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 31 JUNE 1966 NUMBER 6

MEASUREMENTS AND STANDARDS AT RADIO FREQUENCIES

Over five years ago an article in this Journal by the then Chairman of the Institution's Technical Committee, Mr. F. G. Diver, drew attention to the problems which faced the radio and electronic engineer because of the limited nature of the facilities available at the National Physical Laboratory for measurements at frequencies higher than 5 Mc/s.† The vast extension of the usable frequency spectrum calls for measurements of many electrical quantities—voltage, power, attenuation, etc.—but truly authoritative standards have existed only for frequency itself.

This absence of guarantees by an international standards laboratory of the extent to which electronic instruments are within specification has been felt not least by British exporters. The announcement that the Ministry of Technology has undertaken an appraisal of calibration facilities and as a result is establishing a British Calibration Service will therefore be welcomed.

It is intended that the Service will mobilize laboratories in industry, government establishments, research associations, universities, colleges and elsewhere which comply with the high standards to be laid down. It is foreseen that some fifty laboratories will be approved for the scheme in the first year or two, but as the scheme develops this number may increase fairly rapidly to 100 or more. Growth will depend on the support of industry and the success achieved by industry, both in exports and in replacing imports as a result of use of the Service.

The headquarters staff, initially small, will undertake regular inspection of laboratories, provide a central information service, promote development of new methods of measurement, give advice to approved laboratories, and generally encourage education and training in measurement science.

For many years the National Physical Laboratory has been the British custodian of numerous basic international standards of measurement, and certification by N.P.L. has long been a requirement in many branches of science and engineering. Under the new scheme the N.P.L. will remain responsible for the basic international standards of length, mass, time, electrical current, temperature and luminous intensity and about fifty other standards derived from them.

Industry will be strongly represented in the development of the Service. An Advisory Council on Calibration Measurement is to be set up and will have members drawn from industry. There will also be specialist Technical Committees, each to advise on a class of measurement, namely engineering metrology, low-frequency measurements, high-frequency and microwave measurements, temperature and volumetric measurements. The British Standards Institution will be closely associated with all this work.

During the past few months the I.E.R.E. Programme and Papers Committee has been drawing up plans for major conferences to be held during 1967. These include a Conference on 'Radio Frequency Measurements and Standards' which will enable engineers to meet and discuss this vital subject at a particularly appropriate time in the early stages of the establishment of the British Calibration Service.

F. W. S.

† 'Radio Frequency Standards', *J.Brit.I.R.E.*, **21**, p. 105, February 1961.

World Radio History

New Zealand National Electronics Conference

AUCKLAND, 15th–19th AUGUST 1966

Sponsored jointly by the New Zealand Section of the I.E.R.E. and the New Zealand Electronics Institute, Inc.

This will be the first Conference of its kind to be held in New Zealand and it is expected that about 60 papers will be presented. Information on the Conference may be obtained from Mr. C. W. Salmon, I.E.R.E. Auckland Office, P.O. Box 3381, Auckland.

The following are abstracts of some of the papers to be presented.

Profile Milling by Programmed Numerical Control

M. J. Beale, S. Lattimore, W. P. Gabriel and R. A. Morris (*D.S.I.R., Physics and Engineering Laboratory, Lower Hutt*).

This paper describes an electronic control system for a two-dimensional vertical milling machine. Information on the desired profile of the finished work piece is placed on eight-hole paper tape in the form of a series of commands and measurements. This information is processed in the control circuits to produce two pulse trains with variable repetition rates. It is the rates of these pulse trains which determine the X and Y velocities of the table of the machine. At the present stage of development the system is capable of machining straight lines and sections of circular arc.

A Daylight Integrator

R. A. Morris (*Physics and Engineering Laboratory, Lower Hutt*) and R. Robotham (*Plant Physiology Division, Palmerston North*).

This paper describes a small portable instrument which indicates the amount of daylight it has received over a period which may be as long as six months. A high vacuum photocell is used to measure the light intensity. The photocell current charges a capacitor which when the voltage across it exceeds a certain value is discharged by a cold cathode trigger tube. Each discharge is counted on an electro-mechanical register. The instrument with the exception of the battery compartment is hermetically sealed. Power is supplied by a 12-volt battery; the high voltages for the trigger tube and photocell are developed by an invertor. Battery drain is less than 1 mA. Some information on stability is presented.

A Portable Alpha Counter for Measuring Very Low Levels of Activity in Food, Tobacco etc.

F. H. Ellis and R. A. Morris (*Physics and Engineering Laboratory, Lower Hutt*).

This paper describes a portable instrument designed to measure the alpha activity of food-stuffs and other materials. It does this by detecting the scintillations produced in a zinc sulphide screen by the alpha particles. The screen is viewed by a 5 in photomultiplier tube and the output pulses are amplified and counted by transistor circuitry. Some information on the source of the particles is obtained from a measurement of the time separation of double pulses. The spurious background count is less than one per hour. The instrument is powered by four lantern batteries which give continuous operation for several days. Normally these are 'floated' across a small mains unit giving an indefinite life.

A Transistor Diode-Feedback Type Logic Circuit

J. B. Earnshaw and P. M. Fenwick (*Physics Department, University of Auckland*).

The paper describes an improved form of saturated transistor logic circuit in which rectifier diodes are used to provide non-linear current feedback from collector to base of the transistor. The resulting arrangement has many advantages over the basic saturated circuit, particularly with respect to speed and logical versatility. Due to the ability of the circuit to handle two levels of logic in each stage, mean delays approaching 1 ns per logical decision can be obtained using non-selected components.

World Radio History

A Spectrographic Receiver for V.L.F. Transmissions

G. L. Jones, R. A. Morris and N. R. Poletti (*Physics and Engineering Laboratory, Lower Hutt*).

This paper describes a multiple filter spectrograph for recording the spectrum of a radio signal in the v.l.f. band. The primary use of the equipment is in recording the Doppler shift of v.l.f. signals received via the whistler mode. The instrument has 25 narrow band filters spaced symmetrically about the centre frequency which may be shifted to coincide with the carrier frequency of any transmitter in the v.l.f. band. The spectrograph has a channel spacing of 0·05 c/s and the bandwidth of each filter is 0·007 c/s. The amplitude of the voltage at the output of each filter is recorded continuously on slow moving 35 mm film. Examples of typical records are given.

A Novel Frequency Synthesizer with Multiple Outputs

I. R. Richards and R. A. Morris (*Physics and Engineering Laboratory, Lower Hutt*).

A digital frequency synthesizer having ten independent outputs is described. The output frequency range of each channel is 10 kc/s to 100 kc/s and is adjustable in steps of 100 c/s. An alternative output having a non-uniform frequency is also available from each channel with a frequency range of 1 c/s to 100 kc/s, adjustable in steps of 1 c/s. A basic timing unit consisting of cascaded binary coded decimal rate multipliers supplies sets of pulses to frequency selector units which select and combine the pulses to provide the non-uniform output. Any number of units may be driven from the timing unit subject to loading considerations. The normal output is obtained from the non-uniform output by the use of an active filter of high Q which reduces the non-uniformity to negligible proportions.

A Tunable Phase Locking V.L.F. Receiver

N. R. Poletti and R. A. Morris (*Physics and Engineering Laboratory, Lower Hutt*).

This paper describes a receiver developed for v.l.f. propagation research. The receiver is tunable over the range of 10 to 30 kc/s in 1 c/s increments. The required frequency is selected by setting five decade switches on the digital frequency synthesizer described in a companion paper. A digital servo with 3·6 ft resolution tracks the phase of the incoming signal. The phase increments are summed in a reversible counter with two alternative outputs. A digital output in binary coded decimal format is suitable for recording by digital data processing equipment, and an analogue voltage output can be used to drive a pen recorder. A logarithmic field strength indication is also available.

A New System for the Digital Setting of Temperature and Humidity Controllers

W. P. Gabriel and R. A. Morris (*Physics and Engineering Laboratory, Lower Hutt*) and R. W. Robotham (*Plant Physiology Division, Palmerston North*).

This paper describes a digital system for programming the temperature and humidity controllers used in an environmental control installation. The originality of the system lies in the binary switching of Wheatstone bridges by transistors. The programming input signal consists of two d.c. levels and a train of timing pulses. The circuit employed can program at least ten controllers if required.

The bridge switching transistors which define a controller set-point are controlled by the states of a reversible counter, which, in turn, is set to add or subtract by the input d.c. levels. Each timing pulse can initiate a new set point. Each set point corresponds to a state of the counter.

Power Control Using Silicon Controlled Rectifiers in an Environmental Control Application with a Switching Capability of 0·5 MW.

W. P. Gabriel and R. A. Morris (*Physics and Engineering Laboratory, Lower Hutt*) and R. W. Robotham (*Plant Physiology Division, Palmerston North*).

Research at the Plant Physiology Division, D.S.I.R., is to be assisted by the production of controlled climates in 25 rooms. The controlled quantities are temperature, humidity, lighting level, and carbon dioxide content.

The use of silicon controlled rectifiers (thyristors) in controlling some of these quantities is described. The large number of control units in the project, and the high total power to be handled, have led to some interesting problems in s.c.r. technology. The application of s.c.r.'s in climate control is described, with some discussion of malfunctions and interference. The performance of pilot plant currently in operation indicates that control by solid-state devices gives reliability and versatility.

World Radio History

A Precise Digital Intervalometer for Aerial Photography

W. P. GABRIEL and R. A. MORRIS (*Physics and Engineering Laboratory, Lower Hutt*).

An intervalometer is a semi-automatic instrument used to compute and generate the exposure sequence of an aerial camera so as to produce the correct overlap between successive frames as required for stereoscopic pairs.

The application of solid-state circuitry to an intervalometer produces better accuracy, versatility and operational comfort than is normally achieved with conventional mechanical units. The system includes a predetermined binary counter, solid state switching, and a closed-loop motor speed regulator.

Digital Voltmeter and Tape Punch for the Automatic Recording of Data

M. J. BEALE and R. A. MORRIS (*Physics and Engineering Laboratory, Lower Hutt*).

This paper describes equipment for the automatic recording on eight-level paper tape of experimental results gathered by a three-decade digital voltmeter. The output tape is suitable for direct assimilation by a digital computer.

Either of two recording codes may be selected. One digit per tape column (computer code) for normal operation and two binary coded decimal digits per tape column for high-speed operation. The latter mode, while sacrificing one significant figure, takes full advantage of the high-speed perforator used.

Read and record times are 27 and 9 ms respectively.

A Portable Integrating Omni-directional Anemometer

R. A. MORRIS (*Physics and Engineering Laboratory, Lower Hutt*) and A. R. MORMAN and D. SCOTT (*Plant Physiology Division, Palmerston North*).

A portable lightweight instrument developed to investigate the effect of total wind flow on plant growth is described. The sensing device is a directly heated thermistor bead. From the cooling effect of the wind on the bead the electronic circuits produce a voltage that varies linearly with respect to speed over the range of one to thirty miles/hour. This voltage is integrated by a d.c. motor with an attached revolution counter to give the wind flow over a time determined by the operator. The reading is sufficiently independent of direction to be useful in the intended application.

Proceedings of the Conference

Although it is hoped to publish a representative selection of the papers presented at this Conference in *The Radio and Electronic Engineer* over the course of the next few months, inquiries about the papers should in the first instance be directed either to the authors concerned or to Mr. Salmon at the address given on page 326.

World Radio History

U.D.C. 621.382.029.64

Microwave Semiconductor Devices

By

Professor G. D. SIMS, M.Sc., Ph.D., C.Eng.†

Reprinted from the Proceedings of the Joint I.E.R.E.–I.E.E. Symposium on 'Microwave Applications of Semiconductors' held in London from 30th June to 2nd July 1965.

Summary: The paper examines some of the classes of microwave devices which are of current interest. Devices described are: frequency multipliers, up-converters, parametric amplifiers, tunnel diodes, backward diodes, hot electron devices and Gunn effect oscillators. Some of the problems and developments which the future may hold are discussed.

1. Introduction

Since the beginnings of the microwave industry, early in the second world war, we have seen not only a remarkable growth in microwave applications but also a startlingly rapid change during the last six years in the active components available to us. To some extent the growth in diversity of the applications has resulted from the new components available and it is worth remembering that in this as in many other branches of electronics it is not always the demand for a component that initiates its development but sometimes it is the existence of a new component which initiates a new application.

We have seen this kind of situation develop, for example, in relation to satellite communication where, if we had not developed the maser—very largely for its own sake—we should not have had this form of communication at such an early date. In a similar way we should not have found ourselves in a position to apply microwaves for cooking, or bulk timber drying, had it not been for the high-power tubes developed for radar purposes. It is also probably true that many of the current tube applications would not have been found had it not been for the impact on the tube industry of semiconductor devices during the last few years.

Despite the general reduction in defence expenditure in the U.S.A. (see Fig. 1) during 1964, estimated sales of microwave equipment range between \$1000M and \$2000M in 1965 of which only some 15% of the market is accounted for by the Department of Defence and of which some 48% is concerned with exports to Europe. This latter market has grown at a rate some 2½ times that in the U.S.A. where the expansion rate in 1963 was still of the order of 8%. It is clear then that the future for microwaves is still bright and it is instructive to examine the major fields of application with a view to assessing the relevance of the new devices which are continually being developed.

Without doubt the communications industry will be the biggest user in the future and even some of the current statistics are quite surprising to the unfamiliar ear: for example in the U.S.A. petroleum pipelines are among the biggest users of microwave com-

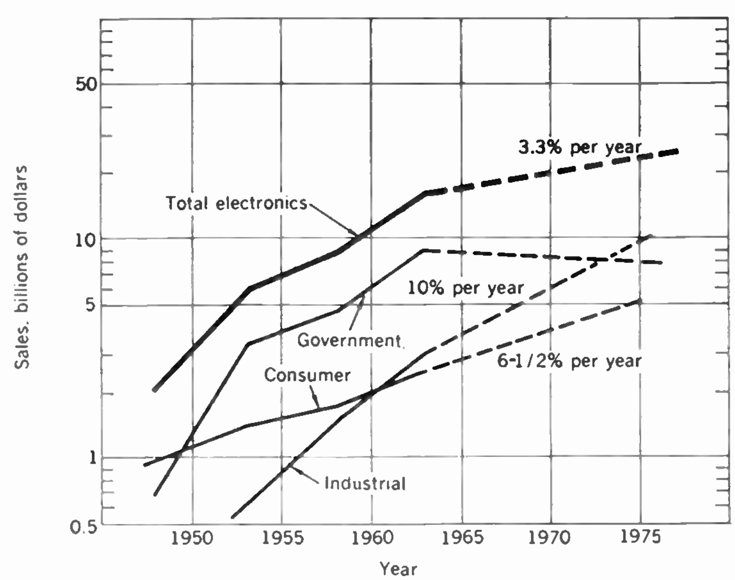


Fig. 1. The solid curves show actual sales of electronic equipment in the U.S.A. as a function of various users. The broken curves represent predictions for the future. The decline in Government expenditure is noteworthy. (After Herold.[1])

† Department of Electronics, University of Southampton.

World Radio History

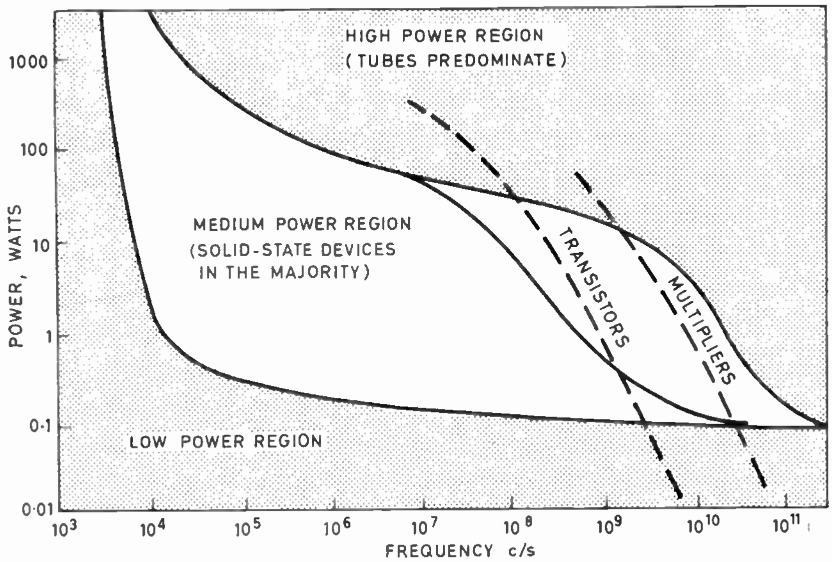


Fig. 2. The three approximate divisions of the power-frequency domain early in 1965. The enclosed area to the right of the medium power region is one where some uncertainty as to the future of the electron tube still exists. (After Hoffins[2] (Fig. 1).)

munications and operate more than 50 000 transmitters among 800 firms. A. T. & T. alone operates of the order of half a million miles of wide band microwave channels and so on. Viewed against a background where the expansion of digital data traffic is expected to increase by a factor of around 1000 over the next ten years and telephone traffic by a factor of 100 it is clear that demand for microwave equipment is unlikely to fall in these areas. There is also much work to be done in replacement of currently installed low capacity equipment by high capacity equipment as well as the provision of more channels, better quality will also be needed to cope with the expansion of colour television systems abroad and ultimately at home too.

Quite clearly the market in navigational aids will continue to prosper although there seems little reason to suppose that there will be any startling increase in the expansion rate unless radar develops a greater range of terrestrial uses.

We have mentioned the rapid expansion which is already occurring in the use of microwaves for drying purposes and in this range we find everything from food and paint, through timber and printer's ink to potato crisps, coming in for treatment.

There are many other minor markets too, for example that in e.s.r. spectrometers, and various millimetre wave equipments which are fast becoming large enough to be commercially tempting.

Finally, we find that the computer industry is approaching the time when 1-nanosecond logic systems are near reality and this will bring computers also within our field (if only just—since the B.S.S. definition makes microwaves those whose frequency exceeds 1 Gc/s).

These applications have been singled out for attention, since between them they comprehend the whole range of microwave techniques and in so doing enable us to see in context and in perspective the consequences of the impact of semiconductors on the microwave industry both now and to some extent in the future.

The active devices necessary to accomplish the tasks outlined above fall into three main categories: (1) high power—such as are required in radar transmitters and materials-processing installations, (2) medium power such as are used in radar receivers and some communications links, and (3) low power such as may be used in future computers.

So far semiconductors have left the first category untouched and it is still the stronghold (though no stronghold is to-day completely unassailable) of the electron tube.

The second category has benefited most, in a variety of ways, for example:

Low noise devices have increased radar range and communications system sensitivity.

Greater reliability is promised (if not always obtained) with solid-state devices.

Lower d.c. powers are needed.

The third area is the one most likely to be of

World Radio History

interest in the computer industry and in some respects constitutes a completely new branch of the subject.

The possibilities currently available[1] are summarized in Fig. 2 which shows the ranges of frequency and power in which semiconductor and thermionic tubes have found their greatest use.

The object of this paper will be to examine some of the classes of device which are of current interest with a view to pin-pointing some of the presently interesting aspects of the field as well as some of the problems and developments which the future may hold.

2. Frequency Multipliers

The varactor doubler has now become a standard piece of microwave hardware but the general quest for improved efficiency and higher frequency continues. However efficient a practical varactor multiplier may be in practice, the spectral Manley-Rowe 100% still hovers above us and although we have made close approaches to this maximum in practice with doublers, efficient higher-order multiplications still defy us. If we wish to make triplers we can, by the provision of a loss-free idler circuit at the second harmonic frequency achieve substantial efficiency and although to go to higher order would only seem to require the provision of further high Q idlers at each of the other intermediate unwanted frequencies involved, the degree of circuit complication involved has so far defeated us.

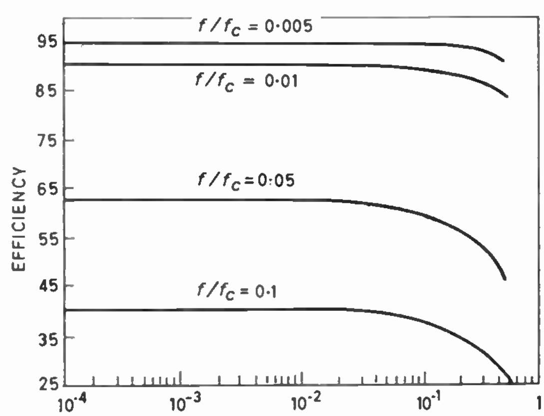


Fig. 3. Variation of varactor doubler efficiency with γ and f_c. (After Grayzel.[6])

Despite this we have seen this year a $\times 11$ multiplier,[2] producing 9900 Mc/s from a 900 Mc/s input with $1\frac{1}{2}$% efficiency (saturating with an input power of 150 mW). For high order multiplication however, the step recovery diode[3] still offers the best prospects. In this design a profiled field in the neighbourhood of the barrier confines injected minority carriers to the barrier region thus ensuring the maximum possible switching speed (and consequent non-linearity) on going into the reverse biased condition. Driven with a sinusoidal voltage the step recovery diode produces harmonics whose amplitude decreases as $1/n$ where n is the harmonic order, as compared with the varactor for which amplitudes decrease as $1/n^2$. A $\times 20$ multiplier with 100 Mc/s input gave an output of 15 mW at 2 Gc/s at 30% efficiency or as much as 80 mW with 10% efficiency.[4] The device is a saturable one therefore giving a stable output and will stand reverse voltages of 100 V. The present record in terms of frequency still seems to be the 420 Gc/s generated and detected as the 6th harmonic of a 70 Gc/s input signal.[5]

As regards the mechanism of the harmonic generation process, the question of the role of charge storage is to be given detailed consideration later in the conference,† and will not be dwelt upon further here except to say that substantial differences of opinion still appear to exist.

A recent theoretical analysis by Grayzel[6] suggests that greater doubler efficiency can be achieved by using varactors with small values of γ

$$C = K/(V-V_0)^{\gamma}$$

and the general position is summarized in Figs 3 and 4 which indicate that best results as regards efficiency are achieved with $\gamma \simeq 0{\cdot}1$ or less while power output remains constant at a minimum value for $\gamma \simeq 0{\cdot}1$ and then increases with increasing γ. This

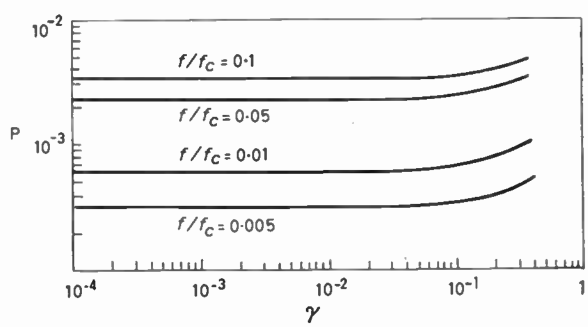


Fig. 4. Variation of varactor doubler output with γ and f_c. (After Grayzel.[6])

situation could have important implications in multiplier chains where in the earlier stages low γ varactors would seem to be indicated, but where output power was the only consideration a higher γ might be worth considering.

† B. C. Heap, 'An investigation into the effects of charge storage on the efficiency of a varactor diode doubler', *The Radio and Electronic Engineer*, **31,** No. 4. pp. 225–33, April 1966.

World Radio History

It is yet possible that some of the experimental hot carrier devices may prove to be important at the highest frequencies and the hot carrier thermoelectric device which would seem to come into its own at frequencies in excess of the reciprocal of the recombination time ($\simeq 10^{-11}$ s for Ge) may still have some possibilities, although a doubling of 71·5 to 143 Gc/s with 21·5 mW in and 1 μW out is scarcely encouraging.[7]

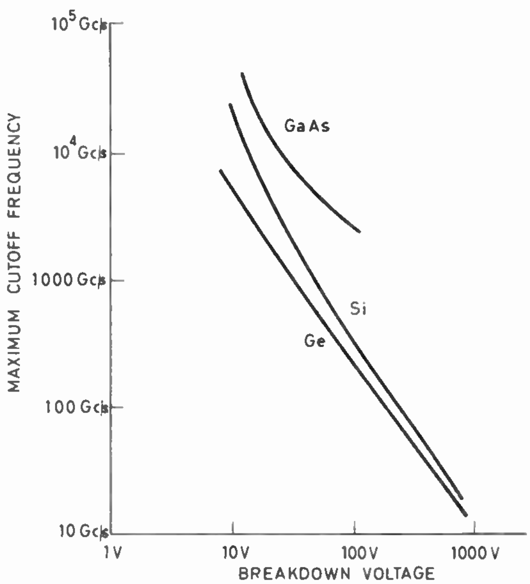


Fig. 5. Maximum predicted cut-off frequency for a given breakdown voltage of varactors in various n-type materials. (After Penfield.[8])

Reference to Grayzel's curves also seems to underline the importance of cut-off frequency of the varactors used and here it is worth noting a recent communication by Penfield.[8] He shows that the maximum cut-off frequency which can be achieved with varactors is given by

$$f_c \simeq 1/\pi\varepsilon\rho$$

where ε and ρ are the bulk permittivity and resistivity of the material and is thus strongly dependent on the dielectric relaxation time $\rho\varepsilon$ of the material used (see Fig. 5). He points out that if the relaxation time and hence f_c is improved by decreasing the value of ρ, this gives a narrower depletion region for a given breakdown field (because the doping density is increased) and hence to some extent there is a compromise between f_c and V_B.

In practice this upper cut-off limit is closely approached in high voltage silicon epitaxial devices although gallium arsenide offers greater promise but is not yet fully exploited.

A further point which has also received recent comment is that of high-level spurious signals in multipliers which seem not to be attributable to up-conversion. Burckhardt[9] suggests that oscillations in the mean value of the varactor capacitance actuated by the input carrier amplitude variation (in the case of an a.m. wave) provide a possible gain mechanism which will exist in the case of all, but abrupt junction, varactors.

In the meantime[10] a report of an X-band all-solid-state parametric amplifier pumped by a varactor chain at 30 Gc/s has appeared and this will be the forerunner of many such devices, the compactness of which will improve rapidly as the power and frequency of transistor driver stages increase.

3. Parametric Amplifiers, Up-converters and Materials

With techniques in this field now largely stabilized we find ourselves still without a really satisfactory answer to the problem of replacement of the circulator in devices where bandwidth or tunability is important. Noise figures for cooled amplifiers seem to be fairly much as they were a year ago and although more consistent results are being achieved with cooled devices[11] some problems still remain. For example, it is known that with silicon varactors the series

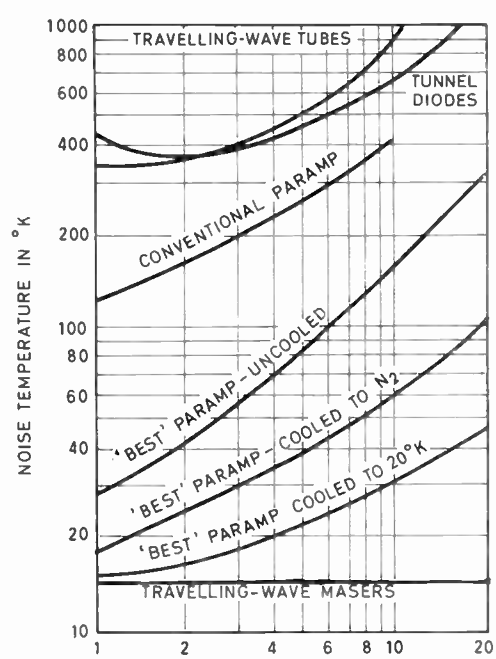


Fig. 6. A comparison of various low-noise devices as a function of frequency. (The 'conventional' parametric amplifier is one in which the idler circuit is provided by self-resonance of the varactor.) (From *Microwave Journal*, July 1964.)

World Radio History

resistance tends to increase at low temperatures whereas for gallium arsenide it does not. Thus the potential of cooled gallium arsenide devices is substantially greater than that of silicon devices. In devices to handle high power a high voltage swing with a large diffusion capacitance in the forward bias direction is required and here silicon possesses some advantage over gallium arsenide. Thus at the moment there is some degree of conflict between low noise and high power requirements. Clearly at the highest frequencies gallium arsenide remains the best material since, as is well known, the cut-off frequency

$$f_c = \frac{1}{2\pi R_s C} \propto \mu\varepsilon^{-1/2}$$

and μ, the mobility, is at least five or six times larger for gallium arsenide than it is for silicon. This again has to be remembered when considering low noise, since as well as requiring low temperature, the achievement of low noise also requires a high value of f_c. Figure 6 gives an approximate indication of the comparative performance of various low noise devices as a function of frequency as things stood at the end of 1964.

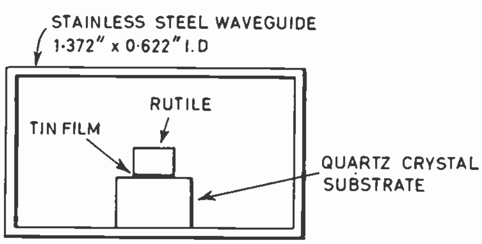


Fig. 7. Clorefine's modified dielectric resonator mounted in waveguide.[12]

No review of parametric devices is complete at the present time without the mention of Clorefine's[12] superconducting device—shown in Fig. 7—which could lay claim to be the most promising variable inductance device—though some are still inclined to dispute whether its operation mechanism is as suggested by its inventor.[13] The simplicity of its construction—a simple tin film ~ 250 Å thick deposited on a quartz crystal substrate and surmounted by a rutile slab mounted in waveguide—lends itself to cheap production, and the fact that it uses bulk material properties and is not dependent on minority carrier effects makes it of great potential interest at the highest frequencies. The results reported nearly a year ago showed 11 dB gain at 6·06 Gc/s operated in a doubly degenerate mode. With a pump power of 0·2 μW—an order of magnitude less than for a comparable varactor device—satisfactory frequency doubling was achieved. Further, frequency shifts at 54 Gc/s were recorded in a resonator containing a similar film. It appears that the main limitation may be in the low saturation level but the potential at the highest frequencies is substantial.

4. Tunnel Diodes and Backward Diodes

In many ways progress in this sector has been slow although this may be a harsh judgment, explicable by the argument that many of us have never quite overcome the feeling of anticlimax which followed the great expectations which the tunnel diode aroused when it was first shown to work.

Burrus[14] has produced point contact diodes which have oscillated at 103 Gc/s and have produced 200 μW at 50 Gc/s. A commercial solid-state counter —counting up to 6·4 Gc/s at a cost of roughly one dollar per megacycle has appeared and moreover the University of Michigan has an 85-ft radio telescope employing a tunnel diode head amplifier at 7·5–8·5 Gc/s which provides 10·6 dB gain at 5 dB noise figure. This latter constitutes a substantial 'seal of approval' and the fact that plots have been made of the radio star 3C295 which produces a rise in aerial temperature of only 0·15°K against an overall noise figure at the centre of the dish of 1300°K, is a significant comment on the stability of the system. The only servicing required is the occasional replacement of the mercury cell which provides the bias voltage on the diode. Transistor stabilization of tunnel diode oscillators has been used elsewhere and the transistor/tunnel diode circuit in this form would seem to offer another good case for the application of integrated circuit techniques.

The last year has also seen the emergence of multi-stage tunnel diode circuits and the curves shown[15] in Fig. 8 show how use has been made of the greater dynamic range of a gallium arsenide diode coupled to a low-noise germanium diode input stage to produce an amplifier with high dynamic range and an overall noise figure substantially the same as that of the input stage.

The development of strip-line tunnel diode and transistor circuits has proceeded rapidly and we are now quite accustomed to compact oscillators and amplifiers produced in this form. The main point of making this comment at this point is to suggest that much more effort could usefully be devoted to the circuit side. For too long has the packaging of all devices in most parts of the spectrum been a cinderella operation and the disk-packaged tunnel diodes and transistors now coming into use represent a sensible step in the right direction.

Substantial amounts of work have been carried out on mixing applications of both backward and tunnel

World Radio History

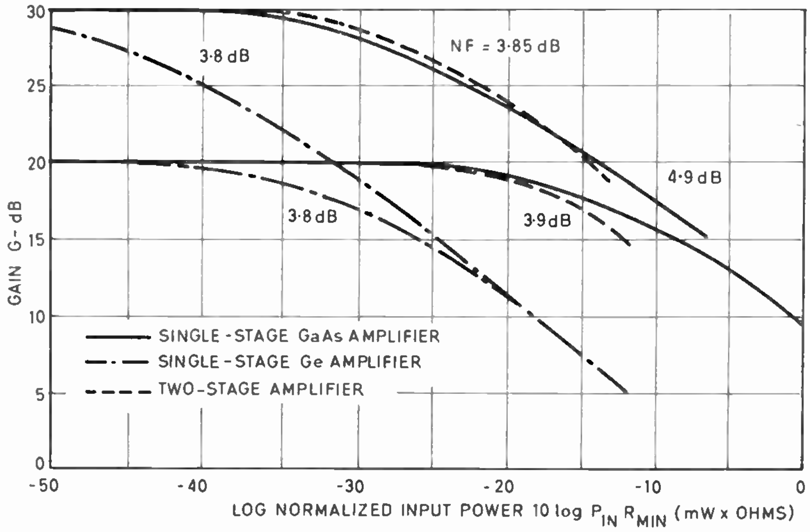


Fig. 8. Performance of single stage and two stage tunnel diode amplifiers. The upper broken curve in each case shows that with a combination of a low noise Ge input stage and a large range GaAs output stage, performance combining the best features of either type of device can be achieved. (From *Microwave Journal*, July 1964.)

diodes.[16] The restrictions on dynamic range and inherent instability of the latter are an obvious limitation but the very low local oscillator powers required, coupled to the temperature independence of the characteristics of tunnel devices, give them an edge on conventional mixers even at the present time. This is of course still only true at the lower microwave frequencies but one suspects that it is only a matter of time before it will be generally true also.

It is remarkable how a device will respond to competition and the fact that we are to hear papers quoting 5 dB noise figures for conventional mixers—which were stuck for so long in the 8–10 dB region—is one evidence of this and is a tribute to the tenacity of the conventional mixer designers. That we should have also seen this month the announcement[17] of a 2·7 dB noise figure for a travelling-wave tube at 12 Gc/s and a 3 dB crossed field amplifier[18] also shows that electron tubes are not going to lie down without a fight!

5. Hot Electron Devices

As the use of the microwave region for ordinary circuit applications grows, so will the need for really fast low-power switching devices. We have already discussed the inherent limitations of minority carrier diodes for this purpose and although the problem of charge storage can to some extent be overcome by building in a minority carrier confining field in the vicinity of the barrier—as in the 'step recovery diode'—it would seem unlikely that any major improvement will result in devices of this type over and above what can currently be achieved.

Attention has been turned for some time in the direction of various 'hot carrier' devices of one kind and another of which the 'hot electron' diodes are the most highly developed so far. Since these devices depend entirely on the motion of majority carriers, charge storage is no longer a problem and they will be subject only to the same kinds of frequency limitation which apply to tunnel diodes, i.e. mainly considerations of geometry.

The hot electron diode is in fact little more than a sophisticated form of point contact diode, comprising a metal film deposited (usually by evaporation or sputtering) on to a thin epitaxial layer of semiconductor. Clearly we require very pure semiconductor materials, careful surface cleaning and uniformity and adequate passivation techniques and most diodes so far have thus been made using n-type silicon.

The Fermi level in the semiconductor exceeds that in the metal and hence under forward bias conditions electrons are injected over the potential barrier (which is thus of the old-fashioned Schottky type) into the metal. These electrons are 'hot' in the sense that their energy greatly exceeds that of those in equilibrium in the metal although once into the metal they quickly achieve equilibrium through interaction with the metal lattice and electrons. In the reverse bias condition electrons cannot surmount the barrier and thus an ideal 'un-charge-storage-limited' characteristic should be achievable.

In that a wide variety of metals can be used for the contact material (e.g. Ag, Pt, Pa, Au have been used[19])

World Radio History

the barrier height can be varied between about 0·3 and 0·8 eV while at the same time the reverse saturation current may lie between 10^{-11} and 10^{-3} A.

The method of construction is such that full advantage may be taken of modern microelectronics techniques and capacitances $\ll 1$ pF are easy to achieve consistent with currents of 20–30 mA, a breakdown voltage of 25 V and saturation current of 1 nA. Switching times of 50 picoseconds have been reported.

As detectors hot electron diodes offer larger stable contact areas than do point contact diodes and thus can handle higher powers at the lower microwave frequencies. However tangential sensitivities so far are at best comparable with those of point contact devices and loss and noise factors are substantially worse. As is often the case this is a matter of fabrication technology and, in view of the promise of these devices, will be overcome in time. Diodes used in mixer applications have already shown conversion losses of about 6 dB at 2 Gc/s with an overall noise figure of about 6 dB which is distinctly promising.[20]

Consideration of the hot electron diode leads naturally into a discussion of majority carrier transistors enabling us to take a look *en passant* at the present state of power transistor development. This is in order at this meeting since not only are transistors fast invading the 'B.S.S. microwave region' but high power transistors are rapidly becoming available in the 'few hundred Mc/s' region enabling us greatly to improve the compactness and simplicity of harmonic generators generally.

The most commonly employed figure of merit for a high frequency transistor is constituted by the expression (see, e.g. Ref. 21),

$$f_{\max} = \left(\frac{f_T}{8\pi^2 r_{bb'} C_c}\right)^{1/2}$$

where $r_{bb'}$ represents the base resistance and C_c the collector capacitance. An alternative and slightly more informative criterion is

$$f_{\max} = \frac{1}{4\pi}\left(\frac{\alpha_0}{r_{bb'} C_c \tau_{ec}}\right)^{1/2}$$

where τ_{ec} is the emitter to collector transit time which is made up of the sum of three terms, namely the emitter charging time, the base transit time and the collector attenuation time. The first expression represents the frequency at which the power gain falls to unity and thus represents a measure of the extreme frequency at which oscillations may be generated. Clearly in order to obtain performance at the highest frequencies we must maximize f_T which is currently ~ 2 Gc/s for silicon and ~ 4 Gc/s for germanium, while reducing $r_{bb'}$, C_c and τ_{ec}.

Reduction of the depletion layer capacitance is largely a matter of improving the transistor geometry (or, crudely, making the device smaller). It should be pointed out that to maximize f_T we must minimize the emitter capacitance and thus in achieving the maximum value of f_T we have already limited the device size substantially. Many other considerations enter also, e.g. the emitter perimeter must be maximized while minimizing the area (see, e.g. Ref. 21). Reduction of $r_{bb'}$—the resistance between the active region of the base and the outside world—is not so simple. It could be reduced by increasing the base thickness—but this would then lead to long carrier transit times in the base region with resultant effects comparable to those experienced with electron tubes. Thus the next best procedure would appear to be to reduce the resistivity of the base material. For the diffusion transistor however, this results in a substantial reduction in the emitter injection efficiency, due to reverse injection from base to emitter. This can to some extent be offset by choosing materials so as to give a wide energy gap between emitter and base but if this is to be the criterion it seems logical to go one step further—maximizing the energy gap, minimizing base resistance, and dispensing with minority carriers in one manoeuvre. This can be achieved in theory in the metal base transistor in which the normal p-type base material of the n-p-n structure is replaced by a thin metal layer.

Rather than one transistor now comprising two p-n diodes back-to-back we have two hot electron diodes back-to-back and two Schottky barriers as shown in Fig. 9. Electrons are injected over the emitter base barrier and collected after traversing the reverse biased base-collector junction, control being effected via the base voltage.

In this device the current will not be diffusion limited, but will obey a thermionic emission law, and in consequence the injection efficiency does not depend on base thickness.

A typical base might be 100 Å thick giving transit times of the order of 10^{-14} s. This does not make a significant difference to the overall problem, however, since the total transit time depends upon the emitter charging time (the ultimate limit to which is set by the dielectric relaxation time of the emitter-base region) and also upon the collector attenuation time, both of which will not differ greatly in either form of transistor. However the base resistance will be two orders of magnitude less than in the n-p-n case giving an improvement of a factor of ten in the case of the upper frequency limit.

In the base region electrons move with velocities of some 5–10 eV and the base must be thin in order to minimize electron/phonon losses which randomize

World Radio History

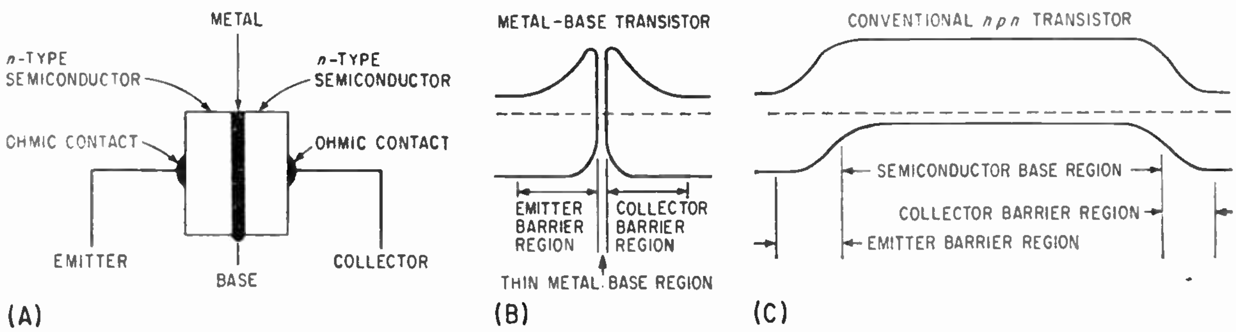


Fig. 9. Construction of the metal-base transistor (A), and comparison between energy diagrams of the new device (B) and a conventional n-p-n structure (C). (After Geppert and Muller.[22])

the electron momentum and cause electrons not to be collected as a result.

The best results so far have given gains of about 10 dB in the region of a few Gc/s with 10 V breakdown.[22] Little investigation of noise has taken place and there are problems of compatibility of materials to be overcome. However there would seem to be no reason why the semiconductors should not be polycrystalline and thus radiation resistant and in consequence no reason why the device should not be ultimately made using ultra-high-vacuum evaporation techniques. This may well be the microwave amplifying element of the future.

It is perhaps worth quoting once again the specification for the design published by Geppert and Muller for their 10 Gc/s metal base transistor which is reproduced in Fig. 10. It is as follows for the grounded-base configuration:

r_e = 2 Ω at 3×10^3 A/cm^2

r_b = 1 Ω

r_c = ∞

C_e = 1 pF for an area $0{\cdot}5 \times 10^{-5}$ cm^2 and forward biased

C_c = 0·1 pF for an area $1{\cdot}0 \times 10^{-5}$ cm^2 and backward biased to 6 V

$\alpha \simeq 0{\cdot}9$ at 10 Gc/s.

With these figures the collector reactance is $-j159$ Ω and for a 1 Gc/s bandwidth a $Q < 10$ is needed and this indicates *either* a parallel 1000 Ω load *or* a 25 Ω series load—the latter obviously being better adapted to strip-line technique. Under these conditions, if circuit losses are ignored a power gain of 25 dB (ignoring circuit losses) is predicted. The grounded-base configuration is preferred because of the natural match provided between the device and the field pattern in the strip line.

The mention of strip-line techniques in connection with transistors perhaps raises one further point which should be mentioned, namely that criteria such as the figure of merit quoted above tend to ignore other considerations which will then become important. Once $r_{bb'}C_c$ has been minimized, we shall still be left with such parameters as C_e in shunt across the transistor input which need to be tuned out by the strip line circuit. Despite the arguments used later in this paper to indicate that, from a circuit point of view, lumped circuit rather than strip-line techniques serve adequately from a purely passive point of view, in dealing with the input and output capacitances of active devices at the low impedance levels pertaining the use of strip-line technique is still favoured.

Above 10 Gc/s strip-line circuits cease to be practical although with waveguide technique there

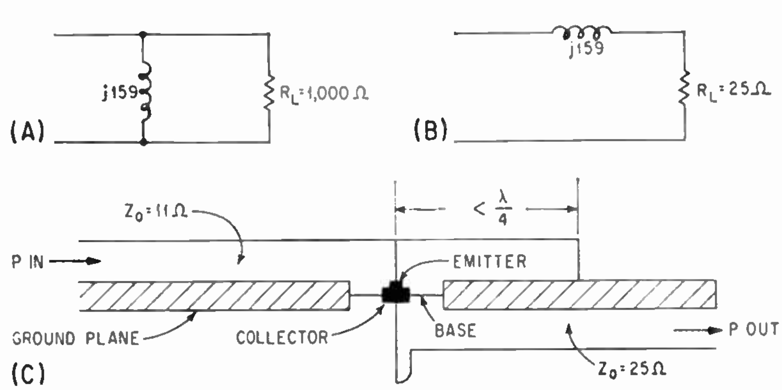


Fig. 10. The metal-base transistor circuit. (A) The collector load must be inductive to cancel capacitance at 10 Gc/s. (B) Alternative load circuit for a strip-line amplifier application. (C) Actual strip-line circuit. (After Geppert and Muller.[22])

World Radio History

seems to be little fundamental reason why oscillations at 100 Gc/s should not be possible. It is worth stressing that, although it has been shown by Rose[23] that the upper frequency limit of all solid-state three-electrode devices is the dielectric relaxation time of the emitter base region, this involves idealization of the other parameters of the device. It may well be that this idealization, and thus the approach to the fundamental limit is more easily achieved with the metal base transistor than with any of its other proposed rivals. On the other hand it must be acknowledged that so far the best high frequency performance (in the X-band region) has been obtained with minority carrier transistors. There are great difficulties in preparing true 'Schottky barriers' and many of the devices employing thin metal films have probably produced their promising performance as a result of a combination of mechanisms rather than by pure Schottky emission. The problem of producing 'continuous' thin metal films approximately 100 Å thick on semiconductor surfaces has yet to be satisfactorily solved and realization of the full potential of the metal base transistor thus awaits the solution of yet another materials problem.

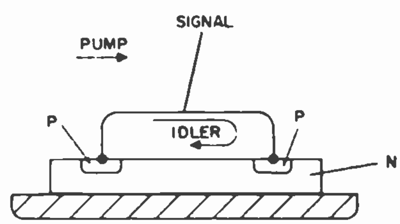
6. Microelectronics Techniques

By now we are becoming fairly well used to revolutions of one kind and another dispossessing us of devices which we had struggled long to understand. Thus many thermionic devices have passed or are rapidly passing into obscurity. At the present time the semiconductor field is presented with an equally difficult challenge, from microelectronics, which is fast causing the dividing line between components and system manufacture to disappear in some fields. The fact that we can now make and interconnect 1000 transistors on a 1-in diameter silicon slice is producing interesting 'fallout' in the microwave region for the simplest of reasons which is so obvious as to be commonly overlooked.

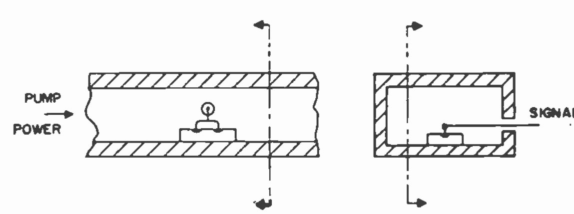
The main reasons for adopting microwave techniques at all in the past have been associated with such phenomena as skin depth, radiation of power from connecting wires, the occurrence of complex active device parameters, parasitic capacitance between components and connections, and so on.[24] The criterion for deciding when we must start to take precaution against these latter phenomena was generally that we must take care as soon as the dimensions of components (in terms of 'electrical' length) and the lengths of connections became 'comparable with the wavelength'. So far as skin depth was concerned, we were concerned almost as soon as we moved away from d.c. With modern thin-film and semiconductor integrated technology however we find ourselves in a very different position. For example, a typical thin film resistor may be 100 Å thick, whereas the skin depth at 10 Gc/s is 6400 Å (for Ag). Thus the average thin film component will behave as if it were operated under d.c. conditions. Similarly, if we were to take as a criterion that radiation effects could be neglected so long as the electrical length of a component was less than $\lambda/100$ (and this is a very safe criterion), then at 10 Gc/s this would allow components of 75 μm dimension to be made. Considering that with electron-beam machining we should be able to cut holes in oxide masks of 1 μm diameter this condition would not seem to be too restrictive.

We are thus in an interesting phase where we are about to see a resurgence of interest in what are essentially lumped circuit techniques in the microwave region. Two examples will suffice to show how these techniques may affect us in the future. The first example can truly be classed as 'fallout' from the microelectronics programme although the second is an intimate part of it.

Example No. 1. In an excellent review paper on integrated circuits at microwave frequencies, Dr. Uhlir[25] described a possible parametric amplifier using a universal varactor element which could also be designed to operate under certain circumstances without unidirectional devices.



(*a*)



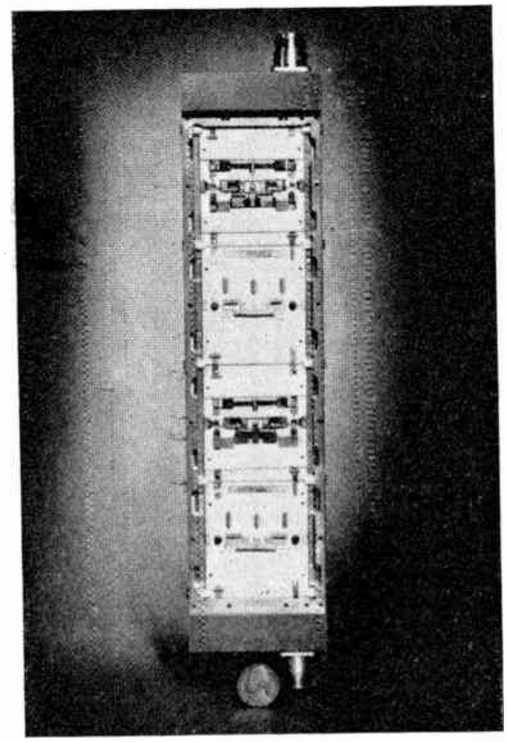
(*b*)

Fig. 11. (a) Balanced parametric amplifier element on single chip with built-in idler circuit. Pumping is by electromagnetic coupling. (b) Balanced parametric amplifier element mounted in pump waveguide. (After Uhlir.[25])

Essentially the amplifier consists of two varactors made in a single silicon chip and coupled by an evaporated connection as indicated in Fig. 11. At these frequencies and with the dimensions possible with microelectronics techniques the connecting strap can provide suitable inductance to resonate with the diode capacitance and thus provide the idler circuit. Such a package then needs only to be coupled to idler and pump resonators which can be tuned at will, subject to a sufficiently high idler frequency having been chosen in the first place. The use of the balanced arrangement prevents the pump and idler signals from exciting the signal input circuit and, although the system shown only provides weak coupling to the pump, tighter coupling could be achieved by use of the system shown in Fig. 12.

Similar techniques of integration of blocking capacitors, resonating inductance and varactors on a single chip are already making for simplification in the design of switching circuits.

Example No. 2. The most advanced integrated circuit amplifier to be described so far does not quite take advantage of the properties just outlined as being characteristic of most of the best of microelectronics. However its performance is challenging both to the travelling-wave tube and to all other semiconductor devices which seek to replace it. The amplifier[26] operates in the range 500–3000 Mc/s with a bandwidth of 1000 Mc/s using common emitter—transistor pair stages coupled via 3-dB strip-line couplers as shown in Fig. 13.



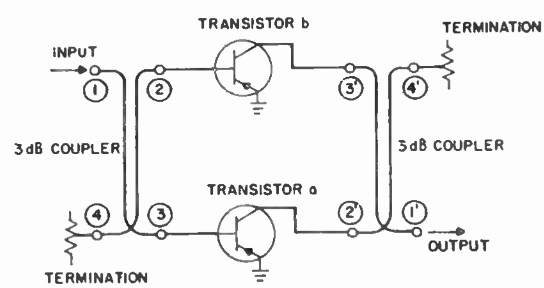


Fig. 13. Englebrecht's amplifier in integrated form possessing characteristics equivalent to those of a typical S-band travelling-wave tube.[26b] (Bell Laboratories Record.)

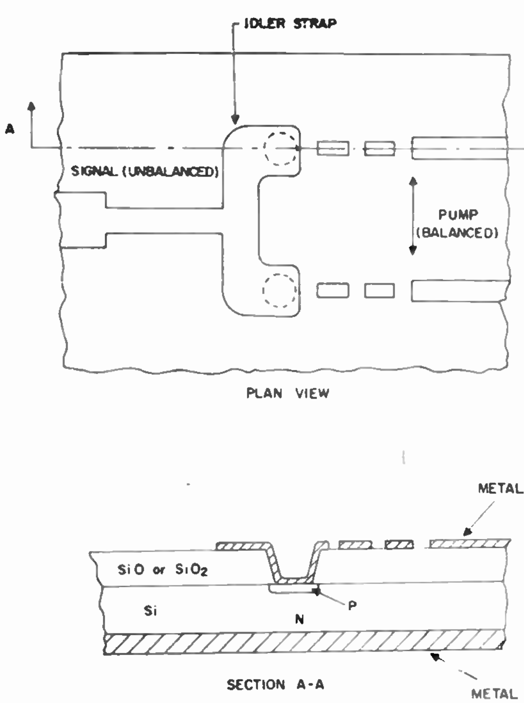


Fig. 12. Monolithic balanced parametric amplifier element. (After Uhlir.[25])

With production transistors, gain of 6 dB/stage and noise figures of ~ 5 dB have been obtained at 1 Gc/s, while 3 dB/stage and 8 dB are more typical at 3 Gc/s. However the gain figures have been doubled, with substantial improvement in the noise figure, using experimental planar germanium transistors and clearly we are only on the threshold of what may be achieved in the future. Basically the technology used is compatible with tantalum thin film circuitry using ceramic substrates.

Both with this and with a second, lumped circuit design[27] which operates with a 1 Gc/s bandwidth from 400–1400 Mc/s, meticulous attention has to be paid to the determination and elimination of parasitics

World Radio History

and computer design has proved essential. However once the design is determined, excellent reproducibility in manufacture is achieved.

7. The Gunn Effect

It is now well known that the application of voltage pulses (giving rise to fields of the order of several kV/cm) across certain semiconductor (e.g. GaAs) specimens gives rise to oscillations in the microwave region.[28]

The quality of the oscillations produced has been somewhat variable and the current stability and coherence has not been consistent from sample to sample. The frequency of oscillation is dependent on the sample size and oscillation on several frequencies simultaneously is often experienced.

Powers of some 17 mW at 6·3 Gc/s—comparable with the output of tunnel diodes in the same frequency range were reported in January 1965[29] using specimens of dimensions $50 \times 250 \times 250$ microns in a microwave cavity, whose function was merely to match the output from the specimen to the measuring device. The line width of the oscillation was some 10 kc/s and the device efficiency 1–2%.

Attempts to make c.w. Gunn oscillators have only recently met with any success[30] and the problem has been in part one of cooling, as specimens require a substantial heat sink. Powers of 2·5 W peak at 3 Gc/s with efficiencies of 7% were reported in March 1965 at room temperature from specimens some 25 μm thick and in June the achievement of 4 W peak at 10% efficiency was reported by Kroemer, while at 1 Gc/s 40 W peak has been recorded.[31] Under pulsed conditions line widths of 2 Mc/s were experienced but it was found that the cavity could be tuned over a 20 Mc/s band before oscillation became indetectable. These figures demonstrate the rapidity with which this field is advancing. It is worth noting too that the only specimens to perform satisfactorily in these latter experiments were those whose low field resistivities, in their mounted state, were consistent with the bulk resistivity of the specimen. Those specimens exhibiting higher resistivity were those which demonstrated instability, incoherence and low power.

This particular class of device makes us rapidly aware that so far as bulk generation semiconductor devices are concerned, we have yet much to learn. It is probably true to say that microwave *semiconductor* devices (as opposed for example to the maser family) have been produced and made to work satisfactorily without the microwave engineer having had to make much investigation of the finer points involved in the understanding of semiconductor phenomena. It would seem that if advances are to be made in future, we may have to pay very much more attention to the physics of the phenomena involved. The suggestion first made by Kroemer[32] in December 1964 that the Gunn effect could be explained in terms of bulk negative resistance of the form discussed by Ridley and Watkins[33] is only tenable if the electron temperature is sufficiently high to explain the filling of the satellite valleys which occur in the conduction bands of GaAs and InP. A superficial knowledge of semiconductor properties is insufficient to enable explanation of how these valleys can be filled at threshold fields of the order of those observed but a full detailed examination shows otherwise.[34] This model allows explanation of many of the observed characteristics associated with the Gunn effect—but since this is the subject of full papers during the Symposium,† it will not be discussed any more here, except to reiterate that we stand greatly in need of more theoretical effort in the solid-state field if we are to be able to deal intelligently with the devices which the future will produce (and this applies to the whole solid-state sphere not only the microwave region).

8. Microwave Acoustics

At a time when there is not much that cannot be achieved in the way of low-noise amplification and frequency conversion in the microwave range using fairly conventional semiconductor devices, and where the possibilities of electron devices are almost fully exploited, it is to be expected that any new advances will be made using hitherto uninvestigated physical properties of materials or by using any unique properties which arise from the development of new materials. The microwave acoustic field has become of interest as a result of the former process and is one which seeks to utilize the properties of the elastic vibrations of solids in active devices and delay lines.

These elastic waves possess the attributes of low attenuation and dispersion at cryogenic temperatures and very short wavelength in the propagating material. As a result of this a 1 mm length of crystal at 3 Gc/s is some 1000 λ long, implying that the acoustic waves will travel with a velocity of approximately 10^{-5} that of the velocity of light. Because of the short wavelength the propagation medium must be a single crystal free from defects to avoid defect scattering, while the supercooling is necessary to avoid attenuation due to phonon interactions. The process is clearly such that only a quantum analysis is likely to yield a satisfactory description although the duality of the quanta of vibrational energy hf—the phonons—and

† P. N. Robson and S. M. Mahrous, 'Some aspects of Gunn effect oscillators', *The Radio and Electronic Engineer*, **30**, No. 6, pp. 345–52, December 1965. J. S. Heeks, A. D. Woode and C. P. Sandbank, 'The mechanism and device applications of high field instabilities in GaAs', *The Radio and Electronic Engineer*, **30**, No. 6, pp. 377–87, December 1965.

World Radio History

their wave nature makes this field one which is a direct bridge between the free particle devices (e.g. the electron tube family) and the true quantum devices such as the masers and lasers.

The properties outlined above suggest that interactions with these waves (or particles) could provide us with a family of compact broad-band devices resembling the travelling-wave classes of tubes, masers and parametric amplifiers. Moreover acoustic phonons propagate at all frequencies up to 1000–10 000 Gc/s thus offering excellent high-frequency possibilities.

Phonon interaction takes place in various ways:

with other phonons via the non-linear elastic coupling provided by lattice distortion

with electrons (which are also coupled to the lattice) when the electron velocity exceeds that of the acoustic waves

with spin (magnetic) waves

and with optical rather than acoustic phonons under electromagnetic stimulation conditions.

In the last year we have seen propagation in crystals up to 70 Gc/s and beyond.

Phonon–phonon interaction resulting in upper sideband frequency to Q band and X band parametric amplification with an acoustic gain of several decibels has been observed.[35]

Phonon-spin wave interaction has also been observed with 30 dB gain at 700 Mc/s when pumped by a 1400 Mc/s spin wave in YIG (yttrium iron garnet).[36]

Phonon–electron interaction has so far been investigated mainly at low frequencies although amplifiers have now been operated in the gigacycles region and the use of high mobility materials such as InSb, perhaps in thin film form, might offer useful gains at higher frequencies still.

This is a field which as yet is only in its infancy and about which we know all too little. Having decided that germanium semiconductor devices would be useful it took about 10 years before *other* semiconductor materials were developed sufficiently to make a substantial impact on the field. At the present time we are very much more materials-minded and perhaps we shall be seeing special families of acoustic materials developing which will bring us new families of devices in this field sooner than we expect at the moment. So far work has been mainly concerned with quartz and cadmium sulphide, although silicon, germanium and ruby as well as YIG have also been used. It is noteworthy though and perhaps only to be expected at this stage of development that all of these materials were already in common use for other purposes.

One of the biggest problems which will have to be faced in the development of acoustic devices is undoubtedly the development of efficient transducers which at the moment cause severe coupling losses and this in the long run may prove to be the greatest limitation on acoustics applications.

9. Conclusion

No attempt has been made to be comprehensive in this review and the author has merely made reference to some of the items in the literature of the past year which have interested him. This selection has inevitably resulted in many omissions, e.g. microwave plasma devices.

It is clear that development of solid-state devices for use in the microwave region is continuing with unabated pace and seldom has there been a time in the history of microwave development where quite so many exciting ideas were all under simultaneous consideration, e.g. Gunn effect, acoustic devices, plasma effects, hot electron transistors. Which of these will establish themselves firmly as commercial prospects only the future can decide. In the meantime they provide fascinating subjects for conference discussion and give the lie to those who were suggesting some years ago that there is little more to be done in microwave electronics!

10. References

1. E. W. Herold, 'The future of the electron tube', *I.E.E.E. Spectrum*, **2**, p. 50, January 1965.
2. C. C. Hoffins, 'Design of X-band varactor diode frequency multipliers', *Proc. Inst. Elect. Electronics Engrs*, **52**, p. 742, June 1964.
3. J. L. Moll, S. Krakauer and R. Shen, 'p-n junction charge storage diodes', *Proc. I.R.E.*, **50**, p. 43, January 1962.
4. R. D. Hall and S. M. Krakauer, 'Microwave harmonic generation and nanosecond pulse generation with step recovery diodes', *Hewlett-Packard J.*, **16**, No. 4, December 1964.
5. F. L. Wentworth, J. W. Dozier and J. D. Rodgers, 'Millimeter wave harmonic generators, mixers and detectors', *Microwave J.*, **7**, p. 69, June 1964.
6. A. I. Grayzel, 'Greater doubler efficiency using diodes with small values of γ', *Proc. I.E.E.E.*, **53**, p. 505, May 1965.
7. R. I. Harrison and J. Zucker, 'Microwave detection and mixing using the thermoelectric effect of hot carriers in semiconductors', *Appl. Phys. Lett.*, **3**, p. 153, November 1964.
8. P. Penfield, 'Maximum cut-off frequency of varactor diodes', *Proc. I.E.E.E.*, **53**, p. 422, April 1965.
9. C. B. Burckhardt, 'Spurious signals in varactor multipliers' *Proc. I.E.E.E.*, **53**, p. 389, April 1965.
10. R. J. Kampf, R. S. Forman, O. J. Hanas and D. H. Knapschaefer, '30 Gc multiplier as an ultrastable parametric amplifier pump', *Proc. I.E.E.E.*, **52**, p. 1074, September 1964.
11. J. C. Sadler and W. E. Wells, Jr., 'Helium cooled X-band parametric amplifier', *Proc. I.E.E.E.*, **52**, p. 1269, October 1964.

World Radio History

12. A. S. Clorefine, 'Microwave amplification with superconductors', *Proc. I.E.E.E.*, **52**, p. 844, July 1964.
13. A. S. Clorefine and A. H. Nethercott, Jr., 'On the mechanism of amplification in superconducting films', *Proc. I.E.E.E.*, **53**, p. 388, April 1965.
14. e.g. D. T. Young, C. A. Burrus and R. C. Shaw, 'High efficiency millimeter-wave tunnel-diode oscillators', *Proc. I.E.E.E.*, **52**, p. 1260, October 1964.
15. R. Steinhoff and F. Sterzer, 'Microwave tunnel diode amplifiers with large dynamic range', *R.C.A. Rev.*, **25**, p. 54, March 1964.
16. e.g. (a) T. H. Oxley, 'Backward diodes as mixers at microwave frequencies', *J. Electronics and Control*, **17**, p. 1, July 1964.
(b) P. E. Chase and K. K. N. Chang, 'Tunnel diodes as millimeter wave detectors and mixers', *Trans. I.E.E.E.* on Microwave Theory and Technique, MTT-11, p. 560, 1963.
17. J. N. Nelson and B. P. Israelson, 'Travelling wave tube with 2·7 dB noise figure at 12 Gc/s', *Proc. I.E.E.E.*, **53**, p. 548, May 1965.
18. R. P. Wadhwa and T. Van Duzer, 'A 3·5 dB noise-figure, S-band medium-power, forward-wave, injected-beam, crossed-field amplifier', *Proc. I.E.E.E.*, **53**, p. 425, April 1965.
19. e.g. S. M. Krakauer and S. W. Soshea, 'Hot carrier diodes switch in picoseconds', *Electronics*, **36**, p. 53, 1963.
20. Hewlett Packard Ltd. Applications Group, 'Hot carrier diodes', *Electronic Components*, **6**, p. 527, June 1965.
21. J. Encinas and A. Minne, 'Design of v.h.f. epitaxial silicon transistors (2–3 and 6W power at 180 Mc/s)', *L'Onde Electrique*, **45**, p. 271, March 1965.
22. D. V. Geppert and R. A. Muller, 'Metal base transistor pushes back the frequency barrier', *Electronics*, **37**, p. 42, March 1964.
23. A. Rose, 'An analysis of gain-bandwidth limitations of solid-state triodes', *R.C.A. Review*, **24**, p. 627, December 1963.
24. G. D. Sims and H. Kemhadjian, 'High frequency limitations of integrated circuits', Proceedings of course on the physics, design and applications of integrated circuits, Oxford, July 1965, to be published.
25. A. Uhlir, Jr., 'Microwave applications of integrated circuit technique', *Proc. I.E.E.E.*, **52**, pp. 1617, December 1964.
26. (a) R. S. Englebrecht and K. Kurokawa, 'L-band transistor amplifier', *Proc. I.E.E.E.*, **53**, p. 237, March 1965.
(b) K. M. Eisele, R. S. Englebrecht and K. Kurokawa, 'Integrated microwave amplifier', *Bell Labs. Record*, p. 134, April 1965.
27. 'Thin-film, lumped-element amplifier', *Bell Labs. Record*, **43**, p. 136, April 1965.
28. J. B. Gunn, 'Instabilities of current in III-V semiconductors', *I.B.M. J. Res. & Dev.*, **8**, p. 141, April 1964.
29. B. W. Hakki and J. C. Irvin, 'CW microwave oscillations in GaAs', *Proc. I.E.E.E.*, **53**, p. 80, January 1965.
30. T. M. Quist and A. G. Foyt, 'S band Gunn effect oscillators', *Proc. I.E.E.E.*, **53**, p. 303, March 1965.
31. 'Gunn effect', *Electronic Components*, **6**, p. 544, June 1965.
32. H. Kroemer, 'Theory of the Gunn effect', *Proc. I.E.E.E.*, **52**, p. 1736, December 1964.
33. B. K. Ridley and T. B. Watkins, 'The possibility of negative resistance in semiconductors', *Proc. Phys. Soc.* (London), **78**, p. 293, 1961.
34. C. Hilsum, 'Transferred electron amplifiers and oscillators', *Proc. I.R.E.*, **50**, p. 185, 1962.
35. e.g. N. Shiren, 'Phonon interactions in crystals', I.B.M. Corpn. Reports 1 & 2, Contract No. DA-36-039-AMC-0228(E) May and August 1963.
36. e.g. H. Matthews, 'Elastic wave amplification in YIG at microwave frequencies', *Phys. Rev. Lett.*, **12**, p. 325, 1964.

Manuscript received by the Institution on 11th August 1965. (Paper No. 1047.)

© The Institution of Electronic and Radio Engineers, 1966

Proceedings of the Symposium on

MICROWAVE APPLICATIONS OF SEMICONDUCTORS

Publication in *The Radio and Electronic Engineer* of a representative selection from the papers read at the Symposium held in June–July 1965 has now been completed. The full record of the Symposium is contained in the revised edition of I.E.R.E. Conference Proceedings No. 5, copies of which are available from the Institution, price £6. The volume contains 40 papers and reports of the discussions which followed their presentation.

World Radio History

U.D.C. 621.37/9.004.2

The Economics of Integrated Circuits

By

G. C. PADWICK, B.Sc., C.Eng.†

AND

R. MATTHEWMAN, B.Sc.†

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: This paper presents the concept of the 'cost of ownership' of electronic equipment, and shows that equipment designed on the basis of integrated circuits often has a very low cost of ownership.

1. Introduction

The past year (1964–5) has been a very significant period in the history of integrated circuits in Europe. Whereas previously integrated circuits had been bought in relatively small quantities for evaluation and development purposes, the last year has seen several equipments go into manufacture using large quantities of integrated circuits. The year has also seen a significant expansion in the areas of usage of integrated circuits. Previously the great majority of integrated circuits had been used in military and related projects. During this year integrated circuits have moved into the civil aviation, computer, instrumentation, and general industrial electronics fields.

With very few exceptions the present and projected usage of integrated circuits is in applications where they are chosen for their economic advantages. In the past the use of integrated circuits had often been justified by their small size or high reliability. Now equipment manufacturers are choosing to use integrated circuits because the overall cost of ownership of an equipment made with them is less than the same equipment made with conventional components.

2. The Situation Today

Engineering can be defined as the application of scientific principles and knowledge to produce an economic solution to problems. The engineer must always consider both the economics and the technology of his work. Cost is an important, and often the most important parameter in engineering.

The cost of an electronic equipment is the sum of many parts:

(*a*) Design cost, including research and development, design and construction of prototype, preparation of drawings, etc.

(*b*) Construction cost, including procurement and assembly of component parts, testing, commissioning, etc.

(*c*) Running costs, including power consumption, supervision, routine maintenance, replacement parts, down-time cost, etc.

The proportion of the cost of ownership attributable to each of these factors varies greatly according to the type of equipment. For example, the most significant factor in the cost of a complex special purpose machine, of which only one is to be built, may well be the design cost, whereas the most significant cost of a machine which is an essential part of a large system may be the loss involved when the machine fails.

It has been shown[1] that integrated circuits contribute to a cost reduction in all these factors, and that the total cost of ownership of an electronic equipment can often be significantly reduced by the use of integrated circuits.

A detailed breakdown of the cost of electronic equipment has been given by Lennon.[2] After showing how each factor in the cost of equipment is reduced by using integrated circuits, Lennon takes a particular example of an airborne v.h.f. receiver, four of which are used in a modern airliner.

Lennon shows that, although the integrated circuit receiver costs slightly more than its conventional component rival (i.e. \$3500 against \$3000), the many advantages of the integrated circuit equipment give it a total cost of ownership over its life of \$9200 less than the conventional component equipment. The principal economy in this case is in the reduced maintenance costs of the integrated equipment.

Another example of the cost reduction brought about by the use of integrated circuits was published recently.[3] A peripheral equipment for a commercial computer was designed both in separate component form and in integrated circuit form. A comparison between the number of components in the two forms of the same equipment (Fig. 1) and the comparison between the cost of the two machines (Fig. 2) show the advantages of the integrated circuit equipment.

The interesting points about this cost comparison are that, whereas the component cost of the integrated circuit equipment is only slightly less than that of the separate component machine, the labour cost is greatly reduced by adopting integrated circuits. It is common to find that the relative component costs of an integrated circuit or separate component equipment are approximately the same today. In the near future the integrated circuit components will be

† SGS-Fairchild Ltd., Ruislip, Middlesex.

World Radio History

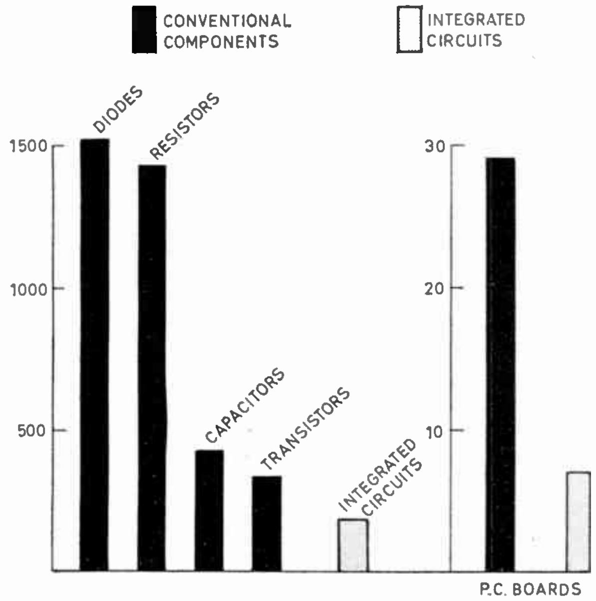


Fig. 1. Number of components comparison.

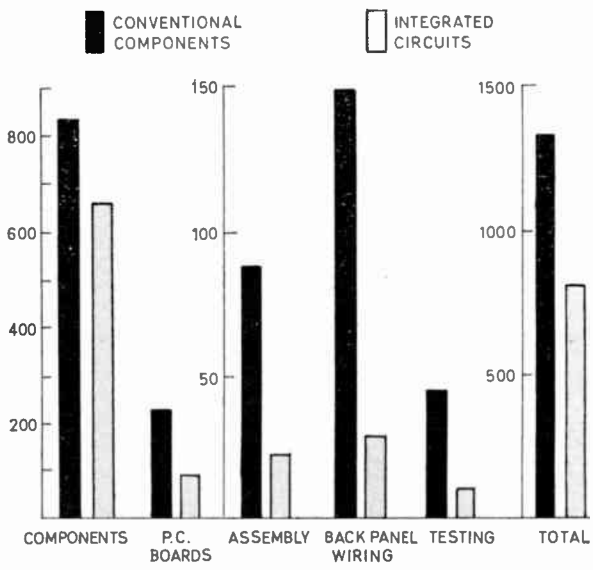


Fig. 2. Cost comparison.

much cheaper. But, even today, the greatly reduced labour cost of the integrated circuit machine makes it the best choice.

It is also relevant to note that the cost of the printed circuit boards for the integrated circuit version of the machine is approximately half that of the separate component counterpart, despite the fact that a more expensive type of board is used. In general, an equipment using integrated circuits occupies between one-quarter and one-tenth the board area that would be occupied by conventional components.

Where cost is more important than space, double-sided printed circuit boards with plated-through holes are normally used. Although these are more expensive than conventional boards, the reduced board area causes the total cost of boards for integrated circuits to cost between half and three-quarters the cost of boards for equivalent separate components.

Neither of these examples take account of the great reduction in electronic and mechanical design cost and time saving made possible by the use of integrated circuits.

3. The Future

In the future the economic advantages of integrated circuits will increase rapidly. Advances in manufacturing technology are rapidly reducing manufacturing costs. Improved packages and packaging methods are decreasing the cost of assembling integrated circuits into equipment.

The manufacturing processes of a monolithic semiconductor integrated circuit are:

(*a*) Manufacture of masks, jigs, etc.
(*b*) Photolithographic and diffusion processes.
(*c*) Mounting on headers and connection of leads.
(*d*) Testing.

Although the cost of manufacturing masks can be high, providing a large number of a particular type of integrated circuit is manufactured, the mask cost per finished device is very small, and is becoming insignificant as the production quantities increase.

During the photolithographic and diffusion processes a very large number of devices are processed simultaneously. The cost of this part of the process per finished device depends on the number of devices per wafer, and the yield of these devices to their final test specification. At present a typical wafer is rather less than $1\frac{1}{2}$ in diameter and contains between 500 and 1000 devices. The yield varies according to the complexity and type of circuit between 30% and 60%.

After diffusion the wafers are cut into individual dies and the dies are individually mounted on headers and connections are made between the header leads and the bonding pads on the die. The cost incurred at this stage consists of the cost of the package and the labour cost.

Testing also involves individual handling of devices. Providing large quantities of a particular device are manufactured, automatic test equipment can be used to minimize the cost of testing.

In a recent paper, Noyce has given details of the anticipated manufacturing costs of integrated circuits and transistors which are expected within the next year.[4]

World Radio History

Table 1

Hypothetical cost comparison for transistors and integrated circuits of comparable specifications.

	Transistors	Integrated Circuits		
		3 transistors	10 transistors	30 transistors
Wafer cost	$15·00	$20·00	$20·00	$20·00
Units/wafer	2,000	1,000	500	200
Yield	75%	60%	50%	30%
Cost/die	$·01	$·003	$·08	$·33
Package	$·03	$·05	$·06	$·07
Assembly and test	$·15	$·19	$·20	$·25
Factory cost	$·19	$·273	$·34	$·65
Cost/transistor	$·19	$·091	$·034	$·022

The figures shown in Table 1 clearly illustrate the fact that in large quantity production integrated circuits have definite advantages over transistors, and that the cost per transistor of an integrated circuit decreases as its complexity increases, providing manufacturing technology is capable of maintaining reasonable yields.

The above figures may be queried on the grounds that integrated circuits tend to use more transistors to perform a certain function than would be the case if separate components were used. This is compensated by the fact that the cost of passive components and the cost of interconnecting the components must be added to the cost of the separate transistors.

On the basis of these figures it is not unreasonable to suppose that we can look forward to the possibility of common digital functions being available at one-tenth the cost of comparable functions, in separate component form, and to linear functions and less common digital functions costing about one-half the separate component cost.

Looking further into the future there are many good reasons to suppose that integrated circuits are going to bring about great changes in the electronic industry. The cost advantage of integrated circuits will increase as the technology evolves towards the production of very complex functions on a single silicon chip, and towards more chips per wafer.

During the diffusion and photolithographic stages of fabrication the production cost is virtually independent of the size of the wafer and the number of circuits on the wafer. Each year sees larger diameter wafers being used, resulting in lower fabrication costs per chip. At the same time technology improvements are allowing the size of integrated circuit components to decrease and yield to increase, again lowering the fabrication costs per chip. The same factors also allow more components per chip without loss of yield.

Combining the effect of cost reduction per function decreasing as the complexity increases with that of yield reduction as complexity increases gives, at any given time, an optimum number of components per chip. At the present time this optimum number is approximately fifty. Looking five years ahead this number will increase to 1000, providing such a circuit can be used in large quantities.[5] Whether or not devices of this sort will ever be manufactured depends on the degree of co-operation that can be achieved between the semiconductor industry and equipment manufacturers, and on the degree of agreement that can be achieved among equipment manufacturers to use common complex functions.

As the table of cost comparisons shows, the cost of packaging and testing an integrated circuit element is a very significant part of the overall manufacturing cost. Improvement in production yields will increase the proportion of manufacturing cost incurred at the packaging and test stages. The interconnection of complex functions at the wafer stage[6] and improved packaging techniques[7] will play their part in making the cost of an integrated circuit function very low compared with separate component costs.

4. Conclusions

This paper has demonstrated that the principal reason for manufacturing and using integrated circuits is that they enable the manufacturing cost and total cost of ownership of many electronic equipments to be greatly reduced. It has also shown that the many advances in technology are enabling the manufacturing costs of integrated circuits to be reduced and that, during the next few years, we shall see the cost of integrated circuits continuing to fall, and the complexity of circuits that can be manufactured economically, increasing rapidly.

5. References

1. G. C. Padwick, 'The economics of integrated circuits', *Proceedings of the Symposium on Microminiaturization*, Edinburgh, 1964.
2. J. R. Lennon, 'A reliability/cost of ownership approach to microavionics', *Proceedings of the Eleventh National Symposium on Reliability and Quality Control*, Florida, 1965.
3. J. A. Rose, 'Industrial integrated circuits reduce system cost', *Electronic Design News*, August 1964.
4. R. N. Noyce, 'Integrated circuits—past, present and future indicative', *Proceedings of the Seventeenth Annual Convention, Financial Analysts Federation*, Houston, 1964.
5. G. E. Moore, 'Cramming more components onto integrated circuits', *Electronics*, **38**, No. 8, pp. 114–17, April 19, 1965.
6. I. M. Mackintosh and D. Green, 'Programmed interconnections—a release from tyranny", *Proc. Inst. Elect. Electronic Engrs*, **52**, pp. 1648–51, December 1964.
7. G. C. Padwick 'High-speed, cost orientated integrated circuits', *Proceedings of the Symposium on 'Applications of Microelectronics'*, Southampton, 1964. (Paper No. 29.)

Manuscript received by the Institution on 23rd September 1965. (Contribution No. 88.)

World Radio History

U.D.C. 621.397.334 : 621.391.837

A Subjective Investigation of some Errors in the Chrominance Signal Decoding Circuits of Colour Television Receivers

By

K. E. JOHNSON, B.A.,

(*Graduate*)†

Summary: This paper describes some experimental work in which the visual effects of incorrect decoding of the chrominance signals in colour television receivers were assessed by panels of observers. There are difficulties in changing circuit parameters in an actual receiver during a viewing session to simulate decoding errors. Apparatus was therefore constructed to produce exactly similar effects to those given by an incorrect receiver.

Preliminary tests were conducted using this apparatus to show the effects of several different kinds of decoding error. For these tests the N.T.S.C. X/Z decoding system was used. Following these tests, the effects of amplitude errors in the output colour difference signals used to drive the display tube were investigated in greater detail. The later tests are therefore relevant to all of the three currently proposed systems. The results of these experiments were compared with other workers' theoretical calculations, thereby allowing some assessment to be made of the visual importance of the calculated colour errors.

1. Introduction

Several types of decoding system for three-gun N.T.S.C. colour television receivers are in current use. In a recent paper,[1] Birt and Freeman have compared and appraised these systems to determine the most suitable one for a domestic receiver. This work was based on a mainly theoretical approach to the circuit design problems. In particular, calculations were made of the effect on certain colours of changing component values.

These calculations simulated the effect of component tolerances. The change in the reproduced colour was given in the form of a change in chromaticity co-ordinates on the 1931 C.I.E. diagram and a number giving the ratio of reproduced to intended luminance. These data are valuable in comparing circuits for sensitivity to component tolerances, but offer little help in assessing the importance of the resulting visual effects on a viewer's rating of the colour quality of the displayed picture.

The problem of how signal errors produced by the receiver affect picture quality has already received attention elsewhere. For example, Weiss[2] has performed similar calculations to those described above and has translated his results for certain defined colours into visual terms, using MacAdam's data on equally noticeable chromaticity differences.[3] He also carried out subjective tests in which viewers were asked to rate the visual effect on the picture of changes in various signals. In this way some check of the theoretical results, though not an extensive one, was obtained.

The present report describes a similar experimental approach to that of Weiss with the object of supplementing the theoretical work of Birt and Freeman. Panels of viewers were invited to assess the effect on still pictures of changing component values and signal levels in a typical decoding system, in this case one using an X/Z type of decoder. This particular decoder was chosen because the theoretical work indicated that it could yield errors as large as any other type and also because it is a well-known circuit, which has been used extensively in domestic receivers.

In the design of this subjective experiment conflicting requirements had to be met. Information was needed about a circuit which was part of a real colour television receiver design. It would seem that this information could best be obtained by using the actual circuit in question. However, on the other hand it was important to avoid any spurious effects arising from other parts of the receiver. In practice it would be difficult to vary the circuit parameters of an actual receiver in a known independent manner during the course of a subjective test.

The problem was solved by building a closed-circuit high quality monitor chain which could include models of real decoder circuits. Controls for varying the circuit parameters under investigation were built into this chain. Through the use of circuits of high stability and good linearity, and with no bandwidth

† Mullard Research Laboratories, Redhill, Surrey.

World Radio History

limitations, it was possible with all the controls set to their 'zero error' positions to display pictures which were indistinguishable from those produced using the direct signals from the picture source. This apparatus is described in the following section.

2. Apparatus to Simulate Effects of Decoder Errors

2.1. *Input Circuits*

The various systems of decoding considered by Birt and Freeman determine mainly the method of colour difference signal matrixing to be employed. Restrictions are only placed on the synchronous demodulators in respect of the gains and phases which must be used. It was not therefore necessary in these experiments to construct actual synchronous demodulators but only to provide means for producing the same effects as would be produced by varying these gains and phases.

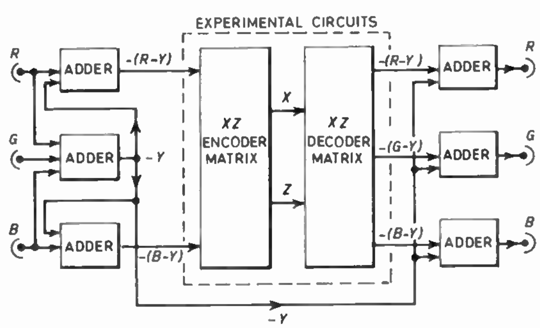


Fig. 1. Block diagram of apparatus.

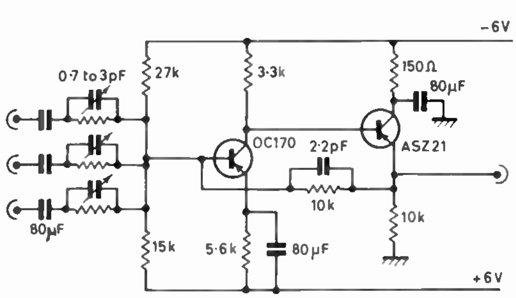


Fig. 2. Diagram of adding circuit.

A block diagram of the complete apparatus is shown in Fig. 1. Red, green and blue colour signals from a 625-line 35-mm slide scanner were matrixed using the adder circuit of Fig. 2 to give a luminance signal $-Y$, and colour difference signals $-(R-Y)$ and $-(B-Y)$. The colour difference signals were then fed to the block marked 'experimental circuits' in Fig. 1. Here it was possible to include the decoder matrix circuits described by Birt and Freeman.

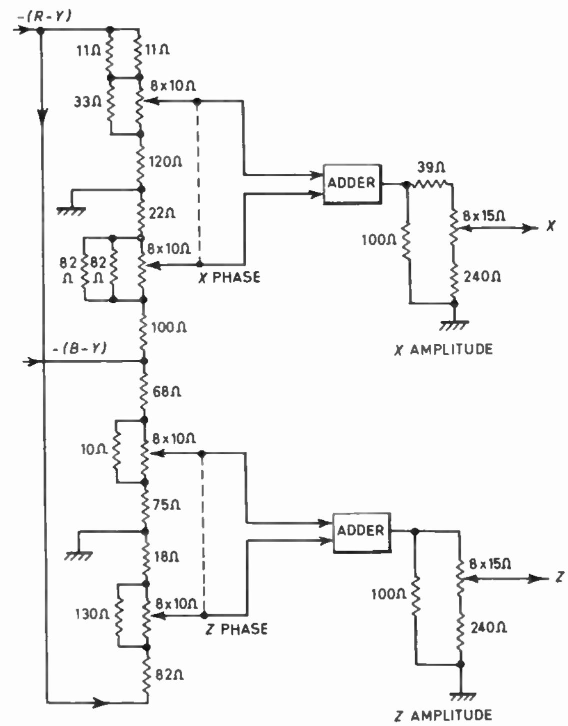


Fig. 3. Diagram of X/Z encoder matrix block.

2.2. *The X/Z Encoder Matrix Block*

For the present experiment, in which the X/Z type of decoder was investigated, it was necessary to include additional matrix circuits in the 'experimental circuits' block to produce the colour signals X and Z. These circuits are shown in Fig. 3. The X and Z signals would normally be produced directly by the synchronous demodulators of an ordinary receiver. In this case they were produced by adding proportions of the red and blue colour difference signals together.

The relative amplitudes of the signals required to give correct X and Z signals (as determined by the Birt and Freeman design for a decoder for a nominal phosphate tube) are shown in the table of Appendix 1. Also given are additional values which correspond to incorrect phases of the X and Z demodulation axes in the synchronous demodulators. These values cover a range of ± 10 deg in $2\frac{1}{2}$ deg steps for each axis, and form the basis for the design of the two pairs of ganged potentiometers shown in Fig. 3. These potentiometers therefore simulated the effect of incorrect demodulation phases.

Two more potentiometers were included after the X and Z adders to give control of the amplitudes of the X and Z signals. In this way the effect of an incorrect ratio of X/Z demodulation gain was simulated. The range of these controls was $\pm 20\%$ about the correct values, in 5% steps.

World Radio History

2.3. *The X/Z Decoder Matrix Block*

The X and Z signals were next fed to the second part of the 'experimental circuits' block, the X/Z decoder matrix, shown in Fig. 4. This was based on the design for a decoder matrix by Birt and Freeman already referred to. (See Appendix 1.) This circuit contained four controls.

Three of these controls varied the amplitudes of the resultant decoded colour difference signals over a range of $\pm 50\%$ about the correct values, in 10% steps. The controls were arranged as shown in Fig. 5. They formed the load resistances of three emitter followers, which were necessary to match the output impedances of the triode stages to the input impedances of the following adder stages.

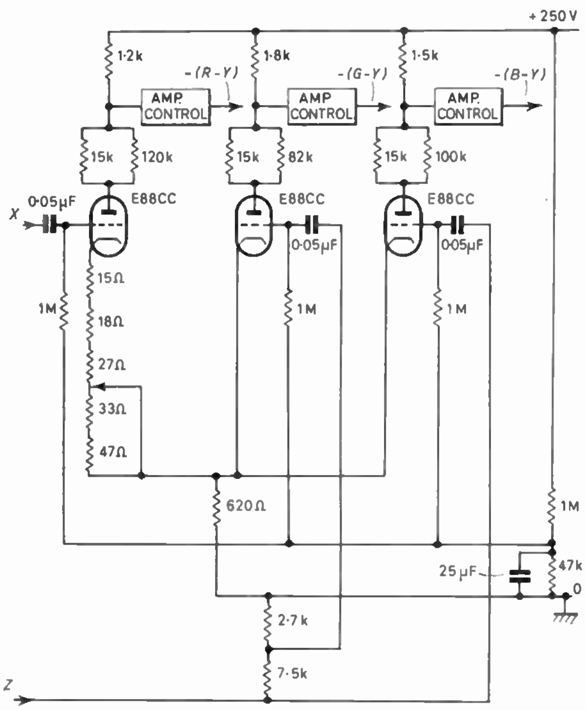


Fig. 4. Diagram of X/Z decoder matrix block.

The fourth control was a variable resistance which could be connected in the cathode lead of any of the three triodes to reduce its effective mutual conductance g_m', according to the relation:

$$g_m' = g_m(1 + g_m R_K)$$

where g_m is the mutual conductance of the valve and R_K is the value of the resistance inserted in the cathode lead. This produced a visual effect very similar to that calculated by Birt and Freeman for an error in the amplification factor (μ) of one of the triodes. The

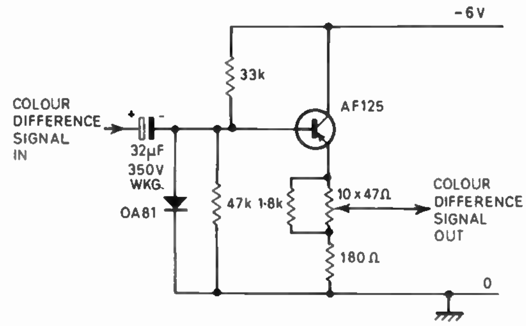


Fig. 5. Colour difference signal amplitude control.

mutual conductance could be reduced by up to 50% in 10% steps.

2.4. *The Output Circuits*

The decoded colour difference signals were each matrixed with the luminance signal in the final three adder stages (circuits as in Fig. 2) to regenerate red, green and blue signals. These were then fed to a 625-line colour monitor fitted with a 21-in 'Shadow-mask' tube which had phosphors of the 'phosphate' type whose colour co-ordinates were measured to be close to the nominal F.C.C. primaries. Errors due to transmission and display primaries not being the same were therefore reduced to a minimum.

High stability resistors of 1% tolerance were used in the construction of the apparatus to ensure a high degree of accuracy throughout. The potentiometers were made from switched chains of resistors. Where necessary, pairs of resistors in parallel were used to give the closest approximation to design values. Matched high quality valves were used in the actual decoding circuit. In spite of these precautions a very small difference, undetectable with an oscilloscope, was just perceptible visually between the white balance with all the controls set to their correct positions and that with the equipment by-passed. This difference was removed by inserting small continuously variable 2 kΩ potentiometers in series with the feedback loops of the three final adder stages.

3. Some Preliminary Subjective Experiments using the Apparatus

3.1. *Organization of Preliminary Experiments*

A full and rigorous subjective examination of the ten variables associated with this apparatus (namely, three colour difference signal amplitude controls, X and Z amplitude and phase controls, and the g_m control for each of the three triodes) would have taken a considerable amount of time and effort. Some preliminary experiments were therefore performed on each of the controls using a limited number of

World Radio History

observers and slides. These experiments were designed to demonstrate the range of each control in visual terms. More rigorous experiments were then conducted for those controls which showed the largest visual effects.

The preliminary experiments took place during two viewing sessions, each lasting about forty minutes. Five controls were investigated in each session, each control being dealt with in turn during a session. To limit the number of judgements to be made by the observers, only two slides were used—'Basket of Fruit', a still life showing oranges, apples, bananas, grapefruit, etc., and 'Girl in Headscarf', a close-up portrait of a model. Both slides were from the series taken by Dr. Hunt of Kodak Ltd.

For one of the pictures and one of the controls, the viewers judged each setting according to whether they regarded the colour quality of the reproduced picture as acceptable or not. The settings were arranged in a random order, each setting occurring twice during the sequence. The procedure was then repeated for the same control and the second picture, and finally for all the remaining controls in the same way as for the first.

The adjustment of the slide scanner and monitor, and the arrangements in the viewing room were as described in a following section (Sect. 4.1.1). The observers were three experienced television engineers, who were each seated at a distance of six times picture height from the screen. The screen was blacked out between each picture condition.

3.2. *Results of Preliminary Experiments*

With such a restricted number of observations for each picture condition it was difficult to fix a limit of acceptability. As the viewers were in very good agreement with one another for almost all the picture conditions, the limit chosen was the condition for which all the viewers rejected the colour quality of the picture. A summary of the results is given in Table 1. Where the maximum introduced error still gave an acceptable result, this is indicated in brackets.

3.3. *Discussion of Results*

The first remark to be made about these results is that, being based on few observers and pictures, they should be treated with reservation. Secondly, in practice more than one error, corresponding to moving more than one control, might occur at any time.

Referring to Table 1, the general observation may be made that the viewers were much less critical of the colour quality produced by the control variations than had been expected from the theoretical results of Birt and Freeman. This, however, was not felt to be due to non-critical observers, as the viewers in this case were technical staff. Nor was it felt to be due entirely to the picture material used, as previous experience with these slides had indicated that they were quite sensitive to variations of their reproduced colour quality.

In more detail, we may note that the maximum positive and negative errors of the $G-Y$ signal amplitude, the X and Z signal amplitudes, and the maximum negative error of the g_m of the green and blue triodes did not result in all the observers rejecting the colour quality of the pictures. Furthermore, the phase of the X or Z demodulation axes had to be at least 7·5 deg in error in either direction before the viewers rejected the colour quality. In normal practice in a receiver the X and Z demodulation axes are very unlikely to vary in phase independently of one another by as much as this.

The viewers were much less tolerant of both positive and negative errors in the $B-Y$ signal amplitude than they were for the $G-Y$ signal. The results show a variation between the pictures for this control. The viewers were also less tolerant of a reduction in $R-Y$ signal amplitude than they were for the $G-Y$ control. Increasing the amplitude of the $R-Y$ signal was, however, more acceptable to them.

In view of these results it was decided to carry out a more detailed investigation of the effects of the colour difference signal amplitude controls. The effect of wrong levels of these drive signals for a particular Shadowmask tube applies to the design of any type of decoder and is therefore of general interest. These experiments are described in the following sections.

4. Investigation of Visual Effects due to Changes in Drive Signal Amplitudes

4.1. *Organization of Tests*

The experimental approach described in Section 3 was felt to be inadequate for a further investigation of the visual effects due to changes in the amplitudes of the signals driving the three grids of a 'Shadowmask' tube. More specifically, apart from the very limited numbers of observers and slides used, it was felt that the viewers should not be asked just whether the colour quality was acceptable. Instead they were asked to give their assessment of the colour quality according to a standard scale, the E.B.U. six-point quality scale:

1. Excellent
2. Good
3. Fairly good
4. Rather poor
5. Poor
6. Very poor

The observers were not provided with a reference picture for comparison purposes. The method adopted

Table 1
Results of preliminary experiments
Minimum conditions for rejection of colour quality by all observers

Control	Slide	Positive-going limit	Negative-going limit
$R-Y$ amplitude	Girl in Headscarf	(+50% acceptable)	−30%
	Basket of Fruit	+50%	−30%
$G-Y$ amplitude	Girl in Headscarf	(+50% acceptable)	(−50% acceptable)
	Basket of Fruit	(+50% acceptable)	(−50% acceptable)
$B-Y$ amplitude	Girl in Headscarf	+20%	−40%
	Basket of Fruit	+50%	−20%
X amplitude	Girl in Headscarf	(+20% acceptable)	(−20% acceptable)
	Basket of Fruit	(+20% acceptable)	(−20% acceptable)
Z amplitude	Girl in Headscarf	(+20% acceptable)	(−20% acceptable)
	Basket of Fruit	(+20% acceptable)	(−20% acceptable)
X phase	Girl in Headscarf	+10°	− 7·5°
	Basket of Fruit	+10°	−10°
Z phase	Girl in Headscarf	+ 7·5°	(−10° acceptable)
	Basket of Fruit	+10°	−10°
g_m of red triode	Girl in Headscarf	—	−50%
	Basket of Fruit	—	(—50% acceptable)
g_m of green triode	Girl in Headscarf	—	(−50% acceptable)
	Basket of Fruit	—	(−50% acceptable)
g_m of blue triode	Girl in Headscarf	—	(−50% acceptable)
	Basket of Fruit	—	(−50% acceptable)

was to show them a random sequence of pictures and errors (including the correctly reproduced pictures) on a single monitor. This method was felt to be a closer approximation to home viewing conditions than would have been given by a comparison method in which two monitors were used, since normally viewers in the home would have only one picture to judge. It had the disadvantage, however, that the observers had no common defined reference standard such as the correctly-reproduced picture. A description of the viewing conditions, observers, slide material and test procedure used in the experiments follows.

4.1.1. Viewing conditions

The colour television monitor was set up on Test Card 'C'. Its white balance was normalized to Illuminant C using a special matching box containing a standardized source of illumination. The colour co-ordinates of the white balance produced in this way were then determined accurately using a tri-colorimeter to ensure correct setting. This was carried out with the input channels to the monitor uncommoned.

The black level of the monitor was set to make the black square of the test card just visible. The highlight brightness was adjusted to be 16 foot-lamberts as measured on the peak white square of the grey scale.

The normal procedure as laid down in the handbook of instructions was followed in setting up the slide scanner. The output signal level was adjusted manually during the test to be 1 volt for all the slides used.

All the lighting in the viewing room was adjusted to be metameric with Illuminant C. White background curtains were used behind the monitor, the brightness level on the curtain surface being 1 foot-lambert. Further local lighting (also Illuminant C) was used to enable viewers to see their answer sheets adequately. The average level of illumination on these sheets of paper was measured to be about 1 foot-lambert. No lighting other than that on the background curtains and the answer sheets was used.

The comment scale was printed on the bottom of the viewers' answer sheets, and also on a placard placed near the monitor and illuminated by the light falling on the background curtains.

A maximum of six viewers at a time was used, the viewers being seated in a straight line at an average distance of eight times picture height (8 ft 8 in) from the screen.

World Radio History

4.1.2. Observers

The viewers were chosen in the following way. Each division of the Laboratories was asked to provide volunteers to take part in subjective tests, of which this experiment was only one example. One hundred and thirty-two people volunteered, including both technical and non-technical (e.g. secretarial) staff. Television Laboratory staff were excluded. Alphabetical lists of these volunteers, divided into men and women, were drawn up. These lists were divided into panels of ten members, each panel consisting of eight men and two women.

Six observers drawn at random from a panel were invited to attend each viewing session. In all, 29 observers attended six viewing sessions (a total of seven viewers being unable to come). These observers were divided in the proportion of 23 technical to 6 non-technical staff, and in the proportion of 22 men to 7 women.

The observers were asked to state at the bottom of their answer sheets if they had any defects of colour vision. Where this was unknown the person concerned was tested using the Farnsworth-Munsell 100-hue colour test.[4] One male observer had slightly defective colour vision, but his results were not excluded from the test as there was no evidence, judging from the results, for doing so.

4.1.3. Slide material

Ten slides of a varied nature were used in this test. These were:

1. 'Girl in Hat' 2. 'Girl in Headscarf' 3. 'Girl on Couch' 4. 'Biscuits and Cheese' 5. 'Basket of Fruit' 6. 'Guardsman'	From the collection taken by Dr. Hunt of Kodak
7. 'Greenhouse Bulbs' 8. 'Park View' 9. 'Lake Scene' 10. 'Thatched Cottages'	From the collection of a Television Laboratory member

Of these slides, four contained flesh tones ('Girl in Hat', 'Girl in Headscarf', 'Girl on Couch' and 'Guardsman', the former two containing large areas of flesh tone as they were close-up portraits) and three contained a considerable area of vegetation ('Greenhouse Bulbs', 'Park View' and 'Lake Scene'). All the slides containing flesh tones were thought to be critical for this test.

4.1.4. Test procedure

Two viewing sessions were devoted to each control. To keep the time for a session down to about half an hour only a limited number of pictures and conditions were shown in each session. For each control, viewers in the first session were shown slides 1, 3, 5, 6, 7 and 9 (the numbers are from the above list) and control positions -50%, -30%, -10%, 0, $+20\%$, $+40\%$; while viewers attending the second session saw slides 1, 2, 3, 4, 8 and 10 and control positions -40%, -20%, 0, $+10\%$, $+30\%$, $+50\%$. This gave a total of 72 picture conditions (i.e. pictures $\times$ conditions per picture) per session as each condition for each picture was seen twice by the viewers. A short break of about 2 minutes after 36 picture conditions gave the viewers (and those operating the apparatus) a rest. All the picture conditions were randomly ordered in a given session, and the screen was blacked out between conditions. Slides 1 and 3 provided a link between sessions, being common to all of them.

The viewers were given a few minutes at the beginning of each session to become adapted to the viewing conditions. During this time forms were distributed and a short introductory talk about the tests was given. The viewers were told only that the object of the test was to obtain their reactions to colour television pictures in which some of the colours might have been changed slightly. They were asked to assess only the colour quality of the pictures. Their attention was drawn to the quality scale to be used, but no explanation or comments about the scale were made. The organization of the viewing session was explained to them; they were told that they would see a total of 72 picture conditions in a random order, each picture being held for about 10 seconds, with a 10-seconds break between pictures. They were also told that there would be a short rest in the middle of the session.

4.2. *Presentation of Results*

A total of 2160 comments was received from the six sessions. These comments were analysed in the following way. Results from the two sessions during which a particular control was investigated were combined firstly by grouping the slides into two categories, one containing the four 'flesh tone' slides, the other containing the remaining six slides. These two categories are referred to in the following as 'flesh tone' and 'still life, etc.' respectively. Secondly, the results in each category of slides were grouped as follows. Results for control positions -50% and -40% were considered collectively to be due to a control setting of '-45%'. Similarly, results for positions -30% and -20% were considered to be due to a setting of '-25%'. Results for positions -10%, 0 and $+10\%$ were grouped under '0'; the remaining groups '$+25\%$' and '$+45\%$' were as for the corresponding negative going positions. This procedure gave a satisfactory smoothing of the results.

The complete results grouped in this way are shown

World Radio History

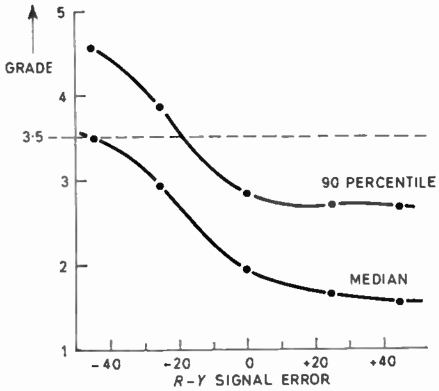


Fig. 6. Median and 90 percentile curves for $R-Y$ signal and 'flesh tone' slides.

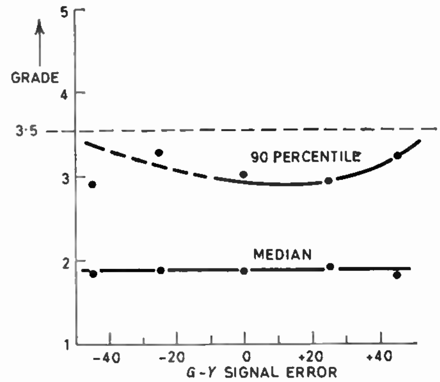


Fig. 9. Median and 90 percentile curves for $G-Y$ signal and 'still life, etc.' slides.

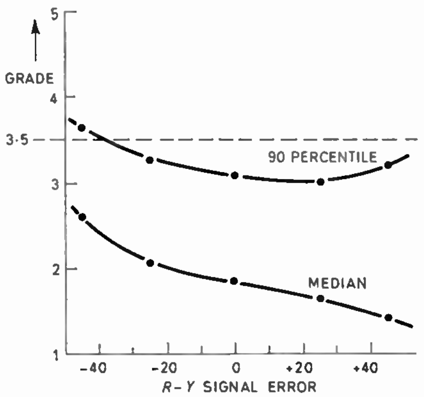


Fig. 7. Median and 90 percentile curves for $R-Y$ signal and 'still life, etc.' slides.

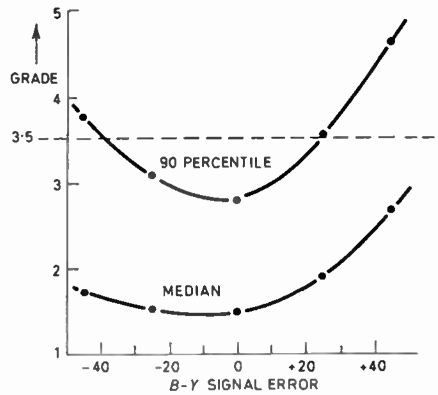


Fig. 10. Median and 90 percentile curves for $B-Y$ signal and 'flesh tone' slides.

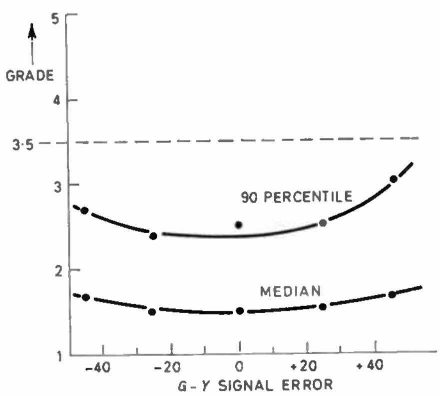


Fig. 8. Median and 90 percentile curves for $G-Y$ signal and 'flesh tone' slides.

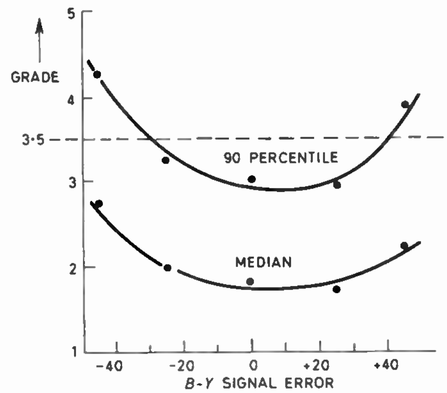


Fig. 11. Median and 90 percentile curves for $B-Y$ signal and 'still life, etc.' slides.

in Appendix 2. The distributions of comments in each grade for each condition are tabulated. The column giving the total number of comments for each condition shows that the total number of comments for each control were divided equally between the two categories of slides.

From the distributions of comments, curves (or ogives) of cumulative frequency against grade were plotted. These curves gave the proportion of comments recorded in a given grade or higher grades. The curves were used to find the median (50 percentile) and 90 percentile for each distribution, i.e. for each condition. In Figs. 6 to 11 curves are shown of the medians and 90 percentiles plotted against the control setting for each control and each group of slides.

4.3. *Calculation of Tolerance Limits of Acceptability*

The value 3·5 on the quality scale can be considered to be the limit of acceptability as measured by this scale, since it divides the scale symmetrically into two parts, the 'good' grades 1, 2 and 3, and the 'poor' grades 4, 5 and 6. If horizontal lines are drawn on the curves of Figs. 6 to 11 through the grade 3·5, then the points in which these lines intersect the curves for the 90 percentiles give the errors in the amplitude of any of the colour difference signals for which 10% of the viewers' comments lie in the lower, or unacceptable, half of the comment scale. That is, these intersections give the maximum error in either direction of the amplitude of any of the colour difference signals (assuming all other signals to be correct) for only one person in ten to reject the colour quality of the displayed picture.

Table 2

Tolerance limits

Control	Group of slides	Positive-going limit (per cent of correct amplitude)	Negative-going limit (per cent of correct amplitude)
$R-Y$	Flesh tone	—	−18
	Still life, etc.	+58 (by extrapolation)	−37
$G-Y$	Flesh tone	+58 (by extrapolation)	—
	Still life, etc.	+56 (by extrapolation)	—
$B-Y$	Flesh tone	+24	−38
	Still life, etc.	+41	−30

These tolerance limits, where they can be found, are shown in Table 2. It was impossible to set a positive-going limit for the $R-Y$ signal and a negative-going limit for the $G-Y$ signal, both on 'flesh tone' slides, because the number of adverse comments was never large enough even to allow extrapolation to the '10% at 3·5' level. No limit was set for the negative-going $G-Y$ signal on 'still life, etc.' slides because the results were too inconclusive in this region (see Fig. 9).

It is interesting to note that the viewers were so critical of a reduction of $R-Y$ signal amplitude for the 'flesh tone' slides that a limit can also be set at an error of −45% at which half the viewers rejected the colour quality of the displayed pictures.

4.3.1. Differences between sessions

If there were any obvious differences in the performance of either observers or apparatus between sessions this should have been reflected in the results for zero signal error.

When the results for the nominally correct positions of all the controls were compared, with the object of showing up any such differences, it was found that the medians and 90 percentile values agreed very satisfactorily for the 'still life, etc.' category of slides (see Table 3(b)), but less well for the 'flesh tone' category (see Table 3(a)).

Table 3(a)

Ratings for 'flesh tone' slides for zero signal error

Control	Median	90 percentile
$R-Y$	1·94	2·83
$G-Y$	1·50	2·50
$B-Y$	1·49	2·79

Table 3(b)

Ratings for 'still life, etc.' slides for zero signal error

Control	Median	90 percentile
$R-Y$	1·88	3·12
$G-Y$	1·88	3·00
$B-Y$	1·83	3·03

5. Discussion

5.1. *Comparison of Two Series of Tests*

Two experiments have been described in this paper for finding tolerance limits on the amplitude of the colour difference signals used to drive a three-gun 'Shadowmask' colour television tube. Although the former of these experiments, described in Section 3, was limited in its approach to the problem, it is interesting to compare the results of this experiment with those obtained in the more comprehensive tests described in Section 4. This can be done by comparing, for example, the limits obtained in the first experiment for one of the signals and for the flesh tone slide 'Girl in Headscarf' with those for the group of four flesh tone slides of the second experiment (of

World Radio History

which this slide was one). Similarly, the results for the still life slide 'Basket of Fruit' may be compared with those for the group of slides 'still life, etc.' of the second experiment (of which this slide was again a member).

The comparison shows that there is a fairly close similarity between the results of the two experiments. The conditions when the three expert observers (television engineers) used in the first experiment were in complete agreement in rejecting the reproduced colour quality agree fairly closely with the conditions at which the critical 10% of the non-expert observers used in the second experiment (see Sect. 4.1.2) also rated the colour quality as unacceptable (i.e. below grade 3·5 on the quality scale). The viewing conditions in each case were almost the same. Thus it would seem that, in this experiment, the 'expert' observers reacted in approximately the same way as the critical 10% of a random selection of non-expert personnel.

5.2. *Comparison of Experimental Results with the Theory of Birt and Freeman*

As was stated in the Introduction the object of these experiments was to supplement the theoretical work of Birt and Freeman.[1] The experimental results were therefore compared with this theory. In their paper, Birt and Freeman gave several diagrams showing, in the form of numbers and arrows on the 1931 C.I.E. diagram respectively, the luminance and chromaticity changes produced by errors in an X/Z decoder circuit.

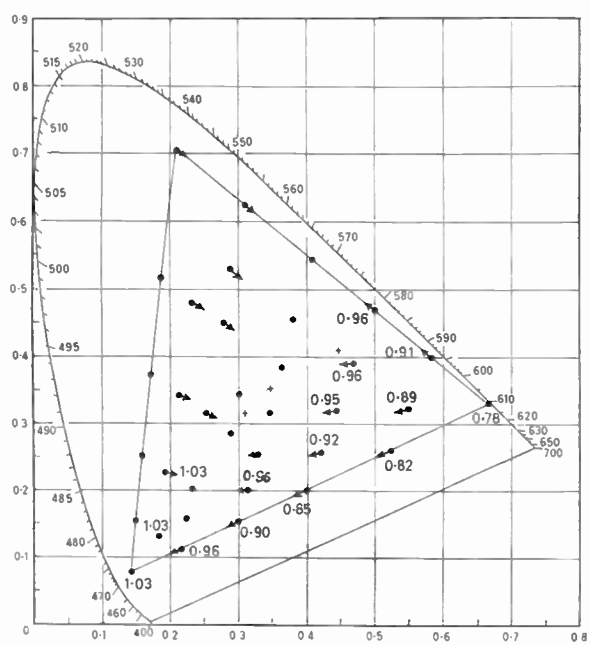


Fig. 12. Effect of a reduction in $R-Y$ drive signal of 17%. (Errors just acceptable for 'flesh tone' pictures.)

These changes were calculated for errors of +5 deg in either X or Z phase, and +10% in either X or Z amplitude. In these four cases the expert observers of the preliminary experiments all accepted the reproduced colour quality of the two critical test slides when the errors were introduced. Positive errors in the triode amplification factors (μ), the theoretical effects of which were given for a +20% error in each valve, were not investigated practically. However, a similar negative error in the mutual conductance (g_m) of each of the valves produced visual effects which were accepted by the observers.

Birt and Freeman also showed the errors that would be produced by a 'Shadowmask' tube (with phosphate phosphors) which requires 20% more colour difference drive signal on one gun than it actually receives. These errors were given for each of the three guns. (The same authors have shown elsewhere the effect of 50% more required drive signal for each gun.[5]) A tube requiring 20% more drive signal on one gun corresponds to a nominal tube which is driven with a signal whose amplitude has been reduced by 17%. Similarly, 50% more drive signal corresponds to a reduction of 33% in the signal driving a nominal tube.

In the first experiment all the expert observers accepted the reproduced colour quality when the $G-Y$ signal was reduced by more than 33%. However, this was not the case for the $R-Y$ and $B-Y$ signals. For both slides all observers rejected the colour quality when the $R-Y$ signal amplitude was reduced by 30%. They also rejected the colour quality of the slide 'Basket of Fruit' when the $B-Y$ signal amplitude was reduced by 20%, although the slide showing flesh tones ('Girl in Headscarf') stood up to this treatment rather better. Its colour quality was not rejected until the $B-Y$ signal amplitude was reduced by 40%.

More detailed information on the acceptability of the errors calculated by Birt and Freeman can be obtained from the results of the second experiment described in this paper. Referring back to Section 4.3 and Table 2 it is clear that when the $R-Y$ signal was reduced by 17% nearly 10% of the observers rejected the colour quality of the 'flesh tone' group of slides. The calculated errors for this condition are shown in Fig. 12 in which the arrows show the changes in chromaticity of certain colours and the numbers show the ratio of reproduced to intended luminance of these colours. This diagram therefore represents errors which were only just acceptable for the flesh tone slides (using the definition of the limit of acceptability given in Section 4.3).

In Fig. 13 are shown the calculated errors for a reduction in $R-Y$ signal of 33%. These errors were judged to be unacceptable by more than 10% of the observers for flesh tone slides but were just acceptable for the 'still life' group of slides.

World Radio History

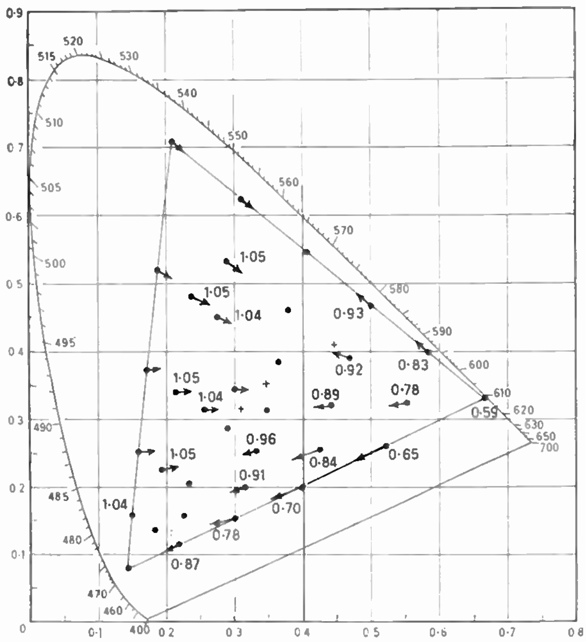


Fig. 13. Effect of a reduction in $R-Y$ drive signal of 33%. (Errors not acceptable for 'flesh tone' pictures, but just acceptable for 'still life' pictures.)

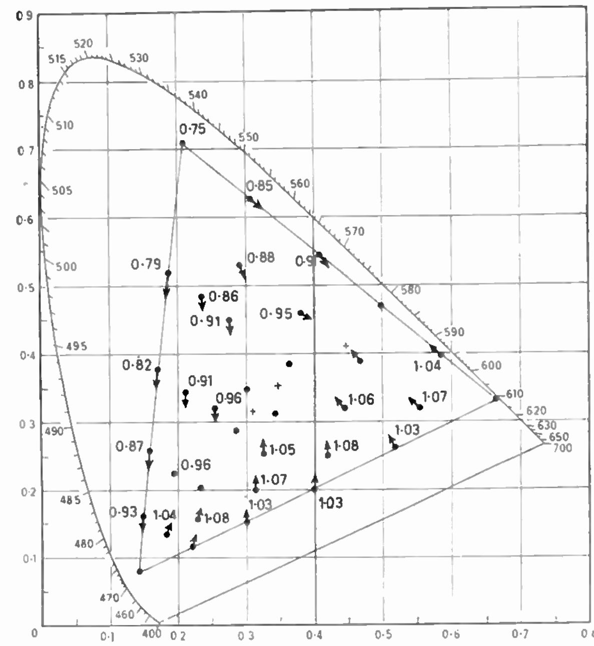


Fig. 14. Effect of a reduction in $G-Y$ drive signal of 33%. (Errors acceptable for 'flesh tone' and 'still life' pictures.)

Similar comparisons between calculated errors and subjective test results can be made for the remaining two signals. The calculated errors for a reduction in $G-Y$ signal of 33% are shown in Fig. 14. These were rejected by fewer than 10% of the observers. Figures 15 and 16 show the errors for reductions of 17% and 33% respectively in the $B-Y$ signal. The former of these was judged to be acceptable for both groups of slides. However, a reduction of 33% in this signal was unacceptable to more than 10% of the observers for the 'still life' slides although just acceptable for the 'flesh tone' slides.

There are two important points influencing this comparison of subjective test results with theoretical error diagrams. Firstly, the viewers judged many colours at the same time and could not help being influenced by the arrangement of the colours in the pictures they saw. Judgements based on single-coloured objects shown against a neutral background could be expected to give rather different results.

Secondly, the actual errors introduced in the practical test could have differed somewhat from those computed. In this experiment, the colour co-ordinates of the phosphors in the 'Shadowmask' tube of the video monitor were close to the primaries of the slide scanner (which in turn were the nominal F.C.C. primaries). Other errors due to the three guns of the tube (such as differences in drive signal required and differences in 'gamma' between individual guns) were eliminated in the setting-up of the particular monitor that was used. However, there was a difference between the value of the overall 'gamma' assumed in the theory ($\gamma = 1$) and that which applied throughout the tests (of the order of 1·4).

5.3. *Comparison with Weiss's Work*

A direct comparison of the work reported here with Weiss's experimental work[2] was not possible. The comment scale used in his experiments was the standard N.T.S.C. quality scale which is unfortunately sufficiently different from the standard E.B.U. quality scale used in the experiments described in Section 4 to make it difficult to compare the two sets of results. His viewing conditions (not quoted in detail) may also have been different from those described above. A further point of difference is that his viewers (classed as 'technical') were asked to 'judge chromaticity differences while ignoring brightness changes'. The observers used in the tests of Section 4 were told to judge the absolute colour quality, according to the quality scale, no special reference to luminance changes being made.

However the calculations carried out by Weiss may be compared directly with those of Birt and Freeman. According to Weiss (Fig. 7 of his paper), the error in the reproduction of flesh tone for a 17% reduction in the $R-Y$ drive signal is approximately 8 just noticeable colour difference units (j.n.c.d. units),

World Radio History

INSTITUTION OF ELECTRONIC AND RADIO ENGINEERS

Conference on 'Electronic Engineering in Oceanography'

at the UNIVERSITY OF SOUTHAMPTON

Monday, 12th September to Thursday, 15th September, 1966

Conference Arrangements

The final programme will include about 40 papers and short contributions, more than a third of which will be by authors from countries outside Great Britain, e.g. from Canada, France, Germany, Japan, Norway and the U.S.A. Working and static demonstrations in the laboratory and afloat will support a number of the papers. The sessions will take place in the Lanchester Building of the University's Engineering Department. The full list of papers with synopses will be published in the I.E.R.E. Journal, *The Radio and Electronic Engineer*, in August 1966.

The Organizing Committee for the Conference is as follows: P. W. Warden (Chairman); R. Bowers; L. P. Frith; E. J. Grimoldby; M. J. Tucker; V. G. Welsby.

Papers. Preprints of the formal papers to be read at the Conference will be prepared and posted to delegates who have registered *before* 1st September. A supplementary volume containing reports of the discussions etc., will be sent subsequently to all who attend the Conference. The full set of papers and discussion will be produced as a complete *Proceedings* of the Conference (I.E.R.E. Conference Proceedings No. 8) and will be available at the end of the year for general sale, price £6 per copy. (Orders can be accepted now.)

Registration. Registration charges cover attendance at all Sessions, preprints and supplement, and lunch and refreshments in the morning and afternoon on each day.

The registration charge is £10 for the full 3½ days of the Conference. Research Students whose applications are supported by their Heads of Departments may register to attend at a charge of £5.

Registration forms for the Conference may be obtained by application to the Institution, 8-9 Bedford Square, London, W.C.1.

Accommodation. There will be residential accommodation in the University's Connaught Hall of Residence for those attending the Conference. The charge will be 40s. per night covering dinner, bed and breakfast. Accommodation will be available for married couples.

[P.T.O.

....

To the Secretary, Conference Secretariat,
I.E.R.E., 8-9 Bedford Square, London, W.C.1.

Conference on Oceanography

Please send me .... registration forms and further information as it becomes available.

*Name..

*Address..

....

(*Block letters please) Date.....................................

World Radio History

Conference on 'Electronic Engineering in Oceanography'

OUTLINE PROGRAMME

Monday, 12th September.

Opening of Conference by Dr. G. E. R. Deacon, F.R.S., *Director of the National Institute of Oceanography.*

Session I: 'Measurements concerned with the Body of the Sea'.

(a) Physical and Chemical Properties.

Papers will deal with: Properties of the marine environment; Conductivity, salinity, temperature, pressure, velocity of sound, dissolved oxygen and carbon dioxide, and current recording.

CONFERENCE RECEPTION

Tuesday, 13th September.

Session I(a): *continued.*

Session I(b): Fisheries Applications.

Papers will deal with: Biological observations under the sea; Undersea light measuring techniques; Fish echo counting; Low-frequency sound sources.

Wednesday, 14th September.

Session II: 'Measurements concerned with the Sea Surface and Sea Bed'.

Papers will deal with: Detection of sediment distribution; Measurement of magnetic micropulsations; Magnetic fields of ocean waves; Free floating wave meter; Sea wave recording; Side-ways looking sonar; Digital readout echo sounder; Acoustic reflectivity of the sea bed and scattering layers; Underwater seismic survey; Sea gravimeters.

CONFERENCE DINNER

Thursday, 15th September.

Session III: 'Instrument Platforms
(ships, buoys and submerged vehicles, telemetry, data handling and navigation).'

Papers will deal with: Underwater communications; Data collection and logging and telemetering; A free-falling deep sea instrument capsule; Expendable buoys; Oceanographic data acquisition systems; V.L.F. navigation.

Friday, 16th September.

Visit to the National Institute of Oceanography, near Godalming, Surrey.

World Radio History

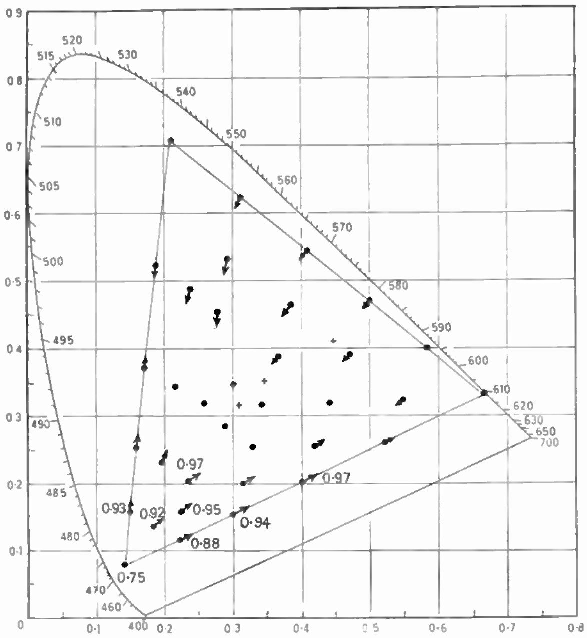


Fig. 15. Effect of a reduction in $B-Y$ drive signal of 17%. (Errors acceptable for 'flesh tone' and 'still life' pictures.)

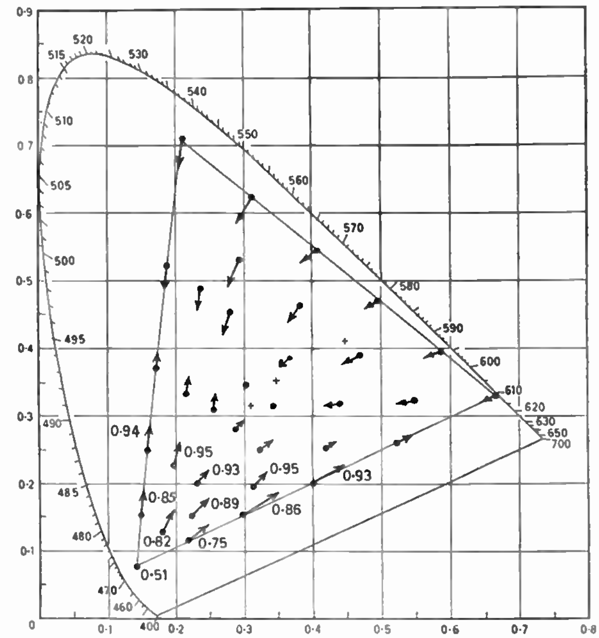


Fig. 16. Effect of a reduction in $B-Y$ drive signal of 33%. (Errors not acceptable for 'still life' pictures, but just acceptable for 'flesh tone' pictures.)

while for a 33% reduction in this signal the error is 15 j.n.c.d. units. Comparable figures for similar reductions in the $B-Y$ signal are 6 j.n.c.d. units and 13 j.n.c.d. units respectively. The errors arising in the reproduction of flesh tones from a reduction in either $R-Y$ or $B-Y$ signals are therefore almost equal in terms of j.n.c.d. units. In contrast to this, the results of the subjective tests reported in this paper clearly show (see Table 2 and also Figs. 6 and 10) that the observers found a reduction of $R-Y$ signal for flesh tone slides very much less acceptable than a similar reduction in $B-Y$ signal. This may partly be accounted for by the design of the subjective tests (see, for example, the results quoted in Section 4.3.1), but it is in agreement with the variations found by Weiss in the constant of proportionality connecting just noticeable and just acceptable colour differences.

5.4. *Practical Implications of the Results*

It appears from the results described above that the use of normal tolerance components in a chrominance signal decoding circuit of the X/Z type is unlikely to produce unacceptable colour errors. However, the results apply only in those cases where not more than one circuit error occurs at once. This is a restriction which may be exceeded in practice and so it seems good policy to follow the recommendations of Birt and Freeman in choosing a decoding circuit which has the least susceptibility to errors, namely one of the $(R-Y)/(B-Y)$ type.[1] This type of decoder would also have the desirable property of allowing the provision of pre-set control of the amplitude of the signals driving the display tube. These adjustments cannot easily be provided in the X/Z type of decoder.

6. Acknowledgments

The author would like to thank those members of the Research Laboratories staff who acted as observers in these experiments. He is also indebted to those members of the Television Display Laboratory who helped in many ways, in particular J. A. Matthews who assisted in running the tests, and P. A. Smithers who supplied the circuit for the adders. Acknowledgment is also due to the Director of Mullard Research Laboratories for permission to publish this paper.

7. References

1. D. R. Birt and K. G. Freeman, 'An appraisal of some decoding systems for three-gun N.T.S.C. colour television receivers', *The Radio and Electronic Engineer*, **28**, No. 1, p. 33, July 1964.
2. H. Weiss, 'Significance of some receiver errors to colour reproduction', *Proc. Inst. Radio Engrs*, **42**, p. 1380, September 1954.
3. D. L. MacAdam, 'Quality of colour reproduction', *Proc. I.R.E.*, **39**, p. 468, May 1951.
4. Dean Farnsworth, 'The Farnsworth-Munsell 100-hue and dichotomous tests for color vision', *J. Opt. Soc. Amer.*, **33**, No. 10, p. 568, October 1943.
5. D. R. Birt and K. G. Freeman, 'Appraisal of the X/Z Decoder for N.T.S.C. Colour Receivers', Mullard Research Laboratories Report No. 488, August 1963.

8. Appendix 1

Design of phase controls for X and Z signals

If an incoming chrominance signal E, where E is given by

$$E = \frac{1}{1{\cdot}14}\left[(R-Y)\cos\omega t + \frac{1}{1{\cdot}78}(B-Y)\sin\omega t\right] \quad \ldots(1)$$

is synchronously detected by a reference signal of $a\sin(\omega t+\theta)$, the filtered resultant E_1 is given by

$$E_1 = \frac{a}{2{\cdot}28}\left[(R-Y)\sin\theta + \frac{1}{1{\cdot}78}(B-Y)\cos\theta\right] \quad \ldots(2)$$

The parameters of an exact design of X/Z decoder for a nominal phosphate tube, as given in Section 3.1.2, part (b), of Birt and Freeman's paper,[1] are:

X demodulation phase	=	244·2°
Z demodulation phase	=	197·0°
Relative X/Z demodulation gain	=	0·91
Ratio of signal on green triode grid to that on blue triode grid	=	0·27
Value of common cathode load	=	612 ohms

Omitting the common gain factor $a/2{\cdot}28$ in eqn. (2) above, we have that the correct fractions of the red and blue colour difference signals to form the exact X signal of this design are respectively sin 244·2° and 1/1·78 cos 244·2°. In the following table, the fractions of the signals $R-Y$ and $B-Y$ to give X and Z over a range of values of θ of ± 10° about the nominal values are given.

Table 4

θ°	Fraction of $R-Y$	Fraction of $B-Y$
187·0	0·122	0·558
189·5	0·165	0·554
192·0	0·208	0·549
194·5	0·250	0·544
197·0	0·292	0·537
199·5	0·339	0·530
202·0	0·375	0·521
204·5	0·415	0·511
207·0	0·454	0·501
234·2	0·811	0·329
236·7	0·836	0·308
239·2	0·859	0·288
241·7	0·881	0·266
244·2	0·900	0·244
246·7	0·918	0·222
249·2	0·935	0·200
251·7	0·949	0·176
254·2	0·962	0·153

9. Appendix 2

Distribution of results

9.1. *R − Y Amplitude Control*

Table 5(a)

'Flesh tone' slides

Signal error (%)	Percentage in grade						Total no. of comments
	1	2	3	4	5	6	
−45	0	7·6	21·2	43·9	24·2	3·0	66
−25	1·7	11·7	40·0	40·0	6·7	0	60
0	8·7	44·9	39·9	6·5	0	0	138
25	13·6	54·6	27·3	4·5	0	0	66
45	30·3	39·4	25·8	4·5	0	0	66

Table 5(b)

'Still life, etc.' slides

Signal error (%)	Percentage in grade						Total no. of comments
	1	2	3	4	5	6	
−45	4·5	21·2	47·0	22·7	4·5	0	66
−25	3·0	43·9	37·9	13·6	1·5	0	66
0	17·4	37·1	33·3	11·4	0·8	0	132
25	18·2	45·4	25·8	10·6	0	0	66
45	24·2	43·9	18·2	12·1	1·5	0	66

9.2. *G − Y Amplitude Control*

Table 6(a)

'Flesh tone' slides

Signal error (%)	Percentage in grade						Total no. of comments
	1	2	3	4	5	6	
−45	15·0	53·3	26·6	5·0	0	0	60
−25	11·7	68·3	18·3	1·7	0	0	60
0	20·0	55·0	22·5	2·5	0	0	120
25	25·0	46·7	26·6	1·7	0	0	60
45	23·3	40·0	26·7	8·3	1·7	0	60

Table 6(b)

'Still life, etc.' slides

Signal error (%)	Percentage in grade						Total no. of comments
	1	2	3	4	5	6	
−45	16·7	40·0	35·0	5·0	3·3	0	60
−25	18·3	35·0	31·7	13·3	1·7	0	60
0	13·3	42·5	34·2	8·3	0·8	0·8	120
25	20·0	33·3	38·3	5·0	1·7	1·7	60
45	18·3	38·3	30·0	8·3	5·0	0	60

World Radio History

9.3. *B − Y Amplitude Control*

Table 7(a)

'Flesh tone' slides

Signal error (%)	Percentage in grade						Total no. of comments
	1	2	3	4	5	6	
−45	16·7	42·6	18·5	14·8	7·4	0	54
−25	25·9	40·7	22·2	11·1	0	0	54
0	25·9	42·0	25·9	4·5	1·8	0	112
25	16·7	37·0	25·9	16·7	1·9	1·8	54
45	7·4	25·9	24·1	22·2	14·8	5·6	54

Table 7(b)

'Still life, etc.' slides

Signal error (%)	Percentage in grade						Total no. of comments
	1	2	3	4	5	6	
−45	3·7	22·2	33·3	25·9	13·0	1·9	54
−25	9·3	40·7	35·2	11·1	3·7	0	54
0	9·6	48·1	31·7	9·6	1·0	0	104
25	16·7	44·4	29·6	7·4	1·9	0	54
45	23·1	21·2	32·7	13·5	5·8	3·8	52

Manuscript first received by the Institution on 13th October 1965, and in final form on 16th February 1966. (Paper No. 1048/T34.)

© The Institution of Electronic and Radio Engineers, 1966.

STANDARD FREQUENCY TRANSMISSIONS

(Communication from the National Physical Laboratory)

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

May 1966	GBZ 19·6 kc/s 24-hour mean centred on 0300 U.T.	MSF 60 kc/s 1430–1530 U.T.	Droitwich 200 kc/s 24-hour mean centred on 0300 U.T.	May 1966	GBZ 19·6 kc/s 24-hour mean centred on 0300 U.T.	MSF 60 kc/s 1430–1530 U.T.	Droitwich 200 kc/s 24-hour mean centred on 0300 U.T.
1	− 300·8	− 301·1	− 0·1	17	− 301·8	− 300·2	− 0·5
2	− 300·2	− 301·8	− 0·6	18	− 301·5	− 300·5	+ 0·2
3	− 300·2	− 300·7	− 0·7	19	− 299·8	− 300·5	+ 0·4
4	—	− 300·5	− 1·1	20	− 299·9	− 300·8	+ 0·4
5	− 300·7	− 300·4	—	21	—	− 300·9	—
6	− 300·5	− 300·6	+ 0·3	22	− 300·6	− 301·9	—
7	− 300·5	− 300·9	− 0·6	23	− 301·1	− 299·6	+ 1·1
8	− 299·5	− 301·3	− 0·5	24	− 301·4	− 300·4	+ 1·2
9	− 300·8	− 300·9	− 0·4	25	− 301·2	− 300·9	+ 0·4
10	− 301·0	− 301·1	− 0·2	26	—	− 300·4	—
11	− 299·9	− 301·0	− 0·3	27	− 300·2	− 300·4	+ 0·6
12	− 300·0	− 301·0	− 0·5	28	− 300·4	− 300·6	+ 0·6
13	− 300·9	− 299·7	− 0·4	29	− 300·6	− 300·5	+ 0·9
14	− 300·3	− 299·9	—	30	− 300·5	− 300·1	+ 1·0
15	− 301·4	− 300·7	—	31	—	− 301·1	+ 0·8
16	− 302·0	− 300·5	0				

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

Note: GBR Rugby has been replaced temporarily by GBZ Criggion.

EXTENSION OF MSF STANDARD TIME AND FREQUENCY TRANSMISSIONS ON 60 kc/s

From the 1st June 1966 the MSF standard time and frequency transmissions on 60 kc/s will be radiated for 24 hours a day instead of the present daily one hour period from 14.30–15.30 GMT.

For an interim period the present one hour schedule of A2 modulation pulses consisting of 5 cycles of a 1000 c/s tone for the seconds pulses and 100 cycles of 1000 c/s tone for the minute pulses will be retained.

The new extension to the service for the remaining 23 hours of the day will operate with a modulation of interrupted carrier pulses of 100 milliseconds duration for the seconds pulses and 500 milliseconds duration at the minute; the epoch of the pulses in each case will be at the start of the break in the carrier signal.

The estimated decay time of the pulses is 0·5 milliseconds.

A shut-down period for maintenance will be observed on the first Sunday of every month between 13.00 and 16.00 G.M.T.

World Radio History

INSTITUTION NOTICES

Earl Mountbatten of Burma, F.R.S.

The Royal Society has announced that Admiral of the Fleet the Earl Mountbatten of Burma, K.G., O.M., (Honorary Member), has been elected a Fellow of the Society. The election is made under the Statute of the Royal Society which provides for the election of persons who have rendered conspicuous service to the cause of science.

Lord Mountbatten's participation in scientific work dates back to his service as a signals specialist in the Royal Navy and his appointment at H.M. Signal School, referred to in the Institution's History, 'A Twentieth-Century Institution'. Although the Royal Navy has made many contributions to scientific progress, few naval officers have been elected to Fellowship of the Royal Society. One of the first was Captain William Bligh (of the *Bounty*), a distinguished navigator elected in 1801. Exactly 100 years later the pioneer of wireless telegraphy in the Navy, Admiral of the Fleet Sir Henry Jackson, was similarly honoured.

As chairman of the National Electronics Research Council since its formation, Lord Mountbatten encouraged the development of one of its outstanding achievements—the S.D.I. project (Selective Dissemination of Information). Lord Mountbatten referred to such a computer-controlled system in his Presidential Address to the Institution in 1946.†

Elected a Member of the Institution in 1935, Lord Mountbatten served as President in 1946–48 and as Charter President in 1961–63. His election as an Honorary Member in 1965 demonstrated the Institution's appreciation of his services to radio and electronic science and engineering.

Institution Conferences in 1967

The Institution is planning to hold two major conferences next year, each of which will last three or four days. The Organizing Committees for both Conferences are now being appointed, and offers of papers and other assistance are welcomed. Offers should be addressed to the Secretary, Programme and Papers Committee, I.E.R.E., 9 Bedford Square, London, W.C.1.

Design for Production

The first Conference, on 'Design for Production', is to be held in early summer; it forms part of the Institution's contribution to 'Quality and Reliability Year' (October 1966 to October 1967), which is being co-ordinated by the National Council for Quality and Reliability, on which the Institution is represented by Mr. F. G. Diver, M.B.E. (Member).

The Conference will bring together design engineers and production engineers in the electronics industry, and the programme will cover quality control, value analysis, programming for automatic test procedures and production engineering (tooling and automation). Mass production, batch production and capital market (i.e. 'one off') production will all be considered.

Radio Frequency Measurements and Standards

The second Conference, to be held in the late autumn of 1967, will be on the subject 'Radio Frequency Measurements and Standards' (see Editorial in this issue). The main subdivisions for this Conference will cover measurements and standards for power, impedance, attenuation, noise, and voltage and current (including field strength).

Israeli Section

At its meeting on 19th May, the Institution's Council formally approved the establishment of a Section of the Institution in Israel. Reference was made in the March *Journal* to the petition which members were submitting to the Council, and the following members now comprise the local Committee:

Chairman: Mr. Z. Gladstein (Associate Member).
Hon. Secretary: Mr. R. Danziger (Member).
Treasurer: Mr. J. Ben-Nun (Associate Member).
Membership Secretary:
Lt.-Col. M. Parann (Associate Member).
Programme Secretary:
Mr. H. Langholtz (Associate Member).
Mr. H. Avni (Associate Member).
Mr. I. Jacobs.

Information about the activities of the new Section, may be obtained from Mr. R. Danziger, whose address is: P.O. Box 21045, Tel-Aviv.

Conference on Microwave and Optical Generation and Amplification

The Sixth International Conference on Microwave and Optical Generation and Amplification (formerly the Conference on Microwave Tubes), sponsored jointly by the Electronics Division of the Institution of Electrical Engineers and the Institution of Electronic and Radio Engineers, will take place at the University of Cambridge from 12th to 16th September 1966.

The aim of the conference is to stimulate the exchange of ideas on the research that has taken place since the Fifth Conference, held in Paris in 1964.

The programme will cover microwave and optical generation by solid state, plasma and other active devices, as well as by the classical forms of microwave tubes. In addition it will deal with the application of various principles and phenomena relevant to the generation and amplification of coherent electromagnetic waves, without strict frequency limitations.

Accommodation will be available in colleges, if required. Further information about the conference and registration forms may be obtained from either of the Sponsoring Institutions.

† *J.Brit.I.R.E.*, **6**, No. 6, pp. 221–25, December 1946.

World Radio History

U.D.C. 621.373.7

Analysis of Varactor Multipliers with Idlers

By

J. O. SCANLAN, M.E.†

AND

P. J. R. LAYBOURN, B.A.†

Based on a contribution presented at the Joint I.E.R.E.–I.E.E. Symposium on Microwave Applications of Semiconductors held in London from 30th June to 2nd July 1965.

Summary: A large-signal analysis of varactor harmonic generators with idlers is presented. The analysis is valid for both abrupt junction and graded junction varactors and any value of Q-factor. Numerical results are given for the cases of triplers and quadruplers operating with a second harmonic idler. The analysis uses a Fourier expansion for the voltage across the varactor in terms of the currents flowing (restricted to input, output and selected harmonic frequencies). The results include optimum generator and load resistances as a function of the operating efficiency. The advantage in terms of efficiency of using idlers is demonstrated.

List of Principal Symbols

C	capacitance of varactor
C_0	capacitance of varactor at reverse breakdown voltage
C_0'	capacitance of varactor at zero voltage
E_g	generator e.m.f.
f_c	cut-off frequency at V_R
i_1	amplitude of current through varactor at ω
i_m	amplitude of current through varactor at $m\omega$
i_N	amplitude of current through varactor at $N\omega$
L_1	tuning inductance in input circuit
L_m	tuning inductance in idler circuit
L_N	tuning inductance in output circuit
m	order of multiplication in idler circuit
N	order of multiplication in output circuit
p_1	normalized amplitude of Q_1
p_m	normalized amplitude of Q_m
p_N	normalized amplitude of Q_N
Q	charge on varactor
Q_1	component of Q at fundamental frequency ω
Q_D	mean charge on diode
Q_F	fundamental-frequency Q-factor of biased varactor
Q_m	component of Q at $m\omega$
Q_N	component of Q at $N\omega$
q_1	amplitude of Q_1
q_m	amplitude of Q_m
q_N	amplitude of Q_N
R_g	generator series resistance
$\bar{R}_g$	normalized R_g
R_L	load resistance
$\bar{R}_L$	normalized R_L
r_s	varactor series resistance
v	voltage across varactor
V	diffusion voltage
$V_0 = V - V_R$	
V_B	bias voltage
V_R	reverse breakdown voltage of diode
Z	equivalent output impedance of varactor
γ	index, depending on junction doping
η	efficiency of multiplier
ϕ_m	initial phase angle between Q_1 and Q_m
ϕ_N	initial phase angle between Q_1 and Q_N
ω	fundamental angular frequency

1. Introduction

Varactor harmonic generators are coming into increasing use for the generation of microwave power. The usual practice is to use a low-frequency transistor oscillator followed by a varactor harmonic generator of appropriate order. Two distinct modes of operation are possible. In one case only the fundamental and the desired harmonic frequency currents are allowed to flow in the varactor, while in the second case other selected harmonic currents also flow. The latter are known as idling currents, or simply 'idlers'. The efficiencies obtainable using idlers are known to be higher than corresponding efficiencies without idlers, this increase in efficiency being obtained at the expense of increased circuit complexity. The increase in

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

World Radio History

efficiency obtainable through the use of idlers is particularly noticeable when high order harmonics are required, but in this case it must be compared with the efficiency obtainable using a cascade of low order multipliers without idlers. The non-linear capacitance-voltage law of the varactor is described by an equation of the type

$$C = \frac{C'_0}{\left(1 - \frac{v}{V}\right)^m}$$

V is the diffusion voltage and the index, m, depends on the grading of the doping in the semiconductor, being equal to $\frac{1}{2}$ for an abrupt junction and less than this for other junctions.

Penfield and Rafuse[1] have analysed the abrupt junction ($m = \frac{1}{2}$) varactor multiplier with idlers on a rigorous basis, while Utsonomiya and Yuan[2] have treated the general case in an approximate manner. The approximation involved was that only a limited number of harmonic voltages were assumed to exist, whereas all harmonic voltages are present. The analysis presented in this paper considers all harmonic voltages. Analytic solutions are given for some typical cases with $m = \frac{1}{2}$ and corresponding numerical results for other values of m.

2. Capacitance-voltage Relationships

The capacitance-voltage relationship is described by eqn. (1):

$$C(v) = \frac{C'_0}{\left(1 - \frac{v}{V}\right)^m} \qquad \ldots\ldots(1)$$

where v is the applied voltage and V is the diffusion voltage.

At the reverse breakdown voltage V_R we have

$$C(v) = C_0 = \frac{C'_0}{\left(1 - \frac{V_R}{V}\right)^m}$$

so that

$$C = C_0\left(\frac{V_0}{V - v}\right)^m \qquad \ldots\ldots(2)$$

now

$$Q = \int C\,dV$$

$$= \frac{-C_0 V_0^m}{1-m}(V-v)^{1-m} \qquad \ldots\ldots(3)$$

where

$$V_0 = V - V_R$$

Rearranging and letting

$$\gamma = \frac{1}{1-m}$$

gives

$$v - V_R = V_0\left[1 - \left(\frac{Q}{Q_R}\right)^\gamma\right] \qquad \ldots\ldots(4)$$

where Q_R is the charge at the reverse breakdown voltage.

3. Harmonic Generator with Idlers

Figure 1 shows the equivalent circuit of a harmonic generator with idlers. The filters are such that the mth filter allows current only at the mth harmonic to flow.

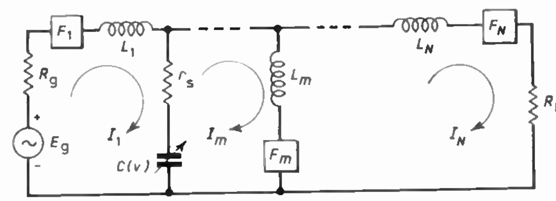


Fig. 1. Circuit of a varactor harmonic generator with idlers.

Let the mean charge on the diode be Q_D and the mth harmonic charge

$$Q_m = q_m \cos(m\omega t + \phi_m)$$

so that

$$I_m = -m\omega q_m \sin(m\omega t + \phi_m)$$

$$= i_m \sin(m\omega t + \phi_m)$$

and for reference $\phi_1 = 0$, the output being at the Nth harmonic. Equation (4) then gives

$$v - V_R = V_0\left[1 - \frac{1}{2^\gamma}\left(\frac{Q_D}{Q_R/2} + \frac{q_1 \cos\omega t}{Q_R/2} - \sum \frac{q_m \cos(m\omega t + \phi_m)}{Q_R/2}\right)^\gamma\right] \qquad \ldots\ldots(5)$$

where the summation is taken over all harmonic currents which are allowed to flow, including the output frequency. Expanding this expression as a Fourier series we have

$$v - V_R = \frac{V_0}{2^\gamma}\left[a_0 + \sum_{n=1}^{\infty}(a_n \cos n\omega t + b_n \sin n\omega t)\right] \qquad \ldots\ldots(6)$$

a_n and b_n are given by

$$a_n = -\frac{1}{\pi}\int_{-\pi}^{\pi}\left(\frac{Q_D}{Q_R/2} + \frac{q_1\cos\omega t}{Q_R/2} - \sum\frac{q_m\cos(m\omega t + \phi_m)}{Q_R/2}\right)^\gamma \cos n\omega t\, d(\omega t) \qquad \ldots\ldots(7)$$

World Radio History

and

$$b_n = -\frac{1}{\pi}\int_{-\pi}^{\pi}\left(\frac{Q_D}{Q_R/2}+\frac{q_1\cos\omega t}{Q_R/2}-\right.$$
$$\left.-\sum\frac{q_m\cos(m\omega t+\phi_m)}{Q_R/2}\right)^{\gamma}\sin n\omega t\,d(\omega t) \qquad \ldots\ldots(8)$$

3.1. *Circuit Equations*

The circuit equations for Fig. 1 may be written as follows: At the fundamental frequency the input circuit gives

$$E_g\sin(\omega t+\alpha) = i_1(R_g+r_s)\sin\omega t+L_1\omega i_1\cos\omega t+ + \frac{V_0}{2^{\gamma}}a_1\cos\omega t+\frac{V_0}{2^{\gamma}}b_1\sin\omega t \qquad \ldots\ldots(9)$$

At the output frequency

$$\frac{V_0}{2^{\gamma}}(a_N\cos N\omega t+b_N\sin N\omega t) = i_N(R_L+r_s)\sin(N\omega t+\phi_N)+i_NL_NN\omega\cos(N\omega t+\phi_N) \qquad \ldots\ldots(10)$$

while at each idler frequency

$$\frac{V_0}{2^{\gamma}}(a_m\cos m\omega t+b_m\sin m\omega t) = i_mr_s\sin(m\omega t+\phi_m)+i_mL_mm\omega t\cos(m\omega t+\phi_m) \qquad \ldots\ldots(11)$$

Resonating the input circuit for maximum efficiency requires

$$L_1 = \frac{-V_0a_1}{2^{\gamma}\omega i_1}$$

so that

$$E_g = i_1(R_g+r_s)+\frac{V_0b_1}{2^{\gamma}} \qquad \ldots\ldots(12)$$

Introducing the following normalizations

$$\bar{E}_g = \frac{2^{\gamma}E_g}{V_0}$$

$$\bar{R}_g = \frac{R_g}{r_s}$$

$$\bar{R}_L = \frac{R_L}{r_s}$$

$$Q_F = \frac{1}{r_sC_0\omega}$$

$$p_1 = \frac{q_1}{Q_R/2}$$

$$q_N = \frac{q_N}{Q_R/2}$$

$$p_m = \frac{q_m}{Q_R/2}$$

eqn. (12) becomes

$$\bar{E}_g = p_1(1+\bar{R}_g)\frac{\gamma 2^{\gamma-1}}{Q_F}+b_1 \qquad \ldots\ldots(13)$$

Now, for optimum power transfer, at a given power level, from the generator to the varactor, the generator resistance should be made equal to the (non-linear) input 'impedance' at the particular drive level. Therefore, from eqn. (13)

$$\bar{R}_g = 1+\frac{Q_Fb_1}{\gamma 2^{\gamma-1}p_1} \qquad \ldots\ldots(14)$$

and eqn. (13) becomes

$$\bar{E}_g = \frac{\gamma 2^{\gamma}}{Q_F}p_1\left(1+\frac{Q_Fb_1}{\gamma 2^{\gamma-1}p_1}\right) \qquad \ldots\ldots(15)$$

Using the normalization given above, eqns. (10) and (11) become

$$\frac{1}{\gamma 2^{\gamma-1}}(a_N\cos N\omega t+b_N\sin N\omega t) = \frac{Np_N}{Q_F}(\bar{R}_L+1)\sin(N\omega t+\phi_N)+ + \frac{L_NN^2\omega}{Q_Fr_s}p_N\cos(N\omega t+\phi_N) \qquad \ldots\ldots(16)$$

and

$$\frac{1}{2^{\gamma-1}}(a_m\cos m\omega t+b_m\sin m\omega t) = \frac{mp_m}{Q_F}\sin(m\omega t+\phi_m)+ + \frac{L_mm^2\omega}{Q_Fr_s}p_m\cos(m\omega t+\phi_m) \qquad \ldots\ldots(17)$$

Rearranging eqn. (17) gives

$$p_m = \frac{Q_F}{m\gamma 2^{\gamma-1}}(a_m\sin\phi_m+b_m\cos\phi_m) \qquad \ldots\ldots(18)$$

$$\frac{L_mm\omega}{r_s} = \frac{Q_F}{m\gamma 2^{\gamma-1}p_m}(a_m\cos\phi_m-b_m\sin\phi_m) \qquad \ldots\ldots(19)$$

3.2. *Breakdown Curve*

In order that the diode is not driven into forward conduction or beyond the reverse breakdown voltage V_R, the instantaneous charge must lie between the limits

$$0 \geqslant Q(t) \geqslant Q_R$$

or

$$0 \geqslant Q_D+q_1\cos\omega t-\sum q_m\cos(m\omega t+\phi_m) \geqslant Q_R$$

Using the normalization previously given this becomes

$$2 \geqslant p_0+p_1\cos\omega t-\sum p_m\cos(m\omega t+\phi_m) \geqslant 0 \qquad \ldots\ldots(20)$$

World Radio History

3.3. *Efficiency and Output Power*

Defining efficiency as

$$\eta = \frac{\text{output power}}{\text{available power from source}}$$

we have

$$\text{output power} = \frac{i_N^2 R_L}{2}$$

$$\text{available power from source} = \frac{E_g^2}{8R_g}$$

But since it is assumed that the input is matched

$$E_g = 2i_1 R_g$$

and

$$\eta = \frac{i_N^2 R_L}{i_1^2 R_g} = \frac{N^2 p_N^2 \bar{R}_L}{p_1^2 \bar{R}_g}$$

From eqn. (16) this becomes

$$\eta = \frac{\left(\frac{Q_F}{\gamma 2^{\gamma-1}}\right)^2 (a_N \sin\phi_N + b_N \cos\phi_N)^2 \bar{R}_L}{p_1^2(1+\bar{R}_L)^2 \bar{R}_g}$$

and substituting the matched value for $\bar{R}_g$ from eqn. (14)

$$\eta = \frac{\left(\frac{Q_F}{\gamma 2^{\gamma-1}}\right)^2 (a_N \sin\phi_N + b_N \cos\phi_N)^2 \bar{R}_L}{p_1^2(1+\bar{R}_L)^2 \left(1 + \frac{Q_F b_1}{\gamma 2^{\gamma-1} p_1}\right)} \qquad \ldots\ldots(21)$$

In principle any circuit may now be analysed using eqns. (7), (8), (15), (16). and (17). However, it is evident that in general this process is extremely difficult even using numerical techniques.

Since maximum efficiency and power output will be obtained by operating on the breakdown curve as given by eqn. (20), this may be used in the analysis and leads to considerable simplification.

4. 'Linearized' Resonance

In order to obtain maximum efficiency or power output some form of resonance in the idler and output circuits is necessary. Since the problem is essentially non-linear no direct method exists for determining the optimum phase conditions in these circuits. However, if a 'linearized' resonance condition, as described below, is used results which differ by only a small amount from the true optima are found, and numerical investigation shows that for any reasonable value of Q_F the difference is completely negligible.

The voltage given by $a_m \cos m\omega t + b_m \sin m\omega t$ (where m may represent an idler or the output frequency) is composed of two parts, one a function of p_m and one independent of p_m. The portion independent of p_m may be regarded as a constant voltage generator as far as the mth harmonic is concerned, and the remainder represents the voltage across a non-linear impedance. If the portion independent of p_m is $k_m \cos m\omega t + l_m \sin m\omega t$, then 'linearized' resonance occurs when the mth harmonic current is in phase with this voltage.

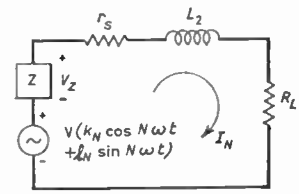


Fig. 2. Output circuit of the harmonic generator.

The arrangement is shown for the output circuit, in Fig. 2 where $m = N$. k_m and l_m are given by setting the appropriate p_m to zero in eqns. (7) and (8), and the other coefficients at the values which result when p_m is zero. This process is, in general, quite difficult but in the case of an abrupt junction varactor ($\gamma = 2$) can readily be carried out.

5. Results for Triplers and Quadruplers

Let us now consider the specific cases of triplers and quadruplers operating with a second harmonic idler. Thus $m = 2$, $N = 3$ or 4. We first evaluate k_2, l_2 for the abrupt junction case. Setting p_2 to zero automatically sets p_3 or p_4 to zero since this diode is only capable of generating second harmonic. Thus

$$k_2 = -\frac{1}{\pi}\int_{-\pi}^{\pi} (p_0 + p_1 \cos\omega t)^2 \cos 2\omega t \, \mathrm{d}(\omega t) = -\frac{p_1^2}{2}$$

$$l_2 = -\frac{1}{\pi}\int_{-\pi}^{\pi} (p_0 + p_1 \cos\omega t)^2 \sin 2\omega t \, \mathrm{d}(\omega t) = 0$$

Thus

$$\phi_2 = \pm\frac{\pi}{2}$$

Evaluating k_3, l_3

$$k_3 = -\frac{1}{\pi}\int_{-\pi}^{\pi} \left\{p_0 + p_1 \cos\omega t + \bar{p}_2 \cos\left(2\omega t \pm \frac{\pi}{2}\right)\right\}^2 \times \times \cos 3\omega t \, \mathrm{d}(\omega t)$$

World Radio History

where $\bar{p}_2$ is the value of p_2 in the absence of a third harmonic output. Hence

$$k_3 = 0$$

$$l_3 = -\frac{1}{\pi}\int_{-\pi}^{\pi}\left\{p_0 + p_1\cos\omega t + \bar{p}_2\cos\left(2\omega t \pm \frac{\pi}{2}\right)\right\}^2 \times$$
$$\times \sin 3\omega t\, \mathrm{d}(\omega t)$$
$$= \pm\, p_1 \bar{p}_2$$

Hence

$$\phi_3 = 0 \text{ or } \pi$$

Similarly for the quadrupler

$$\phi_4 = \pm\frac{\pi}{2}$$

Using eqn. (18)

$$p_2 = \pm\frac{Q_F a_2}{2\gamma 2^{\gamma-1}} \qquad p_2 > 0 \quad \ldots\ldots(22)$$

and similarly for a tripler

$$p_3 = \pm\frac{Q_F b_3}{3\gamma 2^{\gamma-1}(1+\bar{R}_L)} \qquad p_3 > 0 \quad \ldots\ldots(23)$$

and for a quadrupler

$$p_4 = \pm\frac{Q_F a_4}{4\gamma 2^{\gamma-1}(1+\bar{R}_L)} \qquad p_4 > 0 \quad \ldots\ldots(24)$$

Explicit solutions may now be obtained for the abrupt-junction diode since the a and b may be analytically evaluated.

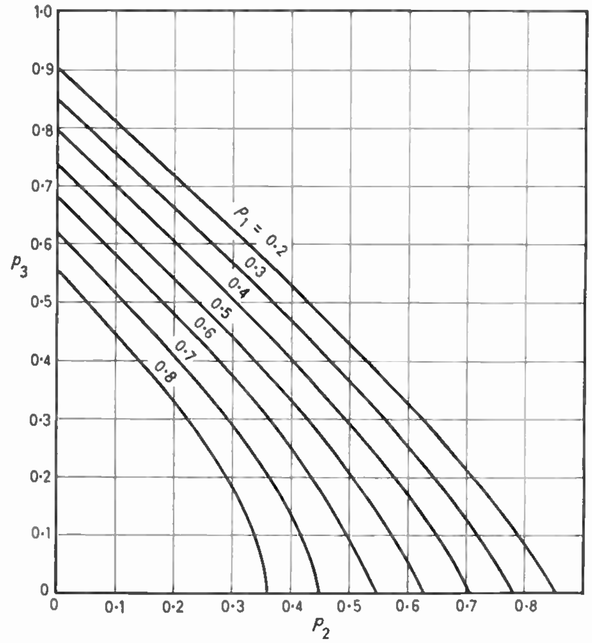
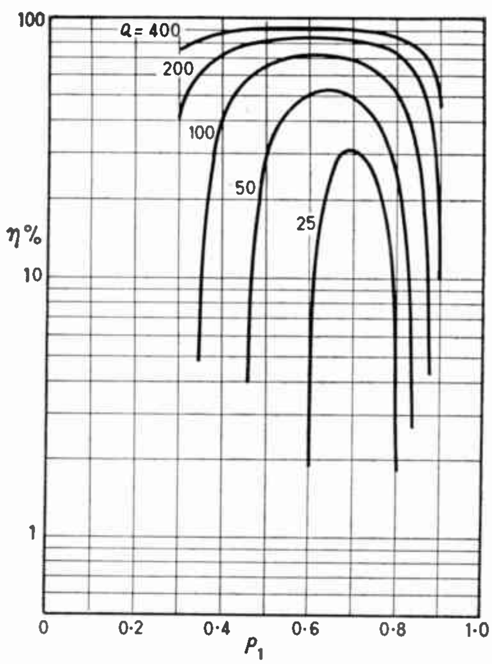
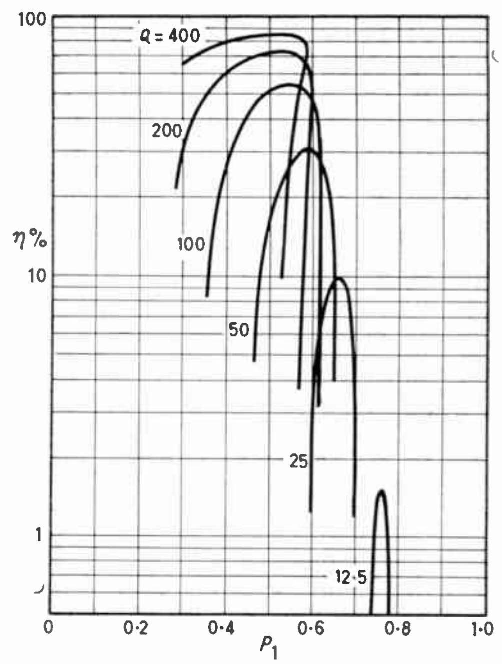


Fig. 3. Breakdown curves for a tripler with second harmonic idler.

In general it is now assumed that the phase angles for a graded junction diode are the same as for the corresponding abrupt junction case and the method of solution is as follows:



(a) Abrupt junction tripler with second harmonic idler.



(b) Abrupt junction quadrupler with second harmonic idler.

Fig. 4. Efficiency as a function of p_1.

World Radio History

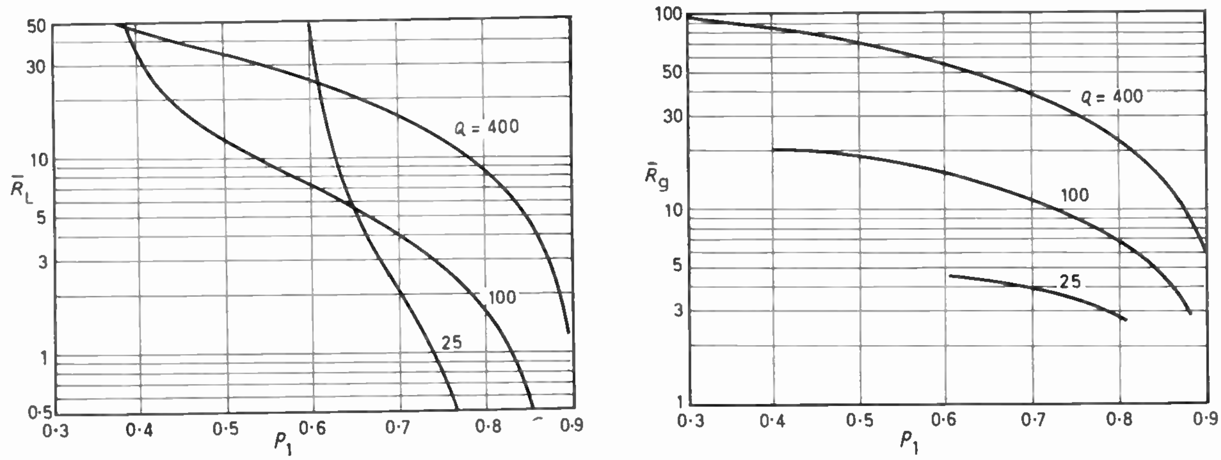


Fig. 5(a). Load and source resistances as a function of p_1 for abrupt junction tripler.

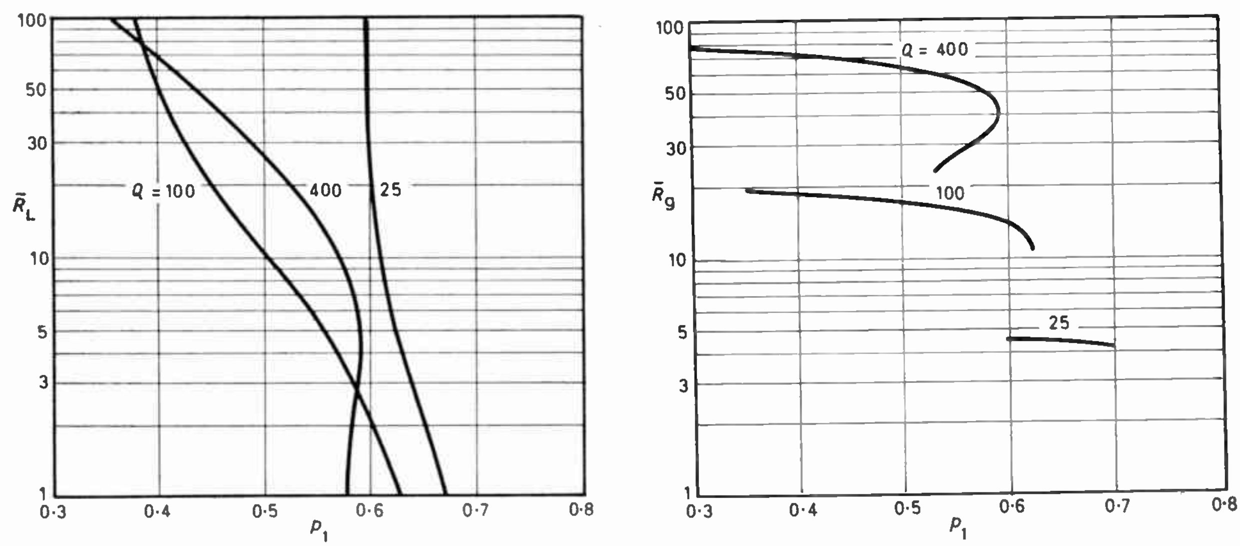


Fig. 5(b). Load and source resistances as a function of p_1 for abrupt junction quadrupler.

Firstly the angles as found above result in symmetrical charge waveform and therefore for maximum efficiency and output power, p_0 is chosen to be unity. The equation for the breakdown curve and eqn. (18) are used, at a particular value of p_1, in order to obtain p_N. Equation (16) is then used to obtain the corresponding value of $\bar{R}_L$, and $\bar{R}_g$ is obtained from eqn. (14). Hence the output power and efficiency are determined. This is repeated for other values of p_1.

Figure 3 shows a set of breakdown curves for the tripler operating with second harmonic idler. Figures 4 (a) and (b) show efficiency as a function of p_1 for the tripler and quadrupler respectively. It is interesting to note that for p_1 greater than a particular value no output is obtained since the condition that operation be confined to the breakdown curve demands a negative load resistance, while for p_1 less than a particular value the equations again cannot be satisfied. This does not of course indicate that no output can be obtained at these values of p_1 but merely that the conditions imposed in order to obtain the point of maximum efficiency may not be fulfilled. The fact that the curves of Fig. 4 (b) may yield two results for efficiency at a particular value of p_1 indicates that there exist two values of $\bar{R}_L$ which will cause operation on the breakdown curve. The corresponding values of load resistance for the tripler and quadrupler are shown in Figs. 5(a) and (b).

Figure 6(a) shows the maximum efficiency and 6(b) the maximum output power as a function of f/f_c for

World Radio History

$\gamma = 2{\cdot}0$ and $\gamma = 1{\cdot}5$ for the tripler. The corresponding load and input resistances are shown in Fig. 7. Figures 8(a) and (b) show maximum efficiency and output power for the quadrupler and Fig. 9 shows the required load and source resistances.

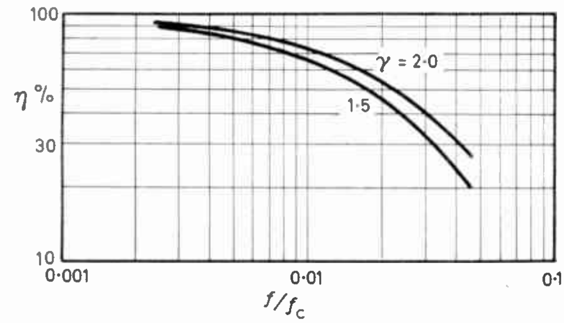


Fig. 6(a). Maximum efficiency as a function of normalized frequency for a tripler.

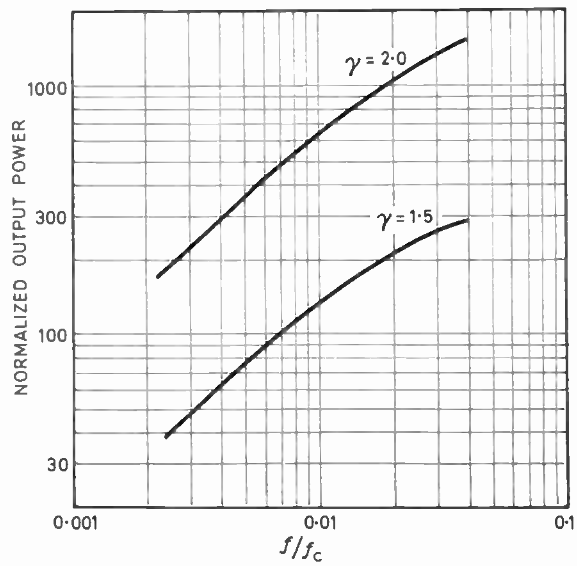


Fig. 6(b). Maximum output power as a function of normalized frequency for a tripler.

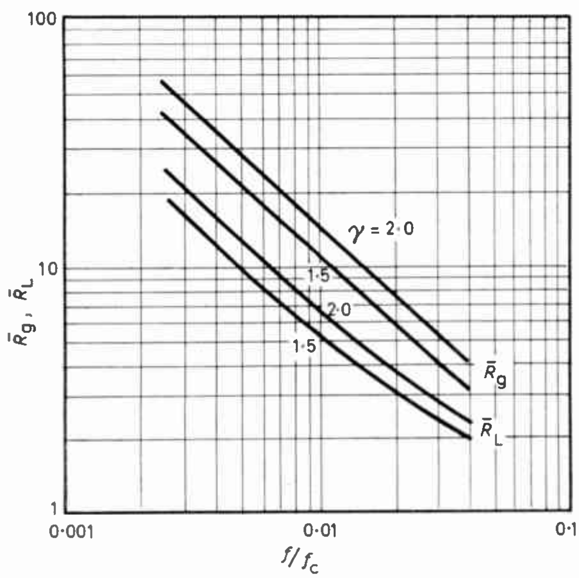


Fig. 7(a). Load and input resistances to produce maximum efficiency in a tripler.

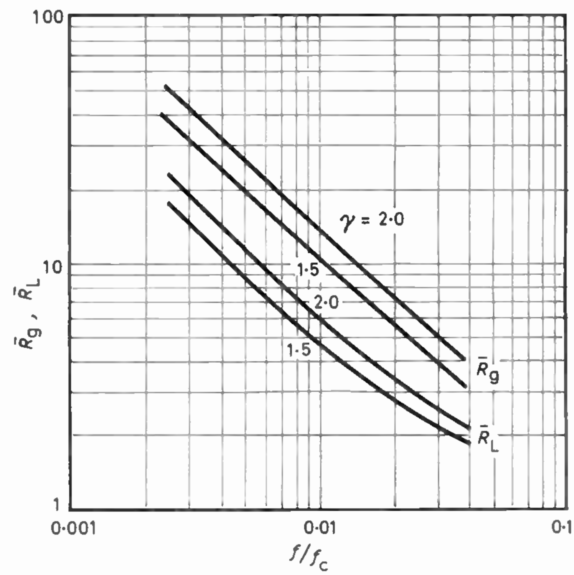


Fig. 7(b). Corresponding resistances to produce maximum output power.

6. Conclusions

A method of analysis for varactor harmonic generators with idlers has been described and results presented for both abrupt and graded junction triplers and quadruplers operating with a second harmonic idler. It is interesting to compare the results obtained[3] in the case of a graded junction diode with and without a second harmonic idler. This comparison is shown in Fig. 10 from which the considerable advantage of using an idler is clear. However, since the use of an idler increases the circuit complexity, it may be sufficient to operate without one in cases where high efficiency is not of prime importance. Another interesting comparison is the possibility of using a

World Radio History

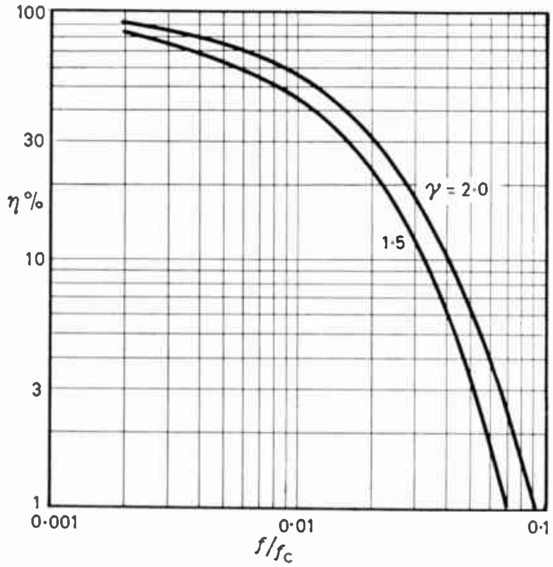


Fig. 8(a). Maximum efficiency as a function of normalized frequency for a quadrupler.

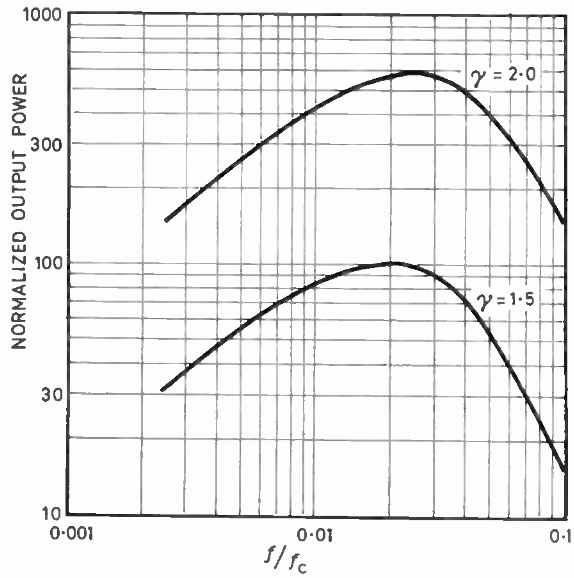


Fig. 8(b). Maximum output power as a function of normalized frequency for a quadrupler.

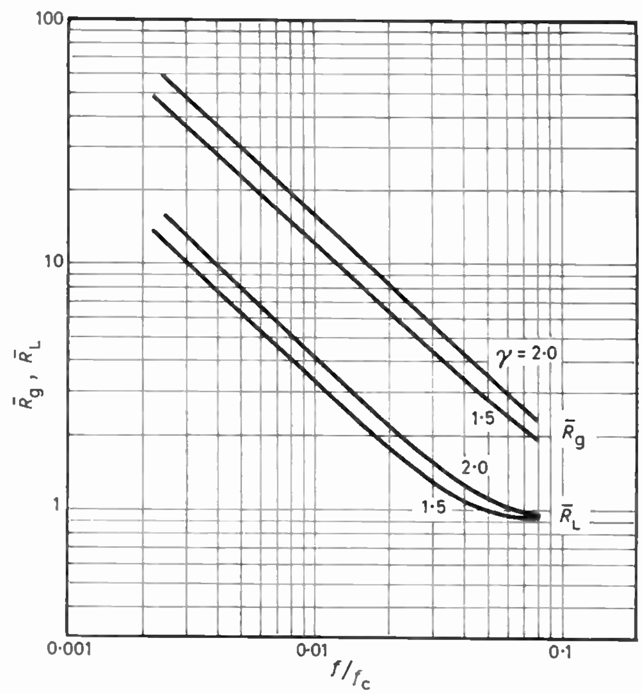


Fig. 9(a). Load and input resistances to produce maximum efficiency in a quadrupler.

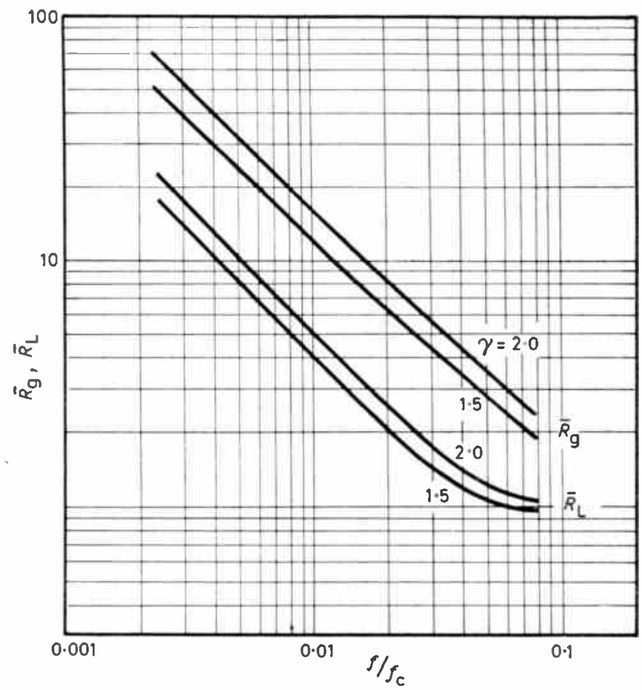


Fig. 9(b). Corresponding resistances to produce maximum output power.

World Radio History

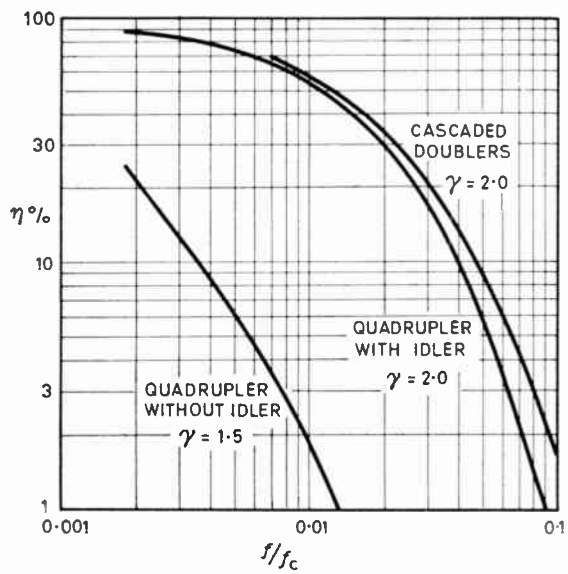


Fig. 10. Comparison of efficiencies obtained with a quadrupled output.

cascade of two doublers in order to obtain a quadrupler output. This curve is also shown in Fig. 10 on the basis of equal input frequencies and fully-driven diodes of equal cut-off frequency.

7. Acknowledgment

This paper is published by the permission of the Ministry of Defence.

8. References

1. P. Penfield and R. P. Rafuse, 'Varactor Applications' (M.I.T. Press, Cambridge, Mass., U.S.A., 1962).
2. T. Utsonomiya and S. Yuan, 'Theory, design and performance of maximum-efficiency variable-reactance frequency multipliers', *Proc. Inst. Radio Engrs*, **50**, No. 1, pp. 57–65, January 1962.
3. J. O. Scanlan and P. J. R. Laybourn, 'Large signal analysis of varactor harmonic generators without idlers', *Proc. Instn Elect. Engrs*, **112**, No. 8, pp. 1515–22, August 1965.

Manuscript first received by the Institution on 24th May 1965 and in final form on 1st March 1966. (Paper No. 1049.)

© The Institution of Electronic and Radio Engineers, 1966

World Radio History

U.D.C. 601.375.7 : 621.372

Parametric Travelling-Wave Amplification along Three Transmission Lines Coupled by Diode Quartets

By

Peter BOBISCH, Dipl.Ing.†

AND

Claus SONDHAUSS, Ing.†

Reprinted from the Proceedings of the Joint I.E.R.E.–I.E.E. Symposium on 'Microwave Applications of Semiconductors' held in London from 30th June to 2nd July 1965.

Summary: Each of the three waves (pump, signal, and idler) is transmitted along an independent line of appropriate characteristic impedance coupled only by equidistant reactance-diode quartets. Such a quad is composed of four diodes mounted in one plane perpendicular to the direction of wave propagation and connected as a ring-type modulator.

The three transmission lines consist of a rectangular waveguide with Lecher wires parallel to its axis. The pump wave (8·45 Gc/s) is transmitted in a waveguide mode through the waveguide, the idling wave (5 Gc/s) in the Lecher mode along the double conductor. The signal wave (3·45 Gc/s) propagates in the coaxial mode, the double line being the inner conductor and the waveguide walls being the outer conductor.

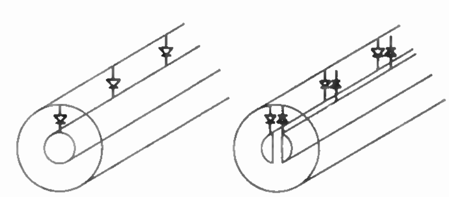
Test results show a signal gain of greater than 10 dB within the instantaneous bandwidth from 3·1 Gc/s to 3·75 Gc/s.

1. Explanation of Travelling-Wave Amplification by Diode Quartets

This paper deals with the parametric travelling-wave amplification with reactance-diodes and presents a novel travelling-wave structure. Travelling-wave amplifier development will first be recalled briefly.

As shown in Fig. 1(a), the simplest parametric travelling-wave amplifier consists of a single transmission line, loaded by equidistantly inserted individual reactance diodes. Since all the three waves, pump, signal, and idler, propagate along the same transmission line, its pass-band must include the frequencies of all the three waves. Therefore it is impossible in practice to meet the phase conditions required for maximal and broad-band signal amplification. A more suitable device, shown in Fig. 1(b), is the amplifier invented by Engelbrecht,[1] which differs in having a split inner conductor and in the arrangement of diode pairs. By this split construction a Lecher system can be made to transmit the pump wave, whereas the signal and idling waves propagate in the coaxial mode as before. As the result of the separate pump line the optimum phase relations between the three waves can more easily be satisfield. Furthermore, the formation of diode pairs leads to a push-pull arrangement, so that the pump wave is decoupled from the signal and idler wave.

But this type of travelling-wave amplifier still presents some disadvantages. Because the signal and idler wave run along the same transmission line, it is impossible on account of line dispersion to separate the signal and idler frequencies too far. In any case the signal and idler cross over at the point of half the pump frequency, so that no more than half of the whole transmission line bandwidth is useful for signal amplification. The signal frequency adjoins the idler frequency, and therefore the noise of the idler termination appears with full magnitude at the signal input, so that the noise factor cannot be reduced below the 3-dB level. The capacitance of all the diodes alters with amplitude and frequency of the pump wave. Therefore it varies the capacitive load for the pump line as well as for the signal and idler line respectively, which results in considerable fluctuations of the cut-off frequency and dispersion



(*a*) (*b*)

Fig. 1. Travelling-wave amplification with reactance-diodes.

properties of the transmission lines. These disadvantages vanish if more than two diodes are joined together in the same manner as in the so-called ring-modulator. This four-diode circuit and its use for special work are the substance of this paper.

† Telefunken AG., Ulm (Donau), Germany.

World Radio History

Let us imagine a development of the Wheatstone bridge, in which the four impedances in the bridge arms are replaced by four reactance diodes connected in a ring-circuit manner, and moreover a third port is established, as is shown in Fig. 2. Then, if the four diodes are identical, any port is perfectly decoupled from the two remaining ports. This means that if a single oscillation is applied to any port, it cannot appear at both the remaining ports. But if the three-port is fed with two different oscillations at two of the ports, the third port will deliver the combination frequencies, namely the sum frequency and the difference frequency only. In case of parametric amplification the difference oscillation at the third port represents the idler.

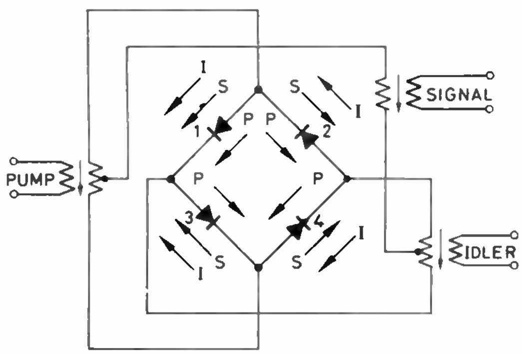


Fig. 2. Circuit of a diode quartet.

Provided that the pump oscillation and the signal oscillation exist at two of the ports, marked by the assumed voltage arrows shown on Fig. 2, then it is quite possible to draw the arrows of the originating idling voltage likewise. Thus the four idler voltages of the quad combine correctly to build up the sum voltage of the idler arrows at the third port. The physical action is described by the fact that the pump and signal voltages across the diodes excite idler currents, which cause the rise of an idler voltage across the resistor in the idling port. Obviously the idler port has this idler voltage in common with the diodes. The pump, signal, and idler ports are freely interchangeable.

In the like manner the power transfer from the pump oscillation to the signal and idler oscillations can be demonstrated. The acceptance of pump power by the signal oscillation within any diode can only occur, if the arrows of signal and idler are equal directed. In the state shown in Fig. 2 the diodes 1 and 3 meet this condition. On the diodes 2 and 4 the signal voltage is just compensated by the idler voltage. Provided that the pump arrow, directed in forward direction of the diode, indicates a decrease of the diode capacitance, and vice versa, the first and third diode are just amplifying, whereas the second and fourth diode, affected by an increase of capacitance, cannot contribute to amplification of the signal, because any power transfer is impossible since the sum of signal and idler charges is zero. In the next state, when the pump or the signal arrow will have turned in the opposite direction, the second and fourth diodes are amplifying, whereas the first and third diodes do not allow any power transfer. Such a quad is consequently able to transfer four times the power out of the pump into the signal that a single diode can. In contrast to a single diode, which is resting during a half-period of every pump swing, there are always two diodes in the quad amplifying in either half-period of the pump swing.

Such a diode quad, or a successive chain of diode quads, has also another qualification for use for parametric travelling-wave amplification. This results from the fact that the capacitances of the four single diodes of a quad are balanced in such a way that the input reactance of each port remains nearly constant, even though the diodes are driven by the pump wave. For example, if the capacitances of the first and third diodes are decreased by the pump wave, the capacitances of the second and fourth diodes are simultaneously increased. Referring to travelling-wave amplification, the dispersion of any transmission line cannot be disturbed by the time dependent loading of all the several diodes.

An assembly is to be devised consisting of three independent transmission lines, which makes possible coupling by the diode quads in certain intervals. The arrangement considered to have the best chance of success is shown on Fig. 3.[2,3] This is a waveguide containing a twin conductor in its axis. The special line system is equidistantly loaded by the diode quads and can be operated in three different modes

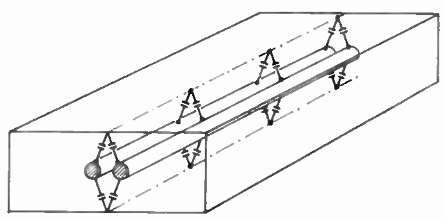


Fig. 3. Parametric travelling-wave amplification along three transmission lines coupled by diode quartets.

at the same time. Firstly, the transmission line can be operated in the waveguide mode, the waveguide becoming a ridge waveguide because of the double line. Secondly, a coaxial mode can propagate, the

World Radio History

double line being the inner conductor and the waveguide walls being the outer conductor of the line. Finally, it can be operated in the Lecher mode, where only the double line transmits one of the waves, the waveguide itself acting as an outer shield. The diodes to couple all the waves are placed between the waveguide walls and the double line in it, with four diodes in each plane perpendicular to the propagation direction.

In principle each of the three waves can propagate in each of the three modes. It seems to be the best combination, however, to transmit the signal wave, being the wave of the lowest frequency, in the coaxial mode, the pump wave in the waveguide mode and the idler wave in the Lecher mode.

Compared with the existing types of parametric amplifiers the advantages of such a travelling-wave amplifier result in using three separate propagation circuits coupled only by diodes. As mentioned before, no variation of the dispersive transmission line properties due to the pump dependent capacitive diode swing may occur. But nevertheless the power flow into the signal wave is effected by the diode quads. Each transmission line needs only to be designed for the usable frequency range of one of the three waves, so that it is quite easy to verify the proper phase and group velocities. Hence it might be possible to achieve a theoretical signal bandwidth of an octave. Using such diode quads no other frequencies than the pump and signal and their sum and difference are generated, nor can any disturbance of the waves being transmitted theoretically occur.

2. Experimental Results

Following theoretical investigations an experimental amplifier for a signal frequency of about 3·4 Gc/s and a pump frequency of 8·5 Gc/s using the travelling-wave structure reported here has been built.

Figure 4, which represents the travelling-wave amplifier, shows the through line for the signal wave (terminated by the coaxial connectors), and the two waveguide flanges which feed the amplifier with pump power as well as absorb the residual pump power. The terminations of the idler wave are internal elements of the amplifier.

Inside the amplifier the inner conductor of the coaxial transmission line is split into a twin conductor in order to enable the idling wave to propagate in the Lecher mode. The absorption of the idler power takes place in cast wedge-shaped dissipative elements, which are fitted in the tapered ends of the slot, terminating the Lecher line with very low reflection.

This test model of a travelling-wave amplifier can be fitted with as many as six diode-quads. Each diode can individually be adjusted to an average junction capacitance of about 0·3 pF. With case capacitance each diode has capacitance load of about 0·5 pF. The separate bias adjustment for all the diodes was accomplished by a d.c. network fed by a battery. The d.c. was supplied to the diodes along special lead-in insulators within the diode fixing-screws.

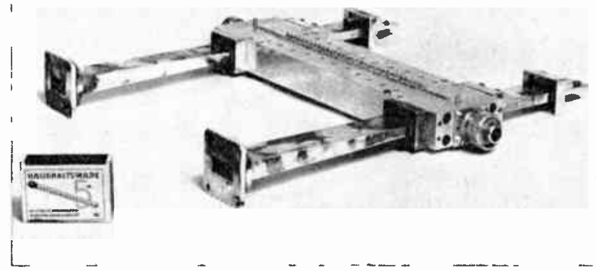


Fig. 4. Parametric travelling-wave amplifier.

Difficulties did arise in meeting the proper matching conditions between the active signal line of the amplification section, periodically loaded and provided with a certain phase velocity by the diode quads at 0·5 cm intervals, and the characteristic impedance of commonly-used 50- or 60-ohm lines. Additional problems arose because the decoupling of the pump- and the signal waves from one another must not be disturbed. These problems are still not wholly solved. This amplifier design provides decoupling between pump and signal of the order of 20 dB only. Therefore a pump absorption filter is inserted behind the signal output to prevent the pump wave magnitude becoming higher than the signal magnitude.

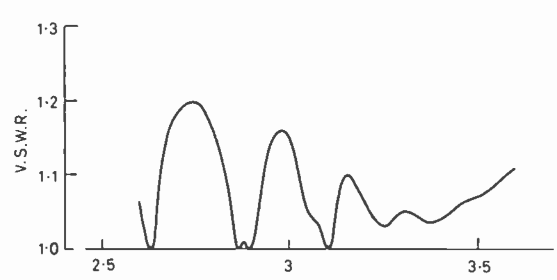


Fig. 5. Matching of the signal input of the parametric travelling-wave amplifier.

The impedance matching and the transition from the periodically-loaded transmission line system to the homogeneous coaxial line or waveguide result from the continuous decrement of the equidistantly mounted loadings in steps like a geometrical progression. Figure 5 shows the signal port input matching of the amplifier as the voltage-standing-wave-ratio against frequency.

World Radio History

In operation this well-matched input impedance will certainly not keep constant. This is especially true in all those cases where some of the pretuned diodes require an individual bias voltage to produce a smoothed gain characteristic.

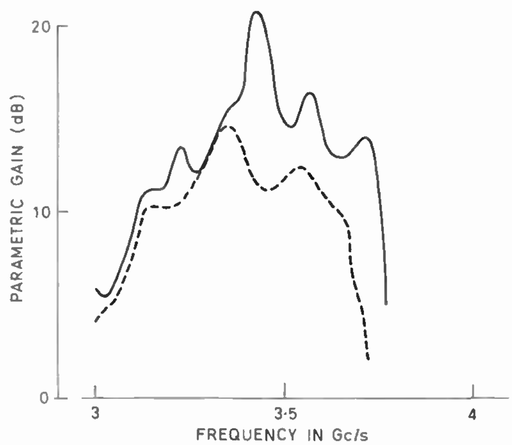


Fig. 6. Signal gain of the parametric travelling-wave amplifier. Six diode quartets; pump frequency: 8·45 Gc/s; pump power: 1W; signal power: 2 μW.

Naturally, the behaviour of this amplifier is to a high degree dependent upon the bias adjustments of all the diodes. Figure 6 shows two gain curves, measured with swept signal at a pump power of 1 W at 8·45 Gc/s and a signal input power of 2 μW. Six diode-quartets have been used, because this quantity has proved to be the optimum. Operating the amplifier with more or fewer quartets has not been advantageous and did not improve the amplifier characteristics.

The curves represent the parametric gain or the ratio between pumped and unpumped signal performance and result from two different bias distributions to all the diodes. It is obvious that the bias adjustment leading to higher signal gain causes undesirable peaks whereby oscillation starts as soon as the gain exceeds the critical value.

The insertion loss through the signal line of the unpumped amplifier is nearly constant and amounts to 2 dB. The parametric gain minus this insertion loss is equal to the insertion gain, which is equivalent to the ratio of power delivered to a signal consuming device after and before the pumped amplifier is inserted. An instantaneous bandwidth of 20% is obtained relative to an insertion gain of 10 dB and a centre frequency of 3·4 Gc/s.

It did not prove possible to damp out the gain ripple without lessening the wide bandwidth. Setting was most critical and delicate due to the numerous mechanical and electrical adjustments, such as the tuning screws, which simulate the periodical shunted diodes in the transition zone by gradually decreasing to where the undisturbed coaxial line begins, or the coupling screws to balance the pump and signal ports, or the diode bias voltages. Most of the adjustments have made necessary individual tuning, observing the transmission properties by oscillographic plotting of the frequency swept signal output power. It was quite out of the question to systematize the tuning procedure and it was therefore impossible to obtain consistency in the measured values. The noise factor has not yet been determined.

Besides this theoretical and experimental work on the wave propagation characteristics of the travelling-wave amplifier at microwave frequencies, extended investigations have to be made into the design of reactance diodes of adequate dimensions and electrical characteristics to meet the requirements. We have reached a state of the art shown in Fig. 7.

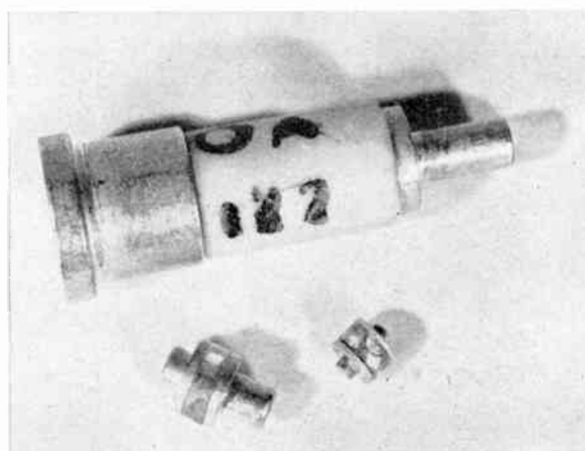


Fig. 7. Reactance diodes compared with a standard radar-type crystal diode.

A cut-off frequency of 200 Gc/s has been achieved with a capacitance of 0·1 pF at a breakdown voltage of 6 V. The capacitance of the smallest diode housing is 0·15 pF and the lead inductance is 0·2 nH.

It is evident that the smallest possible diode housing is needed for use in such iterative amplifiers in the microwave region. The picture shows this diode compared with the standard radar cartridge to give the relative sizes. Recent work shows that it is possible to realize a diode housing of about half the size of the smallest one pictured on Fig. 7.

3. Conclusions and Future Work

Summarizing the findings of our studies it may be said that it is extremely difficult to achieve perfect

World Radio History

symmetry or a fully balanced condition on the three-transmission-line-system. Even if full dimensional symmetry could be obtained in every plane of cross-section, there still exists in some way or other a further unbalance caused by the different dynamical behaviour and negative resistance performance of each single diode. These difficulties might be overcome if the energy exchange between the three waves did not occur in the lumped diodes but rather in an homogeneous dielectric of voltage dependent (and therefore time-dependent) dielectric-constant which acts like a non-linear reactance filling up the volume between the transmission lines. Unfortunately such a material, which at the same time would offer the possibility of an amplifier operating with low pump power, is not yet known.

4. Acknowledgments

The authors would like to acknowledge all those whose efforts have made possible the performing of these experiments. We give particular credit to H. Bendig for fabricating the reactance-diodes and to E. Forner for his technical assistance throughout this work.

This work is sponsored by the Ministry of Defence of the German Federal Republic.

5. References

1. R. S. Engelbrecht, 'A low-noise nonlinear reactance travelling-wave amplifier', *Proc. Inst. Radio Engrs*, **46**, p. 1655, September 1958.
2. P. Bobisch and C. Sondhauss, 'Die parametrische Iterativverstärkung mit Halbleiter-Reaktanzdioden', *Die Telefunken-Röhre*, **42**, pp. 125–44, 1963.
3. P. Bobisch and C. Sondhauss, Deutsches Bundespatent Nr. 1113014.

Manuscript received by the Institution on 26th May 1965. (Paper No. 1050.)

© The Institution of Electronic and Radio Engineers, 1966

World Radio History

U.D.C. 621.319.004.6

Capacitors—Reliability, Life and the Relevance of Circuit Design

By

D. S. GIRLING, C.Eng.†

Originally presented at a Symposium on 'Engineering for Reliability in the Design of Semiconductor Equipment' held at Hatfield College of Technology on 14th May 1965, under the aegis of the I.E.R.E. and the I.E.E.

Summary: Capacitors are made with a wide range of materials to achieve specific types of performance. Their reliability is a function of the voltage and temperature operating conditions of the dielectric and the climatic and mechanical environmental conditions of the encapsulation. About one-seventh of all equipment failures are due to capacitors and half of these are due to improper selection or application.

In choosing a capacitor for a given operation it is important to recognize that the limits given in specifications are statistical quantities. Particular care is necessary to establish the correct operating conditions of voltage and temperature. With proper care in selection, rating and application a reliability of about 0·01% per 1000 hours is achievable but in extreme conditions this may rise to 1% per 1000 hours or more. The conditions for open or short circuit failure are given for various types of capacitor. The other main causes of failure are ingress of moisture and mechanical failure.

New series of specifications are developing which will ensure the provision of more data for the circuit designer. The main need is for regular and systematic testing and the feedback of fault data to improve the control of the manufacturing processes. Reliable components can only be produced by fully optimized manufacturing techniques.

Economic considerations show that it is unreasonable to expect a component to be significantly better than the specification to which it is bought. The reliability of a component is thought to be optimum for a given application when first cost plus replacement cost plus the cost of the loss of equipment time is minimum.

1. Introduction

Modern electronic circuitry requires the smallest possible components for operation at low voltages and the capacitors which are considered in this paper are usually made with the thinnest possible dielectric material. The voltage break-down of an area of such dielectric material is an example of 'weakest link behaviour' since failure usually occurs at the weakest point. A wide range of materials is therefore employed in order to achieve specific types of performance.

There are three broad classes of capacitors:

(i) Capacitors of low loss and good capacitance stability. These are usually of mica, glass, ceramic or a low-loss plastic such as polystyrene. The method of manufacture and construction is determined by the nature of the materials.

(ii) Capacitors of medium loss and medium stability usually required to operate over a fairly wide range of voltages both a.c. and d.c. This need is met by paper, plastic film, or high K ceramic types. The first two of these may have either:

(*a*) electrodes of metal foil, or

(*b*) electrodes of evaporated metal which give a self-healing characteristic.

(iii) Capacitors of the highest possible capacitance per unit volume. These are the electrolytics normally made either with aluminium or tantalum. Both of these metals form extremely thin anodic oxide layers of high dielectric constant and good electrical characteristics. Contact with this oxide layer is normally by means of a liquid electrolyte which has a marked influence on the characteristics of the capacitor. In the case of the solid tantalum the function of the electrolyte is performed by a manganese dioxide semi-conductor.[3]

In all cases the capacitor consists essentially of two parts, the element and the encapsulation. Economic

† Standard Telephones and Cables Ltd., Capacitor Division, Paignton, Devon.

World Radio History

factors dictate that the encapsulation must be adapted to the application and will vary from a simple wrap with plastic film to an hermetically sealed enclosure consisting of metal and ceramic or glass. These two parts must be considered separately from the point of view of performance and reliability. The operating conditions of the dielectric are essentially dependent on voltage and temperature while the climatic and mechanical environmental conditions interact with the capacitor according to the degree of protection and method of mounting employed.

The choice of the most suitable capacitor for a given application is therefore a key step in obtaining a satisfactory performance. The factors affecting this decision are given in Section 2. All other factors affecting the reliability of capacitors are dealt with in Section 3.

2. The Application of Capacitors

2.1. *Selecting a Capacitor*

According to a U.S. Military specification:[1] 'If properly made and properly applied, capacitors have long lives. About a quarter of the parts in any electronic equipment are capacitors, one-seventh of all equipment failures are due to capacitor failures and half of all failures are caused by improper selection and application.'

In selecting a capacitor for a particular application the following should be considered:

(*a*) Variation of capacitance with time, temperature, frequency and possibly voltage (voltage effects are usually only a problem on ceramic capacitors).

(*b*) Loss (usually expressed as $\tan \delta$) under the same conditions.

(*c*) Insulation resistance or leakage current and likely variations due to voltage, temperature and time.

(*d*) Proposed operating conditions of voltage and temperature. If high frequency alternating voltages or impulses are employed the effects of heating must also be considered.

(*e*) Environmental conditions of shock, vibration and humidity.

(*f*) Price.

The electrical characteristics of various types of capacitor are summarized in Tables 1 and 2.

2.2. *Circuit Design from the Component Engineer's Point of View*

In Tables 1 and 2 limits are stated for various parametric changes. These limits should not, however, be regarded as absolute since they are statistical quantities. If a large number of such components are to be used or if a particular parameter is critical it may be necessary to obtain additional data from the manufacturer in order to make a correct assessment.

It is unwise to design a circuit assuming closer control than is given by the specification and to base the design on the actual characteristics of a limited number of samples. The component manufacturer cannot completely eliminate batch to batch variations.

While the need for precise components is recognized as part of economic circuit design, it is best not to specify close limits for every component. The specification to which the component is bought should adequately cover the needs.

2.3. *Types of Capacitor*

Particular characteristics of the types of capacitor referred to in Section 1 are discussed in the following sub-sections.

2.3.1. Foil types

This term is used to refer to all those which are not self-healing. Generally, these are the more reliable, having small parametric changes, and the main probability of failure is open or short circuit. Particularly on paper types this is predictable within certain limits.

2.3.2. Metallized type capacitors

Metallized type capacitors have self-healing characteristics. This is because the extremely thin evaporated metal electrode is sufficiently resistant so that when a fault occurs, it is vaporized by resistance heating. Under suitable conditions this will isolate a fault without adverse effect on the electrical characteristics. A small impulse of current, however, flows during each such fault. The amplitude of such pulses may be 10 mV or less with a duration of 10 μs or less. During manufacture such capacitors are burned out at a voltage in excess of the rated voltage. If the capacitor is rated at less than one-third of the burn-out voltage, the occurrence of self-healing failures should be very rare, but under higher voltage conditions self-healing can be expected to occur at irregular intervals.

If the electrical circuit conditions of voltage and impedance are such that self-healing is not complete each such fault results in a conduction path across the dielectric giving rise to a reduced insulation resistance.[4] For safe circuit design it is unwise to rely upon the stability of the insulation resistance of metallized type capacitors.

2.3.3. Electrolytic types

This type of capacitor employs a liquid electrolyte to make contact with the oxide layer and most of the behaviour characteristics arise from this. There is a

World Radio History

wide variety of conflicting requirements for the electrolytes for these capacitors and it is usually difficult to achieve a high performance over a wide range of temperature and frequencies. In fact, the low frequency equivalent circuit of such a capacitor consists of a capacitor in series with a resistor which usually lies in the range 0·1–10 ohms. Thus electrolytic capacitors should be regarded as capacitors only at low frequencies and as impedances at the higher frequencies. A typical set of impedance characteristics is given in Fig. 1.

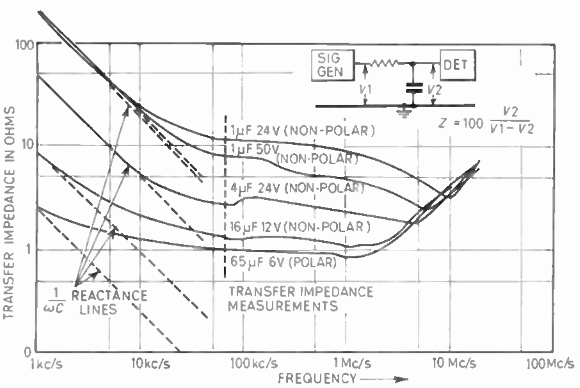


Fig. 1. Variation of impedance with frequency for tantalum foil electrolytic capacitor.

A further problem is that of retaining the electrolyte within the case. Hermetically sealed capacitors are extremely rare and normally sealing is by means of various types of rubber or resins, or both. Diffusion of the electrolyte through such seals, particularly at the highest operating temperature, usually results in a fall in the conductivity of the electrolyte with increase of power factor and reduction of capacitance. Detailed comment as regards individual types of electrolytic capacitor is as follows:

(*a*) *Aluminium electrolytics.* Great care has to be taken to choose electrolytes which will not attack the aluminium oxide layer. The degree of this attack is less the higher the purity of the aluminium. Capacitors using high-purity aluminium are usually classified as high reliability types and have an improved shelf life behaviour. In general, however, the shelf life of such capacitors is limited and hence de-rating for voltage does not result in an extension of life.

(*b*) *Tantalum wet electrolytics.* Tantalum pentoxide is very much more stable and resists attack from most electrolytes. Thus there is a wider choice of electrolytes from which to obtain satisfactory operating characteristics. The shelf life is indefinite and de-rating for voltage results in an extension of life. The upper limit of voltage for these capacitors is determined by the onset of field crystallization which is a function of the purity of the metal. There are two types of construction. A sintered porous anode of tantalum is housed in a case containing sulphuric acid. Alternatively, an anodized tantalum foil is wound together with a tantalum cathode foil separated by paper impregnated with a neutral electrolyte.

(*c*) *Solid tantalum capacitors.* This type of capacitor is made with a sintered porous tantalum body over which the oxide layer is formed and contact with this oxide layer is made with a manganese dioxide semi-conductor. A cathode connection is made over the external surface of the body. Since a liquid electrolyte is not used, it is easy to house this capacitor in an hermetically sealed container. This construction has the advantage of relatively low effective series resistance giving stability of characteristics over a wide range of temperatures and frequencies and good stability of characteristics on life test. Earlier types of this capacitor showed an increase of leakage current on life, a limitation of performance in low impedance circuits and instability under vibration conditions. Most of these defects have been eliminated by careful control of manufacturing processes, but special attention is necessary in these respects.

2.4. *Voltage Ratings*

The application of voltage to a dielectric establishes an electrical field within it and in all commercial capacitors this is far below the intrinsic dielectric strength of the materials. A failure is usually due to the presence of defects or to the occurrence of electrochemical deterioration. This latter is due to ionic conduction resulting in chemical changes and ultimate short circuit failure. These mechanisms are very dependent on the voltage applied to the capacitor. As stated at the beginning of this paper capacitors exhibit 'weakest link' behaviour and the establishment of a life characteristic is a very complex statistical problem. Even today there is no complete agreement as regards the constants of the life equations. A detailed discussion of the statistics of failure is beyond the scope of this paper. A simple approach to the problem is given in the Appendix.

Equation (1) of the Appendix gives a family of curves for various temperatures as shown in Fig. 2. The same equation may also be used to estimate the component hours per failure (or its reciprocal the failure rate) as shown in equation (5). The number of component hours per failure under standard test conditions is a useful figure of merit, 1% failures per 1000 hours corresponding to 100 000 component hours per failure.

World Radio History

Table 1

Limits of parametric change allowed in non-electrolytic capacitor specifications

Type of capacitor	Reference	Highest temperature °C	Humidity severity (days—long-term damp heat)	Initial: Normal capacitance tolerance %	Initial: Power factor	Initial: Insulation resistance Ω-F	Initial: Insulation resistance MΩ	Limits after test: Capacitance %	Limits after test: Power factor	Limits after test: Insulation resistance Ω-F	Limits after test: Insulation resistance MΩ	Remarks
Foil/Paper	I.E.C. 80 1964	55/100	56 21 4	20 20 20	0·01 0·01 0·01	4000 2000 300	12 000 6000 900	5 5 5	1·4 × initial value or initial limits	2000 500 10	6000 1500 30	Insulation resistance after test must not be less than 30% of initial value or limits shown
Chlorinated impregnant types			56 21 4	20 20 20	0·02 0·02 0·02	2000 2000 300	6000 6000 900	5 5 5	,, ,, ,,	1000 500 10	3000 1500 30	
Metallized paper Type 1	I.E.C. 166 1965	55/125	56 21 4	20 20 20	0·015 0·015 0·015	$\frac{250\ \Omega\text{-F}}{c} + 250\ \text{M}\Omega$	3000 3000 3000	5 5 5	,, ,, ,,	$\frac{125\ \Omega\text{-F}}{c} + 125\ \text{M}\Omega$ $\frac{8\ \Omega\text{-F}}{c} + 8\ \text{M}\Omega$	3 000 3 000 100	During endurance test at 2 V insulation resistance shall be not less than 1 MΩ
Type 2			56 21 4	20 20 20	0·015 0·015 0·015	$\frac{100\ \Omega\text{-F}}{c} + 100\ \text{M}\Omega$	1200 1200 1200	10 10 10	,, ,, ,,	$\frac{50\ \Omega\text{-F}}{c} + 50\ \text{M}\Omega$ $\frac{33\ \Omega\text{-F}}{c} + 33\ \text{M}\Omega$	600 600 40	
Foil/Polyester Metallized polyester (non-self-healing)	I.E.C. 202 1965	55/125	56 21 4	5 10 20	0·01 0·01 0·01	10 000 10 000 10 000	30 000 30 000 30 000	5 5 5	,, ,, ,,	5000 5000 5000	15 000 15 000 15 000	
Mica	I.E.C. 116 1959	55/100	56 21 4	1, 2, 5, 10, 20	0·002 (above 1 000 pF)	—	10 000	1% or 1 pF	Initial limits	Initial limits		Temp. coeff. parts in 10^6/deg C: −100 to +100; Cap. drift (pF): $\pm 0{\cdot}003C$
			56 21 4	1, 2, 5, 10, 20	,,	—	10 000	1% or 1 pF	,,	,,		−20 to +100; $0{\cdot}001C + 0{\cdot}1$
			56 21 4	1, 2, 5,	,,	—	10 000	0·5% or 0·5 pF	,,	,,		−20 to +50; $0{\cdot}0005C + 0{\cdot}1$
Polystyrene		40/85	56 21 4	1, 2, 5, 10	0·001	100 000 10 000 10 000	1 000 000 100 000 100 000	1% (1·5% or 1 pF < 500 pF)	1·4 × initial value or initial limits	100 000 5000 5000	1 000 000 50 000 50 000	

(continued on page 377)

World Radio History

Table 1 (*contd.*)

Type of capacitor	Reference	Highest temperature °C	Humidity severity (days long-term damp heat)	Initial: Normal capacitance tolerance %	Initial: Power factor	Initial: Insulation resistance Ω-F	Initial: Insulation resistance MΩ	Limits after test: Capacitance %	Limits after test: Power factor	Limits after test: Insulation resistance Ω-F	Limits after test: Insulation resistance MΩ	Remarks
Ceramic Type I (Low loss high stability)		55/155	56 21 4	1, 2, 5, 10, 20	0·001	—	10 000	1% or 1 pF	1·5 × initial value	—	3000	A range of temperature coefficients from +120 to −1500 parts in 10^6/deg C
Type II (Bypass and coupling)	I.E.C. 187 1965	55/125	56 21 4	10 to +30 −70	0·035 0·035 0·035	250 250 250	10 000 10 000 10 000	10% or 20%, according to class	2 × initial limit	250 125 125	3000 3000 3000	

Table 2

Limits of parametric change allowed in electrolytic capacitor specifications

Type of capacitor	Reference	Highest temperature category °C	Humidity severity (days long-term damp heat)	Initial: Normal capacitance tolerance %	Initial: Power factor	Initial: Leakage current	% Capacitance at min. cat. temp.	Limits after test: Capacitance	Limits after test: Power factor	Limits after test: Leakage current	Remarks
Aluminium		25/85	56 21 4	Between −10 to +50 and −10 to +250, according to voltage	15 to 45%, according to voltage	UC × 10^{-4} + 0·3 mA or 10 mA, whichever is the less	Impedance test only	10% (20% < 150 V working)	1·5 × initial limits	Initial limits	
Tantalum foil axial	DEF 5134-2	55/125	H6 (DEF 5011)	± 20	6 to 12%, according to working voltage	0·02 μA/μFV (> 1μF)	60	± 10%	1·5 × initial limits	Initial limits	
Tantalum sintered solid	DEF 5134 A/1	55/125	H6 (DEF 5011)	± 20	6 to 10%, according to working voltage	0·02 μA/μFV	± 12%	± 30%	3 × initial limits	2 × initial limits	Limits are shown for 125°C. Closer at 85°C

Notes on Tables 1 *and* 2.

1. These are the limits of maximum variation allowed in any of the environmental or endurance tests.
2. It should be noted that the severity of the climatic tests varies according to humidity category. The limits shown therefore assume that the capacitor is operated under its appropriate humidity condition.
3. Changes of capacitance are additive to initial tolerances.
4. The impedance of all electrolytics becomes resistive at low radio frequencies. Liquid types exhibit an increase of impedance at low temperatures.

World Radio History

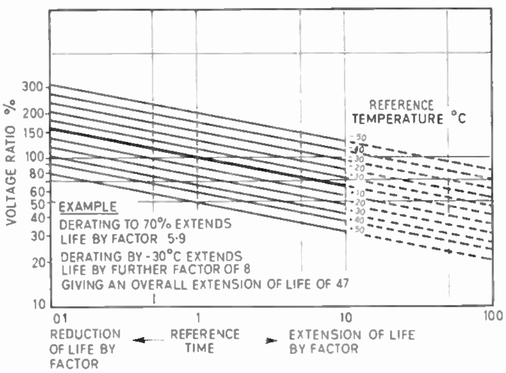


Fig. 2. Graph for calculating the effect of derating for voltage or temperature.

Figure 3 shows the way in which the failure rate curve is affected by derating. In this case derating of either voltage or temperature has been applied to give an extension of life by a factor of 10. Using graphical methods it is possible to construct the new failure rate curve for the derated condition. It will be noted that the onset of wear-out has been shifted by a factor of 10 while, at the same time, the random failure portion has been reduced by the same factor.

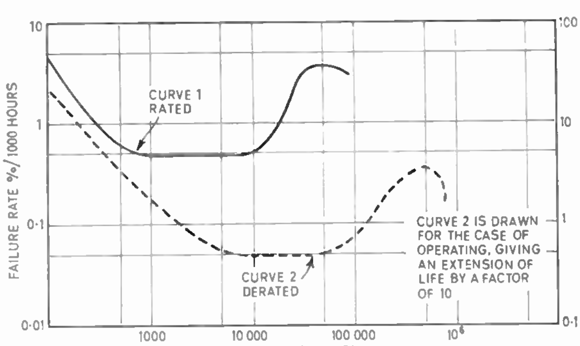


Fig. 3. Effect of derating on the failure rate.

These equations have been established primarily on paper capacitors with foil electrodes. They have also been shown to apply reasonably well to tantalum foil electrolytic capacitors ([5]). It is believed that they can be applied to most capacitors employing foil electrodes. In the case of metallized types the rate of self healing failures and parametric changes will probably follow similar laws. It should, however, be noted that they relate only to that part of the failure rate which is due to the presence of voltage on the capacitor. The failure rate due to all other causes being unaffected.

Because of the effect of voltage on the life of the capacitor it is important that circuit engineers should appreciate fully the significance of the various voltages which are specified for capacitors. These are, in effect, various points on a life characteristic.

2.4.1. Proof test voltage (non-electrolytic types only)

This is a voltage applied for either 5 seconds or 1 minute (according to specification) to eliminate capacitors having defects in the dielectric which will fail at this level. Normally, because the manufacturer is trying to make the smallest possible capacitor at the lowest cost, this voltage will be the highest that the dielectric will withstand consistently.

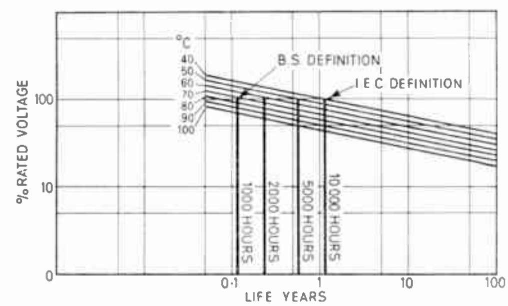


Fig. 4. Approximate life for 1% failures of capacitors to B.S. or I.E.C. specifications.

2.4.2. Rated voltage

This is the nominal voltage at which the capacitor will operate continuously at the rated temperature. Specifications differ in their definition of rated voltage. I.E.C. 80 for paper capacitors specifies a life of 10 000 hours at 40°C but other specifications state 1000 hours at 70°C and it should be noted from Fig. 4 that these two definitions are equivalent. It is, however, generally unwise to assume that the life of a capacitor will be significantly greater than that ensured by the endurance test requirements of the specification. Allowance should, of course, be made for any acceleration resulting from an increased endurance test voltage above the category voltage. It may then be assumed that the life will be the duration of the endurance test multiplied by the acceleration factor. Extension of life beyond this should be achieved by appropriate de-rating on the assumption of a probability of failure under the endurance test conditions of about 1%. The sample size tested does not usually justify the assumption of a lower probability.

2.4.3. Category voltage

This is a voltage lower than the rated voltage at which the capacitor will operate at the maximum category temperature. It varies from the rated voltage by a de-rating factor.

World Radio History

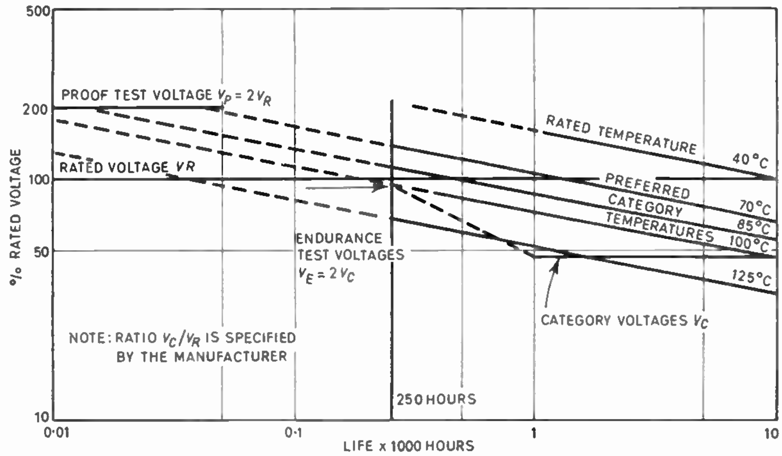


Fig. 5. Specification voltages for paper capacitors I.E.C. 80. Basis of rating 10 000 hours at 40°C.

2.4.4. Surge voltage

This is the highest voltage which the capacitor will withstand under surge conditions. The normal percentage above rated or category voltage is as follows:

Non-electrolytic types	100%
Tantalum or aluminium wet electrolytics	16%
Solid tantalum	30%

These points are indicated diagrammatically in Figs. 5, 6 and 7 for paper, mica and solid tantalum capacitors respectively.

2.5. *Alternating Voltages*

While a discussion of capacitors for a.c. applications is beyond the scope of this paper, some discussion of the heating effects of a.c. is necessary. All specifications state that the peak alternating voltage plus d.c. must be equal to or less than the rated voltage. In the case of electrolytics a ripple current rating is usually given and in others a maximum VA rating.

In all cases the objective is to limit the temperature rise of the capacitor dielectric. This has exactly the same effect as an increase of ambient temperature in reducing life. Generally, a metal-cased capacitor should not operate with a temperature rise of more than about 10°C. It may be assumed that the heat dissipation of such a component is about $0{\cdot}006\ \mathrm{W\,cm^{-2}/deg\ C}$. Insulated cased components may well be limited to much less and encapsulation in a heat insulating medium will also have the same effect.

It is important to examine the voltage surges likely to be present in the system: this applies particularly to electrolytic capacitors and particular attention is necessary to the surge voltage ratings. If the amplitude or repetition rate is high internal heating may again occur. The calculation of this heating effect is difficult, but there are various approximations which may be used.

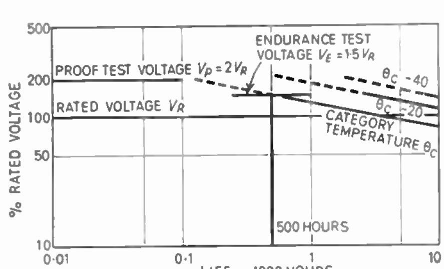


Fig. 6. Specification voltages for mica capacitors I.E.C. 116. Basis of rating 500 hours at $1{\cdot}5V_R$ at maximum category temperature.

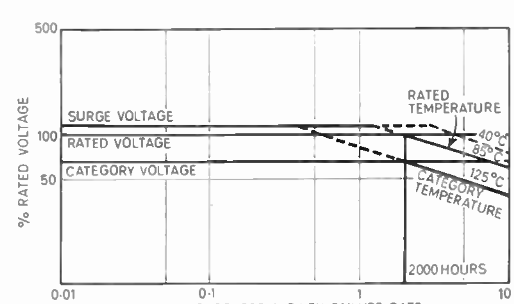


Fig. 7. Specification voltages for solid tantalum I.E.C. Basis of rating 2000 hours at V_R and rated temperature and V_C and category temperature.

World Radio History

3. Reliability

3.1. *Quality Control and Quality Assurance*

It is well known that the measurement of reliability is an expensive and time-consuming process. By the time the results are obtained the data are often of historical interest only. In practice a requirement for a low failure rate can only be translated into a requirement for quality of materials and consistency of product. Sound design together with use of an effective quality control system is capable of giving a product which will meet consistently the requirements of the specification to which it is made. Quality assurance testing (i.e. destructive testing) on samples taken from production is an essential factor of such a control scheme and consistency of product will be achieved only if adequate feedback arrangements exist to correct any failures observed.

Economic pressures are, however, such that it is unreasonable to expect that the component will be significantly better than the specification to which it is purchased. Thus considerable care is necessary in ensuring that the specification covers the reliability requirements. Special quality specifications covering components of higher reliability exist for some products and these are likely to increase in the future. For special applications such as submerged repeater or satellites, completely separate and highly controlled manufacturing units may be necessary.

3.2. *Categories of Failures*

These may be classified as follows:

(*a*) *Catastrophic failure*

Open or short circuit.

(*b*) *Sudden parametric change*

Radical departure from initial characteristics occurring in an unpredictable manner and in too short a period of time to permit detection through normal preventive maintenance practices.[1] Approximate failure rates for various types of capacitors are given in Table 3.

(*c*) *Gradual parametric changes beyond given limits*

The component engineer is bound to consider as a failure a component which exceeds the values permitted in the component specification. In a particular equipment, however, only one of these parameters may be important and this at a lower level at which the circuit will cease to function. Thus, component failure rate analysis under arbitrary test conditions may not be related to the actual conditions.

3.3. *Causes of failure*

It is essential in considering the reliability of a component that *all* operating conditions are considered. Thus, although voltage and temperature may be the dominant conditions in a static equipment, in an aircraft or missile the environmental conditions may be far more important. The circuit engineer must assess the effect of these on his component and the particular application.

The main mechanisms of failure in capacitors are summarized in Table 4. It is assumed that voltage and temperature operating conditions have been correctly assessed in Section 2. Comments on other

Table 3

Summary of failure rate data for various capacitors
(% failures for 1000 hours). Failures are as defined in Sect. 3.2(*a*) and (*b*).

Type	At 40°C and % rated volts			Maximum category temperature	At maximum category temperature and % category volts			Remarks
	10%	50%	100%		10%	50%	100%	
Paper	0·001	0·01	0·1	100°C	0·001	0·005	0·1	
Mica	0·001	0·004	0·05	100°C	0·001	0·007	0·1	Varies according to capacitance
Ceramic	0·001	0·001	0·004	100°C	0·001	0·004	0·035	
Aluminium electrolytic	—	—	0·05†	70°C	—	—	0·3†	†Derating does not improve reliability
Tantalum electrolytic								
(i) Wet slug	0·001	0·005	0·04	125°C	0·001	0·03	0·2	
(ii) Foil	0·002	0·01	0·07	125°C	0·001	0·06	0·5	
(iii) Solid	0·005	0·02	0·03	125°C	0·01	0·1	0·5	

World Radio History

possibilities are as follows:

3.3.1. Short circuit

This is the normal failure mode of non-electrolytic type capacitors employing foil electrodes and normally results from the failure of the dielectric due either to over-voltage or to deterioration of the dielectric at points of weakness. This type of failure may also result from serious over-heating.

Solid tantalum capacitors ultimately fail due to short circuit.

3.3.2. Open circuit

This arises either from faulty connections or from fractures of wires or tapes. The latter usually arise from relative movement under conditions of mechanical stress.

Open circuit conditions may also arise at low voltages between two pieces of metal which are in pressure contact only. In general, capacitors employing such contacts should be avoided.

Metallized type capacitors may also exhibit the same effect between the metallization and the end spray of the winding element and this is dependent upon the choice of materials and the control of the manufacturing process. It may also occur as an extreme case of the corrosion or burn-out of the metallized electrodes.

Wet electrolytic type capacitors ultimately become open circuit due to drying-out of electrolyte.[6]

3.3.3. Ingress of moisture

In the case of capacitors which are not hermetically sealed, ingress of moisture is inevitable since all organic materials have finite moisture transfer characteristics. Thus, such capacitors must be operated in sufficiently dry conditions as to ensure the required performance. The degree of moisture protection is classified in accordance with DEF 5011.

The effect of ingress of moisture is to reduce insulation resistance and to increase power factor. It should be expected that the reduction of life should be proportional to the reduction of insulation resistance.

Generally, H3 type capacitors will give only a limited life under temperate climatic conditions unless they are protected from moisture. H5 capacitors are generally satisfactory to give indefinite lives under temperate climatic conditions but under tropical conditions further protection may be necessary. The hermetically sealed capacitor which is usually H6 or H7 is suitable for use under all climatic conditions and ingress of moisture can only occur as the result of mechanical failure of the seal.

3.3.4. Mechanical failures

Assuming that no damage has been caused in handling or inserting into the equipment, mechanical failure will normally be related to bumping during transport of the equipment, or vibration, acceleration or shock during use.

Methods of mounting depend entirely on the application. For severe mechanical environments capacitors should be supported and care taken to avoid mechanical resonances.

3.4. *Specifications*

The series of internationally agreed component specifications drawn up by the International Electrotechnical Commission which is being adopted by the British Standards Institution is growing rapidly. These may be used as the minimum standard for commercial, industrial and communication purposes.

The Services continue to employ the DEF specifications and these are being coordinated with N.A.T.O. The form of these has been under consideration within the Ministry of Aviation by the Committee on Common Standards for Electronic Parts (the 'Burghard' Committee) and this appears likely to recommend a pattern of specifications consisting of a block of 5:

Block 1—Basic specifications such as DEF 5001–C and DEF 5011.

Block 2—A specification giving test and measurement requirements for a particular family.

Block 3—A detailed specification covering a specific style of electronic parts.

Block 4—This will be as Block 3 but relating to a particular manufacturer's product having discrete characteristics.

Block 5—This will be a certified reliability data sheet giving a record of the performance of the part in accordance with the test requirements.

There is a great deal of common ground in these specifications and it is to be hoped that ultimately the differences can be resolved to give a standard series which covers the majority of applications. The proposals for certified reliability data sheets will be useful in providing more data to the circuit designer. The main need is for regular and systematic testing of components and the feedback of fault data to improve the control of the manufacturing processes. Reliable components can only be provided by fully optimized manufacturing techniques.

World Radio History

Table 4

Mechanisms of failure of capacitors

Mechanism	Cause of change of parameter			Cause of transients		Cause of catastrophic failure		Remarks
	Capacitance (1)	Power factor (2)	Insulance (3)	D.C. (4)	A.C. (5)	D.C. (6)	A.C. (7)	
1. NON-ELECTROLYTIC TYPES:								
1*a*. Moisture:								
(i) Volume absorption	Dipolar contribution to dielectric constant	Dipolar contribution to dielectric loss	Dissociation of impurities resulting in ionic conduction	—	—	Electrochemical deterioration due to ionic conduction	Electrochemical deterioration due to ionic conduction (less severe than on d.c. due to reversal of polarity)	
(ii) Adsorbed on surfaces	Small contribution to dielectric constant and loss					Short circuit due to metal migration across surface		This effect may also occur across terminal surfaces
1*b*. Gas voids	—	—	Increase of leakage current at high stresses	D.c. discharges are infrequent and are generally at a very low level	A.c. discharges may occur several times on each half cycle if the working stress is above the extinction voltage	At normal stresses d.c. discharges are not normally a cause of failure except under low pressure conditions (which may occur in a rigid case at low temperatures)	A.c. discharges are the major cause of a.c. breakdown in service	Gas voids are normally present in wax impregnated types only
1*c*. Poor connections (formation of insulating films or corrosion)	Intermittent connections may cause changes of power factor and/or capacitance. High series resistance		—	Intermittent contact	Current discharges may occur at poor contacts	Open circuit may occur on very stable supplies	Burning at poor contacts may cause overheating and failure	See also 1*d*
1*d*. Dimensional changes and mechanical displacement (including relaxation)	Change of electrode area or thickness of dielectric (normally exponential)	—	—	—	—	In extreme case open or short circuit failure usually resulting from mechanical damage or vibration		
1*e*. Dielectric breakdown due to over voltage	—	—	—	—	—	Short circuit due to carbonization of dielectric		
1*f*. Self-healing (metallized types only)	Change of capacitance due to change of electrode area (only in extreme case)	Self-healing may result in deterioration of contact resistance (see 1*c* above)	Surface conduction at point of burn-out due to incomplete removal of metal and thermal degradation (particularly if self-healing is imperfect)	Self-healing	Self-healing and a.c. discharges	*Short circuit* Repeated burn-out at same point results in carbonization of organic dielectrics *Open circuit* Excessive current at short circuit may destroy connection	As for d.c. Self-healing may be accelerated by a.c. discharge or thermal instability	

World Radio History

1*g*. Thermal instability	—	—	—	—	—	—	Continuous increase of temperature to failure due to heat generated exceeding heat dissipated	
2. ELECTROLYTIC TYPES:								
2*a*. Formation of insulating film on connecting tapes	—	—	—	—	—	Open circuit	—	Mainly on aluminium electrolytics
2*b*. Over or reverse voltage	Decrease due to increased thickness of oxide layer on anode or cathode	No change	—	—	—	If excess voltage is maintained for a long period considerably above forming voltage, gas formation may cause rupture of seal which may be explosive		
2*c*. Field crystallization	Decrease due to increased thickness of crystalline oxide layer	Increase due to poorer properties of crystalline oxide	Increase to cause growth of layer and due to poorer properties of crystalline oxide	Scintillation occurs only near forming voltage	—	Gas formation due to electrolysis of electrolyte to cause growth of oxide layer may rupture seal		Occurs only at highest voltage
2*d*. Drying out of electrolyte due to diffusion through organic seal	Decrease due to increased resistivity of electrolyte and reduced area of contact	Increase due to increased resistivity of electrolyte	In case of gross leakage only a leakage current may flow across the external seal	—	—	Drying-out eventually results in open circuit		Effects are normally very slow, depending on seal efficiency. Much accelerated if physical damage to seal occurs
2*e*. Over-heating (Excess VA)	In extreme case, drying-out of electrolyte (see 2*d*)			—	—	Increased internal pressure may result in rupture of seal		
2*f*. Case made positive to either electrode	—	—	—	—	—	Severe gas formation resulting in explosion		
3. SOLID ELECTROLYTIC TYPES:								
3*a*. Thermal decomposition of manganese dioxide semiconductor	—	—	Increase of leakage current	—	—	Short circuit in extreme case	—	
3*b*. Overvoltage or surge	—	—	Increase of leakage current	—	—	Short circuit in extreme case	—	
3*c*. Overheating	—	—	—	—	—	Mechanical failure		

World Radio History

4. References

1. 'General Characteristics of Military Capacitors', MIL handbook 217.
2. G. W. A. Dummer, 'Component reliability—a survey', *British Communication and Electronics*, **10**, pp. 434–7, June 1963.
3. R. W. Franklin, 'Electrolytic capacitors', *Proc. Instn Elect. Engrs*, **109B**, Supplement 22, pp. 525–36, 1962. (I.E.E. Paper No. 3585.)
4. A. A. New and A. C. Lynch, 'Failure of metallized paper capacitors used with large series resistances', *Proc. I.E.E.*, **109B**, Supplement 22, pp. 496–9, 1962. (I.E.E. Paper No. 3597.)
5. E. E. Smith, 'Reliability of tantalum foil type electrolytic capacitors', *Proc. I.E.E.*, **109B**, Supplement 22, pp. 543–7, 1962. (I.E.E. Paper No. 3621.)
6. A. A. New, 'Electrolytic capacitors and their reliability', *Post Office Elect. Engrs J.*, **55**, No. 2, pp. 115–24, July 1962.

5. Appendix

De-rating as a Means of Improving Reliability

5.1. *Wear-out Characteristic*

It is normally found that, for the wear-out part of the characteristics, the lives of capacitors are related as follows:

$$\frac{L_1}{L_2} = \left(\frac{V_2}{V_1}\right)^n 2^{(T_2 - T_1)/k} \qquad \ldots\ldots(1)$$

where L = time to a given % failure,
V = direct voltage
T = temperature (deg C)
n = power law exponent
k = temperature rate constant

For paper capacitors n is normally taken as 5 and k as 10 and this gives rise to a series of curves as shown in Fig. 2. The value of n is true for paper capacitors at about 100°C and at lower temperatures values up to 14 have been reported. In the absence of more precise information the value of 5 should be assumed. The value of R of 10 derives from the Arrhenius law covering the rate of chemical reactions.

Thus, if the life of a capacitor is known for one set of conditions of voltage and temperature, then it may be calculated for another. Although these data are normally calculated on mean lives it is believed that they can also be applied to other percentages of the sample.

The above equation can readily be rearranged to give an operating life factor A where

$$A = \frac{\text{life under actual operating conditions}}{\text{life under category test conditions}}$$

Subscripts a and c will be used to denote 'actual' and 'category' respectively.

$$L_a = \left\{\left(\frac{V_c}{V_a}\right)^n 2^{(T_c - T_a)/k}\right\} L_c = AL_c \qquad \ldots\ldots(2)$$

5.2. *Random Failure Characteristic*

Before the onset of wear-out it is normally found that failures occur at an approximately uniform rate per unit of time. This provides some very simple relationships.

Let

p_1 = probability of failure in time L_1

r_1 = failure rate (% per 1000 hours)

$$r_1 = \frac{p_1}{L_1} \times 10^5 \qquad \ldots\ldots(3)$$

A very useful figure of merit which may be used as a basis of comparison of various capacitors under standard test conditions is the number of component hours per failure (B),

$$B = \frac{L_1}{p_1} = \frac{10^5}{r} \qquad \ldots\ldots(4)$$

The percentage failures for various times based on given failure rates are given in Table 5.

Table 5

Percentage failures at various times for various failure rates

Failure rate % per 1000 hours	% failures in					
	100 hours	10 000 hours	5 years	10 years	15 years	20 years
0·0001	—	0·001	0·005	0·01	0·015	0·02
0·001	—	0·01	0·05	0·1	0·15	0·2
0·01	—	0·1	0·5	1	1·5	2
0·1	0·1	1·0	5	10	15	20
1·0	0·1	10	—	—	—	—
10·0	1	—	—	—	—	—

If we assume that the random failure part of the characteristic obeys the same laws as does the wear-out, then if we know the component hours per failure under given test conditions it is possible to estimate them for other conditions using eqn. (2).

Thus:

$$\frac{L_a}{p} = A\frac{L_c}{p} \qquad \ldots\ldots(5)$$

It must, however, be made clear that these equations must be used with great caution. They involve a number of assumptions and apply only to that part of the failure rate which is due to the presence of voltage on the capacitor. The failure rate due to all other causes is unaffected.

Manuscript first received by the Institution on 14th May 1965 and in final form on 14th February 1966. (Paper No. 1051.)

© The Institution of Electronic and Radio Engineers, 1966

World Radio History

SUBJECT INDEX

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

World Radio History

World Radio History

World Radio History

World Radio History

World Radio History

INDEX OF PERSONS

Names of authors of papers published in the volume are indicated by bold numerals for the page reference.
Authors of papers which are given in abstract form are denoted by A.
Biographical references are denoted by B.

World Radio History

INDEX OF ABSTRACTS

This index classifies under subject headings the abstracts published throughout the volume in "Radio Engineering Overseas" and the summaries of papers read at the Conference on "Applications of Thin Films in Electronic Engineering" and at the New Zealand National Electronics Conference.

Automatic Control

Broadcasting

Circuits

Communications

Components

World Radio History

JOURNALS FROM WHICH ABSTRACTS HAVE BEEN TAKEN DURING THE FIRST HALF OF 1966

Archiv der Elektrischen Übertragung (Germany)

Ericsson Technics (Sweden)

Nachrichtentechnische Zeitschrift (Germany)

L'Onde Électrique (France)

Philips Technical Review (Holland)

Proceedings of the Institution of Radio and Electronics Engineers Australia

Radio Engineering and Electronic Physics (English language edition of *Radioteknika i Elektronika*) (Russia)

Review of the Electronics Communications Laboratory, N.T.T. (Japan)

Slaboproudý Obzor (Prague)

World Radio History