric amplifier', *J. Appl. Phys.*, **29**, pp. 1321–31, June 1958.
24. C. G. Montgomery, *et al.*, 'Principles of Microwave Circuits', p. 221, Vol. **8**, M.I.T. Radiation Laboratory Series. (McGraw-Hill, New York, 1948.)

7. Appendix

7.1. *List of Symbols*

a	broad dimension of the rectangular waveguide.
A	quality factor of the rod.
b	radius of the dielectric rod.
c	velocity of plane wave in free space.
d	diameter of the dielectric rod.
C_0	capacitance of the dielectric rod at zero bias.
C_V	capacitance of the dielectric rod at V volt/cm.
C_3	amplitude of the capacitance variation from C_0 caused by applied electric field.
C_1, C_2, C_p	effective capacitance of the resonant cavity at the signal, idler, and pump frequencies.
ΔC	change in C_p required to reduce response to half power point.
f	frequency.
f_1, f_2, f_3	signal, idler and pump frequencies.
Δf	total frequency difference between half power points for cavity containing dielectric rod at zero bias.
$\Delta f'$	change in resonant frequency when bias is applied to the dielectric rod.
$f(\theta)$	$\dfrac{J_0(\theta)}{\theta J_1(\theta)}$.
$f_1(\beta_1), f_0(\beta_0)$	functions defined in Appendix 7.3.
G	negative conductance.
G_{T1}	total conductance at signal frequency, without pump.

G_2	conductance of idler circuit.
h, r_0	height, radius of circular resonant cavity.
l	length of waveguide.
P	pump power.
Q	the Q of the cavity containing dielectric rod with some loss.
Q'	the Q of the cavity containing a loss free dielectric rod of the same permittivity.
R	equivalent series resistance of the dielectric rod.
S_0	a complicated function defined in Section 7.2.
S	voltage standing wave ratio.
W	energy stored in resonant cavity.
$-X_a, X_b$	shunt and series reactances of the equivalent circuit of the dielectric rod.
Y	admittance.
Y_0	characteristic admittance of the waveguide.
Z_0	$1/Y_0$.
β_0, β_1	phase constant of plane wave propagating in free space or dielectric.
β_g	phase constant in waveguide.
$\tan\delta$	loss factor in the dielectric.
Δ	skin depth.
ε_0	permittivity of free space.
$\varepsilon = \varepsilon_0[\varepsilon' - j\varepsilon_r]$	complex permittivity of dielectric.
λ_0, λ_1	wavelength of plane wave in free space and dielectric.
λ_g	wavelength in waveguide.
θ	$\frac{\pi d}{\lambda_0}\sqrt{\varepsilon_r'}$.
$\psi = \beta l$	the electrical length of transmission line (Sect. 7.2).

7.2. *Normalized Reactance of a Dielectric Rod*

A cylindrical loss-free dielectric rod of diameter d, placed in the centre of a rectangular waveguide of width a, propagating the H_{01} mode can, if $d/a < 0{\cdot}1$, be represented by the equivalent circuit shown in Fig. 1. The normalized reactance of such a rod is given in the 'Waveguide Handbook'.[16]

$$\frac{X_a}{Z_0} - \frac{X_b}{2Z_0} = \frac{a}{2\lambda_g}\left[\frac{J_0(\theta)}{J_0(\alpha)} \cdot \frac{1}{\theta J_0(\alpha)J_1(\theta) - \alpha J_0(\theta)J_1(\alpha)} - S_0 + \frac{\alpha^2}{4}\right] \quad \ldots\ldots(5)$$

and

$$\frac{X_b}{Z_0} = \frac{\frac{2a}{\lambda_g}\left(\frac{\pi d}{a}\right)^2}{\frac{\alpha^2 J_1(\theta)}{J_1(\alpha)} \cdot \frac{1}{\alpha J_0(\alpha)J_1(\theta) - \theta J_0(\theta)J_1(\alpha)} - 2} \quad \ldots\ldots(6)$$

where

$$\alpha = \frac{\pi d}{\lambda_0}, \qquad \theta = \frac{\pi d}{\lambda_0}\sqrt{\varepsilon_r'} \quad \ldots\ldots(7)$$

and

$$S_0 = \log_e\left(\frac{4a}{\pi d}\right) - 2 + \sum_{n=3,5,7\ldots}^{\infty} 2\left[\frac{1}{\sqrt{n^2 - \left(\frac{2a}{\lambda}\right)^2}} - \frac{1}{n}\right]$$

Assuming that these formulae remain valid for high permittivity materials, a general idea of the effect of the dielectric rod can be obtained. If d/a is not greater than about 0·1 the series reactance is small and so far as this analysis is concerned the rod can be represented simply by a shunt reactance. Then,

$$\frac{X_a}{Z_0} = \frac{a}{2\lambda_g}\left[\frac{J_0(\theta)}{J_0(\alpha)} \cdot \frac{1}{\theta J_0(\alpha)J_1(\theta) - \alpha J_0(\theta)J_1(\alpha)} - S_0 + \frac{\alpha^2}{4}\right] \quad \ldots\ldots(8)$$

If $d = 0{\cdot}2$ cm, $a = 2{\cdot}286$ cm and $f = 9375$ MHz, $\alpha \simeq 0{\cdot}2$ and $S_0 = 0{\cdot}667$, these values being independent of the relative permittivity of the rod. For these values $J_0(\alpha) \simeq 1$ and $J_1(\alpha) \simeq 0{\cdot}1$ and the negative term in the denominator of eqn. (8) is negligible compared to the positive term except when $J_1(\theta)$ approaches zero. Equation (8), then, reduces to:

$$\frac{X_a}{Z_0} = \frac{a}{2\lambda_g}\left[\frac{J_0(\theta)}{\theta J_1(\theta)} - S_0\right] \quad \ldots\ldots(9)$$

Let,

$$\frac{J_0(\theta)}{\theta J_1(\theta)} = f(\theta) \quad \text{say.}$$

The ferroelectric rod is going to be the variable reactance in a resonant cavity and the simplest equivalent circuit for this is shown in Fig. 1(c). This holds provided l is approximately a quarter of a wavelength since the effect of the small series reactances can then be neglected.

The condition for resonance is simply:

$$\frac{1}{-jX_a} + \frac{2}{jZ_0 \tan\beta l} = 0$$

and writing $\beta l = \psi$, and substituting for X_a from eqn. (9) gives:

$$\tan\psi = \frac{a}{\lambda_g}\left[\frac{J_0(\theta)}{\theta J_1(\theta)} - S_0\right] \quad \ldots\ldots(10)$$

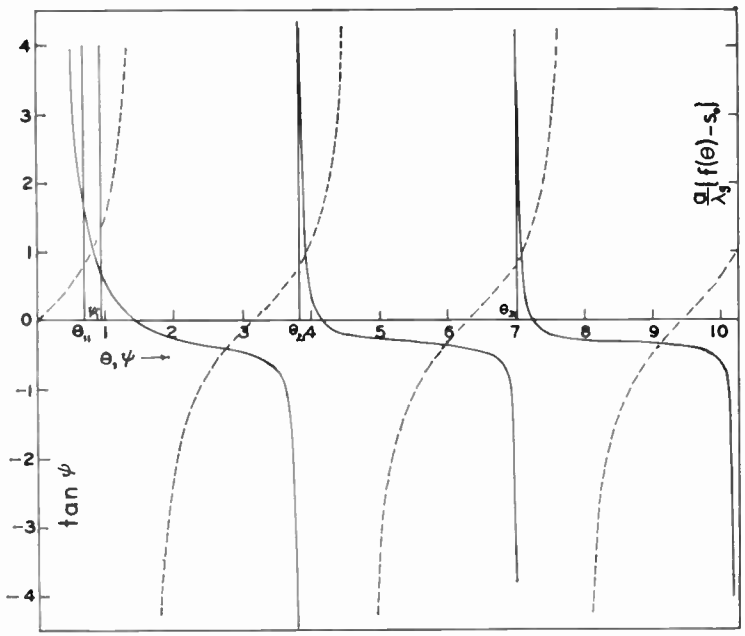


Fig. 8. Graphical determination of ω (or ψ) as a function of θ (eqn. 10).

——— $\frac{a}{\lambda_g}[f(\theta) - S_0]$ vs θ

– – – – – tan ψ vs ψ

The average permittivity of the ferroelectric rod is determined by the applied bias voltage. Under some given bias condition the diameter of the rod can be chosen to make the change in the resonant frequency of the circuit a maximum for any given change in the permittivity. The relation between ω and θ can be derived graphically from eqn. (10) as is indicated in Fig. 8. However, if we assume that S_0 is constant it is straightforward to show that:

$$\frac{\partial\omega}{\partial\theta} \simeq -\frac{f}{\theta}\frac{a}{l}\cos^2\psi\left[1 + \left[\frac{J_0(\theta)}{J_1(\theta)}\right]^2\right]$$

and eliminating the $\cos^2\psi$ from eqn. (10) leads to:

$$\frac{\delta\omega}{\omega} = -\frac{\delta\theta}{\theta}\cdot\frac{a}{2\pi l}\left[\frac{1 + \left(\frac{J_0(\theta)}{J_1(\theta)}\right)^2}{1 + \left(\frac{a}{\lambda_g}\right)^2\left[\frac{J_0(\theta)}{\theta J_1(\theta)} - S_0\right]^2}\right] \quad \ldots\ldots(11)$$

giving the fractional change in the resonant frequency caused by a fractional change in either the diameter of the rod, or of its relative permittivity. Figure 9 shows a plot of $\left(\frac{\delta\omega}{\omega}\right)\Big/\left(\frac{\delta\theta}{\theta}\right)$ vs θ which shows that the smallest optimum value of θ is about 1·0. The subsequent optimum values, which are the roots of $J_1(\theta) = 0$, i.e. 3·83, 7·01, etc., are associated with a very large change in the resonant frequency of the cavity and it is interesting to discuss the physical reason for this. These optimum values of θ correspond to the rod having infinite reactance, i.e. no shunting effect across the waveguide. Any departure from this critical value introduces significant reactance across

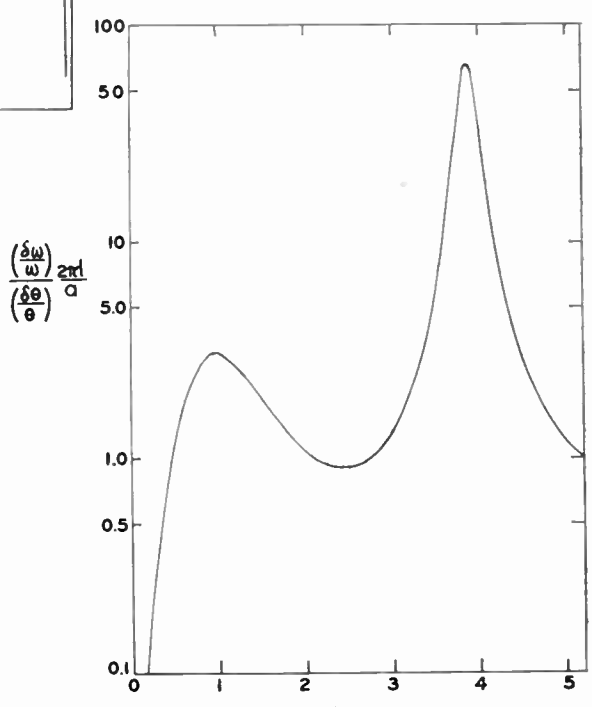


Fig. 9. $\left(\frac{\delta\omega}{\omega}\Big/\frac{\delta\theta}{\theta}\right)\frac{2\pi l}{a}$ vs θ.

the waveguide, and the larger the diameter, for example, the more rapid will be its detuning effect. For the polycrystalline barium titanate† available for the initial experiments the optimum rod diameters are 0·014 cm, 0·055 cm, 0·101 cm, etc.

The situation is changed considerably when the loss

† See footnote to Section 2.1.2.

in the rod is taken into account. For a material with a complex relative permittivity $\varepsilon_r = \varepsilon_r' - j\varepsilon_r''$, eqn. (7) becomes:

$$\theta' = \frac{\pi d}{\lambda}\sqrt{\varepsilon_r' - j\varepsilon_r''} = \frac{\pi d}{\lambda}\sqrt{\varepsilon_r'}\sqrt{1 - j\frac{\varepsilon_r''}{\varepsilon_r'}}$$

The ratio $\varepsilon_r''/\varepsilon_r' = \tan\delta$ and assuming this is small we can write:

$$\theta' = \frac{\pi d}{\lambda}\sqrt{\varepsilon_r'}\left(1 - j\frac{1}{2}\frac{\varepsilon_r''}{\varepsilon_r'}\right) \qquad \text{.....(12)}$$

Writing $\theta' = \theta\, e^{-j\phi}$ we have for small values of ϕ:

$$\theta' = \theta[1 - j\phi] \qquad \text{.....(13)}$$

whence $\phi = \frac{1}{2}\tan\delta$.

The normalized reactance of the rod is now given by:

$$\frac{X_s}{Z_0} = \mathrm{Re}\left[\frac{a}{2\lambda_g}\left[\frac{J_0(\theta\, e^{-j\phi})}{\theta(1-j\phi)J_1(\theta\, e^{-j\phi})} - S_0\right]\right] \qquad \text{.....(14)}$$

and this leads to the condition for resonance:

$$\tan\psi = 2\,\mathrm{Re}\left[\frac{a}{2\lambda_g}\left[\frac{J_0(\theta\, e^{-j\phi})}{\theta(1-j\phi)J_1(\theta\, e^{-j\phi})} - S_0\right]\right] \qquad \text{.....(15)}$$

This is much more easily dealt with graphically. The right-hand side of eqn. (15) has been plotted in Fig. 10 for $\phi = 5°$ corresponding to $\tan\delta = 0{\cdot}175$. (This choice for ϕ was determined by the tables[22] available.)

By comparing the loss free case of Figs. 8 and 9 with the lossy case of Fig. 10 it is clear that to obtain a large change in the reactance for a small change in the dielectric constant, the diameter of the specimen should be lower than the corresponding first resonance value of θ. This shows that the optimum diameter of the rod, for the material under consideration, needs to be below 0·5 mm.

7.3. *Relative Permittivity of Composition Rods*

For the E_{010} mode[17] the real part of the relative permittivity is given by:

$$\varepsilon_r' = \frac{\beta_1}{\beta_0}\cdot\frac{J_0(\beta_1 b).J_1(\beta_0 b)}{J_1(\beta_1 b).J_0(\beta_0 b)}\left[\frac{\dfrac{Y_0(\beta_0 r_0)}{J_0(\beta_0 r_0)} - \dfrac{Y_1(\beta_0 b)}{J_1(\beta_0 b)}}{\dfrac{Y_0(\beta_0 r_0)}{J_0(\beta_0 r_0)} - \dfrac{Y_0(\beta_0 b)}{J_0(\beta_0 b)}}\right] \qquad \text{.....(16)}$$

where

$$\beta_0 = \frac{2\pi}{\lambda_0}, \qquad \beta_1 = \frac{2\pi}{\lambda_1} = \lambda_0\sqrt{\varepsilon_r'}$$

λ_0 is the free-space wavelength corresponding to the resonant frequency of the cavity loaded with the specimen and,

r_0 = radius of the cavity

b = radius of the specimen

Equation (16) is an exact relation for a loss-free system, but is not in a convenient form for the calculation of ε_r'. Substitution of the value of $\varepsilon_r' = \beta_1^2/\beta_0^2$ in eqn. (16) gives:

$$\frac{\beta_1 J_1(\beta_1 b)}{J_0(\beta_1 b)} = \frac{\beta_0 J_1(\beta_0 b)}{J_0(\beta_0 b)}\cdot\left[\frac{\dfrac{Y_0(\beta_0 r_0)}{J_0(\beta_0 r_0)} - \dfrac{Y_1(\beta_0 b)}{J_1(\beta_0 b)}}{\dfrac{Y_0(\beta_0 r_0)}{J_0(\beta_0 r_0)} - \dfrac{Y_0(\beta_0 b)}{J_0(\beta_0 b)}}\right] \qquad \text{.....(17)}$$

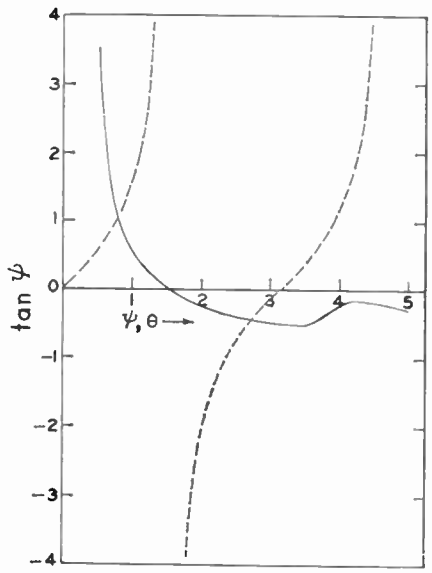


Fig. 10. Graphical determination of ω, (or ψ) as a function θ allowing for loss (eqn. 11).

——— $\frac{a}{\lambda_g}\mathrm{Re}\left[\frac{J_0(\theta_0\, e^{-j\phi})}{\theta(1-j\phi)J_1(\theta\, e^{-j\phi})} - S_0\right]$ vs θ

– – – – – tan ψ vs ψ

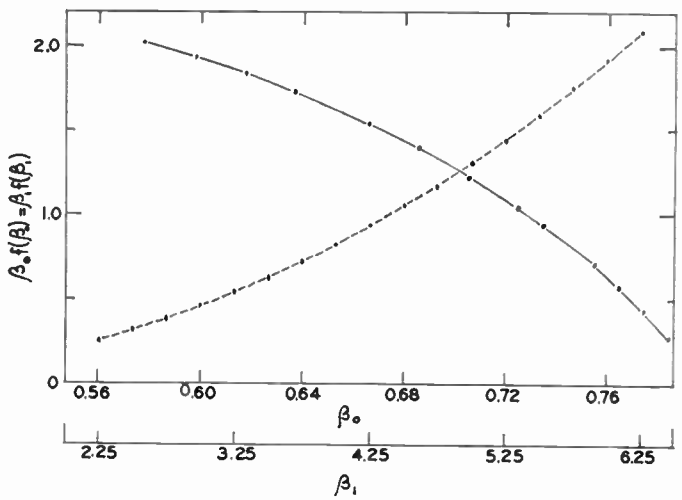


Fig. 11. Graphical determination of ε_r' from the measured resonant frequency (eqn. 18).

Let eqn. (17) be rewritten in the following simplified form:

$$\beta_1 f_1(\beta_1) = \beta_0 f_0(\beta_0) \qquad \text{.....(18)}$$

In Fig. 11, $\beta_0 f_0(\beta_0)$ and $\beta_1 f_1(\beta_1)$ are plotted against β_0 and β_1 respectively. The values of $\beta_0 f_0(\beta_0)$ corresponding to the measured λ_0 is then read off Fig. 11. This determines directly $\beta_1 f_1(\beta_1)$, the value of β_1 is then read off, and the value of ε_r' calculated.

An approximate value of the loss-factor of a very thin specimen is given[17] by:

$$\tan\delta = \frac{0{\cdot}269}{\varepsilon_r'} \cdot \frac{r_0^2}{b^2}\left[\frac{1}{Q} - \frac{1}{Q'}\right] \qquad \text{.....(19)}$$

where Q is the quality factor of the cavity with the lossy sample and Q' that of the same cavity loaded with a loss-free dielectric of the same permittivity.

The measured Q (1750) of the empty cavity was much lower than the theoretical value (5500). The method used by Horner was therefore used for calculating the loss factor of the specimens. The theoretical loss factor of a cylindrical cavity is given[17] by:

$$Q' = \frac{1}{\Delta} \cdot \frac{r_0 h}{r_0 + h} \qquad \text{.....(20)}$$

where h is the height of the cavity and Δ is the depth of current penetration in the walls of the cavity at the resonant frequency. From the measured value of Q' of the empty cavity, an effective value of Δ was computed. The value of the quality factor at any other frequency was calculated by using this value of Δ after taking into account the fact that the skin depth is a function of frequency.

7.4. *Estimated Value of $\Delta f'/\Delta f$ Necessary for Large Gain*

For large gain, the negative conductance presented to the signal circuit is equal to its conductance G_{T1}. Then[23]

$$G_{T1} = \frac{\omega_1 \omega_2 C_3^2}{4G_2} \qquad \text{.....(21)}$$

Let C_p be the capacitance of the cavity at the pump frequency f_3 and let ΔC be the capacitance swing for the half-power points.

Then,

$$\frac{C_3}{\Delta C} = \frac{Q_p}{C_p}\sqrt{\frac{G_2 G_{T1}}{\omega_1 \omega_2}} \qquad \text{.....(22)}$$

where Q_p is the Q of the cavity at the pump frequency.

It is not an easy task to derive a simple expression for the capacitance of a cavity. This has been treated by Beringer.[24] He considers the equivalent circuit of a lossless, unperturbed cavity to be a L–C circuit. From a calculation of the electric and magnetic energies of the system, he derives the following expressions for the equivalent capacitance C_a and inductance L_a of the cavity,

$$C_a = \frac{\varepsilon}{k_a^4 V} \qquad \text{.....(23)}$$

and

$$L_a = \mu k_a^2 V \qquad \text{.....(24)}$$

where V = volume of the cavity

and $k_a = \omega_a\sqrt{\mu\varepsilon}$. .....(25)

The cavity has a number of normal modes which are the periodic solutions of Maxwell's equations. Each of the modes is characterized by a resonant angular frequency ω_a. The expressions for L_a and C_a are consistent in so far as:

$$\omega_a^2 L_a C_a = 1 \qquad \text{.....(26)}$$

Let $f_3 : f_2 : f_1$ be equal to 3 : 2 : 1. Substitution of (26) in (25) results in:

$$C_p = \tfrac{4}{81}\sqrt{C_1 C_2} \qquad \text{.....(27)}$$

where C_1 and C_2 are the capacitance of the cavity at the signal and idler frequencies respectively. Substituting this value of C_p from eqn. (27) in eqn. (22),

$$\frac{C_3}{\Delta C} = 20{\cdot}25\frac{Q_p}{\sqrt{Q_1 Q_2}} \qquad \text{.....(28)}$$

or

$$\frac{\Delta f'}{\Delta f} = 20{\cdot}25\frac{Q_p}{\sqrt{Q_1 Q_2}} \qquad \text{.....(29)}$$

where Q_1 and Q_2 are the quality factors of the cavity at the signal and idler frequencies respectively. Consider a case when:

$$Q_p = \sqrt{Q_1 Q_2}$$

then

$$\frac{\Delta f'}{\Delta f} = 20{\cdot}25 \qquad \text{.....(30)}$$

Equation (30) gives the value of $\Delta f'/\Delta f$ for very large gain approaching oscillation, but some gain is to be expected when the ratio of $\Delta f'/\Delta f$ is lower than 20.

Manuscript first received by the Institution on 15th March 1965 and in final form on 16th November 1965. (Paper No. 1054.)

UDC 621.382.3 : 681.325.6

Metal Oxide Semiconductor Transistors in Digital Logic and Storage

By

J. WOOD, B.Sc.†

AND

R. G. BALL†

Originally presented at a Joint I.E.R.E.-I.E.E. Southern Sections' Symposium on 'Applications of Microelectronics' held at the University of Southampton from 21st to 23rd September 1965.

Summary: The metal-oxide-semiconductor transistor (MOST) offers the possibility of integrating large amounts of switching circuitry on a small area of semiconductor. Potential advantages stemming from this are low cost, reliable logic and storage arrays fabricated from active elements only in a relatively simple production process.

The design of uni-channel and complementary MOST circuits, and the relationship between device parameters and circuit performance, are described in the first part of the paper. The second considers ways in which storage arrays might be assembled, with particular reference to the problems of inter- and intra-connection of circuit elements.

Practical results for circuits using discrete devices are given. It is probable that integration would produce at least a doubling of the speeds obtained and a much lower power consumption.

1. Introduction

It is commonly accepted that monolithic integration of a circuit enhances its reliability to the point where interconnection between semiconductor chips becomes the limiting factor. Increasing the number of logic elements carried by a single chip would, therefore, appear to offer a further improvement in reliability for a given system size, since the number of interconnections would be reduced. However, the yield of a semiconductor manufacturing process falls rapidly when the area of a chip exceeds a certain value and hence an economic balance must be struck between chip yield and system reliability.

Metal - oxide - semiconductor transistors (MOST's) allow a particularly high packing density to be achieved, since there is no need to provide isolation regions between individual devices of like conduction type (p-channel or n-channel) as there is with bipolar transistors. Logic circuits using MOSTS' are simple, being made entirely from active elements. Resistors and capacitors, which are generally extravagant in semiconductor area, are unnecessary. Integrated arrays of a single conduction type (p-channel) are already available and allow the very highest packing densities, offering relatively slow (maximum p.r.f. 1–5 MHz), complex sub-system modules, potentially very cheap in terms of cost per logic function. However, these 'uni-channel' circuits consume power in the quiescent state and the dissipation density in the bulk semiconductor may set limits to the achievable reliability or speed or both. Circuits consuming negligible power in the quiescent state and offering favourable speed performance (between uni-channel and bipolar transistor values) are feasible using complementary MOST's. During a transition the dissipation per logic gate rises from its quiescent value of less than one microwatt by several orders of magnitude, but in a system where the average circuit activity is low the mean power density will still be reasonable. A common example of such a system is a computer store.

This paper considers various circuit and system criteria for the exploitation of the MOST in high density logic and storage arrays. It is divided into two parts. The first considers the design of individual uni-channel and complementary logic circuits. The second deals with the possible application of complementary circuits in storage systems. The use of uni-channel circuits in storage is not considered although there is no reason why they should not be used in systems where the limited speed and higher power dissipation can be tolerated.

Part 1—BASIC 'MOST' CIRCUIT DESIGN

2. MOST Characteristics

The construction and theory of operation of the MOST have been described in detail elsewhere and this will not be repeated here.[1,2] However, as an indication of the orders of magnitude to be expected, Figs. 1 and 2 show typical characteristics of the VX 6505 n-channel MOST. This has a source-drain spacing of 8 μm, a channel width of 1250 μm and input and output capacitances of 3–4 pF. The VX 6504 p-channel MOST uses the same geometry, but conducts about one-third of the current at the same electrode voltages, due to the lower mobility of holes. Although

† Royal Radar Establishment, Great Malvern, Worcestershire.

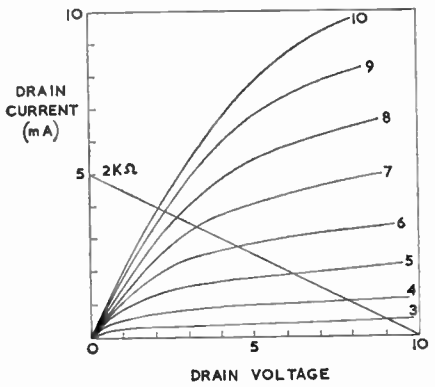


Fig. 1. MOST output characteristics.

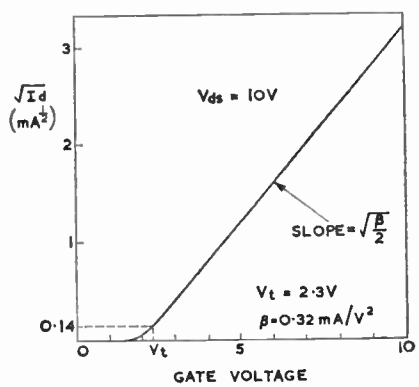


Fig. 2. MOST transfer characteristics.

the threshold voltage (also called pinch-off or starting voltage) is in theory the gate voltage at which drain current just commences to flow, the departure from a true square law characteristic at low drain currents (Fig. 2) makes this point difficult to determine, and the threshold voltage has been arbitrarily defined at a drain current of 20 μA. Its value is typically in the range 1·5–2·5 volts for both types.

3. Simple Logic Circuits

Figure 3 shows the simplest type of MOST logic circuit, directly analogous to direct-coupled transistor logic (d.c.t.l.) as used with bipolar transistors. The circuits are drawn for positive logic, i.e. $+V$ volts represents a '1', and a voltage near earth represents a '0'. In the circuit of Fig. 3(a), if either input A *or* B (or both) is a '1' (i.e. at $+V$ volts), a MOST will be on, and the output voltage will be low, representing a '0'. Only if both inputs are '0' will the output be a '1'. This is the NOR logical function. As MOST's do not bottom to as low a voltage as bipolar transistors, the '0' output voltage may be several volts. To ensure that logic stages can be cascaded with no degradation of the logic levels, a MOST driven by a '0' should be non-conducting, so that its drain voltage can rise fully to $+V$ volts. That is, the '0' output voltage must be less than the MOST threshold voltage. This sets a lower limit to the value of the load resistor R_L. (This design criterion is in fact rather conservative, but has the merit of simplicity and gives some allowance for circuit tolerances.) For example, the transistor shown in Fig. 1 has a threshold voltage of 2·3 V, and a suitable minimum value of load resistor would be 2 kΩ, giving a '0' output voltage of 2·1 V.

Figure 3(b) shows a NAND gate. Only if A *and* B are '1' will the output be a '0'. Since the MOST's are in series, their 'on' resistances add, and the value of the load resistor must be higher than for the NOR gate.

The ability to connect MOST's in series or parallel can be used to perform more complex logical functions in a single stage, a simple example being shown in Fig. 3(c). This is the main reason why MOST logic circuits are more economical in components and area than those based on bipolar transistors, which are difficult to use in series in d.c.t.l. circuits because of the relative values of the base and collector saturation voltages at high temperatures and which require separ-

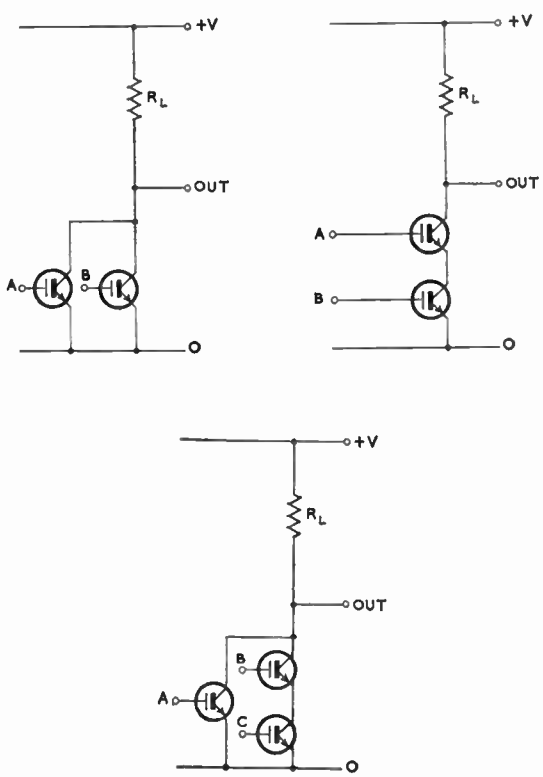


Fig. 3. Simple MOST logic circuits.

ate isolation regions for each series transistor. Further advantages of these simple MOST logic circuits are the high noise immunity, due to the high threshold voltage of the MOST, and a very high fan-out (limited only by reduction of speed) due to the virtually infinite input resistance.

4. Use of MOST as Load Resistor

The circuits shown in Fig. 3 use resistors as drain loads but, as stated previously, resistors take up a great deal of silicon area, particularly in the large values of 10–100 kΩ which would be required with small-geometry MOST's. To avoid this problem, the resistor may be replaced by a MOST with fixed gate voltage as shown in Fig. 4.[3] To ensure a sufficiently low '0' output voltage, the load MOST must be of *smaller* size than the other MOST's in the circuit, and a considerable silicon area is therefore saved. However, with this technique the load resistance is non-linear, since as the output voltage rises, the gate-source voltage of the load MOST is reduced, increasing its resistance. The effect is made worse by the action of the substrate, which is common to all MOST's in the circuit and is normally connected to the most negative supply potential to prevent the source-substrate or drain-substrate diodes from becoming forward biased. Thus, as the output voltage rises, the substrate of the load MOST becomes more negative with respect to its source, which effectively increases its threshold voltage and hence further increases the load resistance. In particular, if the gate of the load MOST is connected to its drain, as is most convenient (i.e. $V_1 = V_2$ in Fig. 4), then the load resistance will become virtually infinite when the output voltage rises to $(V_1 - V_t)$, where V_t is the effective threshold voltage of the load MOST. Thus, the '1' output voltage is no longer equal to the supply voltage V_1, but becomes critically dependent on the parameters of the load MOST. Also, the transition from '0' to '1' becomes much slower, due to the lower average current available. It is much better to connect the gate of the load MOST to a higher potential $+V_2$ such that it remains 'on' for all output voltages, and the '1' output is then $+V$, as before. The separate supply also provides the facility of setting the load resistance to suit the particular application. The additional intraconnection is a nuisance in an integrated circuit, but since it carries negligible current a narrow diffused connection can be used without metallization.

5. Speed and Power Considerations

The mutual conductance of a MOST is constant up to frequencies of several hundred MHz, so circuit switching speed is limited in practice by the device and wiring capacitances and the current available to charge and discharge them. Since the load MOST must have a higher 'on' resistance than the switching MOST's, the output transition from '0' to '1' is always slower than the '1' to '0' transition, by a factor of about 3 to 10 for step inputs. Because of the large voltage swings necessary and the limited current which can be turned on, MOST logic circuits are rather slow, the '0' to '1' transition typically being in the range 0·1 to 0·5 μs. For minimum switching times, the load MOST should conduct as high a current as possible (consistent with a sufficiently low '0' output voltage), but this also results in the highest power dissipation. For example, using MOST's as in Fig. 1 with a 10 V supply and a '0' output voltage of 2 V would give an 'on' dissipation of 40 mW, i.e. 20 mW mean. A three-input gate using this size of transistor could easily be made in an area 0·02 in × 0·02 in (0·5 mm^2), giving a mean power density on the slice of over 20 W/in^2 (3W/cm^2). This is much too high to be dealt with economically, but can only be reduced by sacrificing speed or packing density.

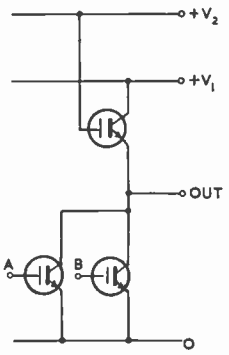


Fig. 4. MOST as load resistor.

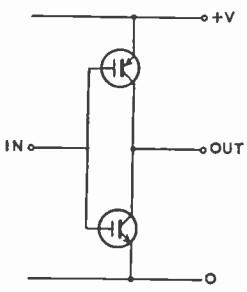


Fig. 5. Complementary MOST inverter.

Summarizing, these simple MOST logic circuits have the advantages of simplicity and economy, using MOST's and no other components, and they require only one diffusion and few other processes in their manufacture, leading to high yields and consequent low cost. However, they have an inherently low switching speed, and cannot always take full advantage of their potentially high packing density due to power dissipation problems.

6. Complementary MOST Logic Circuits

A general technique for reducing the quiescent power dissipation in a switching circuit without sacrificing speed is to use complementary types of active device.[4] The basic principle is illustrated in Fig. 5, which shows a complementary logical inverter using MOST's. If the input voltage is zero (logical '0'), the n-channel MOST is off, with a source-drain resistance generally greater than 100 MΩ. The p-channel MOST has a voltage $-V$ on its gate relative to its

source, and is therefore on, with a source drain resistance typically less than 1 kΩ. Thus, the output voltage is virtually equal to $+V$ (logical '1'). Similarly, if the input voltage is $+V$ ('1'), the n-channel MOST is on, and the p-channel MOST off, giving an output of zero volts ('0').

This circuit technique has many advantages over that using a single polarity of MOST. The two logic levels are virtually equal to the supply voltages, and are independent of the precise characteristics of the MOST's. The quiescent power consumption is negligible (typically less than 1 μW), since one or other MOST is always off, yet the circuit can switch rapidly in both directions, since the output capacitance can charge and discharge via the low resistance of the 'on' MOST. The only energy dissipated is that necessary to change the state of the circuit, and the power consumption is therefore proportional to the *mean* switching frequency. The technique therefore gives the greatest power saving in those applications where any one logical element is quiescent for a large proportion of the time, and yet is required to switch rapidly when the occasion arises. This is typical of most digital applications, and particularly of a store.

7. Performance of Inverter

Since the inverter forms the basis of all complementary MOST logic circuits, it is worth looking at its performance more closely. The graph in Fig. 6 shows its static transfer characteristic. The solid lines are calculated for idealized identical p- and n-channel MOST's, for two values of threshold voltage, and the

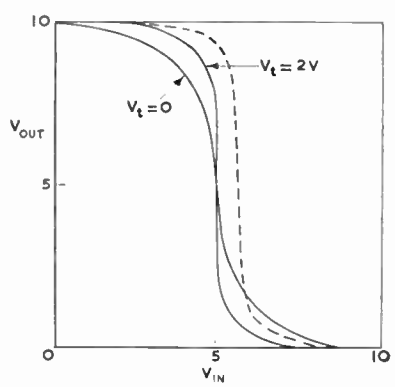


Fig. 6. Inverter static characteristics.

dashed curve is a measured characteristic for a pair of VX 6504/6505 MOST's. It is slightly displaced because the p-channel MOST had a lower threshold voltage, and the central region is not vertical because practical MOST's do not have an infinite output resistance in the saturation region. The graphs show that a large change of input voltage is necessary to give any significant change in output voltage, indicating a high noise immunity, and this is aided by the fact that *all* circuit nodes are low impedance points due to the p- or n-channel MOST's being 'on'.

The dynamic behaviour of an inverting logic element may be investigated by connecting an odd number in cascade to form a ring oscillator. By measuring the frequency of oscillation and the total power consumption, the mean propagation delay per stage and the energy dissipated per transition may be calculated. Figure 7 shows the performance of five complementary inverters in a ring (using individual devices), measured as a function of supply voltage.

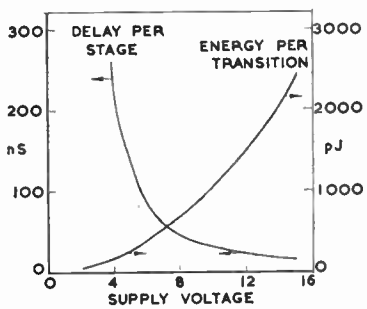


Fig. 7. Inverter dynamic characteristics.

Because the MOST is a square-law device, the circuit speed increases with increased supply voltage, and for MOST's with zero threshold voltage, all switching and delay times should be inversely proportional to the supply voltage. The non-zero threshold voltages and variation from a square law at low drain currents of practical MOST's cause a departure from this behaviour, particularly at low supply voltages.

The energy dissipated per stage for each change of state should be that necessary to change the charge on the output node capacitance, i.e. $\frac{1}{2}CV^2$, where C is the node capacitance and V the supply voltage. The energy is in fact greater than this due to both p- and n-channel MOST's being on simultaneously for a short time during the switching transient, but computer calculations and experimental results both show that this excess energy is usually negligible. Neglecting excess energy, the results in Fig. 7 indicate a node capacitance of 22 pF, which agrees closely with the estimated value.

The choice of supply voltage for complementary MOST circuits is a compromise between speed and power dissipation. A suitable value for most applications is 10 V which, for the inverter, gives a mean propagation delay of 35 ns and a power dissipation of 1·1 mW per MHz.

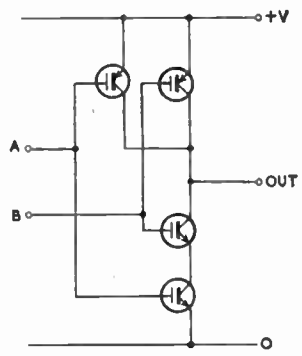


Fig. 8. Complementary 'NAND' gate.

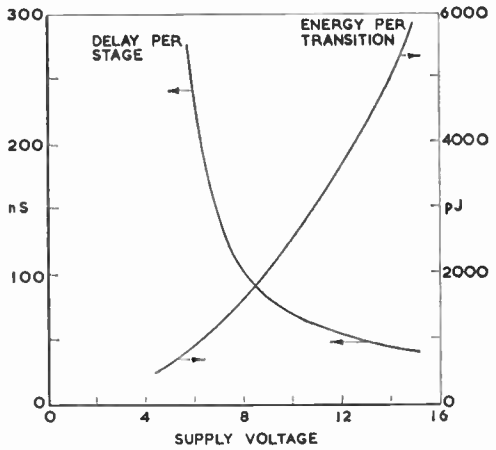


Fig. 9. Dynamic performance of 'NAND' gate.

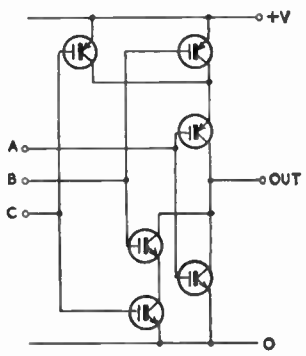


Fig. 10. OUT = (A + BC)'.

8. Complementary Gate Circuits

Figure 8 shows how the complementary technique is applied to a two-input NAND gate (compare Fig. 3(b)). If inputs A *and* B are '1', both n-channel MOST's are on, and both p-channel MOST's off, and the output is at 0 volts ('0'). In all other cases, at least one n-MOST is off, and one p-MOST on, and the output is a '1'. The dynamic performance of this circuit is shown in Fig. 9, for a fan-out of three simulated by added capacitance (again measured using individual devices in a five-stage ring). The shape of the curves is similar to those for the inverter, but the delay time and energy per transition is approximately doubled due to the increased fan-in and fan-out. For example, at a supply voltage of 10 V the mean propagation delay is 70 ns and the power dissipation 2·3 mW per MHz. These results were measured with one input of each gate connected to $+V$. When both inputs are driven simultaneously, the rising output edge is speeded up since both p-channel transistors turn on, and the mean propagation delay is therefore reduced.[5]

In general, any logic function of the form (X)' can be performed in a single stage, where X is any Boolean expression, all of whose variables are available as input voltages. For positive logic, the n-channel MOST network is designed so that it is low impedance when X = 1 and high impedance when X = 0. The p-MOST network is then the dual of this. For example, Fig. 10 shows the gate which performs the function (A + BC)'.

9. Complementary Bistable Circuits

Digital systems require bistable elements of various types. The basic complementary MOST bistable element is produced by cross-coupling two inverter circuits and provides storage of a binary digit with negligible power dissipation. In practice, a means of inserting information is required, and a set/reset bistable may be made by cross-coupling two two-input NOR gates, using the free inputs to insert a '0' or '1' as required. For shift-register applications, a 'clocked' bistable is needed, which can only accept an input in the presence of a clock pulse, and this may be made from two of the gates shown in Fig. 10. A complete shift register stage requires two of these bistables driven by anti-phase clock pulses, used in the 'master-slave' configuration such that one acts as the memory whilst the other is accepting information. To avoid the need for two clock lines, or an additional inverter per stage, one bistable may be made the complement, or dual, of the other. One bistable then accepts information when the clock is at 0 volts, and the other when the clock is at $+V$ volts, and a common clock line may be used. This circuit is shown in Fig. 12, with the outputs of the shift register stage connected back to its inputs so as to form a

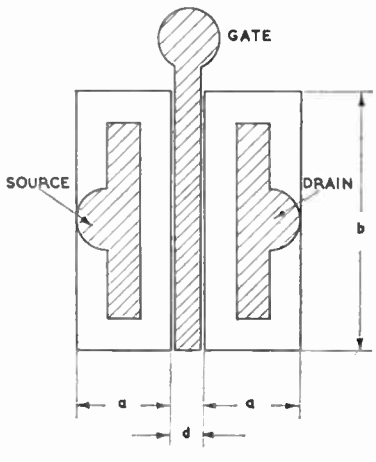


Fig. 11. Plan view of MOST.

divide-by-two element. This circuit has been built using individual MOST's, and with a 10 volts supply had a maximum counting rate of 2·2 MHz, and a power consumption of 11·6 mW per MHz.

10. Optimum Circuit Design

The switching speed of a complementary MOST circuit depends on the gain factors of the MOST's (β in Fig. 2), which should be as large as possible, and the total node capacitance (made up of MOST input and output capacitances and interconnection capacitance), which should be as small as possible. Thus a figure of merit can be used for individual MOST's, given by $\beta/(C_{in}+C_{out})$, where C_{in} is the gate-substrate capacitance, and C_{out} the drain-substrate (depletion layer) capacitance. (The feedback capacitance is generally low and can be neglected in switching circuits.) Individual MOST's should be designed to maximize this figure of merit. Considering first the thickness of oxide beneath the gate, both β and C_{in} are inversely proportional to the thickness, but C_{out} is independent. Hence the figure of merit rises slowly as the oxide thickness is reduced, the practical limit being set by yield and gate voltage breakdown. A typical value is 0·1–0·2 μm, although 0·05 μm has been used.

Other dimensions of a MOST are defined in Fig. 11. The input capacitance C_{in} is proportional to the source-drain spacing d, the gain β is inversely proportional to d, and C_{out} is independent. Hence the figure of merit rises rapidly as d is reduced. The minimum value for d is presently set by manufacturing techniques (at 5–10 μm), but ultimately will be limited by the required source-drain voltage rating. The diffusion width 'a' affects only C_{out}, and should therefore be a minimum. Again, the limit is set by the techniques available, a typical value being 20–30 μm. In an integrated circuit, the node capacitance can often be reduced by making a diffusion common to several devices, so that the whole of its periphery is utilized. In particular, for MOST's in series, if no connection is required to the common diffusion, the metallization can be omitted and the diffusion width reduced considerably, to 10 μm say. These techniques would probably make integrated circuits at least twice as fast as the same circuits using individual devices.

Finally, the dimension b affects β, C_{in} and C_{out} equally, and hence does not alter the figure of merit. However, a reduction of b will reduce the power dissipation per stage, due to the reduced capacitance, and will allow a higher packing density in an integrated circuit. The limit is set by the effect of the residual capacitances, which will begin to affect the switching speed significantly when b falls below a certain value. For a logic element driving only other elements on the same slice, the intraconnection capacitance would be small, and b might be of the order of 100 μm. However, if the element has to drive the capacitance of interconnections between slices, which may be much larger, b must be increased, perhaps to 1000 μm.

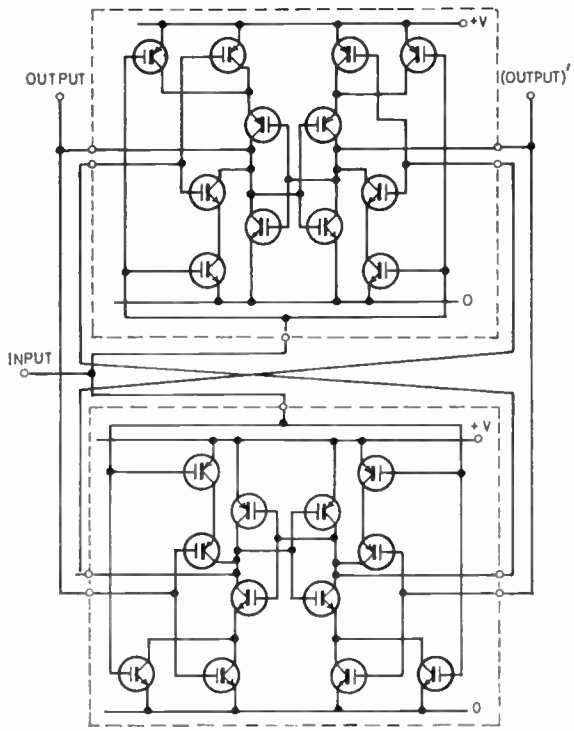


Fig. 12. Binary counter.

The other parameter of the MOST which affects circuit performance is the threshold voltage which, for a given oxide thickness, is determined by the initial doping of the substrate and the precise manufacturing processes. For maximum speed, this should be as low as possible, and ideally should be zero. However, because of manufacturing tolerances, and the departure from a square law at low currents, this would result in an excessive quiescent power dissipation in many cases, and the average threshold voltage must be made rather higher. For example, the VX 6504/6505 specification calls for a drain current at 25°C of less than 0·1 μA for zero gate voltage, and greater than 20 μA for a gate voltage of 3 V. This results in a threshold voltage in the range 1·0–3·0 V.

Part 2—STORAGE APPLICATIONS OF COMPLEMENTARY 'MOST' CIRCUITS

The suitability of complementary MOST circuits for storage arrays has already been demonstrated. Possible ways in which a store might be organized will now be discussed.

11. Parallel or Serial Storage ?

As energy is dissipated only during transitions it is desirable to keep the number of transitions per unit transfer of information to a minimum. If information is handled in parallel form then each storage element involved in the transfer of a block of information will have to change only a sufficient number of times to transfer one bit. If the transfer is non-destructive the sending element need not change at all. If the transfer is non-reset-to-zero (of the receiving element) then the receiving element will change once or not at all, depending on whether the bit to be received differs or not from the bit already being stored. A single-phase non-destructive, non-reset-to-zero parallel system is the fastest transfer mechanism and involves the least transitions per unit transfer. However, such a system, if integrated in a large area array on a single slice, poses quite severe connection problems. On the slice itself a matrix of interconnections, consisting of word, digit and selection lines, is needed. External connections to the slice will consist mainly of address and digit lines and there will be many of these if the storage word is long. The basic cause of the problem is that, in a parallel system, external access must somehow be provided to every storage element. As the yield and reliability of a system may eventually be governed by slice interconnections it may be preferable to employ a system which does not require complete access to every element, thus reducing interconnections between slices and, if possible, intraconnections on the individual slices also.

A serial system reduces the number of digit lines to, and on, the slice by the number of elements included in each serial word (or syllable). Clearly one penalty of this is considerable loss of speed. Another is the fact that to effect each transfer the individual elements in the serial block may have to change many times as the information propagates through, thus reducing the efficiency. Once again a compromise has to be made, this time between speed, efficiency and complexity of connection pattern.

12. Delay Line Storage

The simplest storage element using complementary MOST's is shown in Fig. 13. The long time constant gate circuit of the inverter permits information to be stored for many hundreds of microseconds when S1 is open whilst the state of the element can be read continuously, if so desired, without affecting the state of the input. The storage time is limited by the leakage of S1, the above figure being readily obtainable using presently available devices. For storage over a longer period the information must be regenerated and this can be achieved either by making each element a closed loop, or by circulating the information in a delay line. Of these the latter can be (but not necessarily so) more economical in circuitry and connections.

A section of a complementary MOST digital delay line is shown in Fig. 14(a). This is a two-phase system, each inverter element as shown in Fig. 13

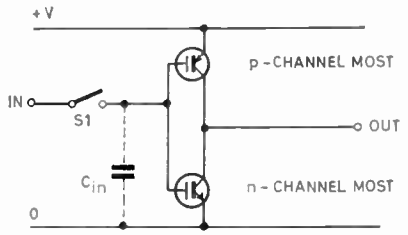
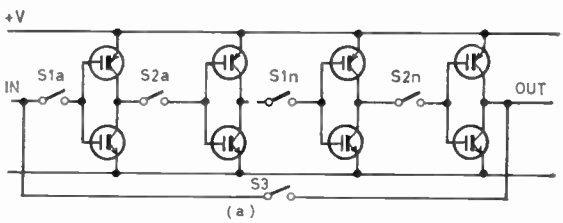
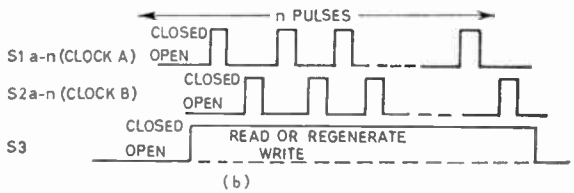


Fig. 13. MOST complementary temporary memory element.



(a) Section of MOST digital delay line.



(b) Control waveforms.

Fig. 14.

constituting a half-clock-period shift element. Figure 14(b) shows the control waveforms for this simple system. The information is presented sequentially at the input of the extreme left-hand element (S1a) and is stepped along the line in a series of half delay steps concurrently with the pulses from clock A and clock B. When a complete block of information has been written into the line the clocks may cease until a read sequence occurs or a regeneration sequence becomes necessary. In either case the regeneration

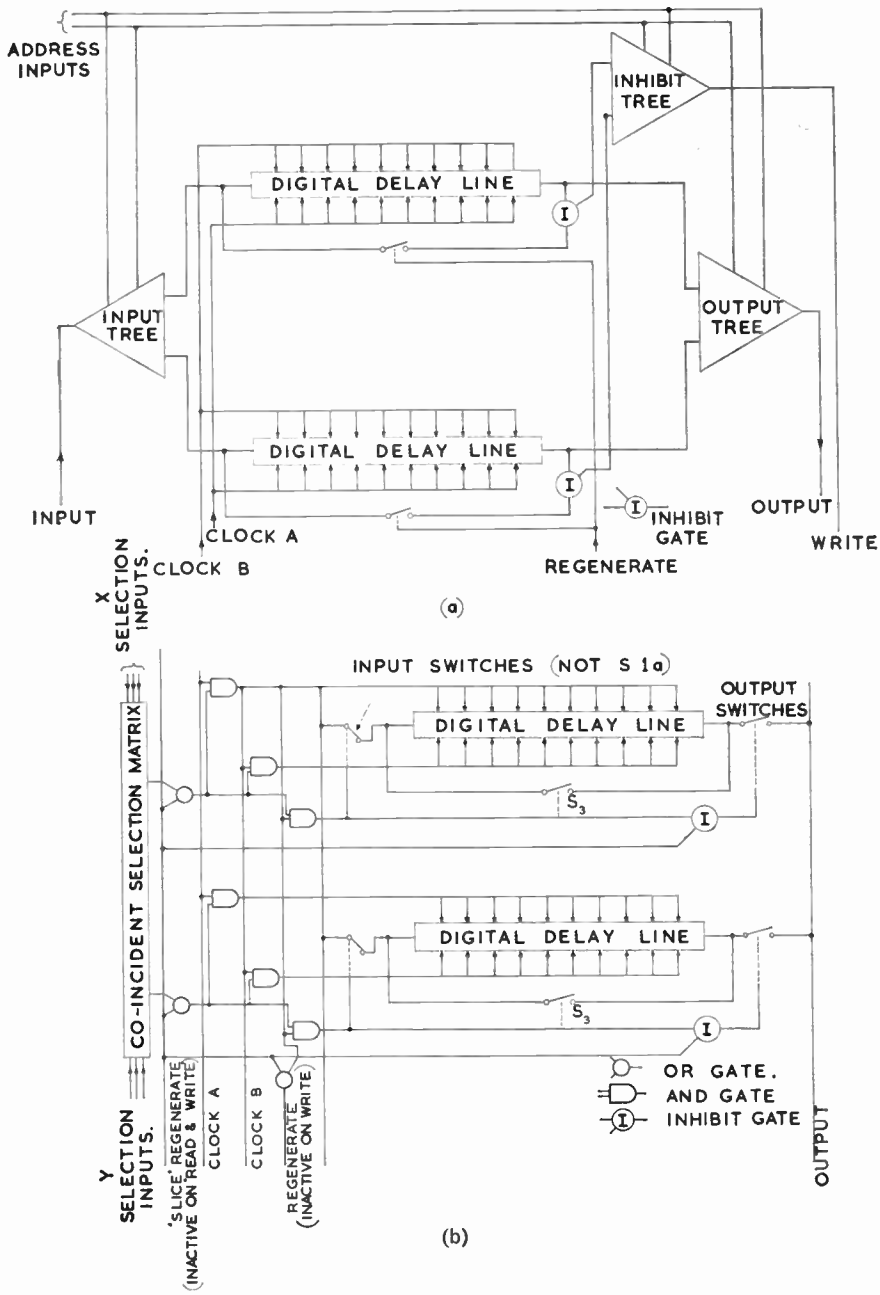


Fig. 15. Store organization.

(a) Delay line store.

(b) Alternative delay line store.

switch (S3) will be closed for the period of the sequence. Regeneration occurs automatically during every read sequence and must be initiated separately if the period between successive read sequences approaches the memory time constant, otherwise the memory capacitors may discharge and the information in the line be lost. Presently available transistors have permitted clock rates up to 5 MHz with memory time constants up to several milliseconds. On this basis a system of 20-digit lines having a cycle time of 5 microseconds and less than 1% of the storage time lost in regeneration seems quite feasible. It is anticipated that monolithic integration using similar transistors would result in an improvement in speed by a factor of about two (due to reduction of nodal capacitance) and there seems no reason why this

figure should not be increased quite considerably with further improvements in device technology. These could include optimization of device geometry, epitaxial deposition on insulating substrates and use of higher mobility semiconductor material in thin film structures.

The ultimate aim in any high density system would be to pack a number of words of storage on a single slice of substrate. Using the delay line technique the external connections required would be:

(a) Power supply lines.

(b) Clock lines.

(c) Regeneration control line(s).

(d) Input and output lines.

(e) Address lines.

The interconnections within the delay lines themselves can be relatively simple as, generally, each element drives only its own immediate neighbour. Hence the metallization pattern would tend to lie all in one direction instead of forming a matrix, as in a parallel system. However, the usefulness of the system in straightforward storage applications rests upon a satisfactory solution to the addressing problem. Figures 15(a) and 15(b) show two of the many possible ways of organizing such a store. Figure 15(a) shows a system having the delay line inputs and outputs selected by means of binary trees. The clock and regeneration pulses are applied to all the words on the slice simultaneously which eliminates the necessity to address these waveforms also. There are, however, two attendant disadvantages:

(a) The power dissipation is substantially higher as the whole slice is cycled when writing, reading or regenerating any single word.

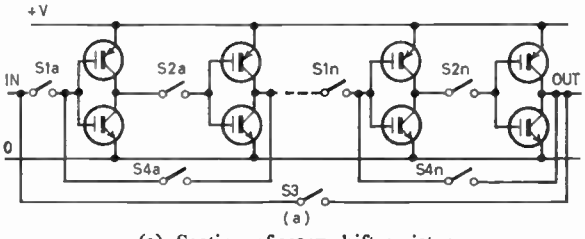
(b) When writing into any single word the selection circuits must inhibit the regeneration switch on the selected word only.

Figure 15(b) shows an alternative system where inputs, outputs, clocks and regeneration are all addressed. This overcomes the disadvantages mentioned in the previous system but at the expense of more complicated selection circuits and the necessity for a separate 'slice regenerate' input which regenerates all the lines in parallel. The addressing in this case is done by a matrix which controls gates for the routing of the separate pulses. For a small system it is probable that the matrix will be faster, but require more MOST's and a more complex interconnection pattern, whilst the tree would only become competitive in speed if a very large number of words were available on each slice. This is because doubling the number of words on the slice calls for an increase of $\sqrt{2}$ in each coordinate of the matrix and would therefore produce a corresponding increase in selection delay. With a tree doubling the size of the store implies that one more level has been added to the tree and each level produces approximately one unit of delay. The relative number of external address lines depends upon whether the ordinates for the matrix are decoded on the slice or externally. If the decoding is done on the slice there is no difference between the number of external inputs for the two selection methods. At present the coincident matrix seems to be preferable and it is not possible to foresee when the tree will become competitive, if ever.

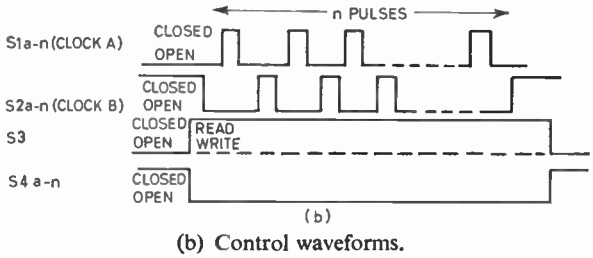
If the usefulness of this system for straightforward storage seems debatable it appears likely that in a 'content addressed' store, or some other auto-correlation system, it may have definite application. Here it may be necessary to cycle complete sections of store simultaneously which is just what is done in Fig. 15(a). The problem of inhibiting regeneration during writing, mentioned above, remains however.

13. Shift Register Storage

A simple modification to the basic delay line which converts each pair of half shift elements into a separate bistable element is shown in Fig. 16(a). If the staticizing switches (S4) are open the system is exactly as shown in Fig. 14 and, under these conditions, control waveforms applied to the S1, S2 and S3 switches are as for the digital delay line. After shifting in, or reading out, a block of information the S2 and S4 switches are closed, the S1 switches are opened and the system then becomes a series of isolated pairs of inverters connected to form bistable elements. This enables the information to be stored indefinitely without regeneration.



(a) Section of MOST shift register.



(b) Control waveforms.

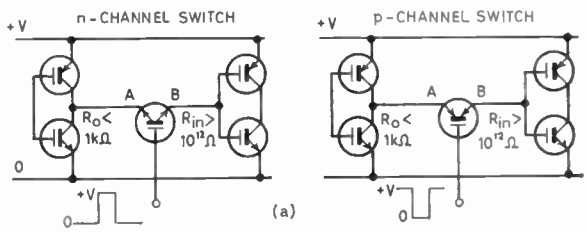
Fig. 16.

The main advantage of this system over the digital delay line is that staticization of the stored pattern allows a practical system to work with much shorter memory time constants. This allows a much broader specification for the switches which may give considerable increase in yield. Further benefits from the lack of need for regeneration are reduced power consumption and the elimination of the 'slice regenerate' input required in Fig. 15(b). Against this an extra 'staticize' input, to every element, is required and this increases the interconnection problem. It is quite probable that this control waveform could be routed to all the elements on the slice without being addressed.

14. MOST's as Switches

As has been indicated in the foregoing sections, any practical system on the lines discussed in this paper must include switches which are compatible with the rest of the circuit. In the circuits so far considered a high open-circuit impedance is essential to prevent rapid discharge of the memory capacitors, and low closed-circuit impedance is desirable to enable the state of the memory to be changed rapidly, when required. Furthermore, the switches have to be truly bi-directional if the number of interconnections is to be minimized.

An obvious choice of compatible switch device is the MOST, and in all practical work to date these have been used. Figures 17(a) and 17(b) show the conditions for a single MOST acting as a switch and having gate control pulse excursions equal to the power supply. It will be seen from Fig. 17(b) that either switch electrode can act as the source—this is why both have arrows on them.



(a) Single MOST switch circuits.

n-CHANNEL SWITCH

INITIAL STATE	A AT +V (DRAIN)	A AT 0 (SOURCE)
B AT +V	B DRIFTS TO $+V-V_T$	B FALLS RAPIDLY TO 0
B AT 0	B RISES TO $+V-V_T$	B STAYS AT 0

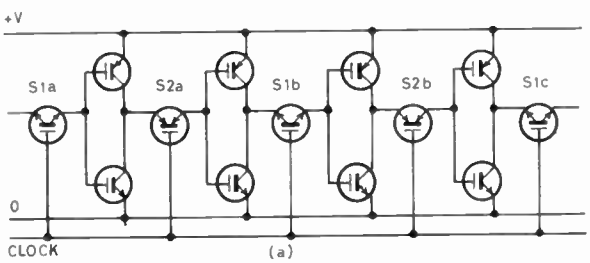
V_T = GATE THRESHOLD VOLTAGE

(b)

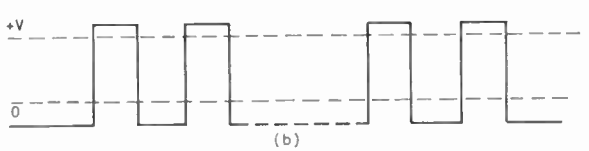
(b) n-channel switch conditions.

Fig. 17.

Considering the n-channel switch in Fig. 17(a) it will be seen that as long as the point A is at 0 V the gate can be driven by a rectangular positive voltage pulse and the gate/source voltage will remain constant for the duration of the pulse. Hence only the drain/source voltage can affect the channel current and it is possible to reduce the voltage across the switch to zero.

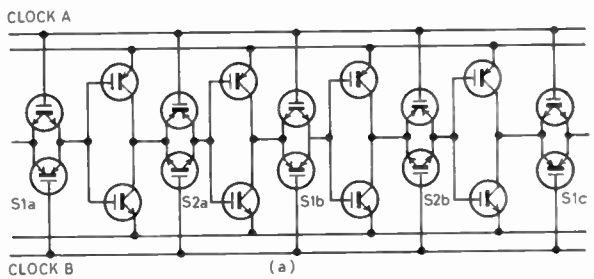


(a) Digital delay line with complementary alternate switches.

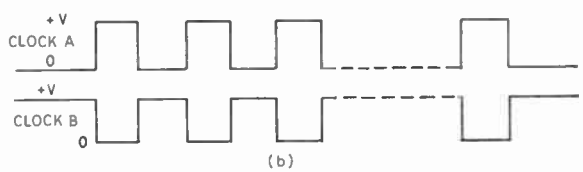


(b) Clock waveform. Note that p-channel switches are normally ON.

Fig. 18.



(a) Digital delay line with complementary alternate parallel switches.



(b) Clock waveforms. Note that S2 switches are normally ON.

Fig. 19.

On the other hand, if the point B starts off at 0 V then the effect of applying a positive gate pulse is to cause the point B to source-follow the gate waveform. This degenerates the channel current and so reduces the rate of change of voltage at B. Eventually, when the gate/source voltage equals the gate threshold voltage (V_T), the channel is cut off and the switch current is then limited to leakage. Practically, this means that, under these conditions, the point B will rise at an ever-decreasing rate towards a value V_T below the gate pulse amplitude. Hence, if it is desired to use a single MOST as a fast switch, it is necessary to use gate pulses providing a considerable overdrive above the level of the power supply. This overdrive should never be less than V_T, preferably much greater.

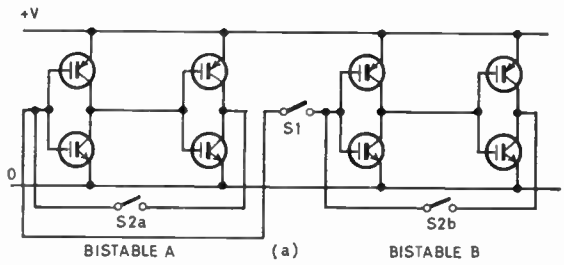
Obviously, similar arguments apply to the p-channel switch with a reversal of all polarities.

A further problem which arises with MOST switches is the d.c. potential of the substrate. Usually this is connected to the source for straightforward amplifier operation, and this is the case with the inverter transistors in the circuits so far discussed. Returning to Fig. 17(a), however, it will be seen that if the substrate is connected to point A then, if point A is positive, the substrate/source (= point B) diode will conduct and short circuit the switch. This can be overcome by connecting the switch substrate to a d.c. potential which ensures that this diode can never become forward biased. This potential will be 0 for an n-channel and $+V$ for a p-channel switch. However, when point B acts as the source it will tend to rise towards $+V$ and reverse-bias the substrate/source diode. This has a depletion effect on the channel analogous to an increase in V_T—hence, increases the amount of overdrive required still further.

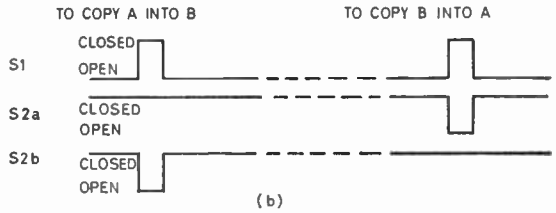
An interesting possibility which arises is that of using a single square wave clock and causing the positive and negative regimes to operate n-channel and p-channel switches alternately in adjacent half-shift elements. This is shown in Fig. 18. This saves one clock line per word but satisfactory operation can only be achieved if the gate pulse overdrive extends outside both power supply potentials. This may be inconvenient in some applications.

All this has referred to a single transistor switch. Satisfactory switch operation can also be obtained using a complementary pair of transistors in parallel and driving the gates simultaneously with anti-phase pulses—which need no longer exceed the power supply in amplitude (Fig. 19). This, however, increases the interconnection problem.

It is clear that the solution of the problem of transfer switches is of prime importance for the development of any viable system.



(a) Storage in bistable elements—switched cross-coupling system.

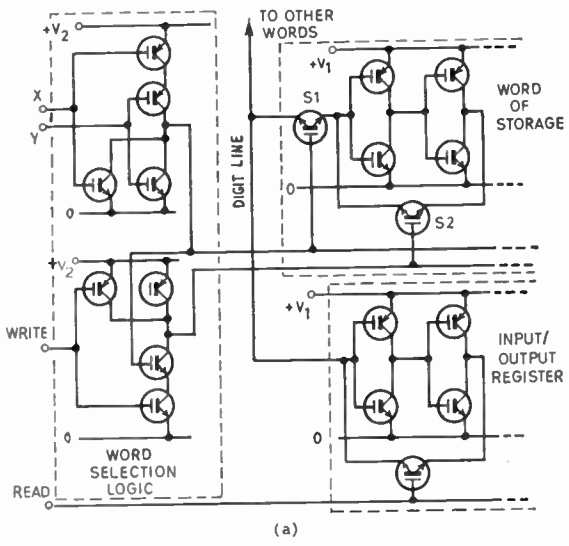


(b) Control waveforms.

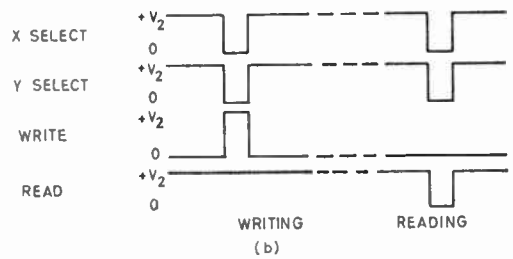
Fig. 20.

15. Parallel Storage in Bistable Elements

Where highest speed and/or minimum power dissipation are required it is necessary to use a parallel system. The shift register element previously discussed makes an element suitable for such a system but, in fact, is more complicated than is necessary as information only has to be transferred from one point in the system to another and does not have to propagate along a line of elements. This dispenses with the need for a temporary memory and two phase clocks. Figure 20(a) shows a simple system which permits bidirectional transfer between two bistable elements. The elements are very similar to those in Fig. 16, the difference being that one switch in the cross-coupling is no longer necessary. The normal state of the system is with the S2 switches closed and S1 open. The conditions for transferring information in the two possible directions are shown in the waveforms (Fig. 20(b)). The nodes connected to S1 are low impedance as long as the S2 switches are closed. If S1 is closed under these conditions, and the states of the two elements differ, it is not possible to say which element will set the other as S1 is connecting together two nodes of comparable (fairly low) impedance. If, however, one of the S2 switches is opened prior to closing S1 then S1 will connect one node consisting only of the input gates of an inverter (in the element having S2 open) to another still having the low output impedance of an element with both cross couplings complete. This allows the element having S2 still closed to set the other unambiguously.



(a) Section of random access parallel store.



(b) Control waveforms. $+V_2 > +V_1$ to provide overdrive to switch MOST's.

Fig. 21.

A small section of a system utilizing this technique—which has been called the 'switched cross-coupling method'—is shown in Fig. 21(a). This includes one digit of one word of store, one digit of a common input/output register and address circuitry to perform coincident selection of one word of storage. The control waveforms are shown in Fig. 21(b) and make allowance for inversion in the address gates.

The most serious disadvantage of this system over the serial ones previously discussed would appear to be the number of digit lines which have to be brought off the slice. Furthermore, access has to be available to each element in both coordinates which may make for complication of the interconnection pattern, although it is doubtful if this system is any worse in this respect than some of the more advanced variants of the serial systems. A limit is placed upon the number of digit selection transistors which can be connected onto a single digit line due to the total capacitance of the digit line which loads the output node of a selected element. Eventually, this may draw such a large charging current that it may destroy the state of the sending element. This can be overcome either by further modifying the storage element, or by treeing the digits into the digit line, which gives a substantial reduction in capacitance.

16. Practical Results

Many of the systems discussed have been tried and a brief résumé of results may be of interest. All work has been done using discrete experimental type transistors type VX 6504 (p-channel) and VX 6505 (n-channel). As previously mentioned it is anticipated that these figures can be improved quite substantially.

Sections of complementary MOST digital delay line have been made operating from a +10 V supply and using n-channel switches with +12 V to +20 V gate pulses. The line had a maximum frequency of about 5 MHz at +20 V gate pulse amplitude. Power consumption is at a maximum when shifting a regular alternate pattern of 0's and 1's and worked out at approximately 3·5 mW/MHz. This is very close to the calculated value of fCV^2 where C is the total half element nodal capacitance, including switch, and should be 30–40 pF per element. Failure at low frequencies, due to discharge of the memory capacitors, occurred from about 50 Hz to 1 kHz, depending upon leakage in individual switch transistors. This failure is heralded by a sharp rise in power consumption due to the fact that the memory partially discharges between clock pulses and conduction takes place through both transistors of an inverter.

Similar results were obtained for a shift register, the main difference being that there was no lower limit of frequency. Power consumption, whilst shifting, was slightly higher than the delay line, due to the capacitance of the extra switch transistor.

Transfer between two digits set up in the 'switched cross-coupling method' has been achieved in less than 150 ns. A simulated store of up to 256 words of 24 bits with tree address decoding and routing of digit input/output has been built. The delay through the trees was approximately 100 ns per level. On writing, the address digits and information can appear in parallel and propagate through the trees coincidentally. On reading, the information from the selected element cannot start propagating through the output tree until the address tree is set up. Hence reading takes longer than writing, although the system could be organized to offset the effects of this. Write times for an 8-level (256 word) system were about 0·9 μs and for read 1·1 μs. An interesting point about this system is that the delay is proportional to *level* and hence a further 10 levels—250 000 words—would only increase the delay by approximately 1 μs. It appears

that the organization of such a store might best be achieved by making each slice carry all elements for a given digit and use as many slices as there are digits in the word. Each slice would then have only one input, one output and address digits as active external connections. Packing densities at the moment approach 10^5 MOST's per square inch so a 12-level store (4096 words) may be practicable some time in the future, but it would only be economic if it offered a speed/cost advantage over other systems. To achieve this it appears that an increase in speed of at least 10 times is required coupled with a finished cost of 1d. or so per bit (£20 per slice of 4096 bits plus trees).

For *storage* applications the writers consider that MOST's will offer only scratchpad-sized memories (100–1000 words) for at least another decade. However, technological breakthroughs have a habit of startling even the most visionary and, eventually, the developments previously referred to may offer some really cheap fabrication process.

17. Conclusions

MOST logic circuits have the general advantages of logical flexibility and economy, high noise immunity and large fan-out. Non-complementary circuits are extremely simple to manufacture, and hence cheap, but are extremely slow, perhaps two orders of magnitude slower than the best bipolar integrated circuits. Complementary circuits are about an order of magnitude faster, and have a lower power consumption, particularly at low *mean* switching rates, which will allow full advantage to be taken of their small size to achieve a high packing density and hence increased system reliability. The technological difficulties of integrating p- and n-channel MOST's on a common substrate have not yet been solved completely, although complementary integrated circuits have been made in the laboratory (e.g. Ref. 6). However, there is every indication that the major problems will be solved within the next year, and that the cost of complementary MOST logic circuits will compare favourably with that of bipolar integrated circuits.

At present MOST's in complementary circuits offer some speed advantage over ferrite cores in small storage arrays but cannot compete with bipolar transistors on speed. However, when the problems of monolithic integration of complementary devices are solved it may well be that they will be cheaper to fabricate than bipolar circuits, allow much higher packing densities, and be faster than the core. Eventually some type of insulated gate field effect device —such as the thin film transistor—may make very large, relatively slow (1–10 μs) stores economic.

What is quite clear is that the circuit designer has to re-orientate his thinking to take advantage of the new technologies being developed and also to bear in mind their inherent limitations. This paper has attempted to extrapolate from present experience to show what sort of design considerations must be borne in mind when exploiting these techniques. At present separate MOST's are available and it is still possible to think in terms of circuits using discrete units. This, whilst being convenient, constitutes the danger to full exploitation. Techniques for overcoming the fabrication problems are well under way. It is not unlikely that the computer systems of the early 1970's will have substantial sections in MOST monolithic integrated form, provided that an adequate appraisal of what is, and is not, feasible can be made in the near future.

18. Acknowledgment

Crown Copyright. Reproduced with the permission of the Controller, Her Majesty's Stationery Office.

19. References

1. S. R. Hofstein and F. P. Heiman, 'The silicon insulated-gate field-effect transistor', *Proc. Inst. Elect. Electronics Engrs*, **51**, No. 9, pp. 1190–1202, September 1963.
2. C. T. Sah, 'Characteristics of the Metal-Oxide-Semiconductor Transistor', *Trans. I.E.E.E. on Electron Devices*, **ED-11**, No. 7, pp. 324–44, July 1964.
3. J. D. Schmidt, 'Integrated MOS transistor random access memory', *Solid State Design*, January 1965.
4. G. E. Moore, C. T. Sah and F. M. Wanlass, 'Metal-oxide-semiconductor field-effect devices for micropower logic circuitry'; 'Micropower Electronics', Agardograph 77. (Pergamon Press, London, 1964.)
5. J. R. Burns, 'Switching response of complementary-symmetry MOS transistor logic circuits', *RCA Review*, **25**, No. 4, pp. 627–61, December 1964.
6. P. J. Coppen, 'FET complementary integrated circuits', *Electronics*, **37**, No. 32, pp. 55–8, 28th December 1964.

Manuscript first received by the Institution on 21st September 1965 and in revised form on 21st March 1966. (Paper No. 1055.)

U.D.C. 551.577: 621.396.962.33

The Use of Doppler Radar in Meteorological Research

By

P. G. F. CATON, M.A., Ph.D.†

Presented at a meeting of the Radar and Navigational Aids Group meeting held in London on 17th November 1965.

Summary: A 3-cm pulsed Doppler radar constructed at the Royal Radar Establishment is described and the use of this equipment to increase knowledge of atmospheric precipitation processes is explained. With the aerial pointing vertically in widespread rain it is possible to deduce the drop size distribution in the free atmosphere and to study its variation with height. In showers it is possible to derive height-time sections of both vertical air motion and maximum particle velocity; examples of these sections are shown and discussed. A new mobile Doppler radar designed for investigation of thunderstorms is briefly mentioned. Finally, the application of Doppler radar operating at low angles of elevation to the measurement of horizontal convergence and vertical velocity of the air during widespread rain is described and initial results shown.

1. Introduction

The Meteorological Office has maintained a radar research unit since 1947. For twelve years the unit was located at East Hill in Bedfordshire and used plan-position and range-height radars primarily to study the physical structure of precipitation systems. Considerable attention was also paid to the investigation of turbulence in cooperation with aircraft, and experience was gained which led later to the use of radar to aid short-period forecasting of precipitation through the filling of gaps in the synoptic reporting network. This period of research culminated in 1959 with the observation of a severe hailstorm near Wokingham by several radars of different wavelength and beam characteristics grouped together at East Hill. From these observations Browning and Ludlam[1] inferred qualitatively the pattern of airflow within the storm which they suggested was characteristic of a certain type or phase of severe travelling storms. In the autumn of 1959 the Meteorological Office radar unit moved to the Royal Radar Establishment at Great Malvern to exploit new radar equipments which were available there. These were an 8·6 mm radar having a 16 ft diameter aerial giving an exceptional degree of resolution, and a 3·2 cm pulsed Doppler radar giving information of received power as functions of range and radial velocity of the scattering particles. This paper is concerned with the meteorological use of the latter equipment and the approach is that of a research scientist seeking to increase our knowledge of atmospheric precipitation processes.

† Meterological Office Research Unit, Royal Radar Establishment, Great Malvern, Worcestershire.

2. The Equipment

The basic parameters of the radar[2] are:

Frequency	9375 MHz
Pulse length	0·8 μs
Peak power	10 kW
Pulse repetition frequency	2426 /s
Aerial beamwidth between half-power points	1·5 deg
Aerial gain (allowing for wave guide losses)	38 dB
Receiver noise factor	13·5 dB

The Doppler frequency shift for meteorological targets is only an extremely small fraction of the transmitted frequency and it is convenient to infer this shift from the change from pulse to pulse in the phase of the returned signal. To measure this it is necessary to 'remember' the phase with which the transmitted pulse starts. This is achieved by the 'stalo-coho' method; a sample of the magnetron pulse, mixed with a very stable local oscillator to reduce it to i.f., is used to pull a coherent oscillator of this i.f. into phase. The returned signal is then compared with the phase of this coherent oscillator and detected in a phase-sensitive detector giving an output proportional to amplitude and phase. After a number of pulses the velocities present can be determined unambiguously within a range 19·4 m/s set by the wavelength 3·2 cm and the pulse repetition frequency 2426 per second. Velocities outside the range of 19·4 m/s 'fold-over' the edges and, in favourable circumstances, may be disentangled. A shift oscillator is available to alter the phase at which

the coherent oscillator starts, enabling zero target velocity to be set anywhere within the range and upward and downward velocities to be distinguished. The video signals from the phase sensitive detector are fed to a storage tube, successive traces being written below one another to form a frame of 80 lines, each line 80 μs long. These signals are read off orthogonally in elements of width 1 μs (corresponding to a range interval of 150 m) over the range 0–12 km. In this write-read process the signal frequencies are multiplied by 320 and now are within the range 0–388 kHz. After mixing with a frequency of 470 kHz the signals are fed into a bank of 20 filters each 20 kHz wide (corresponding to a velocity spread of 1 m/s) centred at 20 kHz intervals from 470 to 850 kHz. The outputs of the filters grid modulate a cathode-ray tube to give a range-velocity display, shown in Fig. 1. If intensity measurements are required these are obtained by r.f. attenuation in 3 dB steps, the value assigned corresponding to the attenuation inserted when the signal is just visible above the background noise. Since the 'write' time on the storage tube is only 33 ms, Doppler frequencies less than a few times 30 Hz (0·5 m/s) are not well represented. However, these frequencies do not necessarily correspond to low target velocities, since the shift oscillator is usually set so that zero velocity is near the middle of the 19·4 m/s range. The photographic exposure time, about 3 seconds, involves the integration of 30 independent frequency analyses, and ensures that the results of the 'threshold' technique of intensity measurement are reproducible within the limit of the steps of the attenuation sequence (3 dB).

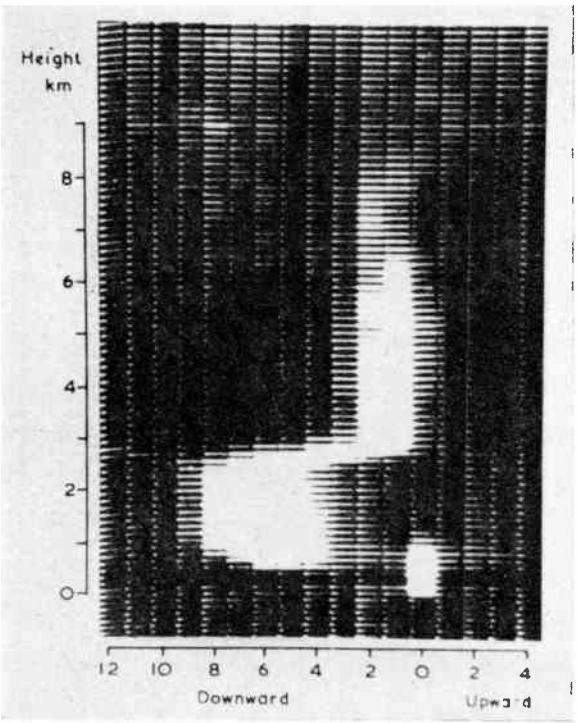


Fig. 1. Photograph of range-velocity display, looking vertically in widespread rain. Each row corresponds to a height interval of 150 m, each column to a velocity band of 1 m/s. The echo from scatterers in each height interval and velocity band is summed and presented as intensity modulation. This example shows, above height 3 km, ice crystals and snowflakes falling mainly at 1 and 2 m/s, and at lower heights raindrops falling mainly at 3 to 9 m/s. The disappearance of echo below 450 m is due to the receiver recovery time. The background represents noise and the strong signal in the zero velocity channel below 1 km break-through of the transmitter pulse and static reflections from side-lobes.

3. Observations in Widespread Precipitation

If the aerial is pointed vertically upwards, the observed fall velocities are those of the particles through still air plus any vertical velocity of the air itself. In widespread 'steady' rain the vertical motion of the air, thought to be in an upward direction at speeds between 5 and 20 cm/s, may usually be neglected in comparison with the fall velocity of the precipitation particles. This velocity lies between 0·5 and 2·0 m/s for ice crystals and snowflakes, and between 1·0 and 9·0 m/s for drizzle and rain drops. In the case of rain there is a well-defined relation between fall velocity in still air and drop diameter (this relation depends on air density and viscosity, and thus varies with height and temperature); thus the velocity channels of the Doppler display correspond to identifiable bands of drop diameter. Drop size distributions may be derived from intensity measurements in each velocity channel and the relation:

$$\text{echo intensity} \propto \sum N_D \Delta D \, D^6,$$

where N_D is the concentration of drops per unit volume and per unit interval of diameter D.

The median drop size distributions observed at height 750 m in rains of intensity between 0·1 and 5·6 mm/h are shown in Fig. 2, and the variability within four of the major categories is indicated in Fig. 3. The variation of size distribution with height gives information on precipitation processes. Thus we find that below the melting level in widespread rain there is relatively little change in drop size distribution; growth of raindrops by the accretion of cloud droplets cannot generally be detected, implying that the water content of the clouds is very low (this conclusion almost certainly does not apply in hilly regions); evaporation of the raindrops below cloud base is also negligible. However, there is evidence of raindrop coalescence, the larger faster-falling drops catching up and coalescing with the smaller slower-falling drops, leading to an increased concentration of large drops and a decreased concentration of small drops towards the ground. Eventually the drop size distribution may depart considerably from the well-

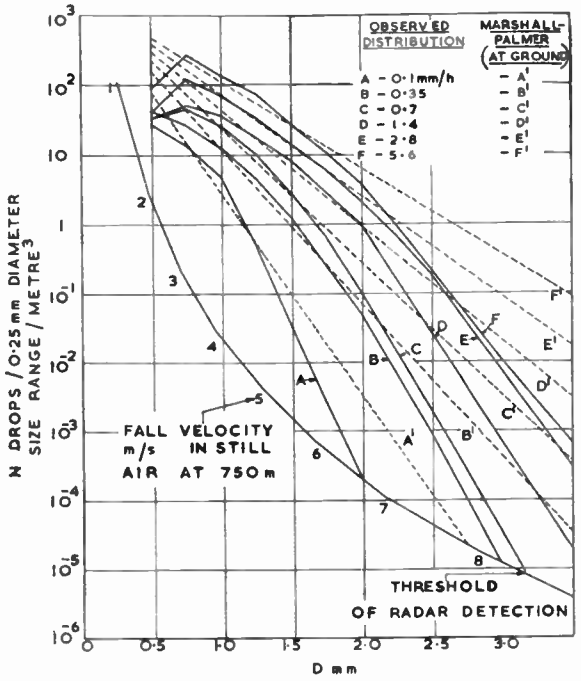


Fig. 2. Median drop size distributions at 750 m in widespread rain. The full lines A to F represent the medians of observed drop concentrations corresponding to rainfall rates $R < 0{\cdot}15$, 0·15–0·5, 0·5–1·0, 1·0–2·0, 2·0–4·0 and $> 4{\cdot}0$ mm/h. The distributions are normalized to R, 0·1, 0·35, 0·7, 1·4, 2·8 and 5·6 mm/h respectively. The dotted lines A′ to F′ show the corresponding Marshall and Palmer exponential distributions. (From Ref. 3.)

known Marshall and Palmer exponential law,[4] $N_D = N_0 e^{-\lambda D}$, with λ a function of rainfall rate, and may show a peak concentration of drops at diameter 0·75–1·00 mm. (See Fig. 2, rainfall rates > 0·5 mm/h.)

A comparable study of processes in the snow region presents much greater difficulties. There are a variety of crystal forms with different fall velocity-size relationships. In general the fall velocity increases only slowly with considerable increase in particle size, and with echo intensity proportional to the sixth power of the melted diameter and complications due to shape, size relative to the radar wavelength and riming of the flakes, it is apparent that velocity channels 1 m/s wide are quite unsuitable. There is provision for writing the information from 80 *alternate* pulses on the storage tube, thus doubling the multiplying factor between 'read' and 'write' and converting the channel widths to $\frac{1}{2}$ (or even $\frac{1}{4}$) m/s. Even so, further difficulties intrude. The spectrum of velocities is broadened by atmospheric turbulence and by a component of the horizontal wind at the edges of a finite beam; for example a wind of 20 m/s will cause a velocity spread of $\pm 0{\cdot}25$ m/s with the present aerial beam. These latter difficulties do not appreciably affect the root mean square Doppler frequency and, possibly, some information might be extracted from this parameter if it was measured with sufficient accuracy.

There remains the 'melting band', the region of enhanced echo some 200 m in depth just below the 0°C isotherm, caused by the melting of snowflakes and their break-up and acceleration during transition to rain. A radar is a rather blunt instrument with which to examine this band, since a single sample involving a time gate of 1·0 μs and a pulse duration of 0·8 μs implies that echo from a zone 270 m deep is superimposed. Nevertheless Lhermitte and Atlas,[5] using profiles of root mean square Doppler frequency and total reflectivity, have concluded that in a rain of rate 0·8 mm/h there was substantial aggregation of snowflakes as well as wetting in the upper portion of the band, followed by break-up of each large wet aggregate snowflake into 4–6 raindrops in the lower portion.

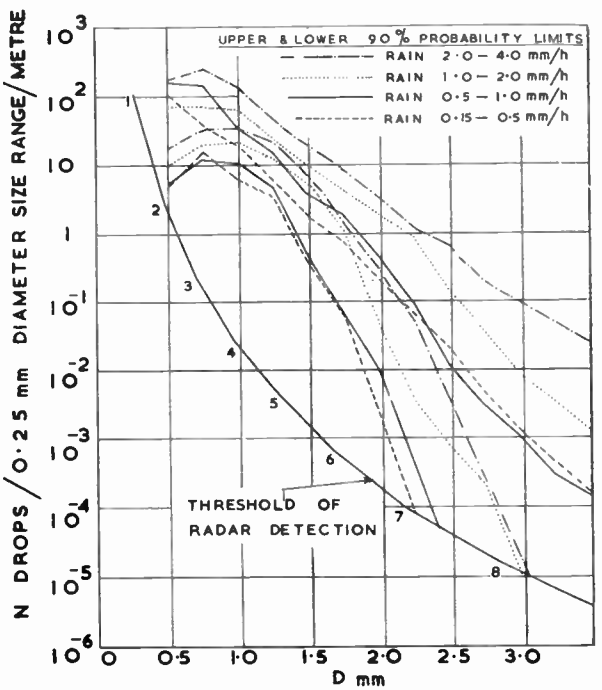


Fig. 3. Variability of drop size distributions at 750 m in widespread rain. The curves show, for four rainfall categories, the range within which approximately 90% of the drop concentrations lie. (From Ref. 3.)

4. Observations in Showers

When the aerial is pointed vertically in showers, the observed velocities are the sum of the downward velocities of the particles through the air and the upward or downward component of air motion (Fig. 4). It is possible to separate these two velocities by making an assumption concerning the fall speed through the air of the slowest falling particles. At

heights where the temperature is below 0°C it seems likely that there are present sufficient ice crystals or supercooled water droplets falling at 0·5–1·0 m/s to give a radar echo. In regions of moderate or heavy rain, it is thought (by analogy with observations in widespread rain) that the concentration of small drops may be much reduced by the process of coalescence so that the minimum velocity detectable is about 1·5–2·0 m/s. Thus displacements of the minimum velocity from the 1 or 2 m/s channels respectively indicate the presence of vertical air motion. By repeated observations as the shower passes over the radar, a height-time section of vertical air motion may be obtained, subject to the validity of the above assumptions.

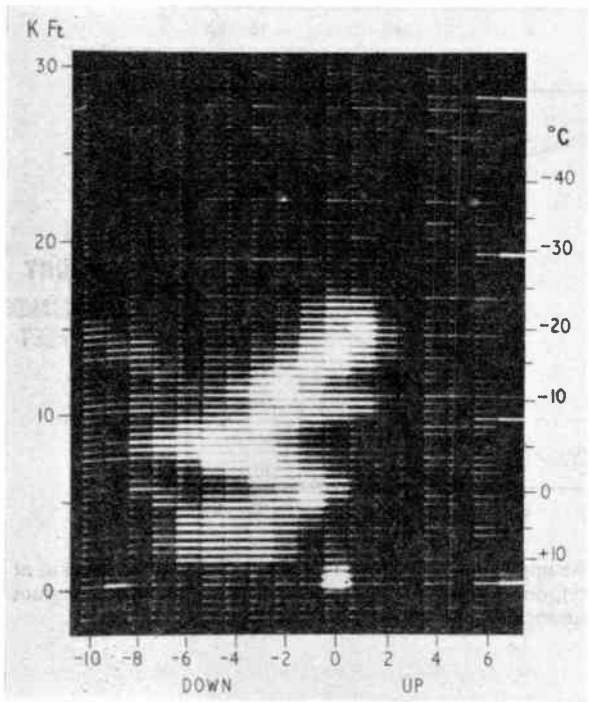


Fig. 4. Photograph of range-velocity display, looking vertically in a shower. The zig-zag pattern is due to the observed velocity being the sum of the downward velocity of the particles through the air and the upward or downward component of air motion. The weak mirror-image signal on the left side of the photograph is a harmonic component and should be disregarded.

Turbulence and variation of the draught air velocity within the pulse volume and the spectrum broadening effect of the horizontal wind lead to an overestimation of up-draughts and an underestimation of down-draughts, but these effects do not appear to be serious except in localized regions of storms. The number of velocity channels containing signal, after correction for the above effects and in combination with the minimum velocity assumptions, leads to an estimate of the fall velocity through the air of the largest particles producing echo. If the nature of these may be inferred, their size may be approximately determined. A check on the air velocity in the rain region is possible since, unless hail is also present, the fall velocity through the air of the largest stable particles does not exceed about 9 m/s near the ground or 10–11 m/s aloft.

A height-time section of vertical air motion observed in a shower which passed over the radar at Pershore, Worcs., on 9th April 1959 is shown in Fig. 5.[6] The left-hand side is the leading edge of the shower and a horizontal distance scale has been added based on the controlling windspeed of 14 knots. The maximum rainfall rate of 8 mm/h occurred at the beginning of the shower and a total amount of 0·7 mm was recorded. It will be noticed that upward air motion is largely concentrated in the upper half of the shower with two broad maxima of velocity of 4 m/s separated by about 2 km. Downward air motion is concentrated in the lower part of the cloud and this also attains a maximum velocity of 4 m/s, just above the 0°C isotherm. The lower diagram shows a simplified pattern of the air motion within this shower, and the authors suggest the possibility of failure of the minimum velocity assumption leading to an underestimation of the up-draught in the front portion of the shower between the 0°C and −10°C isotherms.

Over 20 showers at Pershore have now been recorded and analysed. They show great variety and little generalization is possible regarding the pattern of air motion. Almost all are more complicated than the example above. A second analysis is shown in Fig. 6. This mature shower was quite small, only about 5 km diameter, but due to the slow movement rainfall was prolonged. The maximum rate was about 75 mm/h and the total amount 9·3 mm. In the diagram an attempt has been made to indicate both the vertical air motion and the pattern of spectrum width (corresponding approximately to the fall speed of the largest particles); the vertical air motion is shown by isopleths, full-line for upward motion and dashed line for downward motion, while the spectrum width is indicated in three categories by white (low values), sparse stippling and dense stippling. Above the 0°C isotherm the small white areas probably contain only ice crystals and snowflakes or small supercooled water drops, while the areas densely stippled appear to contain (besides slow-falling particles) large particles of *graupel* (soft hail) or large supercooled raindrops. There is no evidence for large hail in this shower section. Although the pattern of air motion is complicated, the areas of upward motion are again concentrated above 3 km while there is a large area of downward motion below this height. This down-draught reaches a maximum velocity of 7 m/s at low levels (at 1530–1534 G.M.T.)

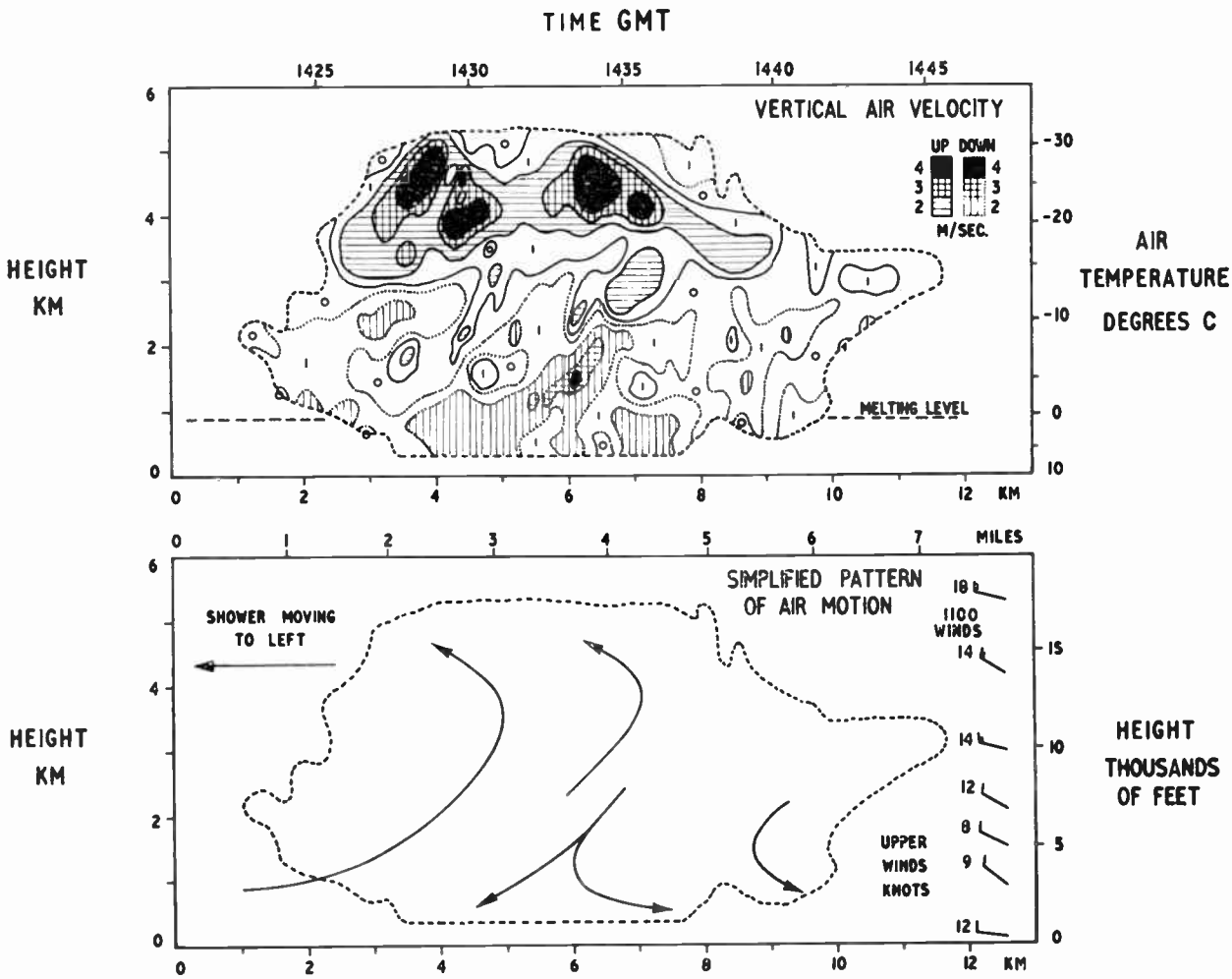


Fig. 5. Pattern of vertical air motion in a shower, 9th April 1959. The upper diagram shows isopleths of vertical air motion at 1 m/s intervals, full-line for upward motion (isopleths ½, 1½ m/s, etc.), dotted lines for downward motion. The lower diagram suggests a simplified pattern of air motion in this shower. (After Probert-Jones and Harper.[6])

and this corresponds with the time of heaviest rain at the ground. An interesting feature is the tower observed at 1526–1528. If the horizontal movement is reliable this tower appears to be only 200–400 m across, of which 150 m is contributed by the width of the beam at this height. Upward motion exceeding 5 m/s is inferred and it is gratifying to record that a range-height radar at Malvern observed this echo top to rise 600 m in two minutes. From the more detailed original analysis it can be seen that the areas of maximum up-draught at height 4 km frequently coincide with areas of sharply increasing spectrum width, suggesting that rapid growth of the particles is occurring while they are suspended in or are falling through high concentrations of supercooled droplets. Columns of high spectrum width extend downwards from several of the up-draught maxima, indicating large particles descending towards the ground; little further growth occurs in the down-draught where cloud water concentrations are probably low due to evaporation.

In future work it is hoped also to deduce patterns of total echo power and in fact power corresponding to each velocity channel. The detailed interpretation of these patterns faces one great difficulty; we do not know sufficient of the nature of the particles above the 0°C isotherm. If these are predominantly snowflakes this is indicated by a rapid increase in spectrum width on melting below the 0°C isotherm; however, if they are *graupel* there is not necessarily an appreciable change of velocity on melting and such particles may not be distinguishable from supercooled raindrops using radar evidence alone.

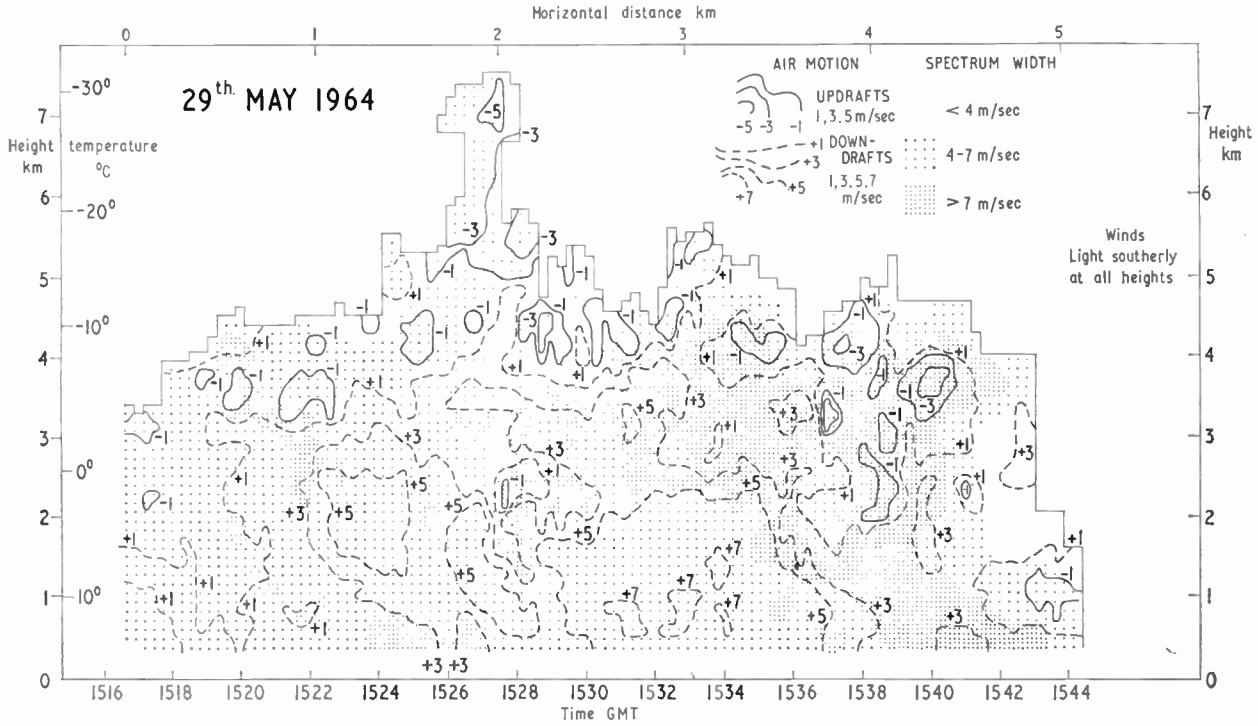


Fig. 6. Pattern of vertical air motion and spectrum width in a shower, 29th May 1964. The vertical air motion is shown by isopleths at 2 m/s intervals, full-line for upward motion (−ve) and dashed line for downward motion (+ve). The spectrum width is shown in three categories by absence of stippling (< 4 m/s), sparse stippling (4–7 m/s) and dense stippling (> 7 m/s).

5. A New Mobile Doppler Radar

The Doppler radar which has been used to date is fixed in a laboratory and has a velocity range of only 19·4 m/s. This is insufficient to enable analysis of the strong up- and down-draughts in thunderstorms and a second radar has been constructed by the Royal Radar Establishment, mounted in a vehicle to permit transport to regions where severe storms are more frequent. Figure 7 shows an exterior view of the vehicle. The basic method of phase measurement is retained, but a higher pulse repetition frequency of 10 kHz provides an unambiguous velocity range of 80 m/s whilst still

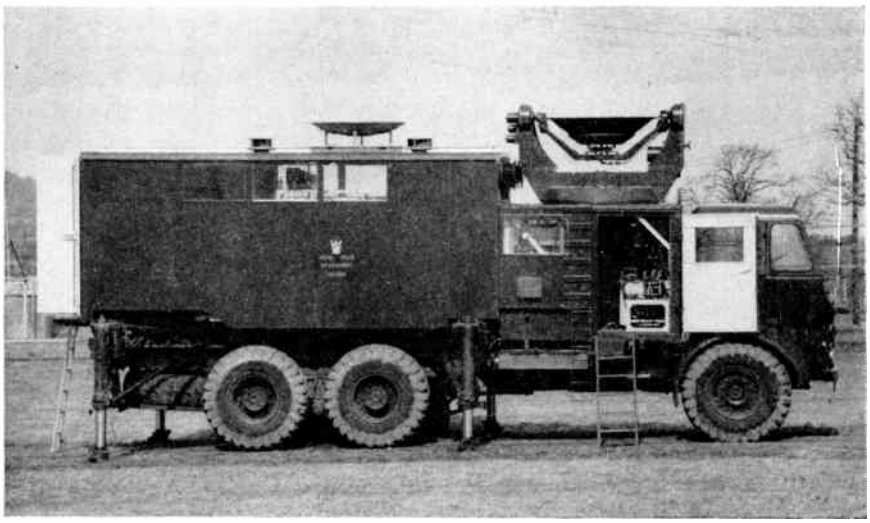


Fig. 7. A mobile Doppler radar, constructed at the Royal Radar Establishment, Great Malvern.

The petrol-electric generator supplying all necessary power, the radar transmitter, the waveguide assembly and the aerial are mounted in the forward compartment of the vehicle. The receiver equipment and the recording system are housed in the rear portion. The frequency analysing equipment is in a separate laboratory.

permitting observations to a range of 12 km. The video signals from the phase-sensitive detector are filtered down to audio frequencies and recorded on a multi-track tape recorder. Seven height elements 150 m wide are examined simultaneously for one second, so that after 11 seconds 77 elements (450 m–12 km) have been recorded. The process is repeated every 15 seconds. The signals are later analysed in the laboratory, one track and one frequency at a time. A linear dynamic range of 40 dB has been achieved throughout the system and this is extended by the inclusion of 30 dB of r.f. attenuation on alternate 15-second samples. This wide range is necessary to cope with the range of reflectivities within severe storms.

6. Studies of the Horizontal Convergence of Air during Widespread Rain

One of the most difficult measurements in meteorology is that of the slow ascent of air leading to widespread precipitation. Average values of 5–20 cm/s have been inferred from typical rates of rainfall [7] and by applying the dynamical equations to radiosonde observations.[8] However, variations of vertical air motion over areas of 'county' size almost certainly occur and these, combined with variations in the humidity of the air, account for the appreciable variations in rainfall which occur within the larger area of precipitation. Doppler radar should contribute to our knowledge of vertical air motion on this 'county' scale through measurement of the horizontal convergence of the air. The technique is based on the principle that any net horizontal inflow of air into a horizontal disk is balanced by excess upward air motion through the upper surface relative to that through the lower surface. A series of observations are made at 15 deg intervals of azimuth using aerial elevations of 30 deg and 15 deg. The radar measures the line-of-sight velocity of the precipitation particles; after correction for the component of fall velocity (observed by the radar at elevation 90 deg) the net inflow or outflow of the horizontal wind over the circle of observation may be determined. Since observations are available at a number of ranges the variations of convergence with height may be examined and vertical velocities inferred. The errors in the measurements have been carefully considered so that we are confident that the results have significance. In addition rapid repetition of the measurements gives results suggesting that the values are representative of a rather larger area than that sampled by the radar (85–400 km^2 at height 3 km). As the rainbelt passes over the radar a height-time section of convergence and vertical motion may be derived. Figure 8 shows results obtained from observations on 1st June 1964.[9] In the left-hand diagram areas of significant convergence and divergence are indicated by letters C and D respectively; in the right-hand portion are shown profiles of vertical velocity obtained by independent observations at elevation 30 deg and 15 deg at times some 25 minutes apart; the agreement between the curves is striking. This investigation is in an early stage, but the results appear most promising. In future work we desire to use both the fixed and mobile Doppler radars some distance apart to obtain parallel sections through the precipitation and demonstrate convincingly the scale of the phenomena that we are observing.

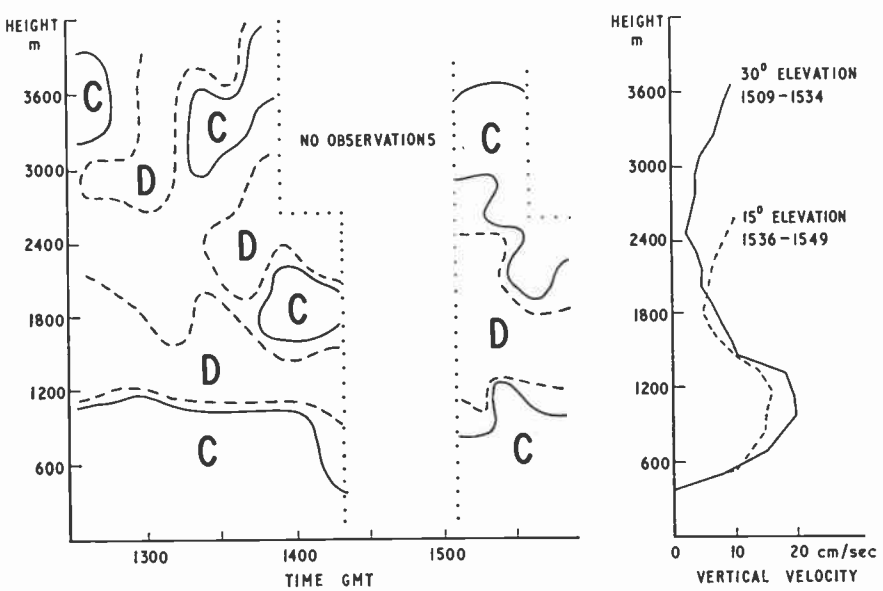


Fig. 8. Pattern of horizontal convergence and vertical velocity during widespread rain, 1st June 1964.

In the left-hand diagram areas of significant convergence and divergence are indicated by letters C and D respectively. In the right-hand diagram profiles of vertical velocity are shown, derived from independent observations at aerial elevation 30 deg and 15 deg at times some 25 minutes apart. (After Harrold.[9])

7. Acknowledgments

In this paper the author has drawn heavily on the work of colleagues, past and present, of the Meteorological Research Unit, Great Malvern. It is a pleasure to thank the Director of the Royal Radar Establishment for the opportunity to use Doppler radar equipment, and the several members of his staff who have contributed to its development. This paper is presented by permission of the Director-General of the Meteorological Office.

8. References

1. K. A. Browning and F. H. Ludlam, 'Airflow in convective storms', *Quart. J. Roy. Meteorol. Soc.*, **88**, pp. 117–35, April 1962.
2. E. H. Boyenval, 'Echoes from precipitation using pulsed Doppler radar', *Proc. 8th Weather Radar Conf., Amer. Met. Soc.*, pp. 57–63, Boston, 1960.
3. P. G. F. Caton, 'A study of raindrop size distributions in the free atmosphere', *Quart. J. Roy. Meteorol. Soc.*, **92**, pp. 15–30, January 1966.
4. J. S. Marshall and W. McK. Palmer, 'The distribution of raindrops with size', *J. Met.*, **5**, No. 4, pp. 165–6, 1948.
5. R. M. Lhermitte and D. Atlas, 'Doppler fall speed and particle growth in stratiform precipitation', *Proc. 10th Weather Radar Conf., Amer. Met. Soc.*, pp. 297–302, Boston, 1963.
6. J. R. Probert-Jones and W. G. Harper, 'Vertical air motion in showers as revealed by Doppler radar', *Meteorological Magazine*, **91**, pp. 273–84, October 1962.
7. J. K. Bannon, 'The estimation of large-scale vertical currents from the rate of rainfall', *Quart. J. Roy. Meteorol. Soc.*, **74**, pp. 57–66, January 1948.
8. E. Knighting, 'Some computations of the variation in vertical velocity with pressure on a synoptic scale', *Quart. J. Roy. Meteorol. Soc.*, **86**, pp. 318–25, July 1960.
9. T. W. Harrold, 'Measurement of horizontal convergence in precipitation using a Doppler radar—a case study', *Quart. J. Roy. Meteorol. Soc.*, **92**, pp. 31–40, January 1966.

Manuscript first received by the Institution on 17th November 1965 and in final form on 7th April 1966. (Paper No. 1056/RNA49.)

DISCUSSION

Under the chairmanship of Mr. R. N. Lord

Mr. T. W. Welch: I would like to comment upon the relationship between radio engineers and the meteorologist. In one sense the weather men are our clients and customers —for their day-to-day work both in research and in forecasting we expect to be able to supply them with radio, radar and other electronic apparatus of great variety. In another sense they are our colleagues and advisors—certainly no radio engineer who has ever been concerned with the propagation of microwaves in atmospheric paths will deny the very important contribution to his work made by the meteorologist, whilst most radar engineers have been grateful for the guidance of meteorological researchers in their endeavours to understand and defeat the clutter and attenuation problems which come with weather.

In my own fairly long association with radio-meteorologists I have come to feel that in many ways meteorology is the Cinderella of our applied sciences. The amount of hard cash which is allotted both for research and for operational equipment, in this country at any rate, is ludicrously small. Perhaps because of this impoverishment the meteorologists and their radio technician assistants have devised some remarkable adaptations of equipment to meet their needs and have achieved some notable results with the most unlikely apparatus. It is at this point that we should, perhaps note the rather different requirements of the researcher in meteorology and his colleague who is concerned with forecasting. The needs of the researchers are rather highly specialized and, in terms of the numbers of pieces of apparatus required, too small to merit very large private venture investment in satisfying them. On the other hand, the requirements of the forecasting branches could be numerically attractive to private industry in many fields, from storm-warning and wind-finding radars to communication networks, but the authorities concerned are a little tardy in standardizing specifications and defining real requirements. Obviously what is numerically an attractive market for private venture investment becomes much less attractive when there is danger that the specifications set by two authorities will be so completely different as not to be satisfied by the constituent elements of a system based on a single philosophy. One important result of this lack of agreed specifications is that, in the radar field at least, the equipment on offer from industry is to some extent a mere rehash of apparatus in quantity production for some quite different field of use.

A really strong lead from the Meteorological Office in this country, were it allowed the necessary funds, could do much to improve the equipment on offer and bring down unit costs thereby making British equipment and systems attractive to the meteorologists throughout the still very considerable areas of the world which look to the United Kingdom for guidance and for equipment.

Coming to the specific matters which Dr. Caton has put before us, it is, of course, of perennial interest to meteorologists and radio engineers to know more about the drop-size distribution in precipitation. Upon our knowledge of the behaviour of this parameter in changing circumstances of rainfall rate, the geographical location and so on, depends our ability to deduce from fairly simple radar presentations quantitative assessments of aerial distribution of precipitation. In those parts of the world subject either to disastrous seasonal flooding or to a chronic inadequacy of water catchment these measurements are becoming more and more important. It is probably too much to hope that a wide network of these specialized and expensive Doppler radars will ever come into service to provide themselves

with the regular quantitative measurements needed. It would, however, be interesting to hear from Dr. Caton whether he considers that there are good chances of being able to derive from a few suitably placed Doppler equipments a satisfactory calibration procedure for converting iso-echo measurements made on conventional p.p.i./h.r.i. radars into rainfall rate with a good probability of accuracy. I would also like Dr. Caton's comments on the modification of results to be expected in other climates, particularly tropical ones, and of the extent to which the radar at Malvern has given indications of precipitation not reaching the ground.

Dr. Caton (*in reply*): I do not think it necessary to use a network of Doppler radar equipments to determine the relation between radar echo intensity and rainfall rate. Very considerable information relating these quantities has already been derived in many parts of the world using techniques of raindrop sampling at the ground. The great majority of workers have deduced relations of the form $Z = aR^b$, where Z is the reflectivity factor in $mm^6\ m^{-3}$, R the rainfall rate in mm/h, with values of a varying between 70 and 700 and b between 1·0 and 2·0 depending on the type of rain.†

My colleague Mr. T. W. Harrold has given a critical review‡ of the possibility of estimating rainfall using radar. He discusses not only the variability of the Z–R relation, but also the effect of attenuation of the radar beam through intervening precipitation and the variation of reflectivity with height which is important both above the 0°C isotherm in widespread rain and at all levels in convective storms. In dry climates the effect of evaporation between cloud base and ground will also be extremely important since, except at short range, the radar beam will be sampling over an appreciable height interval and, normally, considerably above the earth's surface.

† For example, M. Fujiwara, *J. Atmos. Sci.*, **22**, No. 5, pp. 585–91, 1965.

‡ Meteorological Office Scientific Paper No. 21, H.M.S.O., 1965.

STANDARD FREQUENCY TRANSMISSIONS

(Communication from the National Physical Laboratory)

Deviations, in parts in 10^{10}, from nominal frequency for June 1966

June 1966	24-hour mean centred on 0300 U.T.			June 1966	24-hour mean centred on 0300 U.T.		
	GBZ 19·6 kHz	MSF 60 kHz	Droitwich 200 kHz		GBZ 19·6 kHz	MSF 60 kHz	Droitwich 200 kHz
1	— 302·2	— 301·2	+ 0·8	16	— 298·3	— 300·6	— 2·3
2	— 301·3	— 300·4	+ 0·5	17	— 299·3	— 300·8	— 2·2
3	— 300·3	— 300·6	+ 0·3	18	— 298·5	—	— 1·8
4	— 300·0	— 300·7	— 0·1	19	— 300·2	—	— 1·0
5	— 299·7	— 300·6	— 0·2	20	— 300·2	— 299·8	+ 0·1
6	— 300·4	— 301·2	— 0·5	21	— 301·3	— 299·9	— 0·3
7	— 300·0	— 300·4	— 0·9	22	— 301·0	— 300·2	— 0·3
8	— 300·4	— 300·7	— 0·5	23	— 300·4	— 300·8	— 0·5
9	— 299·9	— 301·4	— 0·9	24	— 301·2	— 300·8	— 0·1
10	— 301·3	— 300·0	— 1·3	25	— 300·3	— 300·8	+ 0·2
11	— 300·7	— 300·5	— 1·8	26	— 301·6	— 300·9	0
12	— 300·8	— 301·2	— 2·1	27	— 301·1	— 301·5	— 0·2
13	— 300·1	— 301·2	— 2·1	28	— 300·4	— 300·3	+ 0·1
14	— 301·6	— 300·4	— 2·4	29	— 301·0	— 300·0	+ 1·4
15	— 299·6	— 300·4	— 2·5	30	— 300·2	— 300·3	+ 0·6

Nominal frequency corresponds to a value of 9 192 631 770·0 Hz for the caesium $F,m'_F(4,0)$–F,m (3,0) transition at zero field.

Note: GBR Rugby has been replaced temporarily by GBZ Criggion

U.D.C. 621.372.832.8

A Ferrite Ring Stripline Junction Circulator

By

J. HELSZAJN, M.S.E.E., C.Eng.
(*Associate Member*)†

Summary: An essential feature of Bosma's theory of the three-port stripline junction circulator is the smallness of the splitting of a degenerate pair of normal modes. This results in only a small asymmetric distortion of the isotropic field configuration in the disks. A new junction geometry is described which makes use of a garnet ring around a ceramic disk instead of the conventional garnet disk shape. When the degenerate modes of the two isotropic configurations are identical, the circulation adjustment for each geometry is nearly the same. This new circulator geometry is characterized by a low insertion loss and a large thermal capacity.

1. Introduction

One of the most widely used ferrite devices is the three-port junction circulator. The most general form of the junction consists of a symmetrical distribution of magnetized ferrite material at the junction of three transmission lines. The waveguide junction has been described by Chait.[1] A stripline version was subsequently described by Milano *et al.*[2] A three-port junction circulator is shown schematically in Fig. 1.

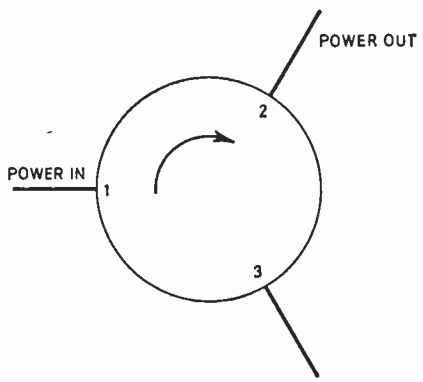


Fig. 1. Schematic of junction circulator.

In this device, power entering port 1 emerges from port 2, and so on in a cyclic manner. In a common application, this allows a single antenna to be used for both transmission and reception. The direction of circulation is reversed when the direction of the ferrite biasing magnetic field is reversed. This makes the junction circulator useful as a prototype switching element. The most important theoretical contributions to the understanding of the symmetrical junction circulator were made by Auld[3] and Bosma.[4] Auld considered the theory in terms of the scattering matrix of the device. An important property of the device is that a perfect circulator is obtained when the three ports are matched. For a three-port junction this requires two independent variables. Bosma analysed the junction in terms of the electromagnetic field to obtain the resonances of the junction under certain simplifying conditions. Theoretical contributions have also been made by Skomal,[5] Fay and Comstock,[6] and Weiss.[7] Butterweck[8] contributed to the description of the waveguide junction theory. Recent work has also included a four-port circulator consisting of a magnetized ferrite at the junction of four transmission lines.[9] Five-port single junctions are also currently under development. At low frequencies lumped constant LC elements have been used to adjust the resonances of the junction thereby reducing the ferrite volume of the junction.[10]

In this paper we will describe a new junction geometry which has a low insertion loss and a large thermal capacity. This new geometry makes use of a garnet ring around a ceramic disk instead of the usual garnet disk shape.

By adjusting the admittance of the junction, a circulator adjustment exists which requires no external matching. This results in an extremely compact circulator.

2. The Ring Geometry

The average power handling capabilities of ferrite devices is determined by the general problem of heat dissipation within the ferrite material. When the power level is high, the total heating effect may be considerable. The resultant temperature rise leads to a decrease in the saturation magnetization and to changes in the linewidth, which result in detuning effects. Temperature stabilized materials with high curie temperatures are therefore recommended.

This new geometry which makes use of a garnet ring around a ceramic disk is shown in Fig. 2. Another possibility is to use an array of posts around the periphery of a ceramic disk as shown in Fig. 3. An essential feature of Bosma's theory is that the splitting of the degenerate resonant modes need only be small. This results in only a small asymmetric distortion of

† Microwave Associates Inc., Burlington, Mass, U.S.A.

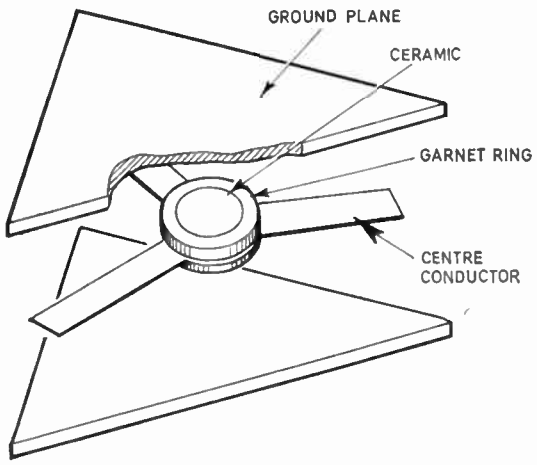


Fig. 2. Ceramic disk with garnet ring.

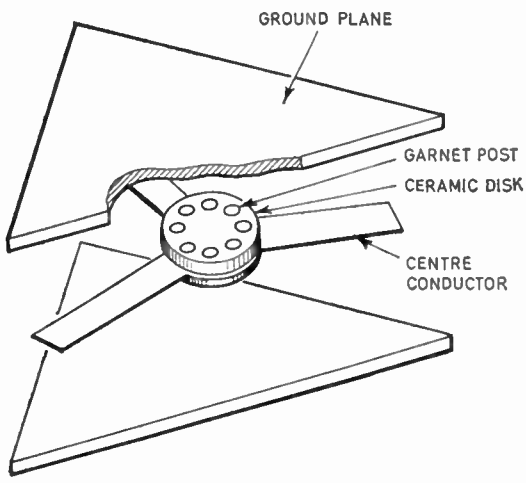


Fig. 3. Ceramic disk with garnet posts.

the isotropic field configuration. By replacing the bulk of the garnet material by a low loss ceramic with the same dielectric constant the magnetic losses associated with the ferrite material are minimized. This is particularly important at u.h.f. The use of such a ceramic insert will therefore reduce the insertion loss and hence minimize the heat dissipation within the device. Because the thermal conductivity of garnet materials is relatively poor compared to ceramic materials the ceramic insert behaves as a heat sink resulting in radial cooling of the garnet material at high average powers. The thermal capacity of the junction is also increased because most of the garnet material has been replaced by the ceramic. In particular, the thermal conductivity of WESCO AL-995 ceramic is 70×10^{-3} cal cm^{-1} s^{-1} and that of beryllia oxide is 525×10^{-3} cal cm^{-1} s^{-1}. This must be compared with the thermal conductivity of ferrites which is about 5×10^{-3} cal cm^{-1} s^{-1}.

We also note that detuning effects, in terms of the resonances of the junction, due to changes in the properties of the garnet material are minimized with such a junction. This is as a result of the stabilizing effect of the ceramic disk.

Good agreement between the experimental work given in this paper and Bosma's theory is obtained.

3. Resonances of Junction Circulators

The stripline junction circulator consists of two ferrite disks placed between the two ground planes and the centre-conductor of the stripline. The ferrite disks are magnetized perpendicularly to the plane of the conductors by a static magnetic field. At equal distances around the edge of the centre conductor, three inner conductors are connected. The theory of the junction is now well understood. Bosma has described the operation of the stripline junction circulator in terms of the electromagnetic field. He assumed the electric field to be perpendicular to the conductors and the magnetic field to be parallel to them and perpendicular to the direction of propagation. By applying the proper boundary conditions required for circulation at the three ports under simplifying conditions Bosma obtained the modes of the junction disks. The dominant mode for such a disk is the dipolar mode, shown in Fig. 4(a) when the disks are excited at the input port and the ferrites are not magnetized. If the standing wave pattern is rotated as shown in Fig. 4(b), then port 3 is situated at

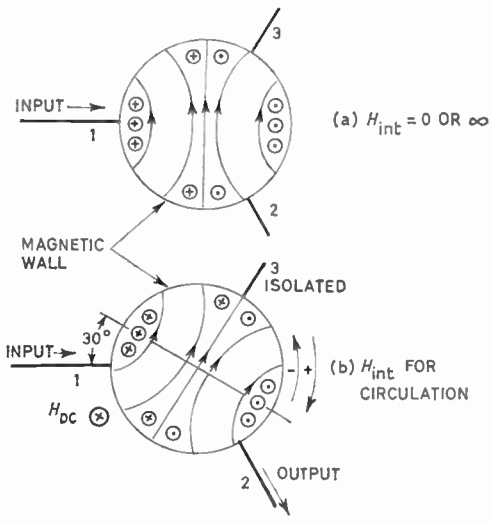


Fig. 4.

(a) Dipolar mode of a dielectric disk. $H_{int} = 0$ or ∞

(b) Analogous pattern of a magnetized disk. H_{int} for circulation. The pattern for the magnetized disk has been rotated to isolate port 3. The magnetic fields of higher modes than the dipolar mode are needed to achieve the pattern shown in (b).

a null of the electric field and is isolated. The last figure is reproduced from Fay's paper. We note that the phase relation between ports 1 and 2 is 180 deg. Circulation is then described in terms of standing wave within the ferrite disks due to the interference of a pair of resonances with field patterns rotating in opposite directions. The resonant frequencies for the uncoupled disks with magnetic short circuits are given by the roots of

$$J_{n-1}(kR) - \left(\frac{nJ_n(kR)}{kR}\right).\left(1 \pm \frac{K}{\mu}\right) = 0$$

where μ and K are the relative on and off diagonal components of the Polder tensor,[11] n is any positive or negative integer. The relative Polder tensor permeability is given by

$$[\mu] = \begin{bmatrix} \mu & -jK & 0 \\ jK & \mu & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

where

$$\mu = 1 - \frac{P\sigma}{1-\sigma^2}$$

$$K = \frac{P}{1-\sigma^2}$$

and

$$\sigma = \frac{|\gamma|H_i}{\omega}$$

$$P = \frac{|\gamma|M_s}{\omega}$$

In the above $|\gamma|$ is $2{\cdot}21 \times 10^5$ rad/s/At/m, M_s is the saturation magnetization of the material, H_i is the internal biasing magnetic field within the ferrite material and ω is the microwave frequency in rad/s. We note in passing that ferromagnetic resonance occurs when $\sigma = 1$ and that the material is saturated when $\sigma \geqslant 0$. The effective propagation constant is

$$k = \omega\sqrt{\mu_{eff}\mu_0\varepsilon_r\varepsilon_0} = \frac{2\pi}{\lambda_0}\sqrt{\mu_{eff}\varepsilon_r}$$

where

$$\mu_{eff} = \frac{\mu^2 - K^2}{\mu}$$

and ε_r is the relative dielectric constant.

When $K/\mu = 0$ the two resonances are degenerate. When the ferrite disks are magnetized the degeneracy is split, giving different values of kR for the two resonances. We also note that the bandwidth is dependent on the degree of splitting and hence on K/μ. The degenerate roots for the dominant pair of resonances are given by:

$$J_0(kR) - \frac{J_1(kR)}{kR} = 0$$

The lowest root is $kR = 1{\cdot}84$, where R is the disk radius. This result coincides very closely with the isotropic disks used in practice for above and below resonance circulators when the proper value for μ_{eff} is used. Because μ_{eff} is largest above resonance, the disk radius for that mode of operation is smaller than for the below resonance biasing condition. This is one of the conditions required for circulation. When this condition is met, the input impedance can be represented according to Butterweck by a lumped constant parallel resonator. The second condition for circulation is obtained by matching the junction to the connector impedance. According to Bosma this imposes a relation between the strip width and the biasing magnetic field. However, if we match the junction to the connector impedance through a quarter-wave transformer no such restriction exists. This is what is normally done in practice.

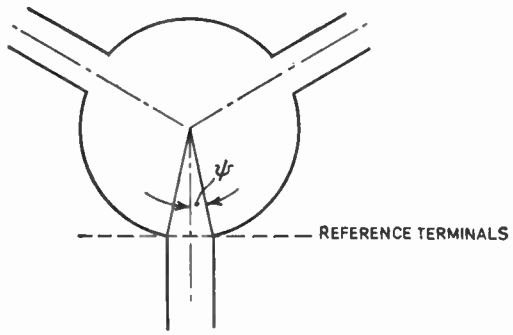


Fig. 5. Reference terminals for admittance of junction.

4. Admittance Measurements

For a three-port junction circulator, the isolation at any port is dependent on the v.s.w.r. of the preceding port provided that the v.s.w.r. is less than about 1·5. When the circulator is matched, we have a perfect circulator. Hence, the experimental adjustment of the stripline junction reduces to an admittance problem which can be best handled with the help of the Smith chart. Once the equivalent circuit of the junction is known at some reference terminals, the circulator can be matched by conventional techniques. Consistant with the resonances of the junction circulator, the reference terminals will be taken at the edge of the ferrite disk where the representation is that of a parallel resonant circuit. The input admittance of the junction at resonance is given approximately by Fay as

$$Y = \frac{Y_{eff}|K/\mu|}{\sin\psi}$$

where

$$Y_{eff} = \sqrt{\frac{\varepsilon_r\varepsilon_0}{\mu_{eff}\mu_0}}$$

and ψ is defined in Fig. 5. The admittance at resonance can be adjusted by varying K/μ or $\sin\psi$. For an arbitrary junction, the v.s.w.r. at port 1 is appreciably modified by multiple reflection between the

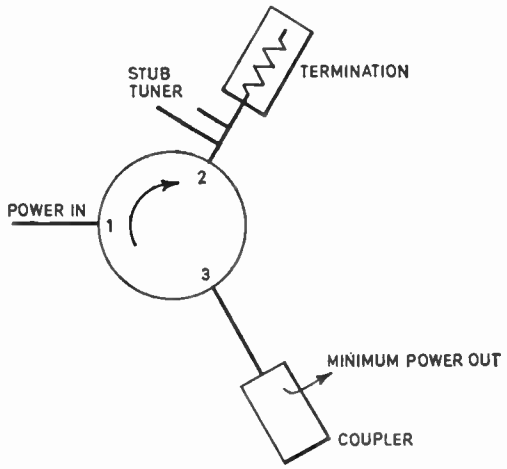


Fig. 6. The use of the double-stub tuner in admittance measurements.

three ports. To overcome this, it is necessary to decouple port 3 from 2 by placing a conjugate match on the latter with the help of a double-stub tuner. This arrangement is shown schematically in Fig. 6. Using this procedure, the junction geometry may then be adjusted for parallel resonance.

5. Experimental Results

In the development of the ring circulator, we calculated the diameter of the garnet disk on the following basis:

Garnet material		Trans Tech G500
M_s	=	45000 At/m
ΔH	=	5650 At/m
kR	=	1·84
μ_{eff}	=	0·75
ε_r	=	14·3
λ_0	=	13·4 cm
External magnetic field	=	18000 At/m
Dielectric constant of ceramic insert ε	=	9

The admittance plots in Figs. 7 and 8 were obtained using the technique described in the previous section. The admittance plot in Fig. 7 shows the effect of varying the internal diameter of the garnet ring with the outside diameter fixed at 2·35 cm. In terms of the resonances of the disk we note that these shifted only slightly as was expected. For the thin rings the bandwidth is decreased because K/μ averaged out over the full disk is now reduced. Lastly, the admittance is proportional to K/μ, the data obtained agree

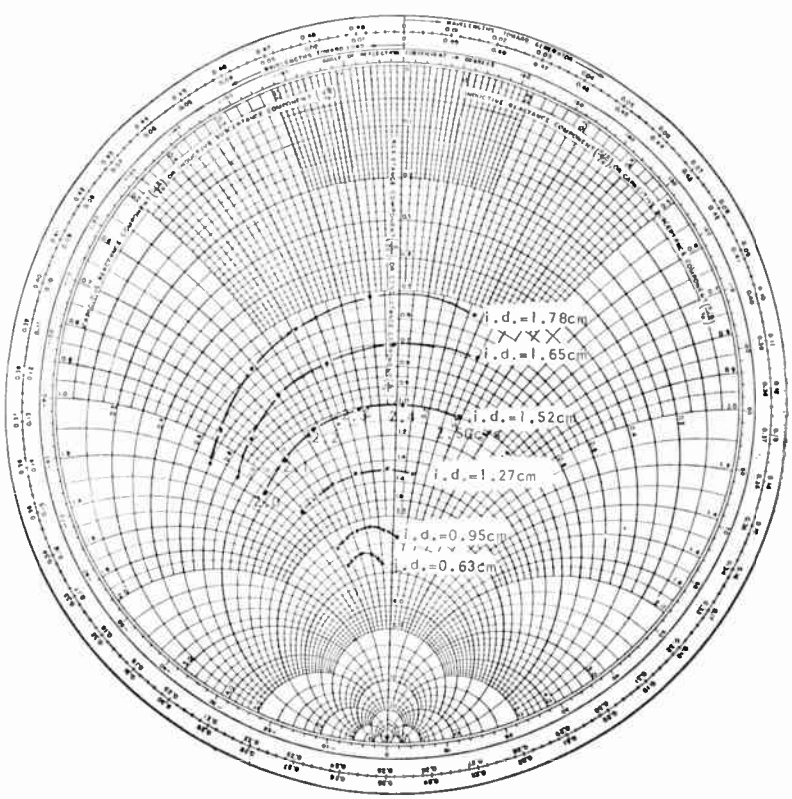


Fig. 7. Admittance of ring circulator for fixed outer diameter (2·35 cm).

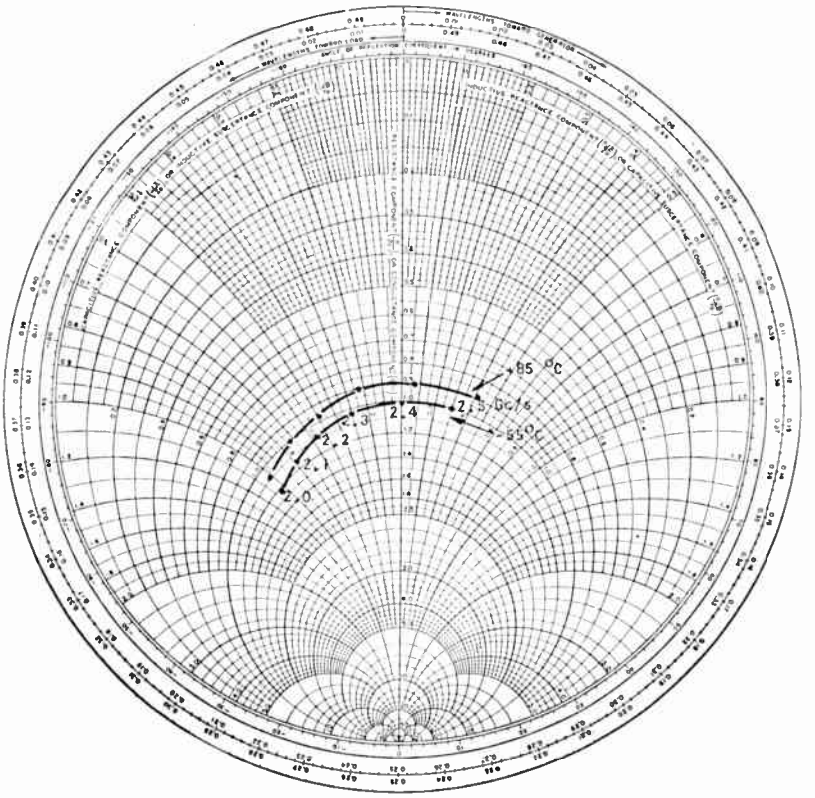


Fig. 8. Temperature behaviour of ring junction.

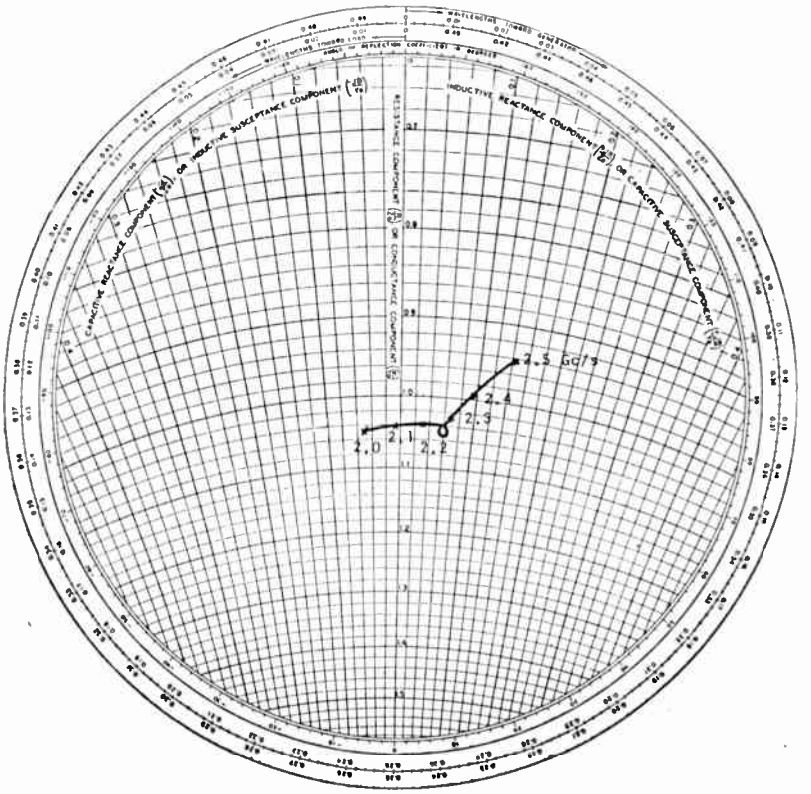


Fig. 9. Admittance of quarter wave coupled junction circulator.

with this result. In Fig. 8, the temperature behaviour of the junction is given at $-55°C$ and at $+85°C$. Good temperature stability results because of the stabilizing effect of the dielectric inserts. For the garnet ring with i.d. of 1·52 cm a circulator adjustment exists which requires no external matching. A compact circulator is therefore possible with this arrangement. The insertion loss for this junction was less than 0·15 dB at 2350 MHz. In Fig. 9 the admittance of a quarter wave-coupled junction circulator using a 2·35 cm diameter disk is shown for comparison.

6. Conclusions

This paper has described a new stripline three-port junction circulator which is characterized by low insertion loss and large thermal capacity. For the isotropic case, it should be possible to obtain the degenerate counter-rotating modes when the dielectric constant of the ceramic insert is not the same as that of the garnet material. For the geometry described a circulator adjustment exists which requires no external matching. By loading the garnet ring by a small dielectric constant insert it should be possible to increase the outside diameter of the garnet ring, thereby increasing the thermal capacity further. The agreement between Bosma's theory and the experimental results is good throughout.

7. Acknowledgment

The author wishes to thank Microwave Associates, Inc. for permission to publish this paper.

8. References

1. H. N. Chait and T. R. Curry, 'A new type Y-circulator', *J. Appl. Phys.*, **30**, Suppl., pp. 152S–153S, April 1959.
2. U. Milano, J. H. Saunders and L. Davis, Jr., 'A Y-junction strip-line circulator', *Trans. Inst. Radio Engrs on Microwave Theory and Techniques*, **MTT-8**, pp. 346–51, May 1960.
3. B. A. Auld, 'The synthesis of symmetrical waveguide circulators', *Trans. I.R.E.*, **MTT-7**, pp. 238–46, April 1959.
4. H. Bosma, 'On stripline Y-circulation at u.h.f.', *Trans. Inst. Elect. Electronics Engrs on Microwave Theory and Techniques*, **MTT-12**, pp. 61–72, January 1964.
5. E. N. Skomal, 'Theory of operation of a 3-port Y-junction ferrite circulator', *Trans. I.E.E.E.*, **MTT-11**, pp. 117–23, March 1963.
6. C. E. Fay and R. L. Comstock, 'Operation of the ferrite junction circulator', *Trans. I.E.E.E.*, **MTT-13**, pp. 15–27, January 1965.
7. J. A. Weiss, 'Circulator synthesis', *Trans. I.E.E.E.*, **MTT-13**, pp. 38–44, January 1965.
8. H. J. Butterweck, 'Der Y-Zirkulator', *Arch. Elekt. Übertragung*, **17**, pp. 163–76, March 1963.
9. L. E. Davis and M. D. Coleman, '4-port crossed junction circulator', Digest of the National PTGMTT Meeting, May 1963, p. 171.
10. Y. Konishi, 'Lumped element Y circulator', International Conference on Microwaves, Circuit Theory, and Information Theory, Tokyo, September 1964.
11. D. Polder, 'On the theory of ferromagnetic resonance', *Phil. Mag.*, **40**, pp. 99–115, February 1949.

Manuscript first received by the Institution on 3rd August 1965 and in final form on 7th March 1966. (Paper No. 1057.)

The Use of SI Units†

The Metric System and the Institution's Journal

Summary: The SI system of units is based on the MKSA system with the addition of specifications of units for temperature and luminous intensity. There are also changes to several derived units.

1. Introduction

In May 1965 the British Government expressed the view that industry should move, sector by sector, to the use of the metric system. This move has already begun in some industries, and it points to the need for all concerned to have a sound knowledge of the units of the metric system.

This change to the metric system is at a time when a newly rationalized set of metric units is coming into international use. This is the Système International d'Unités (SI) and the United Kingdom is able to adopt it from the outset of the change forecast by the Government's announcement.

The SI is a rationalized selection of units in the metric system which individually are not new. It involves the use of a unit of force (the newton) which in some sectors may be less well known than the unit of force in the system of metric technical units (the kilogramme-force). The use of SI units instead of metric technical units will have little effect in everyday life or trade. The metre and the kilogramme will still be the units of linear measure and mass, and the litre will still be commonly used as a unit of volume. However, there are important changes involved in scientific and technical work and this article gives information on the units of the metric system, especially those of the Système International, likely to be of most importance to radio and electronic engineers, both as users, readers and authors.

2. Historical Background

The centimetre and the gramme have been basic units of length and mass since the inception of the Metric System as we know it—it was Talleyrand at the time of the French Revolution who first gave it official status. Measurement of other quantities called for a basic unit of time and the adoption of the second for this purpose gave the centimetre-gramme-second system (c.g.s.). Other combinations are possible and in about 1900 practical measurements in metric units began to be based on the metre, the kilogramme and the second (the MKS system). In 1935, the International Electrotechnical Commission (I.E.C.) accepted the recommendation of Professor Giorgi that this system of units of mechanics should be linked with the electro-magnetic units by the adoption of any one of the latter as a fourth basic unit. In 1950 the I.E.C. adopted the ampere, the unit of electrical current, as the fourth basic unit, giving the MKSA (or Giorgi) system.

The tenth Conférénce Générale des Poids et Mesures (CGMP) in 1954 adopted a rationalized and coherent system of units based on the four MKSA units, the degree Kelvin as the unit of temperature and the candela as the unit of luminous intensity. The eleventh CGPM in 1960 formally gave it the full title 'Système International d'Unités' for which the abbreviation is 'SI' in all languages. SI units have been adopted by the International Organization for Standardization (ISO) and I.E.C. and it is expected that they will become the generally accepted metric units throughout the world.

3. Coherent and Non-coherent Systems of Units

A recurring theme in any discussion of a system of units is whether or not it is coherent, and the meaning of this term must first be explained.

A system of units is coherent if the product or quotient of any two unit quantities in the system is the unit of the resultant quantity. For example, in any coherent system unit area results when unit length is multiplied by unit length, unit velocity when unit length is divided by unit time, and unit force when unit mass is multiplied by unit acceleration. Thus in a coherent system of which the foot is the unit of length, the square foot is the unit of area (but the acre is not). Similarly in a coherent system of which the foot, the pound and the second are the units of length, mass and time, the unit of force is the poundal (and not the pound-weight or the pound-force).

Whatever the system of units and whether it be coherent or non-coherent, magnitudes of some physical quantities must be arbitrarily selected and declared to have unit value. These magnitudes form a set of standards and are called basic units. All other units are derived units, related to the basic units by definitions.

4. The International System of Units (SI)

The International System (SI) is a coherent system with six basic units which are listed in Table 1, together with the unit symbols assigned to them. Special names have been adopted for some of the derived SI units, together with special unit symbols,

† This article is based on the B.S.I. publication PD 5686.

as listed in Table 2. For these units, the definitions show the relationships between them and the basic units, and their definitions are therefore given in Table 4.

The SI units that are not listed in Tables 1 or 2 can only be expressed in terms of the units from which they are derived. Examples are listed in Table 3.

Prefixes by means of which the names of multiples and sub-multiples of units are formed and symbols for them have been agreed internationally. Most of these are already in use in radio and electronic engineering. Table 5 shows these multiples and sub-multiples.

The principal departure of the SI from the more familiar form of the metric system is the use of the newton as the practical unit force instead of the kilogramme or the kilogramme-force. Similarly, the units of quantities derived from force, e.g. pressure, work, power, are combinations of other units with newtons instead of with kilogrammes or kilogrammes-force. It has already been stated that one of the advantages of the SI is that it is coherent. Another, and perhaps more significant advantage from a practical standpoint, is that its unit of force, the newton, is independent of gravitational accelerations and is truly universal. Thus the use of the newton as the practical unit of force obviates both the necessity to introduce the conversion factor g and the necessity to take account of the variations of gravity.

Table 1
Basic SI units†

Quantity	Name of unit	Unit symbol
length	metre	m
mass	kilogramme	kg
time	second	s
electric current	ampere	A
thermodynamic temperature	degree Kelvin‡	°K
luminous intensity	candela	cd

† These units are defined in B.S. 3763 : 1964, 'The International System (SI) units'.

‡ Temperature difference is commonly expressed in degrees Celsius instead of degrees Kelvin. But the unit for Celsius and Kelvin scales is the same: 1 degree C = 1 degree K (see Table 4).

Table 2
Some derived SI units having special names

Physical quantity	SI unit	Unit symbol
force	newton	N = kg m/s²
work, energy, quantity of heat	joule	J = N m
power	watt	W = J/s
electric charge	coulomb	C = A s
electric potential	volt	V = W/A
electric capacitance	farad	F = A s/V
electric resistance	ohm	Ω = V/A
frequency	hertz	Hz = s^{-1}
magnetic flux	weber	Wb = Vs
magnetic flux density	tesla	T = Wb/m²
inductance	henry	H = V s/A
luminous flux	lumen	lm = cd sr§
illumination	lux	lx = lm/m²

§ For definition of steradian (sr), see footnote to Table 4.

Table 3

Some derived SI units with complex names

Physical quantity	SI unit	Unit symbol
area	square metre	m^2
volume	cubic metre	m^3
density (mass density)	kilogramme per cubic metre	kg/m^3
velocity	metre per second	m/s
angular velocity	radian per second	rad/s
acceleration	metre per second squared	m/s^2
angular acceleration	radian per second squared	rad/s^2
pressure	newton per square metre	N/m^2
diffusion coefficient	metre squared per second	m^2/s
thermal conductivity	watt per metre degree Kelvin	W/(m degK)
electric field strength	volt per metre	V/m
magnetic field strength	ampere per metre	A/m
luminance	candela per square metre	cd/m^2

Table 4

Definitions of derived SI units having special names

force	The unit of force called the newton is that force which, when applied to a body having a mass of one kilogramme, gives it an acceleration of one metre per second squared.
energy	The unit of energy called the joule is the work done when the point of application of a force of one newton is displaced through a distance of one metre in the direction of the force.
power	The unit of power called the watt is equal to one joule per second.
electric charge	The unit of electric charge called the coulomb is the quantity of electricity transported in one second by a current of one ampere.
electric potential	The unit of electric potential called the volt is the difference of potential between two points of a conducting wire carrying a constant current of one ampere, when the power dissipated between these points is equal to one watt.
electric capacitanee	The unit of electric capacitance called the farad is the capacitance of a capacitor between the plates of which there appears a difference of potential of one volt when it is charged by a quantity of electricity equal to one coulomb.
electric resistance	The unit of electric resistance called the ohm is the resistance between two points of a conductor when a constant difference of potential of one volt, applied between these two points, produces in this conductor a current of one ampere, this conductor not being the source of any electromotive force.
frequency	The unit of frequency called the hertz is the frequency of a periodic phenomenon of which the periodic time is one second.
magnetic flux	The unit of magnetic flux called the weber is the flux which, linking a circuit of one turn produces in it an electromotive force of one volt as it is reduced to zero at a uniform rate in one second.
magnetic flux density	The unit of magnetic flux density called the tesla is the density of one weber of magnetic flux per square metre.
electric inductance	The unit of electric inductance called the henry is the inductance of a closed circuit in which an electromotive force of one volt is produced when the electric current in the circuit varies uniformly at the rate of one ampere per second.
temperature	The units of Kelvin and Celsius temperature interval are identical. A temperature expressed in degrees Celsius is equal to the temperature expressed in degrees Kelvin less 273·15.†
luminous flux	The unit of luminous flux called the lumen is the flux emitted within unit solid angle of one steradian‡ by a point source having a uniform intensity of one candela.
illumination	The unit of illumination called the lux is an illumination of one lumen per sq. metre.

† This is true for the thermodynamic scale and for the international practical scale of 1948. There are, however, slight differences between thermodynamic scales and practical scales.

‡ One steradian is the solid angle which, having its vertex at the centre of a sphere, cuts off an area of the surface of the sphere equal to that of a square with sides of length equal to the radius of the sphere.

Table 5

Multiples and sub-multiples of units

The names of the multiples and sub-multiples of the units are formed by means of the following prefixes:

Factor by which the unit is multiplied	Prefix	Symbol
$1\,000\,000\,000\,000 = 10^{12}$	tera	T
$1\,000\,000\,000 = 10^{9}$	giga	G
$1\,000\,000 = 10^{6}$	mega	M
$1\,000 = 10^{3}$	kilo	k
$100 = 10^{2}$	hecto	h
$10 = 10^{1}$	deca	da
$0{\cdot}1 = 10^{-1}$	deci	d
$0{\cdot}01 = 10^{-2}$	centi	c
$0{\cdot}001 = 10^{-3}$	milli	m
$0{\cdot}000\,001 = 10^{-6}$	micro	μ
$0{\cdot}000\,000\,001 = 10^{-9}$	nano	n
$0{\cdot}000\,000\,000\,001 = 10^{-12}$	pico	p
$0{\cdot}000\,000\,000\,000\,001 = 10^{-15}$	femto	f
$0{\cdot}000\,000\,000\,000\,000\,001 = 10^{-18}$	atto	a

5. Implications of the SI System in the Journal

From the point of view of the radio and electronic engineer the most noticeable change will be the adoption of a specially-named unit for frequency in place of the long-established 'complex name'. From July 1966 all papers and articles published in the Institution's *Journal* and *Proceedings* will use the hertz (symbol Hz) for the cycle per second (c/s).

Because it has been the practice over recent years generally to follow the MKS system, it will only be necessary to introduce a few other changes. The introduction of the newton, mentioned above, implies the adoption of newtons per square metre (N/m^2) for the measurement of pressure: atmospheres, bars (and millibars), pieze, millimetres of mercury and torrs have all been used according to the particular preference of the discipline or even of the individual. The unit of magnetic flux density, the tesla (T) representing webers per square metre, and its sub-multiples will replace the more familiar c.g.s. unit, the gauss, involving a conversion factor of 10^{-4}. (The 12 000 gauss loudspeaker field will thus become 1·2 tesla while the horizontal intensity of the earth's magnetic field will henceforth be expressed as 18 μT.)

The Institution's Programme and Papers Committee is encouraging authors of papers to employ metric measurements for all quantities and dimensions and in particular for those covered by Tables 1, 2 and 3. It is of course appreciated that, where materials supplied by industry to British units are referred to, it may well introduce a degree of obscurity to translate, for instance, the dimensions of a $\frac{1}{4}$-in bar to read 6·35 mm. This has been a problem for a long time and will not be solved until British Industry has gone over completely to metric measurements. In the meantime, equivalents of this kind of measurement will normally be shown in parentheses whenever it is considered desirable for the guidance of readers to whom the U.K. units would be unfamiliar. Similarly, the former metric units from the MKS and c.g.s. systems will be shown alongside the less familiar S.I. units over an interim period.

6. Bibliography

Some British Standards and other Publications concerned with units and conversion of units

B.S. 350: Conversion factors and tables.

Part 1: 1959 Basis of tables. Conversion factors (*with Amendments* April and Oct., 1963, and Jan. 1965). 15s.

Part 2: 1962 Detailed conversion tables. 25s.

B.S. 1637: 1950 Memorandum on the M.K.S. System of electrical and magnetic units. 3s.

B.S. 1957: 1953 Presentation of numerical values (fineness of expression; rounding of numbers). 3s. 6d.

B.S. 1991: Letter symbols, signs and abbreviations.

Part 1: 1954 General (*with Amendments* July, 1955, Feb., 1957 and Oct., 1960). 7s.

Part 6: 1963 Electrical science and engineering. 12s. 6d.

[Note.—Other Parts of B.S. 1991 cover chemical engineering, nuclear science, fluid mechanics, materials, applied thermodynamics, etc.]

B.S. 2045: 1965 Preferred numbers. 6s.

B.S. 2856: 1957 Precise conversion of inch and metric sizes on engineering drawing. 4s. 6d.

B.S. 2990: 1958 Rationalized and unrationalized formulae in electrical engineering (*with Amendment* PD 3677, Jan., 1960). 4s. 6d.

B.S. 3763: 1964 The International System (SI) units. 6s.

Pamela Anderton and P. H. Bigg, 'Changing to the Metric System: Conversion Factors, Symbols and Definitions', H.M. Stationery Office, London, 1965, revised 1966, 3s. 6d.

P. Vigoureux, 'Electric and magnetic units in the international System', *Bulletin of the Institute of Physics*, **15**, No. 10, p. 241, 1964.

E. Bradshaw, 'Electrical Units: with special reference to the MKS System' (Chapman & Hall, London, 1952).

L. H. Bedford, 'The theory of units', *J. Brit. Instn Radio Engrs*, **3**, No. 3, pp. 82–103, December 1942–January 1943.