JOURNAL OF

The British Institution of Radio Engineers

(FOUNDED IN 1925 – INCORPORATED IN 1932)

"To promote the general advancement of and to facilitate the exchange of information and ideas on Radio Science."

Vol. 6 (New Series) No. 2 MARCH-MAY, 1946

BROADCASTING IN EUROPE

Continued from page 40 of the Journal. Erratum : The figure published as Fig. 4 on page 37 should be transferred to the bottom right hand corner of page 35. The table at present on page 35 is, of course, Fig. 4.)

A. Cross : The principal doubt regarding this plan appears to concern the validity of the formula for service area λ upon 4.

According to the use of this formula the effective service area of Droitwich should be 375 miles, whereas I can personally confirm as a result of experimental observation that quite serious fading of the Droitwich transmission occurred in Newcastle at a distance of 200 miles in 1938.

It has been noted that the plan has been designed upon the basis of a population map published by S.G.I. shortly before the war. It would seem obvious that the plan would need considerable revision to cater for the redistribution of the population density which has occurred during the war.

The authors appear to have overlooked the possibility of using wire networks to serve areas that have a high density of population. If such networks were used, a reduction in the total number of transmitters required, together with a reduction in their power might be effective, and this would enable a better kc/s separation to be established and consequent reduction in mutual interference and an over-all improvement in quality.

H. Whalley : In view of the fact that a number of the extended LW channels may not be available, would it not be wise to reduce channel spacing to 10 kc/s, using the same MW frequencies as the United States? This, admittedly, does not ease the LW position very much, but does give a useful increase in the number of MW channels.

As, with typical receivers, the improvement in quality which is theoretically possible in increasing the bandwidth from 9 to 10 kc/s will hardly be noticed by the average listener, there does not seem to be a great case for increasing bandwidth to 11 kc/s. Listeners who are willing to pay the increased cost for high-fidelity reception would receive better value for money from a UHF high-fidelity relay station.

Interference from U.S. stations should not be very troublesome. I believe the signal strength in this country from Cincinnati, WLW, when operating experimentally on 500 kW was well down on lower power U.S. stations, presumably nearer the east coast. If U.S. stations are going to increase power greatly then surely coastal stations are not going to waste it by radiating over the Atlantic. If, however, tests show that serious U.S. interference is to be expected, then the case for 10 kc/s separation is surely strengthened providing European stations sacrifice the slight advantage of even numbers for frequency and space their channels 5 kc/s from the American channels.

The use of directional aerials can be very valuable in reducing interference, and intelligent use of such systems should be stressed when putting any plan into practice.

M. Kinross : Bearing on the range formula, what difference had been allowed for between day and night conditions, and between maximum and minimum sun spot conditions?

Dr. Hanney : Droitwich was a failure in South Wales and the report did not deal with the special conditions as would apply, for example, in South Wales.

P. Adorian (contributed) : My main criticism of the R.I.C. plan has been the optimistic estimate of service range. Mr. Clark's reply to this criticism was that the C.C.I.R. and F.C.C. figures which are used as the basis of my criticism were known to him and his collaborators. but these figures were rejected by him and his collaborators as being too conservative. Instead of the C.C.I.R. and F.C.C. figures, they have based their $\frac{\lambda}{4}$ assumption on figures published by P. P. Eckersley in 1928, a paper by Gillet and Eager (of Glenn D. Gillet, Inc., of Washington) prepared in 1934, a paper by P. Patrick, of the South African Broadcasting Corporation (1941) and to some extent on figures given by the N.B.C. of the U.S.A.

Even without examining in detail the merits of the different papers published on this subject, I think we should agree that we can place more reliance on the C.C.I.R. Report (signed by Great Britain, the Netherlands, the U.S.A. and the Union Internationale de Radiodiffusion) and on the figures published by the F.C.C., than on the very interesting but now somewhat superseded 17 year old report by Eckersley or on some of the other reports of individual broadcasting organisations.

In any case, careful examination of some of these papers proves that the authors of the R.I.C. plan have only taken convenient points on the various graphs of these various papers and have either not appreciated some of the points made in these papers or have decided to ignore them. The following are some examples :

In Gillet and Eager's paper in the February, 1936, issue of the Proceedings of the Institute of Radio Engineers, Fig. 11 gives the service range at 200 kc/s (soil conductivity 10^{-13} e.m.u.) as 280 miles, as against the $\frac{\lambda}{4}$ value, which would be 375 miles. The value given in my Fig. 3 for this wavelength is 175 miles.

Next, Patrick's paper in the September, 1942, issue of the Proceedings of the Institute of Radio Engineers gives ranges of about $\frac{\lambda}{4}$ (soil conductivity of 10^{-13} e.m.u.) between 500 and 1,500 kc/s (Fig. 5), and even if we assume that these same relations will apply to frequencies down to, say, 150 kc/s, *these conditions only apply, according to Patrick's article, for* $\frac{1}{2}\lambda$ *antennas.*

Mr. Clark has also referred several times to $\frac{5}{8}\lambda$ antennas. In actual practice, of course, it is not practicable to erect $\frac{1}{2}\lambda$ or $\frac{5}{8}\lambda$ antennas for long wave transmitters, as the height of these would have to be, in the case of 2,000 metre wavelengths, 1,000-1,250 metres. In the case of a 1,600 metre wavelength, the height would have to be 800-1,000 metres. (*Note.*—The height of the Eiffel Tower is 300 metres.) Obviously, Patrick appreciates this, because in his Fig. 4 he shows the service range for $\frac{1}{4}\lambda$ and shorter aerials and in this case the service range works out at about $\frac{\lambda}{8}$. Thus, this authority quoted by the authors of the R.I.C. plan clearly shows that the $\frac{\lambda}{4}$ figures for range can only be achieved if, on long waves, aerials of quite impractical heights were erected, and as this seems quite impracticable for the immediate future, the service range of the long wave stations is very much over-estimated in the R.I.C. plan.

As most of Europe, according to the R.I.C. plan, is to be covered by long wave national programmes, the plan obviously fails.

Mr. T. D. Humphreys : The suggestion that 2 mV/M is the minimum field strength aimed at is ideal—for non-industrial areas—but as the lowest value permissible is 1 mV/M would it not be better to raise this suggested minimum in industrial areas, or, alternatively, use frequency modulation for such areas, since this system has proved invaluable in the U.S.A. ?

The average value of soil conductivity appears to have been arbitrarily assumed, and the better conduction of water, which cannot be neglected in view of the geography of Europe, seems not to have been taken into account. The suggested range in miles of wavelength/4 does not appear to agree with that obtained from the Austin-Cohen formula, which shows that the ground wave is attenuated logarithmically by reason of absorption of energy by the earth.

It is stated that each nation must have one programme on a national basis, but I would suggest that this may not always be possible without having several synchronised transmitters with possible areas of " no capture."

I agree with the remarks about the choice of intermediate frequency but suggest that if double heterodyning were used, as in a normal short-wave set, no break would be required in the frequency band and short wave working would be provided as a matter of course.

Perhaps the authors of the plan could state why four of the regional transmitters for the British Isles require allocations in the long wave band ? The situation of the regional transmitters, to my mind, leaves a great deal to be desired. For example, Hull is not covered by any of the stations. An apparently superfluous station is located at Plymouth to serve a very small area which is not densely populated. There is no Welsh regional transmitter, and half of Wales is not covered in any way. The present sites of stations seem to have been entirely ignored (in the Scottish area, two transmitters are situated at Westerglen and Burghead, but no use is made of these ; instead, a transmitter is to be set up at Glasgow which, as a built-up area, will give relatively poor ground-wave propagation).

Regarding the remark that long waves are unsuitable in latitudes lower than 40° N. because of atmospheric noise, perhaps the authors could elaborate this point as it is an unusual statement.

The plan, as a whole, seems to have been proposed by the Radio Industry without sufficient regard to practical transmission difficulties and without sufficient reference to the European authorities (the bibliography quotes one British article and ten American). In view of the allocation of wave bands on most present-day receivers, i.e., 200 to 550 metres in the medium wave band and 1,000 to 2,000 metres in the long wave band, it appears that the adoption of the suggested frequency allocations would render many receivers useless, and a state of affairs would exist similar to that experienced with present utility receivers. This, I feel, is very undesirable unless alternative wavelengths are provided during the transition period.

Commander Turnbull : I feel that the R.I.C. Committee has worked hard to produce a sound plan for broadcasting, but I take exception to their attitude to the problem indicated by the very first sentence on page 6. This says: "The Report confines itself to considering only existing transmitting means." Now that, to my mind, is a retrograde step. It is appreciated, of course, that new types of transmission or new frequencies make necessary new sets, but at the present time, when there is a tremendous market for new sets, it seems opportune to consider all the possibilities. In addition, short wave converters can now be provided with circuits adaptable to frequency modulation. Considering the areas proposed to be covered, the North West of Scotland will be poorly served, likewise Western Wales. I would also suggest that the coast bordering the North Sea and English Channel receives very much better treatment than that bordering the Atlantic and Irish Sea.

Lieut. Dickson : It was stated that every region should have the opportunity of listening to radiation from other regions presumably with the idea of an exchange of ideas. I presume that this will be left entirely to directional ultra high frequency transmissions. Furthermore, it appears that the broadcast listener is only going to have made available to him those portions of the frequencies spectrum which he now enjoys; there is no consideration given to reallocation of frequencies of existing services.

Mr. Hay : I agree generally with the proposed plan which, while not complete, does provide a reasoned method of approach to the problems involved. Commander Turnbull referred to frequency modulation, and other developments which will, no doubt, come into use eventually in some fields of broadcasting. I do not think, however, that at the present time the public would tolerate enforced complete replacement of its receiver; although the manufacturers would not be disturbed by such a situation, I think that they would produce whatever was required to cover the broadcast receiving requirements in this country or elsewhere.

As regards the European situation, Russia is, of course, an unknown quantity, and I agree with the "n" figure.

The complaint of a previous speaker that the North of Scotland is poorly served would be obviated, of course, by the short wave transmitter working purely in the local area.

S. G. Button : After a careful study of the report, I consider certain fundamental details have been overlooked, and that there are other points which must be regarded as open to question.

Firstly, I cannot agree with the authors' contention that the range of a station is equal to wavelengths divided by 4. If this be so, then it does not matter whether the aerial is 1 ft. or 1,000 ft. high; the range of the station is assumed to be the same, which is surely misleading.

Secondly, the minimum field strength of 2 mV per metre appears to me to be excessively high. A signal of 100 uV/M would be ample to overcome man-made static except in the very worst cases, and surely the whole plan is not being based on requirements under exceptional circumstances.

Thirdly, in one sentence, the report stresses the importance of good quality reproduction, but later mentions that by allowing 11 kc/s separation between stations "the heterodyne whistle is far less audible." Surely these two statements are contradictory.

Fourthly, the first paragraph stated "other methods rendering obsolete all existing receivers have not been considered," and yet the plan suggests employment of frequencies of 330 to 450 kc/s, a band which I suggest is not provided in 1 per cent. of receivers in use in this country to-day.

Lastly, the report is concerned throughout with the employment of omni-directional aerials. I hope to show that the use of uni-directional aerials under certain circumstances has much to recommend it.

Considering these points more fully, the fading radius of any station depends upon :—

(1) The height of the receiving and transmitting aerials.
(2) The transmitter power.
(3) The ground conductivity.
(4) The nature of the terrain.
(5) The frequency of operation.

How can it possibly be suggested that all these points are taken into account in the formula Range = Wavelength divided by 4. Naturally, this question and the second one are very closely related.

Regarding the third point, a separation between channels of 11 kc/s is certainly a step in the right direction, but it is insufficient to make any material contribution towards improved quality of reproduction. As I will show later, many channels now suggested could be eliminated and I would expect that, assuming it is finally decided to adhere to the range of frequencies outlined in the report, a separation of at least 15 kc/s to be attainable.

On the fourth question, I quite agree that we should avoid, if at all possible, anything which renders present receivers obsolete, and yet the report seems to suggest just that obsolescence. If it is decided to adhere to a frequency separation between stations of 11 kc/s, and the suggestion made in my fifth point were adopted, the band of 330 to 450 kc/s would be unnecessary.

The most important point, namely, the employment of uni-directional aerials, now requires attention. I strongly suggest the employment of uni-directional radiation for both National and Regional services, where the advantages of directive radiation are worth while.

Let us consider the proposed plan, dealing firstly with the National services, and see how uni-directional radiation will be of benefit. Take Italy as an example. In the proposed plan there is at least a 60 per cent. waste of radiation, the toe and heel of the mainland, together with Sicily, and the apparently thickly populated northern region not being adequately served. One station situated, say, north of Milan, beamed in a direction 50° east of south, could adequately cover the whole country, with a better average signal than would be provided by the proposed plan, and what is even more important, a second station situated at, say, Marseilles, with its radiation beamed 20° west of north, could cover the whole of France, and the two stations could operate on the same frequency without mutual interference.

Of the 36 national channels proposed, at least 12 could be released for other services.

Other countries which could be similarly treated are Norway, Sweden, Great Britain, Spain, Eire and Iceland, to mention the most obvious examples.

Now consider the Regional scheme suggested for this country. Certain points are noticeable :—

(*a*) Considerable overlap would exist in the Somerset and Devon areas.

(*b*) Non-coverage of many areas such as Lindsey, the East Riding of Yorkshire, Westmorland, Cumberland, approximately one-half of Wales and the North of Scotland.

(*c*) Wasted radiation from the Plymouth, Norwich, Newcastle and Aberdeen transmitters.

All the above criticisms could be eliminated by careful employment of uni-directional aerials, and with the proposed number of stations the whole of the country could be adequately served. Of the proposed 93 Regional channels, it would be possible to dispense with one-third.

W. W. Smith : The report seems to be based on a simple formula, and it is interesting to consider how it has been derived. The range is given, expressed in miles, the wave length in metres, divided by four. This is based on the assumption of a certain conductivity of the soil and a ratio of space to ground wave of 1 to 3. This means that if one takes any point relative to the transmitter the wave reflected from the Heaviside layer returns at one-third the strength of the direct ray. This is all very well if the reflected ray remains absolutely constant, but this varies considerably in magnitude, and the figure of 1 to 3 is only an average. It is quite possible that where the ground wave intensity is only 2 millivolts per metre, the reflected wave would be of the same order, and it would not be right to use the average value of 1 to 3. The wave reflected from the Heaviside layer and the direct ray might be in phase or out of phase with one another, in which case the signal strength could be 4 millivolts per metre, or zero. That would cause severe distortion.

I would expect considerably lower service values than those suggested in the R.I.C. plan.

It is stated on page 9 of the report that, taking Great Britain as representative, and comparing roughly the areas and population distribution of other countries, a list could be made which showed the number of regional wave lengths required in every country. According to Appendix 5, Great Britain is to be allocated 10, plus the other two stations, and the same allocation is made for France. The population of Germany is twice that of Great Britain, and the former country is allocated 10 wavelengths. Presumably, Britain should have only 5 stations if Germany were to have 10. Northern Italy, near the Alps, is virtually uncovered, and this is the area which has the most dense population. Again, why should Iceland need such a large station ? Further, all the Russian stations are to share their wave lengths with someone else. It is suggested that interference would be negligible, and that each country would also have its national station. Argument would certainly result from this on the grounds that during the war Britain had managed quite well with two wave lengths, and now, apparently, must have 12 : a reasonable argument from the foreign point of view.

On the technical side, I do not consider that the adoption of 11 kc/s is going to improve matters. A minimum signal strength of 2 millivolts per metre is, in my view, a reasonable figure nowadays.

With reference to directional aerials, I recall that Start Point was able to cover the whole of Cornwall and South Wales, and I suggest that five such transmitters could be arranged in England to give adequate coverage. Other small areas might be covered with lower frequency stations, using aerials with high-angle radiation.

M. Exwood : In suggesting the use of small booster stations in areas with low soil conductivity, reliance seems to have been placed on the reciprocity of transmission and reception. I disagree on that point because, fortunately, while the ground radiation might be very poor indeed, the sky wave would be radiated from those areas quite effectively, as it is not to any large extent affected by the soil conductivity. As to the use of directional aerials, it is probably a mistake to suggest that the service area could be considerably extended by that means. I agree that energy could be saved by concentrating the radiation where it was wanted, but, assuming that the formula derived for the range of a station was correct, it could not be expected that even with directional aerials it would be possible to extend the service area beyond that given by the formula. I am of the opinion that a signal strength of 2 millivolts/metre is not high. There are areas where the signal strength is approximately 1 millivolt/metre and interference is quite considerable. The present Droitwich transmitter, operating on 1,500 metres, and which should have a range, therefore, of 375 miles, faded seriously in the Newcastle area. There is also fading in the Plymouth area due, probably, to the exceptionally low soil conductivity of Dartmoor. Further, the Midland Regional transmitter fades at Nottingham, only 62 miles away, while the range would be calculated at 75 miles.

H. Kinross : It would be interesting to know if the authors of the plan have considered the difference between day and night conditions, and also between maximum and minimum sun spot conditions in adopting their formula for the effective range of broadcast transmitters.

REPLY TO THE DISCUSSION

(Submitted by the Authors of the R.I.C. Report)

Before proceeding to deal with the detailed points of criticism raised by various speakers, the authors feel that it is necessary to establish that the R.I.C.'s plan for post-war European broadcasting sets out to deal with the problem of medium and long waves only, and, as stated on page 6 in the first paragraph, restricts itself to the use of existing transmitting means only with normal amplitude modulation, as otherwise the vast majority of broadcasting receivers in Europe would be rendered obsolete. Furthermore, it is generally accepted that the existing means provide the most economical method of broadcasting so far as the cost of the receiving equipment is concerned.

The authors visualised that there would inevitably be locations where it would not be possible in broad plan to provide a satisfactory broadcasting service, and provision was made for dealing with this situation by common channel working at low power in the high frequency end of the band, thus making use of a well-established practice.

So far as broadcasting in the V.H.F. bands is concerned, it can be stated that this is being studied actively by the various technical committees of the Radio Industry Council, but this problem is less urgent, to the extent that the means for such broadcasting are not yet generally available in Europe.

Dealing first with the remarks of Mr. P. Adorian. who takes exception to the validity of our assumed factor $\lambda/4$ (see page 35 of the January-February issue, 1946, of the Institution's Journal). That this factor is an arbitrary approximation has been repeatedly stressed in the paper, but as it necessarily depends on a number of conditions varying between wide limits over the area of Europe, it is considered to be justified. In support of the authors' views, the following instances are recalled :—

(*a*) The reception of North Regional over a wide area of the Midlands and Home Counties.

(*b*) The reception of Fécamp in the Midlands.

(*c*) Athlone, as received in East Anglia.

(*d*) Luxembourg—reception in the North of England.

(*e*) The generally satisfactory performance, from the fading point of view, of Droitwich over England and the Scottish Lowlands.

These remarks refer to night-time reception and are instances of the ground ray being maintained sufficiently to give reception reasonably free from fading difficulties. Local difficulties due to high ambient noise level can be overcome with increased transmitter power.

Several of the instances quoted show workable radii considerably in excess of $\frac{\lambda}{4}$ and if it be remarked that these are known to be due to well-sited transmitters, then the reply must be that transmitters should be well-sited !

On page 35 of the issue of the Journal, Mr. Adorian joins issue with us on the proportion of time during which the quasi-maximum value of the indirect ray equals or exceeds that of the direct ray. The fact that broadcasting in the past has been reasonably satisfactory at the limit of the service radius as defined seems to us to disprove his contention and makes more likely a much lower figure, which we have assumed.

On page 36, paragraph 4, Mr. Adorian considers that our figure of 10^{-13} E.M.U. is too optimistic. Again, this is a controversial point. We consider that we have quoted an average value. There are obviously places where better propagation exists, for instance, where the path is generally over water or fertile, cultivated land, and others where the conditions are worse, for instance, in Switzerland, but reference to the Plan in detail will show that where such difficulties and advantages are obvious we have recognised the fact by allocating very much lower or higher frequencies than are indicated either at the figure 10^{-13} E.M.U. or $\lambda/4$. This point was also raised by Mr. Humphries in the Scottish discussion, and we feel we should point out that in allocating a comparatively high frequency, for example, to Denmark, we have taken into account the better transmission conditions that exist where the path is largely over water.

On page 35 on the Table shown, Mr. Adorian considers that common frequency working as is proposed on certain channels in the long waveband is impracticable. He quotes figures that are intended to serve as proof of the point for Iceland and Russia which, in two cases, have been allocated common channel working. We feel we should point out that for a number of years past a pair of stations in these countries have shared a long-wave channel, and, so far as any evidence to the contrary goes, this has been entirely satisfactory. It is, therefore, reasonable to suppose that corresponding examples of geographical separation should be equally workable.

On page 37, Mr. Adorian considers that the Plan should take account of broadcasting locations in the whole of Russia. We must point out that we have information on good authority that, in the past, European broadcasting schemes have been restricted to European Russia as defined by the limit of longitude East 40 deg. In any case, it will be possible to make use of lower frequency medium wave channels with restricted common channel working, thus affording a satisfactory solution to this problem.

We feel that Mr. Adorian does not make a good case against the main principles of the Plan, even though he may be interested in doing so. We consider that there are two false assumptions at the basis of his argument—

1. That in the past European broadcasting has been fundamentally unsatisfactory, and
2. That if the war had not intervened, Montreux would never have been put into effect, and, even had it been implemented, would not have worked satisfactorily.

Mr. Scroggie and Mr. St. John Jones refer to interfering heterodynes from U.S.A. transmitters. If, in point of fact, a new factor is introduced into broadcasting problems in Europe by the use of higher power transmitters in U.S.A., it is certainly possible that synchronisation with the U.S.A. in the medium waveband may have to be considered.

Colonel Parker, Flight Lieutenant Bovill, Mr. Wardman and Mr. Humphries all make reference to the limited use envisaged of frequencies between 300 and 500 kc/s. We must point out that we have only allocated these assignments in a manner similar to the Montreux Plan, where the geographical location of the transmitters involved was restricted so that interference with existing services was a minimum. It is perhaps worth pointing out that it is recognised that the vast majority of these services do not need channels of 10 kc/s, but can operate with complete satisfaction when separated by as little as 1 kc. There are indications that not only has this been recognised but that the principle is likely to be adopted. The use of these frequencies for broadcasting, therefore, becomes even more possible.

Mr. Humphries, in the Scottish discussion, considers that long wave allocations nearer the Equator than 40 deg. North are practicable. The authors would like to point out that there is plenty of evidence that this is not so, due to the high level of natural static in these latitudes. He is also dissatisfied with the apparent disregard of the broadcasting needs of Scotland and Wales. The map, which appears as Appendix VIII, gives transmitter locations which are intended to indicate the industrial area served rather than the precise location of the transmitter. Perhaps the meaning would have been clearer to him if the wording "transmitter localities" in preference to "location" had been used. For instance, the transmitter at Washford was visualised as providing the South Welsh service, as it does to-day, and a large transmitter near Plymouth was expected to serve the whole West region, as now, i.e., Start point. So far as Scotland is concerned, it was not expected that the sites chosen would be other than those which exist at present.

On page 10 of the Institution's Journal, Mr. Sargrove quotes common frequency working in Germany in 1939 as an example to show that such practice in unworkable in general. The authors would like to point out that the idea as described in the Plan was to separate transmitters operating on common frequencies by large geographical distances. The problem is, therefore, of an entirely different kind to that quoted by Mr. Sargrove. We feel that a study of such allocations in the Montreux Plan will establish the point.

TRANSFERS AND ELECTIONS TO MEMBERSHIP

The following elections and transfers were recommended by the Membership Committee at their meetings held on the 26th February and 12th and 29th March, 1946. At these three meetings the Committee considered a total of 136 proposals for election or transfer to Graduateship or higher grade membership.

The full analysis of proposals received and accepted during the year ended the 31st March, 1946, will be given in the Annual Report of the General Council which will probably be published in the June-July, 1946, Journal.

Transferred from Associate Member to Full Member

BARNETT, Henry Cameron Leslie	Potters Bar
LEVY, Maurice Moise	Earls Court, S.W.5

Transferred from Associate to Associate Member

BENTLEY, Joseph William Cyril, Capt.	Kidderminster
BROWN, John Isaac, Major	London, N.4
HARVIE, Ronald James Martin	Dover, Kent
HOLFORD, Frank Alan, B.Sc.	Purleigh, Essex
HOPE, Richard Arnold Hoyle, M.Sc.	Westhoughton, Lancs
JONES, Arthur Allerston	Sheffield
JONES, Harry	Leicester
LAMPITT, Robert Alfred	Wolverhampton
LAYZELL, David Robert Rayner	Sutherland
WARNER, Clifford	Lincoln

Transferred from Student to Associate Member

RICHARDSON, Leonard Geoffrey	N.S. Wales, Australia
ZELINGER, Geza	Haifa, Palestine

Transferred from Associate to Companion

SNOOK, Wilfred	Dorchester

Transferred from Graduate to Associate

SHAPLAND, Albert John	Swansea, Glam

Transferred from Student to Associate

BRADFORD, Alexander	Edinburgh, 8
FARLEY, William Morrison	London, S.W.11
HUTSON, Geoffrey Henry	Canterbury, Kent
KING, John Christopher	Bromley, Kent
McCORMICK, John Marriott	Worcester
SQUIRES, Terence Leighton	Wolverhampton
STIBBE, Harry	London, S.E.20
TURNER, Lewis Edgar	Deal, Kent

Transferred from Student to Graduate

BETTS, Charles Anthony Blundell	Leigh, Lancs
FERGUSON, Fergus Esler	Belfast, N.I.
FORD, Edmund Alfred	Nottingham
TOLL, John Walker	Hitchin, Herts
WARD, Geoffrey Martin	Lincoln
WIESNER, Adolf Leopold	London, W.9

Elected to Full Member

HUGHES, James Robinson	London, N.W.7
SYKES, Jack	Croydon
WIDGER, Leslie	Tolworth, Surrey

Elected to Associate Member

ATKINSON, David	Falkirk
BREAKELL, Charles Clayton, Lt.-Cmdr.	Preston
CARTER, Donald Walter, B.Sc., Lt.-Cmdr.	Southampton
CORNISH, Walter Leslie	Plympton, Devon
DAVIES, Benjamin Howell	Cardiff, S. Wales
DAY, George Arthur	N.S. Wales, Australia
LOCKE, Dick Norton	West Wickham, Kent
McCULLAGH, John Charles	London, W.3
McCUSKER, Joseph Bernard	Roby, Lancs
PARSONS, David Reginald	Shipley, Yorks
RAIKES, Robert Martin	Newent, Glos.
SMITH, Edward Thomas, B.Sc.	Hatch End
WILKINS, Edwin John, Major, M.B.E.	Cape Town, S. Africa
WILLIAMS, Ernest Leslie, S/Ldr.	Cosham, Hants

Elections (*contd*).

Elected to Associate

ARMSTRONG, George Kenneth, B.Sc.(Hons.)	Stockport
BAXTER, Kenneth Arthur	Beckenham
BILLIN, James John	London, S.E.15
BONNEFOI, Jean Yves Francois	Paris, 17
BRANDT, Joseph Charles	Purley, Surrey
BROADLEY, Charles Eastley	London, N.8
BUCK, Peter James	Paignton, Devon
BYFORD, Thomas William David	Leeds, 6
COE, Clifford Lionel	London, W.13
DATT, Krishan Sharma	Nairobi
DAY, George Frederick	Bournemouth
DE CLERK, Jan Raoul Francois	London N.W.11
DE JONG, Leslie John Bernard	Southbourne, Hants
DENNY, William Guy Cecil	Andover
DUCKELS, Leslie	Goole, Yorks
EADES, Robert George, B.Sc.	Crieff, Perthshire
EXLEY, Donald	Sheffield
GAMMON, Michael Morley Johnston, B.Sc.	Kingston-on-Thames
GRANT, Francis, Major	Beckenham
GRIFFITHS, Leslie	Wrexham
HALL, John Russell	Auckland, S.E.3, N.Z.
HARRY, Glyn Lewis	Leamington Spa, Warwicks
HARVEY, Laurence Marshall, M.A.(Hons.)Cantab.	Ickenham, Middx
HENNELL, Ian Alexander, M.A.(Cantab.).	London, S.W.15
HENNESSEY, Gerard	Preston, Lancs
HOPKIN, Thomas Arthur	Rhondda, Glam
JOHN, Stuart Wilton	London, S.W.20
KONGSHAUG, Arthur L.	Oslo
LAMBERT, Ernest Eric	Liverpool, 6
LOVELOCK, Peter Albert	Loudwater, Bucks
McCHEYNE, Alfred	Orpington, Kent
McLACHLAN, Donald	Thornton Heath
METCALFE, Christopher James	Bury, Lancs
MUIR, Leslie Malcolm, B.Eng.	Birkenhead
NASH, Frank Joseph, B.Sc.(Hons.)	Liverpool, 19
OWEN, Thomas Mervyn Gray	Stretford, Lancs
PARDOE, Sidney	Stourbridge, Worcs
PARKER, Norman Reginald	London, S.E.6
REDER, Nathan	London, E.3
REED, Clifford Albert	Whitton, Middx
REGNAULT, Pierre Louis	London, S.W.4
RICHARDS, John Charles	Aberdeen
RICHENS, Sidney Arthur	Carlisle
RODICK, Ralph Stanley	London, N.10
SCHMIDT, Alan Robert	Wellington, N.Z.
SHEPHERD, Robert Henry	Leeds, 7
SMAIL, James Elliott	Edinburgh, 9
SPENCER, Eric, B.Sc.(Hons.)	Oldham, Lancs
STEWART, Athol Gordon	Wellington, N.Z.
THOMASON, John Bernard	Oldham, Lancs
TIMMINGS, Harold Edward	London, N.W.8
TRAVIS, Philip Edward B.Sc.(Hons.)	Calne, Wiltshire
TREACHER, George Augustus	Plymouth
TURNER, George Norman	Alfreton
TWEEDLE, Alan Kay	Nottingham
VERSEY, Hedley Varlow	Southampton
WALTER, Norman Archibald	Portsmouth
WHITTAKER, George William Keeton	Southsea
WILLIAMS, Brian Arthur	Bristol, 6
WILSON, George Raymond	Bradford

STUDENTSHIP REGISTRATIONS

The following were registered as Student members of the Institution at meetings of the Membership Committee held on the 26th February and 12th and 29th March, 1946. At these three meetings the Membership Committee received a total of 44 proposals for Studentship Registration.

ASHWORTH, Stanley	Salford, 7, Lancs
BROWN, Albert Geoffrey	Nottingham
BROWN, Frederick Arthur	Barkingside
BUDDS, James Hamilton	Chelmsford
CASPERD, Stanley George	Littlehampton
CHENEY, Harry Seymour	Wallingford
CORKILL, William Robert	Glasgow, S.1
DAVIES, Joseph Henry	Middlesbrough
DIXON, Howard Francis	Theydon Bois, Essex
EDE, Robert	Jarrow, Co. Durham
FRANCE, George	Bamburgh, Northumberland
FURZE, John David Rodrick	Victoria, Australia
GARD, Ronald Lione.	Swansea
GRUNBERGER, Hanus	Praha, 2, Czechoslovakia
HARVEY, Trevor Frederick	Wolverton
HAZELDINE, Frederick William	Ilford
HEISLER, George Henry, B.Sc.	Southport, Lancs
HODGES, Hugh Walter Charles	Bridgwater
HODGSON, Alfred Ernest Musgrove	York
HUDSON, Roland Frederick	Lincoln
JACOBS, Jack	Ipswich
JAMES, Antony Bracher	Welwyn Garden City
JONES, Stephen Brocklebank	London, N.19
KANAGALINGAM, Suppiah	Ceylon
LEVY, Franz	Jerusalem
LIMBRICK, Clive Albert	Exeter, Devon
MACDONALD, Neil Fraser, M.A. (Hons.)	Glasgow, W.2
MACKIE, James Alexander	Inverness
MAUNSELL, Thomas	Tralee, Co. Kerry
McHARG, Norman	Newcastle-on-Tyne
MOTT, Richard John	London, N.12
NARENDRANATHAN, Ponndurai	Ceylon
PALFREEMAN, Alfred	Darlington
PLANT, Victor Ernest	Walthamstow
PROVAN, William McLean	Motherwell
RICHARDSON, Leslie	Barrow-in-Furness
RITCHIE, Roy	Rutherglen
ROBB, Ronald Laidlow	Kenton, Middx
RODGER, Ridley	Hebburn
SANYAL, Gitindra Saran, M.Sc.	Assam, India
SUNDARAM, Alampallam, Venkatasubramaniam	S. Malabar
THOMASSON, Donald William	Exeter, Devon
WALPOLE, Walter Leslie	Lowestoft
WILLIAMS, John Harding	Mosman, N.S. Wales

GOLD FILM ELECTRODES FOR HIGH FREQUENCY QUARTZ PLATES

by

R. A. Spears (Associate Member)*

(A paper to be read before the Midlands Section of the Institution on January 17th, 1946, repeated before the London Section on January 29th, 1946, and the North-Western Section on February 26th, 1946.)

This paper has been written to correlate information relative to the manufacture of Quartz Crystals of the plated-electrode type.

It is essential, to facilitate manufacture, that the various effects of plating upon frequency, activity and stability should be predictable.

The first object of the paper is to develop a simple practical theory to account for and predetermine the frequency changes resultant upon varying amounts of plating.

Closely associated with the design of electrodes and indeed inseparably linked, is the subject of clamping and mounting of the quartz plate. This is dealt with and reasons for selecting the location of clamping points are discussed at some length.

Inferior plating is known to cause trouble, not merely immediately, but also over a long period. Some precautions to be observed in respect of plating quality are suggested.

Throughout the paper the aim has been to provide information with a minimum amount of indirectly associated theory.

The extension of the theory not only to sputtering of gold, but also to the electro-plating method, increases its value and enables an extremely fine adjustment of frequency to be made.

Contents

List of Symbols used

α	= Acceleration, cms./sec.^{2}
$\sigma_{xx}, \sigma_{yy}, \sigma_{zz}$	= Tensile stresses, dynes/cm^2
$\sigma_{xy}, \sigma_{zx}, \sigma_{yz}$	= Shear stresses, dynes/cm^2
P	= Polarisation, coulombes/dyne/cm^2
d_{11}, etc.	= Piezo-electric constants, coulombes/dyne.
E	= Potential, volts.
I	= Current, amperes.
C	= Capacity, farads.
C_g	= Capacity, grid to earth, farads.
q	= Charge, coulombes.
r	= Ratio $\frac{\text{Electrode Area.}}{\text{Quartz Area.}}$
ρ	= Density, grams/cm^3.
X	= Reactance, ohms.
ω	= $2\pi f$.
ϵ	= Elastic modulus.
t	= Time, seconds.
A	= Area, cm^2
T_0	= Thickness of quartz plate, cms.
T_q	= Equivalent thickness of quartz, cms.
M	= Mass in grams.
Q	= Quality factor of resonant circuit.
R	= Resistance, ohms.
K	= Constant = $3{\cdot}83 \times 10^{-12}$.
γ	= Constant = $1{\cdot}35 \times 10^{-6}$.
β	= Constant relating velocity and current.
B	= Constant denoting particle displacement.
D	= Constant denoting conductance of diode.

* Automatic Telephone and Electric Co., Ltd.

1. Introduction

During recent years and particularly since the commencement of hostilities in Europe, the use of quartz crystals to control high-frequency oscillators has become universal.

It was soon realised that, for mobile and aircraft equipment, the simple crystal assembly units in use some years prior to the war left much to be desired.

Particularly was this applicable to mounting methods. The severe shocks and multiple vibrations set up in tanks and aircraft caused a severe strain upon the assembly as a whole.

It appeared that one solution to this major problem might lie in the direction of an extremely light, semi-resilient construction. Accordingly, a crystal unit was developed having a minimum of metal in the assembly and its electrode system comprising an extremely thin film of metal.

The provision of the film and its attendant modifying influence upon the frequency, activity and permanence of the unit so produced formed a subject for thorough investigation.

This paper has been written to put some of the facts on record and, it is hoped, to assist others in some degree.

Although the paper deals only with the specific case of gold film electrodes, it is thought that it should in general apply equally to other metallic deposits.

2. Electrode Design

2.1. *Function of Metallic Film Electrodes*

In order that we may utilise the mechanical resonance properties of natural crystalline quartz to control electrical circuits, it is first essential to have at our disposal the means of converting an electric alternating potential into a mechanical stress.

There is a series of equations [1] which illustrate the relationship between polarisation P, and the resultant stress in given directions.

$$\left.\begin{aligned} P_x &= d_{11}\sigma_{xx} + d_{12}\sigma_{yy} + d_{13}\sigma_{zz} + d_{14}\sigma_{yz} + d_{15}\sigma_{zx} + d_{16}\sigma_{xy} \\ P_y &= d_{21}\sigma_{xx} + d_{22}\sigma_{yy} + d_{23}\sigma_{zz} + d_{24}\sigma_{yz} + d_{25}\sigma_{zx} + d_{26}\sigma_{xy} \\ P_z &= d_{31}\sigma_{xx} + d_{32}\sigma_{yy} + d_{33}\sigma_{zz} + d_{34}\sigma_{yz} + d_{35}\sigma_{zx} + d_{36}\sigma_{xy} \end{aligned}\right\} \quad (1)$$

In this series, the suffixes xx, yy, and zz indicate tensile stress in the direction of the axes of natural quartz[2]. Those suffixes yz, zx and xy similarly designate shear stresses. Since many of the piezo-electric constants d_{12}, d_{13}, etc., are zero, Eq.1 many be simplified to :

$$\left.\begin{aligned} P_x &= d_{11}\sigma_{xx} + d_{12}\sigma_{yy} + d_{14}\sigma_{yz} \\ P_y &= d_{25}\sigma_{zx} + d_{26}\sigma_{xy} \\ P_z &= 0 \end{aligned}\right\} \quad (2)$$

Suppose now that we desire to excite an "xy" shear stress in a + 35° cut crystal plate. From the series above we obtain, by eliminating the terms containing zero values of piezo-electric constants .

$$\sigma_{xy} = \frac{P_y - d'_{25}\sigma_{zx}}{d'_{26}} \quad \ldots\ldots (3)$$

Since both d'_{25} and d'_{26} are constants in this instance, then the maximum stress σ_{xy} will be brought about by a maximum value of Py, the polarisation component along the y or mechanical axis of the crystal plate.

The second term on the right hand side of Eq.3 defines the "zx" shear stress created simultaneously with the "xy" shear stress. They may therefore be considered to be coupled with electro-mechanically. The strain or displacement will be proportional to the stress in practical cases.

The problem of the conversion from an electrical potential to a mechanical stress with the highest possible efficiency has then resolved itself into the provision of the maximum value of Py. Consequently, the electric field through the quartz plate in the direction of the y axis must be a maximum.

Fig. 1 illustrates a pair of metal electrodes having, as dielectric, the quartz wafer and a certain amount of air spacing.

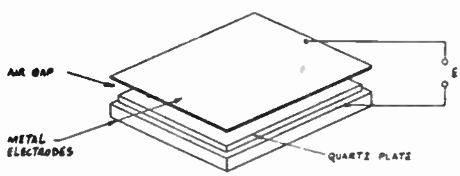


Fig. 1. Essentials of Air-gap Type Assembly

It will simplify explanation if the quartz is deemed to have virtual electrodes integral with its major faces. It will then be quite obvious that we have, in effect, two condensers in series, the first having a dielectric constant of 4·5 for quartz and the second of unity, for air. This results in a higher field intensity per millimetre for the air in inverse ratio to the respective dielectric constants.

So we see that the presence of the air gap materially weakens the effective field acting upon the piezo-electric element.

We find, then, that mechanical damping being for the moment ignored, the absolute minimum of airgap becomes desirable. It is possible to place the metal electrodes in contact with the quartz surfaces and thus obtain a very small airgap and an increased field intensity.

There are at least two serious objections to this procedure. Firstly, since the electrode has considerable mass which tends to oppose the stress created by the applied potential, this is reflected as a "loss" and may actually be represented in an equivalent circuit by a resistance. Added to this resistance is that caused by friction between quartz and electrode.

The foregoing has primarily been concerned with the static conditions of electrodes with a piezo-electric dielectric. The second objection to contacting metal electrodes is the possibility of their being disturbed by mechanical vibration; either that due to the resonant vibration of the quartz plate itself or, alternatively, to external causes. The full effects of the electrode mass upon the dynamic state of a high frequency crystal plate are dealt with subsequently in section 4.

Therefore it would appear that a desirable line of attack would be to use a comparatively light, durable, but thoroughly integral metallic film, leaving no air-gap whatever. It was along these lines that high-frequency plated crystals were developed.

2.2. *Area of Electrodes*

In Fig. 2 is represented a simple Pierce crystal oscillator circuit. When the circuit is about to commence oscillations, a very small initial pulse or fluctuation in the plate current of the valve is passed, via the anode to grid valve capacity, to the electrodes energising the piezo-electric element.

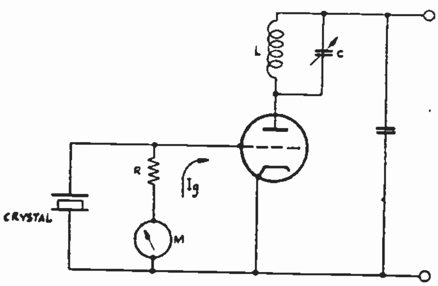


Fig. 2. Pierce Crystal Oscillator

These electrons, flowing into the capacitance provided by the crystal electrodes with the quartz element as dielectric, build up into a charge q coulombes.

For any given quartz plate, the maximum polarisation Py will be obtained when $E = q/C_g$ is a maximum. Since q is limited by the magnitude of the plate current fluctuation and also by the magnitude of the valve inter-electrode capacities, it will be apparent that in these circumstances an increase in E is obtainable only by a decrease of C_g. The numerical value of C_g is given by the sum of the crystal electrode capacity and the input capacity of the valve.

Superficially, it would appear disadvantageous to attempt to reduce the area of the electrodes, since this would proportionately decrease the area of crystal activated under static conditions.

In the dynamic state, however, this does not hold good except for very low frequencies.

It has been shown* that a material reduction in the area of one or both of the metal electrodes will produce an increase in the crystal activity under practical dynamic conditions.

The important feature is the relationship between frequency of resonance and the ratio $r = \frac{\text{Electrode Area}}{\text{Quartz Area.}}$

For example, the optimum value of "r" for a quartz plate designed to operate at 1 Mc/s is given as 0·85. Higher up the frequency spectrum at 10 Mc/s this ratio become 0·25.

The reason for this is difficult to understand and still more difficult to prove mathematically. A reduction in electrode area below that of the quartz plate does not merely result in a given charge being collected and used to build up a potential in a relatively small capacity.

To be able to find a reason for this behaviour, we must examine the dynamic state of a quartz plate qualitatively.

2.3. *Surface Displacement*

There are three common methods for the examination of surface displacement under dynamic conditions at high-frequency resonance.

The first is that originated by Dr. Dye and utilises the stroboscopic principle [3]. This is excellent for the purpose, but requires considerable equipment.

Secondly, there is the well-known Lycopodium powder method [4]. This is excellent for lower frequencies, but hardly suitable for use with high-frequency plates.

Lastly, there is a simple method which only requires an insulated, sharp-pointed "probe." The crystal plate to be examined is plated over its major surfaces by the means outlined in Section 3.

Referring to Fig. 5, the series arm of the crystal equivalent circuit L, C_1, R passes a current I_s which is proportional to the mean velocity of the plated zone of the crystal.

That is:

$$I_s = \frac{\beta}{x_2 - x_1}\int_{x_1}^{x_2} v.\ dx \qquad \ldots\ldots\ldots\ldots(4)$$

in which β is a constant, relating velocity and current.

The radio frequency potential developed across the terminals A B will then be:

$$E_{RF} = Z_s I_s = I_s\sqrt{R^2 + \left(\omega L - \frac{1}{\omega c_1}\right)^2} \qquad \ldots\ldots(5)$$

for a plated crystal, as Cp is infinite.

As is obvious from Fig. 2, the resultant grid current

* British Patent 548517.

Ig will be directly dependent upon E_{RF} and, therefore, upon the mean velocity of the quartz surface particles. The modification of the mean velocity due to the application of a localised restrictive force by the "Probe" therefore is reflected in the behaviour of the "grid" current Ig as read on the meter M.

Further, ignoring the comparative frequency instability, the crystal may be replaced by a parallel tuned inductance and capacity, the grid current being brought to the same value as that obtained with the crystal, by means of added resistance in the tuned circuit.

The dynamic impedance of the substituted circuit may then be calculated and recorded for future reference.

It should be borne in mind that the effective Q for a crystal oscillating in a Pierce circuit is given by:

$$Q = \frac{X}{R} = \frac{R^2 + \left[\left(\omega L - \frac{1}{\omega c_1}\right)\left(\omega L - \frac{1}{\omega c_1} - \frac{1}{\omega c}\right)\right]}{-\frac{1}{\omega c} R} \quad (6)$$

which may be simplified to:

$$Q = \frac{\omega c}{R}\left[R^2 + \left(\omega L - \frac{1}{\omega c_1}\right)^2\right] \quad \ldots\ldots\ldots\ldots\ldots (7)$$

for the typical crystal circuit wherein C is much larger than C_1. The magnitude of Q will depend upon ω, but will rarely exceed 100.

The exploration of a typical 5Mc/s "BT cut" quartz plate by a "probe" is relatively simple. The electrodes are connected to the oscillator by means of thin wires or spring contacts. The "probe" is then applied to the surfaces of the plate, systematically and in such a way as to explore the whole of the surface area in very small steps.

The variations in grid current Ig or in the equivalent parallel impedance Zp are then recorded, together with the location of the "probe" for that particular reading.

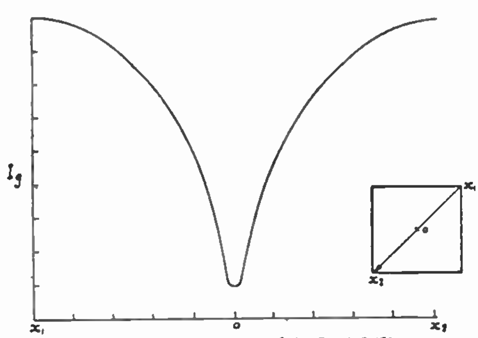


Fig. 3. Effect of Probe upon Grid Current

Fig. 3 is drawn from readings obtained in the exploration of the surface of a 5 Mc/s plate. The relative quiescence of the peripheral regions of the plate is obvious, the decrease in grid drive as the probe progresses toward the central zone being quite pronounced.

2.4. *Location of Electrodes*

In the case where the optimum activity is available whilst still using the whole surface area of the crystal for electrodes; obviously, the location of the metallic plate or film is decided for us.

However, should we desire to utilise the principle laid out in British Patent 548517 the location and form of the electrode system becomes more arbitrary.

It is certainly desirable to apply the energising electric field only to the relatively active region of a quartz plate. This results not merely in the reduction of the static capacity of the crystal unit, but also in the material reduction of losses. Consequently, the general activity will be higher.

So it appears that we have decided upon two ideals worthy of achievement. Firstly, the provision of a film of high conductivity metal which must be in firm physical union with the quartz surfaces. In the second instance the area and location of the electrode in a manner which will induce the optimum degree of activity.

3. Gold Sputtering

3.1. *Equipment*

We are presented with various possibilities for the deposition of a metallic film on quartz. Aluminium, silver, copper, nickel and gold have all been used more or less successfully. Since, however, the specific purpose of this paper is to discuss the use of gold film electrodes, the alternatives will not be referred to again.

The film of gold required for crystal fabrication is extremely thin, being of the order of 20×10^{-6} cms. The most economical and practical scheme is illustrated in Fig. 4; this process is termed "sputtering."

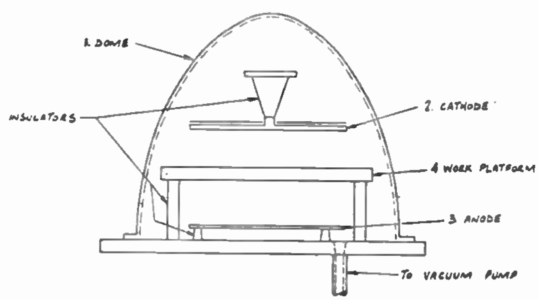


Fig. 4. Gold Sputtering Equipment

A glass or metal dome (1) capable of being hermetically sealed is evacuated by means of a pump until a pressure of about 0·01 millimetre Hg is attained.

The cathode (2) is normally of brass, the lower or working face being heavily gold plated; usually to at least 0·01 cm. in thickness.

An anode, frequently of aluminium, is placed below the cathode. Care should be taken to ensure the plane parallelism of their two opposing surfaces.

Between anode and cathode a potential difference in the region of 1,000 volts should exist. The exact voltage will depend upon the rate of sputtering desired, also upon the cathode-anode spacing and to a lesser extent upon the position of the work platform (4).

This platform is commonly of aluminium, usually having considerable volume in order to conduct much of the heat dissipated in the dome away from the quartz plates themselves.

The action of the equipment is, in short, as follows:

When the gas content of the chamber is sufficiently rarefied, a positive potential of, say, 1,000 volts is applied to the anode (3). Under these conditions, due to electronic collision with gas molecules, the space between cathode and anode becomes ionised.

The relatively heavy ions are accelerated toward the negative cathode and possess considerable kinetic energy by the time they arrive at the gold layer.

Under the impact of these gas ions, the gold molecules are dislodged or bombarded from their residence.

The resultant gold particles necessarily come to rest on the nearest surface, which is arranged to be the work upon which the gold film is to be deposited.

The molecules of gold appear to arrive at the work surface with sufficient force to flatten themselves out and seemingly to "embrace" the surface.

For this reason, a slightly matt surface provides a far more adherent base than a highly polished one.

3.2. *Surface Preparation*

The subject of surface finish and preparation for a gold sputtering process gives rise to varied schools of thought.

Generally, however, it is agreed that a final lapping or grinding should be carried out using carborundum which has been passed through a 600 and 800 mesh.

Some manufacturers then etch their plates in hydrofluoric acid to remove foreign matter and "relieve" the surface of ageing effects.

Perfect chemical cleanliness is essential, and to aid this an immersion in a bath of chromic acid is interposed at this point. The plates are then washed with pure distilled water and dried.

In many designs of crystal, indeed practically all, some form of masking is required to enable the gold deposit to be confined to the desired regions only.

Either a form of water-paint or a metal screen may be used in this connection, the masking being removed after the sputtering is completed.

3.3. *Variability of Gold Film*

Under this heading comes variation of (i) quality and colour, (ii) durability and adherence, (iii) thickness and weight of the film.

(i) The quality and colour are influenced to a great extent by the types of gas present in the chamber during plating. Water vapour particularly causes dark colouring. The presence of carbon due to arcing or brush discharge is also troublesome and avoidance of sharp corners, edges and points is the only solution.

It has been found that the use of Argon gas instead of air in the dome materially improves the quality and colour of the gold deposit at the same time giving a somewhat greater ionic bombardment of the cathode and consequently depositing rather more gold in a given period.

(ii) The durability of a deposited gold film may be considered as the resultant effect of the sputtering anode potential, distance of work from cathode, the type of gas present and the chemical cleanliness of the surface to be sputtered.

These four factors are so interdependent that it is extremely difficult to give any concrete data. Further reference to the relation between spacing of the work from the cathode and durability is made in section 4.2.

Generally, however, good durability is obtained by use of a high anode potential, together with conditions in the dome which are unfavourable to oxidation.

(iii) Variability in the thickness and accordingly the weight of the film is of great importance in the production of high-frequency gold plated crystals. The subject is dealt with fully in the following section.

4. Effect of Gold Deposit

4.1. *Effect on Crystal Resonant Frequency*

In a diode circuit operating well away from voltage saturation, i.e., subject to space-charge effects, the current flowing may be given by [6]:

$$I = DE^{1.5} \qquad \text{.......(8)}$$

where D is a constant determined by the vacuum level and electrode spacing.

This holds good for sputtering with potentials about 700-800v. It has been found, however, that between 900 and 1,200 volts anode potential, the usual rating for the process, a far more accurate value for the current is given by:

$$I = AKE^{2.5} \qquad \text{.......(9)}$$

in which A is the area of the cathode, and K is a constant determined by the electrode spacing, usually of the order of 4×10^{-12}.

It has also been found that the weight M of the deposit is, in this voltage range, practically proportional to $E^{3.5}$, that is, to the dissipation of electrical power in the evacuation chamber, the product of the applied potential E and the cathode current I.

The actual weight of gold deposited M has been found to be given by

$$M = \gamma K E^{3.5} t \text{ grams/cm}^2 \quad \ldots\ldots (10)$$

where M is in grams, γ is a constant determined by the nature of the gases present being $1{\cdot}35 \times 10^{-6}$ for air, and t represents the duration of the sputtering operation in seconds.

The weight of film, normally less than 1 milligram/cm², is very slightly affected by a pause in the action of sputtering, for example, in the event of the necessity for a film on each face of a quartz plate. This will cause a reduction in the deposit of up to 5 per cent.

The addition, in a perfectly adherent state, of a metallic electrode or electrodes on a quartz plate, has the effect of increasing the equivalent inductance value L.

Fig. 5 shows the equivalent electric circuit of a crystal. The value of Cp is infinity in a plate with contiguous electrodes. C represents the electro-static capacity of the two electrodes having the quartz wafer as dielectric. C has a value dependent upon the Elastic Modulus in the direction normal to the major surfaces of the plate.

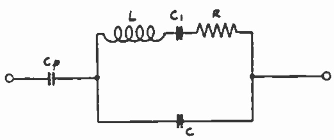


Fig. 5. Equivalent Circuit of Crystal

R will be referred to later.

When the added metal film is extremely small relative to the volume of quartz, only the change in equivalent inductance value has any appreciable effect upon the crystal resonant frequency.

It is permissible, therefore, to substitute for the mass of gold deposited a similar mass of quartz. Consequently the amount of gold may be specified as an equivalent thickness of quartz Tq as follows:

$$T_q = \frac{\gamma K E^{3.5} t}{\rho} \quad \ldots\ldots (11)$$

wherein the constants for the equipment as shown in Fig. 4 become: $\gamma = 1{\cdot}35 \times 10^{-6}$, $K = 3{\cdot}83 \times 10^{-12}$, ρ = density of quartz = 2·65.

The frequency change Δf, brought about by the addition of the metallic film or films, will therefore be

$$\Delta f = f_0 - f_1 = \left(\frac{1}{2T_0}\sqrt{\frac{\epsilon}{\rho}}\right) - \left(\frac{1}{2T_0 + 2T_q}\sqrt{\frac{\epsilon}{\rho}}\right)$$

$$= \frac{T_q\sqrt{\dfrac{\epsilon}{\rho}}}{2T_0^2 + 2T_0T_q} \text{ cycles} \quad \ldots\ldots (12)$$

(Since in practice the term $2T_0T_q$ is negligible compared with $2T_0^2$ this may be omitted.)

Substituting Eq.9 in Eq.10 we have:

$$\Delta f = \frac{\gamma K E^{3.5} t\sqrt{\dfrac{\epsilon}{\rho}}}{2T_0^2\rho} \text{ cycles} \quad \ldots\ldots (13)$$

It will be realised that the electrode system in use for the initial frequency measurement, assuming that the film is not already present, will substantially modify the readings obtained.

The presence of an air-gap will raise the natural frequency of resonance and will consequently appear to increase the effects brought about by sputtering. Conversely, the imperfect cleaning of a plate lowers its frequency, appearing in this case to reduce the modification in frequency.

It is also necessary to load the whole of the active region of a quartz plate by the added electrodes otherwise results are likely to be widely divergent from those predicted by the use of the formulæ.

4.2. *Effect upon Crystal Activity*

There are many factors contributing towards an optimum crystal activity.

The most appreciable contribution is that due to the effects of friction. There is friction between molecules of quartz, usually termed " viscosity losses." Additionally, there is the direct effect of air friction upon the surface layers of the quartz plate. The latter has roughly ten times the damping value of the former.

It is usual, in the equivalent circuit diagram (Fig. 5) to represent all losses, electric and mechanical, by the quantitative value of a pure resistance R.

Air friction is proportional to air pressure and also to the velocity of the surface in contact with the air. Therefore the component of resistance due to air friction will increase as the current through a particular crystal is increased.

The velocity of a surface particle of quartz at resonance may be given as

$$V = B\omega \cos \omega t \text{ cms/sec.} \quad \ldots\ldots (14)$$

in which B is dependent upon the location of a particle, and the crystal current I_s (see section 2.3).

The particle acceleration α will therefore be:

$$\alpha = -B\omega^2 \sin \omega t \text{ cms/sec/sec.} \quad \ldots\ldots (15)$$

and the maximum positive and negative values for α are obtained when the equation

$$-B\omega^3 \cos \omega t = 0 \quad \ldots\ldots (16)$$

is satisfied, which obviously is when $\cos \omega t = 0$, since neither ω nor B can be zero for a vibrating crystal plate.

The maximum acceleration of a surface particle may be given in terms of gravity as follows:

$$\alpha \max = \frac{-B\omega^2}{981} \quad \ldots\ldots (17)$$

The quantitative values attained by this surface acceleration are truly amazing and may in the distant future of electronics be used in some way to impart energy to an electron stream with possibly far-reaching effects.

It is easy to realise that with so large an acceleration, any weakly adherent gold molecules, abrasive or quartz particles will cause considerable loss of energy due to friction.

Some of these poorly bonded particles will inevitably be ejected from the crystal, thus causing a small but appreciable change both in resonant frequency and activity.

We are thus concerned with the provision of a firmly bonded gold deposit. In this respect Fig. 6 illustrates the effects upon the gold film of a variation in the height of the work platform (4) in Fig. 4. Using a stylus and a pressure applied just sufficient to remove the gold film, the graph shown was obtained.

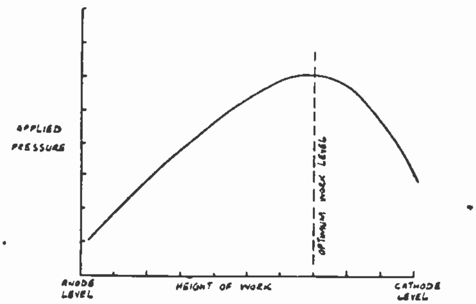


Fig. 6. Hardness of Gold: Optimum Height of Work

This clearly shows an optimum position for the work to achieve the maximum degree of bonding.

Particles of abrasive or detached quartz are commonly removed by brushing or wiping the plate and then etching in a bath of hydrofluoric acid to dissolve other matter.

To summarise, the losses which we have considered as included in R in the crystal equivalent circuit diagram, Fig. 5, may be grouped as follows in order of their importance relative to crystal activity:

(*a*) Quartz-air friction.

(*b*) Surface particle friction.

(*c*) Viscosity of quartz.

(*d*) Mounting restrictive action.

This latter must necessarily vary with the disposition of the clamping points relative to the nodal or quiescent zones of the quartz plate. If clamped on or near an antinode (*d*) will exceed (*a*) in its value. This subject is dealt with in section 5.

4.3. *Fine Adjustment of Frequency by means of Gold Film*

In section 4.1 the author discussed the effect of a gold deposition upon an oscillating quartz crystal.

It was shown that sputtering equipment could be so controlled as to deposit a known weight of gold upon quartz.

Then it was shown that for a known quantity of the metal, the frequency of resonance of the crystal plate would deviate by an amount.

$$\Delta f = \left(\frac{1}{2T_0}\sqrt{\frac{\epsilon}{\rho}}\right) - \left(\frac{1}{2T_0 + 2T_a}\sqrt{\frac{\epsilon}{\rho}}\right)$$

$$\approx \frac{\gamma K E^{3.5} t \sqrt{\frac{\epsilon}{\rho}}}{2 D T_0^2} \qquad (18)$$

We are then given a positive means of obtaining any reasonable value of Δf, so that we shall be able to modify the precise frequency of an otherwise finished crystal.

Assume that the crystal is already cut, lapped and gold plated and that we are to modify its frequency to the extent of n cycles. All the quantities in the equation for Δf are constants with exception of t.

In practice Δf is directly proportional to t after the first few seconds of the sputtering operation. So we may control the frequency of any quartz plate oscillator merely by varying t.

A chart similar to that used in the production of crystals for the final adjustment of frequency is shown in Fig. 7.

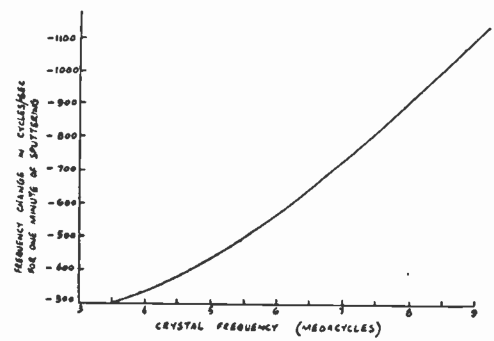


Fig. 7. Fine Frequency Adjustment Chart

In this chart, the frequency change in cycles due to sputtering for one minute is plotted as a function of the nominal frequency of the crystal. From this it will be seen that a crystal nominally at 8·5 megacycles will have its frequency of oscillation lowered by 30 kilocycles if sputtered for 30 minutes.

A variation in frequency in the positive direction is equally practicable although essentially limited in scope by the ultimate thickness of the gold film.

A weak cyanide solution or other gold solvent is used to reduce the weight of gold film, and consequently its thickness. The rate of change of frequency

$\frac{df}{dt}$ may be controlled closely by varying the strength of the cyanide solution; a convenient concentration is found to be 0·3 per cent. It is important to maintain the solution at constant temperature since this affects $\frac{df}{dt}$ to an appreciable extent.

As an alternative to the fine adjustment of frequency by sputtering or dissolution of gold, there is the process of electrolysis.

A "pilot" film of gold or other metal must be deposited primarily. The two or more electrodes of the quartz plate are connected together electrically and become the cathode of an electrolytic bath containing gold cyanide solution.

The anode may be of platinum or pure gold. A small potential is applied to the cathode and electrode via a variable resistor and a milliameter.

The electro-chemical equivalent for gold cyanide is 0·00203, so that for a given current I, through the electrolyte for t seconds, there will be deposited a definite weight of gold:

$$M = \frac{0{\cdot}00203\ It}{A} \text{ grams/cm}^2 \quad \ldots\ldots\ldots\ldots\ldots(19)$$

In which A is the total area of the cathode.

As in the case of sputtered gold, we may equate this quantity of gold to a similar mass of quartz. To do this and to obtain value of Tq corresponding to that in Eq. 11 we write:

$$Tq = \frac{0{\cdot}00203\ It}{A\rho} \text{ cms.} \quad \ldots\ldots\ldots\ldots\ldots(20)$$

and subsequently Eq. 12 becomes by substitution and simplification

$$\Delta f = f_0 - f_1 = \frac{0{\cdot}00203\ I^t \sqrt{\frac{\epsilon}{\rho}}}{2A\rho T_0^2} \text{ cycles} \quad \ldots\ldots\ldots(21)$$

On examining this equation, it is seen that $\frac{df}{dt}$ may be controlled by the current I, or the area A to be electro-plated.

A reversal of the process, making the plated crystal the anode of the electrolytic bath circuit, will result in removal of gold into the electrolyte.

This deplating action is much less consistent, being dependent almost entirely upon the state of the electrolyte. That is to say, if the cyanide becomes saturated with gold, as would be the case if a gold anode were immersed continuously, then gold deposited upon the cathode would tend to be drawn from the electrolyte. Therefore, the weight of gold removed from the anode will be considerably less than the amount deposited upon the cathode for a given current.

It is advisable then, for the sake of consistency in action, to use separate baths for plating and deplating, also to remove the gold anode immediately after use.

In this way we are provided with alternative methods for the final critical adjustments of frequency, either positive or negative in sense. The final frequency of oscillation is stable and permanent.

5. Mounts

The subject of mounting design for quartz crystal work is inseparably connected with the physical displacement of the quartz in its dynamic state.

In a thickness-mode high-frequency plate, given complete freedom of vibration, the molecular displacement p is governed by

$$p = P(\text{Cos } an\ \pi) \text{ Sin } \omega t \quad \ldots\ldots\ldots\ldots\ldots(22)$$

in this, n is the order of harmonic, always an odd integer for a two electrode high frequency plate; a represents the ratio $\frac{T_a}{T_0}$, wherein T_a is the distance in cms. from a major surface to the point under consideration; T_0 is the thickness of the quartz plate.

P is a somewhat vague term, being the maximum displacement of the surface at any one point. This factor is controlled by what could, for lack of a better term, be called "local impedance." If a quartz plate has electrodes over the whole of its major surfaces then the voltage wave applied to the electrodes will set up a uniform field in the quartz. However, as indicated by the probe tests and illustrated in Fig. 3, the displacement of the quartz is far from uniform over the surface.

So we say that the electro-mechanical "local impedance" varies from a very low to a relatively high value toward the periphery.

Mounting a crystal in order to support it only, really amounts to coupling it at some point or points to another more or less rigid and non-resonant system.

A minimum transfer of energy from the crystal to the mounting will obtain when the mount is in contact with the crystal plate at a point or points where the "local impedance" is high. For high frequency shear mode crystals then, the most efficient mounting will be in close proximity to the periphery.

Since lines joining points on the surface having equal velocities are concentric, it follows that the corners of rectangular plates will be inactive relative to the rest of the surface area.

The presence of what is termed "twinning" in a plate produces a shifting of the curve in Fig. 3, such that the crevasse no longer occurs at the exact geometric centre.

Twinning [7] is actually a physical deformation of the atomic structure of the natural quartz.

The section of "twinned" quartz will propagate a wave with a different velocity to "untwinned" quartz. So that if the crystal is oscillating at or near resonant frequency, its "twinned" component will contribute nothing to the resonant piezo-electric effect. It does, in fact, increase the value of the equivalent circuit inductance L (Fig. 5).

5.1. *Spring Contacts*

The provision of gold electrodes allow the use of a small clamping area, since the electrical continuity is catered for by the film itself.

To prevent mechanical shock fracturing the plate and to accommodate plates of varying thickness, it is desirable to use somewhat resilient clamps or springs.

Again, since even at the periphery there is some small degree of displacement, it is necessary to avoid anything which might tend to damage the gold film at the point of contact.

Unfortunately, owing to the small thickness of high frequency plates, it is not possible to couple to the mounting at the true nodal region, since this runs midway through the plate and parallel to its major faces. The true node would have an infinite " local impedance " theoretically, since infinite voltage would produce zero displacement at that point.

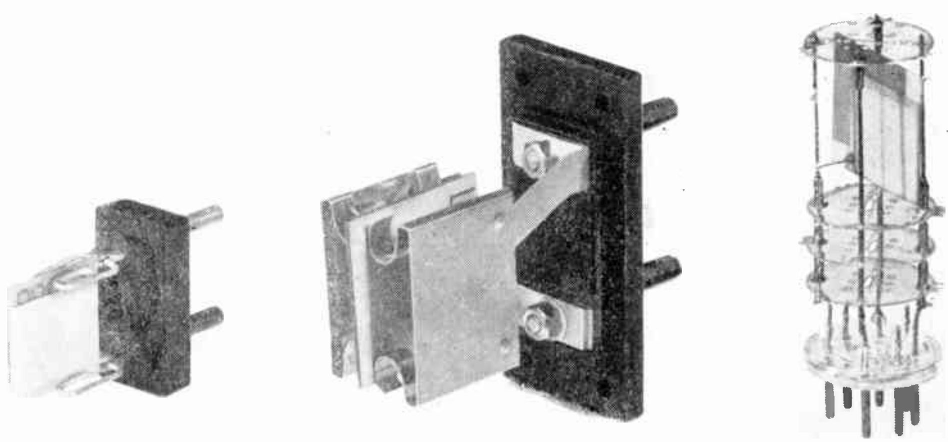
5.2. *Soldered-wire Supports*

Comparatively recent developments in high frequency plated crystal technique have been in connection with fine wire suspension.

A process is adopted which results in a reasonably strong soldered wire to crystal joint.

The same fundamental principles rule, i.e., the point or points of junction for the soldered suspension wires must be at or near the zone of maximum " local impedance."

It is possible to make the actual area of contact quite small and yet, by plating over the soldered junction to obtain high conductivity between the wire and the active regions of the quartz. The contact resistance is quite important, since it is effectively in series with the whole crystal equivalent circuit. Consequently, if it exceeds a few ohms, it will introduce a serious loss.



(*a*) (*b*) (*c*)

Fig. 8. Typical Mountings for Metallic Film Electrodes

A high contact resistance indicates a possible future source of failure, in so far as it denotes a thin or weakly adherent film connecting the wire to the main electrode system.

Fig. 8 illustrates 3 crystal mountings of the type used in radio communications equipment. Fig. 8 (*a*) and (*b*) are representative of the spring contact classes. In assembly (*b*) the quartz plate is held under considerable pressure at the four corners, since the cover of the unit is designed to fit tightly over the spring assembly.

Fig. 8 (*c*) shows the support system of a soldered contact type. The structure is sealed into a miniature valve envelope, making a satisfactory unit for use under exacting tropical conditions.

6. Conclusion

Finally, the author would like to emphasise that though the paper sets out to deal specifically with gold as the deposited metal film, there seems no reason to expect any different behaviour in the main features if materials other than gold are used, provided due substitutions are made for the physical characteristics of the alternative metal or metals. Neither should there be material difference between electro-plating, sputtering and evaporation of a given metal in its final effects upon a quartz plate.

Similarly, with certain reservations, the same thesis should be applicable to other types of crystal.

The exceptions will be in cases where the propagation of vibrations through the quartz are in a direction other than sensibly perpendicular to the deposited metal film. This will be so in X cut bars wherein the propagation is along the Y axis, but the electrode system is normal to the X axis.

The author expresses his thanks to the Automatic Telephone & Electric Company for granting permission to read and publish the paper.

Bibliography

1. Vigoureux. ("Quartz Oscillators"), p. 11. (*H.M.S.O.*, 1939.)
2. National Physical Laboratory. (" Notation of Piezo Electric Quartz.") (*H.M.S.O.*, 1938.)
3. Dye (Proc., Royal Society), 1932.
4. Meissner (Physick. Z.), 1927.
5. Strong. (" Modern Physical Laboratory Practice"), p. 168. (*Blackie* & Son, 1942).
6. Terman. (" Radio Engineering"), p. 114. (*McGraw-Hill*, 1942.)
7. Willard. *B.S.T.J.*, Jan., 1944, p. 11.

LONDON DISCUSSION

Mr. Fairweather : I am interested in the probe method of exploring the various regions of the oscillating plate and thus obtaining a measure of relative amplitude of the vibrations at the centre and periphery. In the case of a slightly convex plate, has the author noted that the frequency change with movement of the probe across the diagonal may be discontinuous ?

No mention has been made of the stability of sputtered electrodes. The sputtered metal appears to be very active chemically and takes up large volumes of gas on exposure to air. The absorbed gas appears to be removed from the electrodes when vacuum pumps capable of 10^{-6} m.m. Hg are used to exhaust the envelope in which the crystal is mounted. The change in the gold on pumping may explain frequency changes of about $-$ 2 kc/s on plated crystals of 8 Mc/s. There seems to be a change of only 50-100 c.p.s. due to air friction losses. This can be checked by opening after evacuation. The permanent or semi-permanent change is 10-20 times this amount.

Incidentally, high frequency crystals of the plated type are very suitable for studying the effects of sputtered electrodes on frequency stability. An investigation of this kind would probably help in explaining long term frequency drifts of small order on low frequency plates used in standards of time and frequency.

I can confirm the figure of 500,000 for the Q of B.T. plates. This usually indicates a series resistance of about 5 ohms. The author's opinion of how much of this figure arises from the plating and soldered joint would be of interest.

This last point reminds me that the use of the word " film " for the sputtered gold rather implies a layer of metal having the resistance of the pure metal. I prefer to think of the metal as a collection of small particles, many in contact.

Perhaps I was not able to follow the author's remarks, but he did not make it clear to me that, for a given amount of gold on a crystal the effect on plate frequency (i.e. difference in frequency between plated and unplated crystal) is proportional to the square of the plate frequency. The same, of course, also applies when gold is added to the sputtered layer by electroplating methods.

In connection with the use of cyanide to remove gold from the crystal, I think the author should mention that silver spots or solder must not be present on the plate or local action will take place during immersion.

It was mentioned that twinning was easily produced but impossible to remove. There has been a report of a change in the twinning boundary being effected by U.S. workers, but I do not know the method used.

Mr. West : Has anything been done in the direction of sintering the sputtered layer in order to reduce gases? Supposing that after the sputtering the whole thing were heated, so that the layer became more or less molten ; it would then solidify and possibly bond better with the quartz and reduce its effective area, for it may or may not have absorbed gases.

Mr. Webster : Can the author say how crystals are prepared for the electroplating process to ensure that gold is deposited ?

Mr. Hamburger : Will the author give a few details of a typical "probe." What form does it take ?

Mr. West: Will the lycodium powder test show the presence of twinning? What would be the frequency difference of oscillations when a crystal is behaving as a series and as a parallel circuit?

Mr. Mitchell: How is the thickness of gold electrode estimated?

AUTHOR'S REPLY TO LONDON DISCUSSION

Replying to Mr. Fairweather, the prime function of the probe was to locate zones of relative quiescence, therefore frequency discontinuities were ignored. In fact, the frequency variations are more pronounced as the nominal frequency of the crystal is lowered and couplings to other modes are increased.

The stability of the sputtered crystal has been the subject of investigation, both in this country and the States, the results of which have not yet been published.

The change in frequency, due to evacuation, has been observed, but I hesitate to agree as to the cause, since a second return of air to the crystal electrodes should have a similar effect to the initial exposure to air, and such is not the case.

A high frequency resistance, due to inferior gold sputtering or preparation, of up to 10 ohms is possible, although the figure usually associated with a good specimen will be less than 2 ohms. In this connection soldered joints are slightly inferior to spring contacts.

I agree entirely with Mr. Fairweather's criticism of the word "film" in the paper. It was merely intended to differentiate between metal plate electrodes, and the extremely thin deposit of gold particles.

In equation 13 for the effect upon frequency of the gold deposit, all the terms will be constants, for a given amount of gold, with the exception of T_0. The nominal frequency of a given type of crystal is inversely proportional to thickness (T_0) so that the amount Δf will be proportional to frequency.

Cyanide dissolution of excess gold is not used with soldered type units, the more rapid reversed electroplating scheme being much safer for this purpose.

Mr. West has reminded me of an experiment made along the lines he suggests. Unfortunately, the result was disastrous, the gold layer became transparent at far below the melting point of gold, which incidentally would change α quartz into β quartz, thus making it inactive. The resulting film was very high resistance, and useless as an electrode. I am unable to account for this behaviour.

Mr. Fairweather: I have heard that it occurs when lithium is present in quartz.

Mr. Spears: Replying to Mr. Webster, I should like to emphasise the fact that electroplating of crystals is only used in the latter stages of production. An initial deposit by means of sputtering is essential to provide the conducting path for use in the electrolyte. This first deposition is normally about 90 per cent. of the total final quantity of metal.

Mr. Hamburger would like a few details of a typical exploring "probe." The essentials are merely that a definite pressure should be applied to the crystal at a very local point. The probe may consist of a pair of rounded platinum contacts under a compression of about 250 grams. Connection to the oscillator is made via these contacts. To prevent damage to the electrode it may be advisable to remove the pressure each time a new location is required for the probe.

Lycopodium powder concentrates at any region of a quartz plate where vibrations are a minimum. Therefore, if twinning were sufficiently widespread to create a zone of relative quiescence the answer to Mr. West's first point would be "yes." Normally, however, twinning accidentally included in a crystal is scattered, and I do not think the practical application of lycopodium to the detection of twinning would be sufficiently positive in its indication.

Mr. West, in his second question, has asked what is the difference between the frequencies of oscillation of a crystal in its series and parallel circuits. In an oscillator of the Pierce type the crystal must present an inductive reactance to the circuit, whilst in certain bridge arrangements the crystal is operated very close to zero reactance. The separation of the positive reactance peak from the zero reactance point varies with types of crystal, and with electrode capacity.

Commonly it is only of the order of a few hundred cycles, so that in these circuits the oscillatory frequency must fall within this small frequency range.

Answering Mr. Mitchell regarding the estimation of electrode thickness, possibly the simplest method is to procure a large number of crystals, weighing before and after plating and obtaining average thickness by means of the known density of gold. In the case of extremely thin electrodes the transparency or colour of the gold may be used as a guide to thickness. Film resistance may also be used for comparison purposes, but is less reliable.

Mr. Beattie has, in the course of his remarks, mentioned a drift of 1 cycle in 10 seconds. Whilst this obviously is dependent upon the frequency in question, if one assumes a frequency of a megacycle this figure is relatively easy to attain, and more important, to maintain over long periods. More recent improvements in crystal controlled time systems show drifts of less than 1 part in 10^8.

I will agree with Mr. Beattie's suggestion for air-conditioning of crystal finishing and assembly departments; some degree of conditioning or filtering of the air is quite common in such workshops.

In conclusion may I thank the Chairman and those present for the obvious interest accorded to the paper.

MIDLANDS SECTION DISCUSSION

Mr. G. F. Knewstub : Regarding the apparatus for sputtering, Fig. 4 gave a very clear diagram, but it did not convey any idea as to the physical size of the apparatus. What is the spacing between the cathode and anode, and approximately how long does the process require to obtain a suitable film ?

Mr. S. G. Button : No reference has been made to circular crystals. Do the various points raised apply to these? As to the change in frequency due to the deposition of the gold film, do formulae given apply irrespective of the thickness of the film ?

Mr. W. W. Smith : It would be interesting to know if small errors in grinding could be corrected by controlling the deposition of gold film.

Mr. F. H. Alston : It has been stated that thorough cleansing is carried out, using a mixture of sulphuric and chromic acids. In the case of soldered types of crystal, would not this damage the joint and wires ?

The Chairman : It is not usually regarded as part of Institute procedure that the vote of thanks should come from the chair, but on the present occasion I feel that Mr. Spears deserves that distinction. It seems to me that some of the ancillary problems to the business of applying electronic equipment are not always understood by engineers. Mr. Spears has enlightened us on a great deal of the research and development work that has been carried out in matters which do not normally come within the purview of radio engineers.

REPLY TO THE MIDLAND SECTION

Mr. Spears : In reply to Mr. Knewstub, a typical equipment set-up would have a cathode of about 8 in. diameter, and a thickness of probably ¼ in. It should be mentioned that the upper (ungilded) surface of the cathode is masked, since in Eq. 9 the area A is the active gold cathode area. The distances from work platform to cathode and anode will be between 2 and 2½ in., and the aluminium anode should be at least the diameter of the cathode.

Contamination of any part of the assembly must be avoided.

A normal working vacuum would be 0·01 mm. of mercury, and for practical purposes the process is completed in 40 to 80 minutes, depending upon the treatment to be expected in service. Mr. Button has raised an interesting point regarding the shape of crystal; the effects quantitively and qualitatively, of a film upon circular plates are not noticeably different from the case for the square or rectangular crystal. His second point, relating to frequency change resultant upon the deposition of an excessive quantity of gold, may be covered by drawing attention to the statement that " the equations only hold good for the case when the gold film thickness is very small compared with that of the quartz."

In actual fact the frequency could be lowered by as much as 200 kc/s before an appreciable fall in activity became apparent.

The possibility of compensation for small errors in grinding or lapping to a specified frequency has been mentioned by Mr. Smith. Provided the errors to which he refers are in thickness and not in surface contour, then the process will correct them.

Mr. Alston has called my attention to a part of the process which may not have been fully explained. When soldered types of crystal are acid-cleaned prior to sputtering, it is essential to limit the time of immersion in the acid bath to a matter of around 20 seconds. This means that the bulk of the cleansing operation has, of necessity, to be done prior to the silver-spotting stage.

NORTH WESTERN SECTION DISCUSSION

The Chairman : Reference has only been made to the square crystal. Are there any modifications so far as circular crystals are concerned? The paper deals in the main with the low temperature coefficient types AT and BT. Will the paper also apply to other types, e.g., X cut plates ?

Mr. Miller : Can the author say whether films other than gold have been used, and, if so, what factors have decided the present use of gold ?

It has been stated that the silver colloid spot is fired on to the quartz and subsequently soldered and sputtered with gold. Is there any objection to silvering the whole electrode area and similarly firing, so obviating the gold deposit ?

I would like to hear more about the higher frequency limit of operation for quartz crystals. Using harmonic types what would be the estimated maximum workable frequency ?

In connection with water vapour and gas impurities in the sputtering process, it seems to me a possible improvement might be obtained by using the evaporation process, since the level of vacuum required will in this case be higher, and probably the undesirable gas contents would be much less troublesome.

Mention has been made of surface preparation by acids followed by rinsing in distilled water. How are the traces of water vapour removed after the operation ?

REPLY TO NORTH WESTERN DISCUSSION

Mr. Spears : The Chairman's comments relating to crystals other than AT and BT, and having boundaries other than square, merit a reply in full. The equations for change of frequency, due to the application of a specified amount of film are in each case dependent upon the factor $\sqrt{\frac{\epsilon}{\rho}}$. In this instance ϵ is the elastic modulus in the direction of propagation of the wave through quartz, which for the purposes of the paper were considered normal to the plane of the crystal. The numerical value of ϵ varies considerably with angle of cut, ranging from $1{,}280 \times 10^9$ at $-45°$ down to 648×10^9 at $+71°$. When the appropriate value ϵ is inserted in Eqs. 12, 13, 18 and 21, the results should be as predicted.

Generally, a circular crystal will behave similarly to a square, in so far as its response to plating is concerned; this also applies to plates of irregular shape.

However, should the shape or dimensions of the plate be so designed that high order face-shear or other desirable modes are excited by the driving frequency, there will be considerable modification of the theoretical behaviour.

Suppose a quartz plate is lapped to oscillate in a given circuit at 1,500 kc/s per second, unplated. This frequency may not be harmonically related to a face-shear or other frequency. When we lower the frequency by plating, a point may be reached at which strong coupling takes place, with a resultant rapid or abrupt change in frequency. This is obviously a function which cannot be catered for except by correct design of the quartz wafer. With this possible exception it will be possible to predict the action of any type of thickness-wave plate.

Mr. Miller asks why use gold in preference to other metals. This is to a great extent a question of choice. Silver has been widely used, and is very effective and aluminium deposition by evaporation is also excellent. This latter has the advantage of a comparatively small frequency modification, but requires considerably more efficient pumping apparatus. An obvious point in favour of gold is its resistance to oxidation, which, if allowed to form, may affect the conditions of operation.

The firing of a layer of silver colloid to provide an electrode has been used with success, the main difficulty being that of obtaining low damping and constant thickness.

The high frequency limit for the overtone (or harmonic) operation of a quartz plate is largely determined by its associated circuit. If the electrode capacity of a high activity crystal is reduced to zero by balance in a bridge circuit it has been found possible to produce crystal oscillators operating as high as about 200 Mc/s. Without this method the probable limit would be approximately 30 to 40 Mc/s.

Mr. Miller's suggestion regarding the reduction of gas and vapour troubles by using the aluminium evaporation method is, no doubt, quite sound, but I have insufficient comparative data upon the method to say whether any improvement would be worth while.

The use of distilled water in the preparation for sputtering was found to introduce water vapour. This was prevented by baking crystals in an oven at 104°C. until ready for the sputtering process. Post-sputtering contamination is equally important, since this affects the stability. The ultimate results can only be obtained by meticulous care and chemical cleanliness at all stages.

NOTICES

Honours

Council tender congratulations to **F/Lt. George Derek Coates** (Associate) and **Stanley Rangi Sydney Mead** (Registered Student) on their having been Mentioned in Despatches.

Obituaries

The Council regret to record the death, after a very brief illness, of **Dr. Patrick Dalton** (Member), of London, at the age of 47 years.

Dr. Dalton was elected a Member of the Institution in October, 1939; he was a founder member of the International Committee for short-wave work and was a consultant in the Eelectrical Department of St. Bartholomew's Hospital.

Dr. Dalton had, since 1933, served on a number of medical and other Committees dealing with radio-therapy and other radio research for medical purposes. He was Chairman in 1940 of the Institution's Radio-therapy Committee, and later that year was elected to the General Council of the Institution.

Council also learn with regret of the recent death of **Herbert John Selves** (Associate), of Eltham, S.E.9, at the age of 22 years.

Mr. Selves was registered as a Student Member of the Institution 1943 and passed the Graduateship Examination in the same year. He was then transferred to the grade of Graduate and in January, 1945, transferred to Associateship.

Twenty-first Anniversary Celebration

Members will have already received the notice dated April 20th regarding the arrangements for celebrating the 21st Anniversary of the Institution on October 31st, 1946.

There has been an excellent response and all grades of members have co-operated in sending early advice of their intention or otherwise to attend the dinner. Those members who have not yet returned their cards should do so immediately in order to facilitate a fair and early allocation of tickets.

For members abroad and members living outside London, hotel accommodation can be reserved if *immediate* application is made to the Institution.

1947 Convention of the Brit.I.R.E.

Arrangements are now being completed for appropriate hotel reservation in Bournemnouth. As far as possible, members will be accommodated in one hotel or an adjacent hotel. Some 200 members have already indicated their intention of attending the Convention and members who have not sent in their cards should do so as soon as possible in order to facilitate the booking of hotel accommodation. Indication should also be given of the period for which hotel accommodation is required. If, as at present intended, the Convention takes place during the week before Whitsun, 1947, one of the hotels is prepared to continue reservation of accommodation over the Whitsun weekend.

1945 Examination Prize Winners

Council has approved the following awards to the most outstanding candidates appearing in the 1945 Graduateship Examinations :—

President's Prize

RIDGERS, Charles Edward S., London, N.W.9.

Mountbatten Medal

HARES, F/O Walter. Coalville, Leics.

S.R. Walker Prize

MURPHY, Kevin Anthony. Dublin.

Measurements Prize

RIDGERS, Charles Edward. London, N.W.9.

Specially Commended

FEHER, George. Haifa, Palestine.

Generous Gift of I.R.E. (Australia)

Typical of the co-operation which exists between the British and Australian Institutions is the cable and subsequent letters received from the Council of the Australian Institution indicting the desire of some 100 members of the Australian Institution to send parcels of provisions to members of the British Institution.

In acknowledging this generous and personal token of good will, the General Council of the Brit.I.R.E. have advised their Australian confreres that the allocation of parcels is being arranged by the Trustees of the Benevolent Fund.

Merger of the Examinations previously conducted by the Radio Trades Examination Board and the City and Guilds of London Institute in Radio Service Work

Arrangements have now been completed for the holding of the joint examination in Radio Service Work and the first examination under the combined bodies will be held on May 17th, 1947. This arrangement now provides a national examination in the subject and in addition to the City and Guilds of London Institute, the contributing organisations are the Radio Industry Council, the English and Scottish Radio Retailers' Associations and the Brit.I.R.E.

All applications to sit the examination must be addressed to the Secretary, the Radio Trades Examination Board, 9 Bedford Square, London, W.C.1.

"A SYMPOSIUM OF MATHEMATICAL METHODS FOR RADIO ENGINEERS"

(A Discussion held by the London Section of the Institution on October 17th, 1945)

FOURIER SERIES, by M. M. Levy (*Member*)*

SUMMARY

A study of the properties of Fourier series with special reference to functions presenting discontinuities and to radio and pulse problems ; summary survey of the methods of harmonic analysis ; scope of these methods ; extension to the calculation of a great number of harmonics and to the analysis of non-periodic curves.

The Fourier series expansions are particularly applicable to functions having discontinuities, and the coefficients of these expansions can be expressed in terms of the discontinuities of the function and its derivatives. The discontinuity formulae thus obtained are very useful in that a great number of expansions can be written straight away. Such is the case of periodic pulses.

The Fourier series can also be applied, by means of a limiting process, to the analysis of non-periodic functions such as a single pulse. In some cases the mathematical difficulties encountered when applying Fourier integrals are so avoided. As examples of application, the frequency characteristics of the unit step and of the impulse functions are calculated, and the response of an ideal low-pass filter to a rectangular pulse and to a double pulse is studied.

A general survey of the methods of harmonic analysis is given. Methods of calculation are approximate since they start from a limited number of points of the curve to analyse. A considerable variety of methods have been suggested, the main tendency being to get the greatest accuracy with a limited number of points. The writer has studied closely all these methods and has found that they are closely correlated and can be reduced to a single one, the method using Lagrange's equations with Runge's schedules, and the writer's corrective coefficients.

Further studies by the writer, particularly on the quick calculation of a great number of harmonics such as 120 or 160, are mentioned.

Contents

(1) FOURIER SERIES

(1.1) Fourier Expansions

The Fourier series expansions enable one to express any periodic function in terms of sinusoidal functions. Since sine waves are easily accessible to analysis in radio, the Fourier series is a powerful method for extending this analysis to periodic curves of any shape, particularly to curves of the pulse type.

Fourier series can be written under two forms : trigonometric and exponential.

(1.1.1) *Trigonometric Form of the Expansion*†

(1.1.1.1) *General Formulae*

If $f(t)$ is a periodic function of radian frequency ω_0, the corresponding Fourier expansion is

$$f(t) = \frac{C_0}{2} + \sum_{n=1}^{\infty} C_n \cos (n\omega_0 t - \phi_n) \quad \ldots\ldots(1)$$

where $\frac{C_0}{2}$ is the average value of the function, C_n the amplitude and ϕ_n the phase angle of the nth harmonic component.

† No demonstration of these formulae is given here. For further details see any treatise. Usually only approximate demonstrations are given. A relatively simple demonstration, and in the same time useful in applications, is to start by studying the properties of the integral $\int_a^b f(x) \sin nx\, dx$ and those of Dirichlet's integral $\int_a^b \frac{f(x)}{\text{Sin}\, x} \sin nx\, dx$. It is then easy to show that there is only one expansion in harmonic components which represents a given periodic function ([4]). If the function has a discontinuity at point x, the expansion's value is $\frac{1}{2}[f(x - o) + f(x + o)]$.

* Research Laboratories of The General Electric Co., Ltd. *Formerly with* Standard Telephones and Cables.

The expansion can also be written

$$f(t) = \frac{C_o}{2} + \sum_{n=1}^{\infty} A_n \sin n\omega_o t + B_n \cos n\omega_o t \quad (2)$$

where

$$A_n = C_n \sin \phi_n = \frac{\omega_o}{\pi}\int_{-\frac{\pi}{\omega_o}}^{+\frac{\pi}{\omega_o}} f(t) \sin n\omega_o t \, dt \quad \ldots.(3)$$

$$B_n = C_n \cos \phi_n = \frac{\omega_o}{\pi}\int_{-\frac{\pi}{\omega_o}}^{+\frac{\pi}{\omega_o}} f(t) \cos n\omega_o t \, dt \quad \ldots.(4)$$

These two formulæ can be written as one

$$B_n \pm iA_n = \frac{\omega_o}{\pi}\int_{-\frac{\pi}{\omega_o}}^{+\frac{\pi}{\omega_o}} f(t)\, e^{\pm in\omega_o t}\, dt \quad \ldots\ldots(5)$$

We have also the relations

$$\tan \phi_n = \frac{A_n}{B_n} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(6)$$

$$C_n = \sqrt{A_n^2 + B_n^2} \quad \ldots\ldots\ldots\ldots\ldots(7)$$

(1.1.1.2) *Simplified Expressions for Symmetrical Functions*

(1) If $f(t)$ is an *even function* (Fig. 1a) so that $f(t) = f(-t)$ for all values of t, we have $A_n = O$ and expansion (2) reduces to

$$f(t) = \frac{C_o}{2} + \sum_{n=1}^{\infty} B_n \cos n\omega_o t \ldots\ldots\ldots\ldots\ldots(8)$$

and inversely.

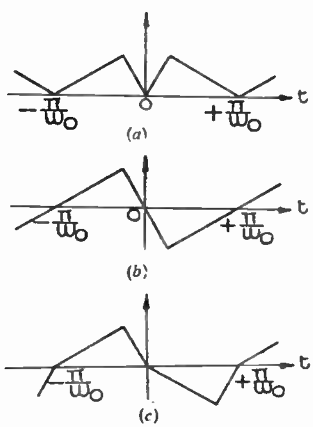


Fig. 1.—Different classes of symmetry for periodic curves —a: $f(-t)=f(t)$; b: $f(-t)=-f(t)$; c: $f(t+\pi)=-f(t)$.

(2) If $f(t)$ is an *odd function* (Fig. 1b), so that $f(t) = -f(-t)$ expansion (2) reduces to

$$f(t) = \frac{C_o}{2} + \sum_{n=1}^{\infty} A_n \sin n\omega_o t \quad \ldots\ldots\ldots\ldots(9)$$

and inversely.

(3) If $f(t)$ is such that $f(t) = -f(t+\pi)$, (Fig. 1c), the Fourier expansion (1) contains only harmonics of odd order (i.e. $Co = A_{2p} = B_{2p} = 0$) and inversely.

(1.1.2) *Exponential Form of the Expansion*

The exponential form of the expansion is compact and very useful particularly when the function lacks any special symmetries[3,4].

It may be derived from the trigonometric form by replacing cos $(n\omega_o t - \phi_n)$ by its corresponding exponential expression

$$\cos(n\omega_o t - \phi_n) = \tfrac{1}{2}[e^{i(n\omega_o t - \phi_n)} + e^{-i(n\omega_o t - \phi_n)}]$$

Then

$$f(t) = \tfrac{1}{2} \sum_{n=-\infty}^{+\infty} C_n\, e^{-i\phi_n}\, e^{in\omega_o t} \ldots\ldots\ldots\ldots\ldots(10)$$

where $\begin{cases} \phi_{-n} = -\phi_n \ldots\ldots\ldots\ldots\ldots\ldots\ldots(11) \\ C_{-n} = C_n \end{cases}$

Putting $\mathbf{C}_n = C_n\, e^{-i\phi_n} \quad \ldots\ldots\ldots\ldots\ldots\ldots(12)$

we get

$$f(t) = \tfrac{1}{2} \sum_{-\infty}^{+\infty} \mathbf{C}_n\, e^{in\omega_o t} \ldots\ldots\ldots\ldots\ldots\ldots(13)$$

To get $\mathbf{C}_m$ multiply both sides of (13) by $e^{-im\omega_o t}$ and integrate from $-\frac{\pi}{\omega_o}$ to $+\frac{\pi}{\omega_o}$; integrals such as

$$\mathbf{C}_o \int_{-\frac{\pi}{\omega_o}}^{+\frac{\pi}{\omega_o}} e^{i(n-m\omega_o)t} dt = \frac{\mathbf{C}_m}{i(n-m)\omega_o}\left[e^{i(n-m)\omega_o t}\right]_{-\frac{\pi}{\omega_o}}^{+\frac{\pi}{\omega_o}}$$

cancel if $n \neq m$. If $n = m$, these integrals are equal to $\frac{2\pi}{\omega_o}\mathbf{C}_m$. Thus replacing m by n in the final result, we get

$$\mathbf{C}_n = \frac{\omega_o}{\pi}\int_{-\frac{\pi}{\omega_o}}^{+\frac{\pi}{\omega_o}} f(t)\, e^{in\omega_o t}\, dt \quad \ldots\ldots\ldots\ldots\ldots\ldots(14)$$

The complex coefficient $\mathbf{C}_n$ contains both the amplitude and phase of the nth harmonic component.

From (12) we deduce also the useful relation

$$\mathbf{C}_n = C_n \cos \phi_n + i\, C_n \sin \phi_n = B_n + iA_n \quad (15)$$

(1.2) Expansion in Terms of Discontinuities

The Fourier series is particularly useful in that it will apply to functions having a finite number of discontinuities within the period. Such is the case of square and saw-tooth shaped periodic pulses. Fig. 2 shows for the square shaped curve how the discontinuities appear progressively when the successive harmonic components are added.

There is a close relationship between the discontinuities of the function and its derivatives and the amplitude and phase of the successive harmonic components. We shall obtain a very general formula for the Fourier constants in terms of the discontinuities from which a great number of expansions can be written straight away.

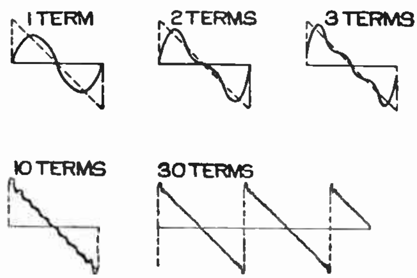


Fig. 2.—Progressive approximation of a periodic curve by addition of successive harmonic components. Fourier expansions are suitable for discontinuous functions.

Let $f(t)$ be a periodic function represented in different portions of its range by straight lines or by arcs of parabola of any finite degree (Fig. 3). Let the function have discontinuities at the points $t_1 \,.\, t_2 \,..\, t_j \,......\, t_{p-1}$.

For simplicity put $-\frac{\pi}{\omega_0} = t_0$ and $+\frac{\pi}{\omega_0} = t_p$ and assume that

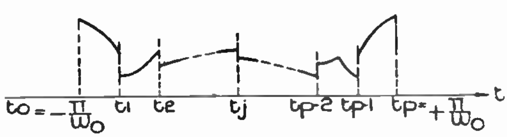


Fig. 3.—Function presenting discontinuities at the points t_0, t_1, t_2, - - - t_j, - - t_{p-2}, t_{p-1}, t_p.

discontinuities occur also at the boundaries. By definition the discontinity I_j at the point t_j is expressed by the relation

$$I_j = f(t_j + o) - f(t_j - o)$$

(1.2.1.) *Discontinuity Formulae.*

Equations (3) and (4) or (5) giving the values of coefficients A_n and B_n are still valid when the function has discontinuities provided that the integrations are done separately in each continuous interval $(t_{j+0} \,.\, t_{j+1-0})$ and the results added together. Now, if we integrate by parts equation (5)

$$B_n + iA_n = \frac{\omega_0}{\pi}\int_{t_0}^{t_p} f(t)\, e^{in\omega_0 t}\, dt \quad ....(5)$$

where for simplicity we replace temporarily $\omega_0 t$ by t' and then t' by t, we get

$$B_n + iA_n = \frac{1}{i\pi n}\sum_{j=0}^{p}\Big[f(t)e^{int}\Big]_{t_j+0}^{t_j+1-0} - \int_{t_0}^{t_p} f'(t)\, e^{int}\, dt$$

$$= \frac{-1}{i\pi n}\,\Sigma\, I_j\, e^{intj} - \int_{t_0}^{t_p} f'(t)\, e^{int}\, dt \quad ......(16)$$

Assume now that the derivatives of the function have discontinuities at the same points $t_0, t_1 \,......\, t_p$. Let

$$I_j^{(k)} = f^{(k)}(t_j+0) - f^{(k)}(t_j-0)$$

be the discontinuity of the derivative of order K at the point t_j. Integrating by parts indefinitely equation (16) and replacing the final result t by $\omega_0 t$ we get

$$\pi(B_n + iA_n) = \frac{i}{n}\sum_j I_j\, e^{in\omega_0 tj} - \frac{1}{n^2}\sum_j I_j'\, e^{in\omega_0 tj}$$
$$-\frac{i}{n^3}\sum_j I_j''\, e^{in\omega_0 tj} + \frac{1}{n^4}\sum_j I_j'''\, e^{in\omega_0 tj}$$
$$+ \;........................(17)$$

The series terminates, for if k is the highest degree of any of the parabolic arcs, $f^{(k+1)}(t)$ is everywhere zero, so all discontinuities of order superior to k are zero.

Separating real and imaginary parts we get

$$\pi A_n = \frac{1}{n}\sum_j I_j \cos n\omega_0 t_j - \frac{1}{n^2}\sum_j I_j' \sin n\omega_0 t_j$$
$$-\frac{1}{n^3}\sum_j I_j'' \cos n\omega_0 t_j + \frac{1}{n^4}\sum_j I_j''' \sin n\omega_0 t_j$$
$$+ \;...............................(18)$$

$$\pi B_n = -\frac{1}{n}\sum_j I_j \sin n\omega_0 t_j - \frac{1}{n^2}\sum_j I_j' \cos n\omega_0 t_j$$
$$+\frac{1}{n^3}\sum_j I_j'' \sin n\omega_0 t_j + \frac{1}{n^4}\sum_j I_j''' \cos n\omega_0 t_j$$
$$- \;...............................(19)$$

(1.2.2.) *Examples of Application.*

(*a*) As first example, consider the function explicitly represented in Fig. 4a. Taking the period as

$-\frac{\pi}{\omega_0} + 0 < t < +\frac{\pi}{\omega_0} + 0$, the function has two discontinuities within the period one at $t = 0$ and one at $t = \frac{\pi}{\omega_0}$. All the derivatives are zero so we can write straight away

$$A_n = \frac{1}{n\pi}\sum_j I_j \cos n\omega_0 t_j = \frac{1}{n\pi}\gamma_0 (1 - \cos n\pi)$$

$$B_n = 0$$

The corresponding Fourier expansion is

$$f(t) = \frac{2\gamma_0}{\pi}\left[\frac{\sin \omega_0 t}{1} + \frac{\sin 3\omega_0 t}{3} + \frac{\sin 5\omega_0 t}{5} + ...\right] (20)$$

The amplitudes of the harmonics decrease as $1/n$.

b — Consider now the periodic pulses of Fig. 4b. Here also we can write straight away

$$An = 0$$

$$Bn = \frac{\omega_o}{\pi} \gamma_o \epsilon \frac{\text{Sin } n\omega_o\epsilon}{n\omega_o\epsilon} \quad \ldots\ldots\ldots\ldots\ldots\ldots (21)$$

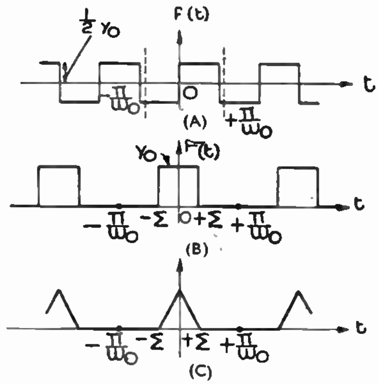


Fig. 4.—The Fourier expansion for these curves can be obtained straight away from the discontinuity formulae.

C — Consider now the triangular pulses of Fig. 4c. The function is continuous, but the first derivative is discontinuous.

We have

$$f'(t)=0 \quad \text{for } -\frac{\pi}{\omega_o}<t<-\epsilon \text{ and } +\epsilon<t<+\frac{\pi}{\omega_o}$$

$$f'(t)=\frac{\gamma_o}{\epsilon} \quad \text{for } -\epsilon<t<0$$

$$f'(t)= -\frac{\gamma_o}{\epsilon} \quad \text{for } 0<t<+\epsilon$$

Thus $\text{I}'(-\epsilon) = \frac{\gamma_o}{\epsilon}, \text{I}'(o)=-2\frac{\gamma_o}{\epsilon}, \text{ I}'(+\epsilon)=\frac{\gamma_o}{\epsilon}$

All other derivatives are zero and we have

$$A_n = 0$$

$$B_n = -\frac{1}{n^2\pi}\frac{\gamma_o}{\epsilon}\Big[\text{Cos } n\epsilon - 2 + \text{Cos } n\epsilon\Big]$$

$$= \frac{\omega_o^2}{\pi}\frac{\gamma_o\epsilon}{\pi}\left(\frac{\text{Sin } n\,\omega_o\epsilon/2}{n\omega_o\epsilon/2}\right)^2 \quad \ldots\ldots\ldots\ldots (22)$$

In this case the amplitudes of the harmonics decrease as $1/n^2$.

3. *Degree of Convergence*

(1) In the Fourier expansion for the curves of Fig. 4 the amplitude of the harmonics decrease as $1/n$ for curves a and b and as $(1/n)^2$ for curve c.

In the first case the function has discontinuities whilst in the second the function is continuous but the first derivative has discontinuities.

If there is only a finite number of discontinuities it is obvious from (18) and (19) that the coefficients of high order decrease as

$\frac{1}{n}$ if the function has discontinuities ;

$\left(\frac{1}{n}\right)^k$ if the function and the first (k-2) derivatives are continuous.

(2) The degree of convergence is decreased by integration and increased by differentiation. This can be seen at once by integrating or differentiating the Fourier expansion

$$f(t) = \frac{C_o}{2} + \sum_{n=1}^{\infty} C_n e^{in\omega_o t}$$

By integration we get

$$\int_o^t f(t)\,dt = K + \frac{C_o}{2}t + \sum_{n=1}^{\infty}\frac{1}{n}\frac{C_n}{i\omega_o}e^{in\omega_o t} \ldots (23)$$

where k is a constant ; and by differentiation we get

$$f'(t) = \sum_{n=1}^{\infty} n\,(iC_n\omega_o)\,e^{in\omega_o t} \quad \ldots\ldots\ldots\ldots\ldots (24)$$

C. Extension to Non-Periodic Functions and Application to Pulse Problems

When the period $2\pi/\omega_o$ is made infinite, the Fourier series formulæ converted by this limiting process become the Fourier transforms. With these new formulæ it is possible to obtain the frequency and phase characteristics of any non-periodic function provided that the integral

$$\int_{-\infty}^{+\infty} |f(t)|dt \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (25)$$

is convergent.

However, the frequency and phase characteristics can be obtained by a reverse process : we consider the non-periodic function as the limit of a periodic function when the period tends to infinity ; we deduce the frequency and phase characteristics of the periodic function and we increase the period to infinity.

This reverse process is usually more elaborate than the straightaway method of applying the Fourier transforms ; but in some cases, when these transforms cannot be applied because integral (25) is not convergent, the reverse process is useful.

We shall give two examples and some applications to pulse problems.

1. *Frequency spectrum of the Unit Pulse and Unit impulse functions*

We shall call "unit pulse function" $P_1(t)$ the function represented in Fig. 5a, and such that $2Y_0\epsilon = 1$. ϵ can have any finite value. In the limiting case where $\Sigma \rightarrow 0$ and $Y_0 \rightarrow \infty$, the pulse becomes very sharp and of infinite amplitude (Fig. 6a); we shall call it a "unit impulse" and $I_1(t)$ will designate the corresponding function.

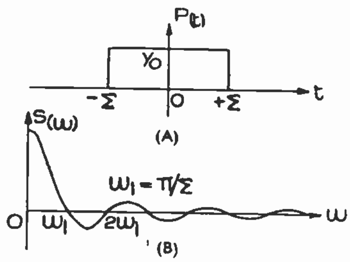


Fig. 5.—Unit impulse function and corresponding frequency spectrum.

From equation (21) we get for the coefficient B_n of the unit pulse function

$$\lim_{\omega_0 \to 0} B_n = \lim_{\omega_0 \to 0} \frac{1}{2\pi} \frac{\sin n\omega_0 \epsilon}{n\omega_0 \epsilon} \omega_0$$

since $2Y_0\epsilon = 1$. The unit pulse function is thus

$$P_1(t) = \lim_{\omega_0 \to 0} \sum_{n=1}^{\infty} \frac{1}{2\pi} \frac{\sin n\omega_0\epsilon}{n\omega_0\epsilon} \cos n\omega_0 t \; \omega_0$$

Putting $\lim_{\omega_0 \to 0} \omega_0 = d\omega$ and $\lim_{\omega_0 \to 0} n\omega_0 = \omega$ we get

$$P_1(t) = \frac{1}{2\pi}\int_0^{\infty} \frac{\sin \omega\epsilon}{\omega\epsilon} \cos \omega t \, d\omega \quad \ldots\ldots\ldots (26)$$

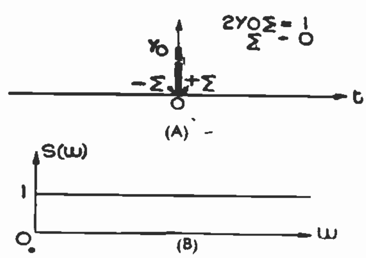


Fig. 6.—Pulse function and corresponding frequency spectrum.

The frequency spectrum of the unit pulse function is thus

$$S_p(\omega) = \frac{\sin \omega\epsilon}{\omega\epsilon} \quad \ldots\ldots\ldots\ldots\ldots\ldots (27)$$

The shape of this function is represented on Fig. 5b. It is a wavy curve with progressively damped oscillations. The first wave occurs between $\omega = 0$ and $\omega_1 = \pi/\epsilon$. Note that π/ϵ is the radian frequency of a sine wave having a period equal to 2ϵ (Fig. 5a). The following oscillations have a much smaller amplitude. Note also that the frequencies having the higher amplitudes are all in the low-pass band 0 to ω_1. We will show further that these are the most important frequencies for the reproduction of the pulse in amplitude

The unit impulse function $I_1(t)$ is

$$I_1(t) = \lim_{\epsilon \to 0} P_1(t) = \frac{1}{2\pi}\int_0^{\infty} \cos \omega t \, d\omega \quad \ldots\ldots\ldots (28)$$

and the corresponding frequency spectrum is

$$S_1(f) = 1 \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (29)$$

since $d\omega = 2\pi \, df$.

Thus, the unit impulse function has a uniform frequency spectrum and contains all frequencies with the same amplitude. We shall see that this function is very useful in the study of the response of electric networks.

In the case of pulse and impulse functions, integral (25) is convergent and Fourier transforms can be used and give the frequency spectrum straight away (see Part II).* We shall give now an example for which integral (25) is infinite and Fourier transforms cannot be used directly (unless care is taken to use a convenient path of integration[3]).

2. *Frequency Spectrum of the Unit Step Function*
The unit step function $S_1(t)$ is represented in Fig. 7a. It can be considered either as the integral of the unit impulse function or as the limit of the periodic pulse function of Fig. 7b when $\omega_0 \rightarrow 0$.

Consider the second case. We have an example of application of Fourier expansions. Comparing Fig. 4a with Fig. 7b, we see that the Fourier expansion corresponding to Fig. 7b can be obtained from equation (20) and is

$$f(t) = \frac{1}{2} + \frac{2}{\pi}\left[\frac{\sin \omega_0 t}{1} + \frac{\sin 3\,\omega_0 t}{3} + \frac{\sin 5\,\omega_0 t}{5} \ldots\right]$$

$$= \frac{1}{2} + \frac{1}{\pi} \sum_{n=1,3}^{\infty} \frac{\sin n\omega_0 t}{n\omega_0} 2\,\omega_0$$

* Part II will be published in the September Journal.

Putting $\lim_{\omega_0 \to 0} 2\,\omega_0 = d\omega$ and $\lim_{\omega_0 \to 0} n\omega_0 = \omega$ as before, we get

$$S_1(t) = \lim_{\omega_0 \to 0} f(t) = \frac{1}{2} + \frac{1}{\pi}\int_0^{\infty} \frac{\sin \omega t}{\omega}\, d\omega \ldots\ldots (30)$$

$$\text{Hence } S_2(\omega) = \frac{1}{\pi}\frac{1}{\omega} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (31)$$

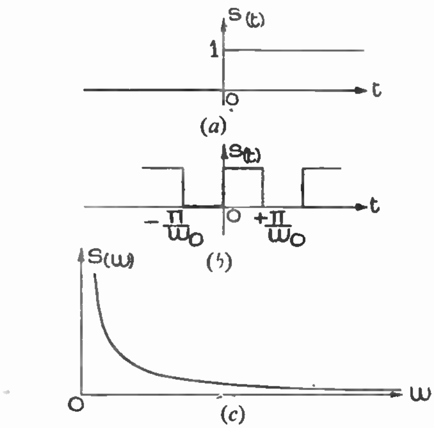


Fig. 7.—Unit step function (a) considered as the limit of a square periodic function (b) when $w_0 \longrightarrow 0$ and corresponding frequency spectrum (c).

We shall give now some applications of the results obtained above.

3. *Indicial Response of an Ideal Low-pass Filter*

An ideal low-pass filter is one having a constant amplitude transmission in the pass-band, a sharp cut-off, an infinite attenuation outside the pass-band (Fig. 8a) and a constant delay for all frequencies, that is a phase-shift ϕ proportional to the frequency

$$\phi = \omega\tau \pm 2\pi$$

We will assume that there is no gain or loss in the pass-band (amplification equal to unity). Then, if we apply a unit step voltage at the input of the filter at time $t = t_0$, the response $R_1(t)$ will be

$$R_1(t) = \frac{1}{2} + \frac{1}{\pi}\int_0^{\omega_c} \frac{\sin \omega\,(t - t_0 - \tau)}{\omega}\, d\omega$$

The integral on the right-hand side is the well-known sine integral, tabulated in many mathematical books. The response $R_1(t)$ is represented in Fig. 8d. It seems in appearance that the response appears before the time at which the step voltage appears at the input. This is due to the fact that we have assumed τ finite. An ideal filter cannot be realised in practice as is explained in Part II. In practice the response starts at $t = t_0$, for instance, and the portion of curve on Fig. 8d for which $t < t_0$ does not appear.

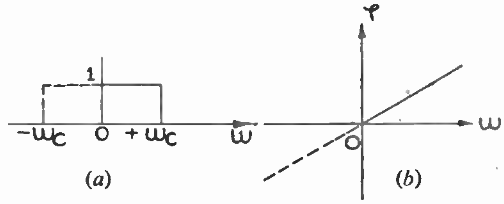


Fig. 8.—(a) and (b)—Frequency and phase characteristics of an ideal low-pass filter.

The centre of symmetry 0 of the response appears at $t = t_0 + \tau$, that is with a delay τ with the step of the unit step voltage applied. If only an approximate representation is required, we can neglect the waves of the lower and upper parts, and the response may be approximated by a lower and higher horizontal portion connected by a slope line. (Fig. 8e.)

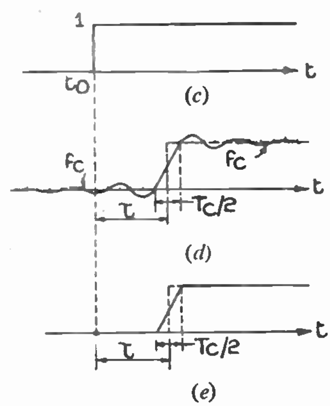


Fig. 8.—(c)—Unit step voltage applied at the input of the ideal filter.
(d) Response of the filter to the unit step voltage.
(e) Response in first approximation.

4. *Response of an Ideal Low-pass Filter to a Pulse of width $\triangle t$.*

To find the response to an impulse of duration $\triangle t$, we will consider the pulse as the difference of two step voltages delayed by $\triangle t$ one from the other. It is clear then that the response is the difference of two sine integral curves also delayed by $\triangle t$. In taking for the sine integral curve the approximate curve of Fig. 8e, one may easily notice that three cases are to be considered :

(1) $\triangle t >> T_c$.—In this case (Fig. 9a) the amplitude

of the response is equal to unity and independent of the value of the ratio $\Delta t/T_c$.

(2) $\Delta t << T_c$. —In this case the maximum amplitude A of the response varies with the ratio $\Delta t/T_c$:

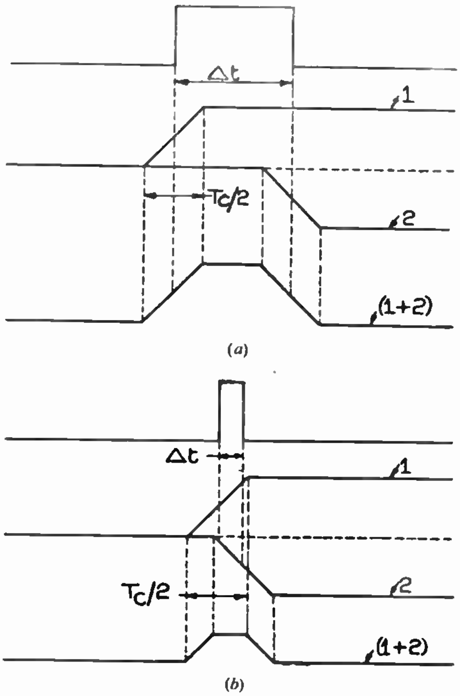


Fig. 9.—Response of a low-pass filter to a pulse of width Δt:

(a)— $\Delta t > T_c$.
(b)— $\Delta t < T_c$.

$$A \cong \frac{\Delta t}{Tc} = \Delta t\, f_c$$

and decreases proportionally to the band width f_c (Fig. 9b.)

Figs. 10 and 11 represent the same construction as in Fig. 9, but in using the correct sine integral curve. In Fig. 10 a pulse of adjustable width is applied at the input of an ideal filter of cut-off period T_c.

In Fig. 11 a pulse of fixed width is applied at the input of a low pass filter whose cut-off period is assumed variable.

One can see that the above conclusions are still applicable.

These conclusions are important in radar receivers since it is important to design the receiver in order to get the maximum signal/noise possible.

As the R.M.S. amplitude of the noise is proportional to $\sqrt{f}$, f being the received frequency, the signal/noise ratio is proportional to

$$\frac{1}{\sqrt{fc}} \quad \text{for } fc > \frac{1}{\Delta t}$$

and

$$\frac{\Delta t\, fc}{\sqrt{fc}} = \Delta t \sqrt{fc} \quad \text{for } fc < \frac{1}{\Delta t}$$

it is clear that the best value is when $fc \simeq \frac{1}{\Delta t}$; thus :

For an ideal low-pass impulse receiver, the signal to noise ratio is maximum when the cut-off frequency of the receiver is equal to $1/\Delta t$ (i.e. when the cut-off period is equal to the duration of the impulse.)

5. *Response of a Low-pass Filter to Two close Pulses*

The same method as above can be used to determine the optimum *band*-width necessary to distinguish two close pulses. Fig. 12 shows that a *good separation of the two pulses is obtained if the cut-off period is equal to about* 1·5 *to twice the distance between the two pulses.*

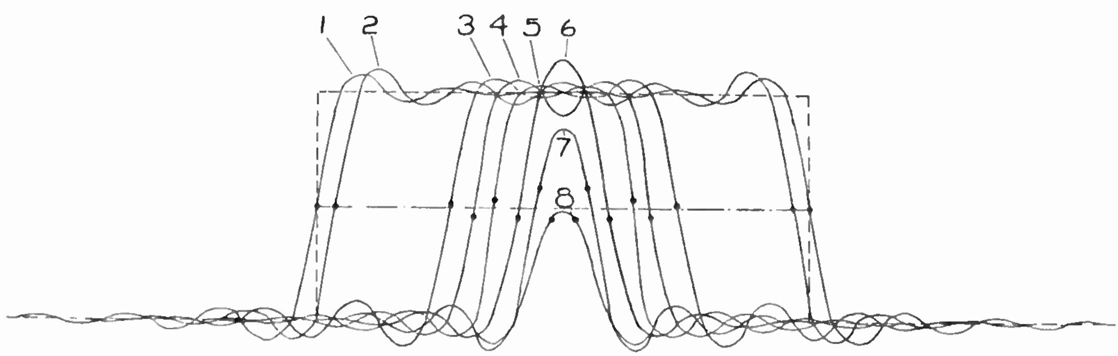


Fig. 10.—Response of an ideal low-pass filter to a rectangular pulse for different values of the pulse-width. (For $\Delta t > T_c$, pulse and response have same width at half the pulse height). On each curve the distance between the two points represented is equal to the width of the original square pulse. From curves 1 *to* 8, Δt *is equal reepectively to* 5·5, 5, 2·5, 2, 1·5, 1, 0·5, 0·25 *times* t_c.

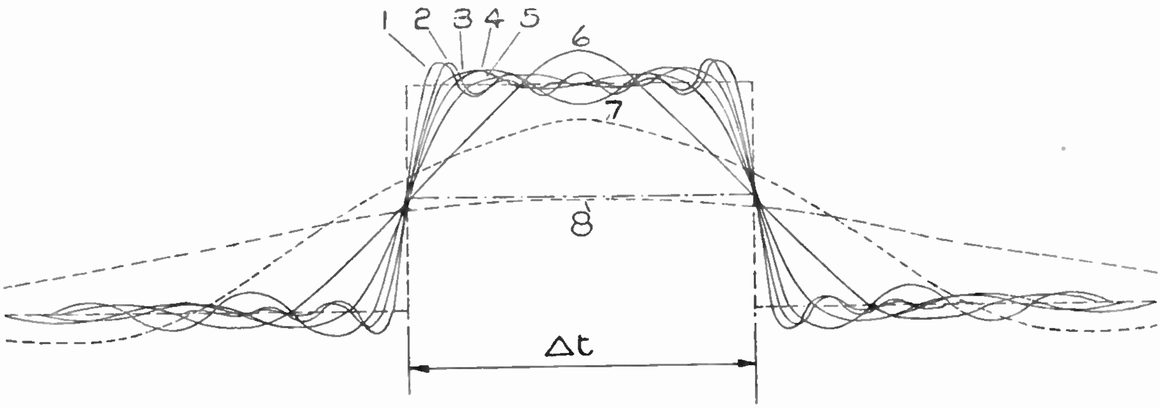


Fig. 11.—Response of an ideal low-pass filter to a rectangular pulse for different values of the cut-off frequency. (For Tc < Δt, pulse and response have same width at half the pulse height). From curves 1 *to* 8, T_c *is equal respectively to* 0·18, 0·2, 0·4, 0·5. 0·67, 1, 2, 4, *times* Δ*t.*

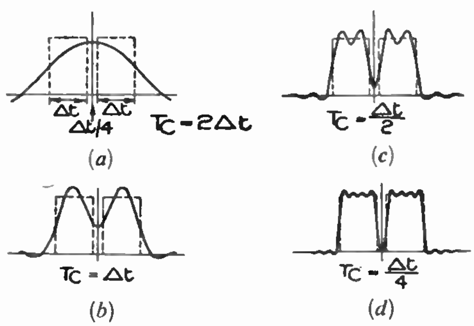


Fig. 12.—Response of an ideal low-pass filter to a double pulse. From (a) to (d) the cut-off frequency is increased progressively.

II. Harmonic Analysis*

Fourier series are essentially used to analyse complex periodic phenomenon in sinusoidal components.

The first method seems to be the one published in 1766, i.e. about 200 years ago. Since then new methods are published periodically, either based on numerical calculations or mechanical, electrical and electronical integrating devices.

To evaluate the harmonic components, the straightforward tendency is to evaluate approximately the integrals

$$\left.\begin{aligned} A_m &= \frac{1}{\pi}\int_{-\pi}^{+\pi} \gamma(x)\sin mx\,dx \qquad (1') \\ B_m &= \frac{1}{\pi}\int_{-\pi}^{+\pi} \gamma(x)\cos mx\,dx \qquad (1'') \end{aligned}\right\} \quad (1)$$

* This section has been written from memory, all original documents having been lost at the time of the German invasion of France.

This evaluation can be done in assuming that dx is a small but finite quantity Δx, and in assuming that in each interval Δx, $\gamma(x)$, $\sin mx$ and $\cos mx$ are constants.

If we divide the period 2π in $2n$ equal parts, equation (1) becomes

$$\left.\begin{aligned} A_m &\cong \frac{1}{n}\sum_{i=0,1}^{2n-1} \gamma(x_i)\sin mx_i \qquad (2') \\ Bm &\cong \frac{1}{n}\sum_{i=0,1}^{2n-1} \gamma(x_i)\cos mx_i \qquad (2'') \end{aligned}\right\} \quad (2)$$

where $x_i = i\,\dfrac{2\pi}{n}$.

Another way of approach is as follows : If we consider the expression

$$Y(x) = Co + A_1 \sin x + A_2 \sin 2x + \dots + A_n \sin nx$$
$$B_1 \cos x + B_2 \cos 2x + \dots + B_n \cos nx$$

which is identical to the Fourier series limited to the n first harmonics, we can try to evaluate the values of the the $(2n + 1)$ unknown coefficients in order that this development represents the curves as closely as possible. To do this we divide the curve in $2n$ parts by equidistant ordinates and we write that the above development gives the same amplitude $Y(x)$ for each subdivision. This gives $(2n + 1)$ equations from which all the coefficients are deduced. Calculation shows that one gets equations similar to (2). They are due to Lagrange, and will be designated by his name in the following.

These equations will thus give the real values of the coefficients if all the harmonics of order higher than n are of negligible amplitude or nil. If such is not the case, the values will be approximate.

Some writers have tried to get a better approximation by replacing the given curve by another as similar as possible to the first one but easily accessible to calculation. Such is the case, for instance, for curves formed by a series of lines or portions of parabolas connected together. Three ways of doing such an approximation are shown in Fig. 13. In each case the formulæ of discontinuity can be used.

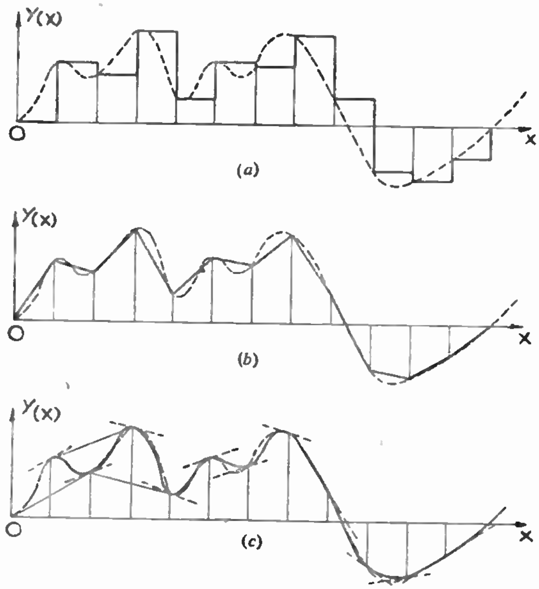


Fig. 13.—Different methods of approximating a curve of which we know only n *points. The curve is approximated in* (a) *by a sum of rectangles, in* (b) *by lines connecting each two successive points, and in* (c) *by arc of parabolas of the third degree connecting each two successive points and having at each point a tangent parallel to the line connecting the preceding to the following points.*

The writer has carefully studied all these methods and has found that they are closely correlated. The values obtained by one method can be deduced from the values obtained from any other one by multiplication by convenient coefficients. The conclusion is that the best is to use Lagrange's formulæ first, which is the simplest, and then to multiply the values obtained by convenient coefficients corresponding to the type of representation which seems the best.

Unfortunately, Lagrange's formulæ, if applied, requires a considerable amount of calculation. To see it, observe that the calculation of the coefficients A_m and B_m of each harmonic component requires $2n$ multiplications and two additions of n values for each. For the n harmonics, $2n^2$ multiplications and $2n$ additions of n terms are required. For $n = 20$, this gives 800 multiplications and 40 additions of 20 terms each.

Many methods have been suggested to reduce the number of calculations. It can be observed that the values of $\sin m\frac{\pi}{n}i$ can be identical for two or more values of i $(i<2_n)$, so that some operations can be connected together and the number of multiplications can be reduced.

Ingenious methods have been suggested by Lohmann, Lahr, Perry, Kintner and others. But it is the merit of Runge to have shown that these connections can be done systematically and to have reduced considerably the amount of operations.

Runge observes that the sum or difference of equidistant ordinates from both ends appear in all the coefficients so that he introduces the new variables

	y_0	y_1	y_2	$y_3 \ldots\ldots y_i \ldots\ldots\ldots$
		y_{n-1}	y_{n-2}	$y_{n-3} \;\cdot\cdot\; \ldots\ldots\ldots$
Sum	a_0	a_1	a_2	$a_3 \ldots\ldots a_i \;\ldots\ldots$
Difference		b_1	b_2	$b_3 \ldots\ldots b_i \;\ldots\ldots$

and he expresses all coefficients in function of these new variables. This nearly halves the number of terms for each coefficient. But Runge observes that the same operation can be done on the new coefficients, thus halving again the number of terms, and so on. By so doing he reduces finally the number of multiplications to the 20th of the original number.

The method found, it was necessary to establish schedules for making the calculations quick and easy. Runge and Emde gave schedules for the calculation of 6 and 12 harmonics. Later on, Running, Taylor, Whittaker, Pollack and others gave schedules more or less similar for the calculation of a number of harmonics usually not greater than 24.

The labour required for computing these tables increases considerably with the number of harmonics. Furthermore, the amount of operations remains still proportional to the square of the total number of harmonics.

It is found that the methods of formation of the schedules can be simplified considerably so that tables for the calculation of as many as 160 harmonics can be established very quickly. Furthermore, for some values of n (total number of harmonics), the number of equations is a minimum, whilst for others it is considerably greater. For instance, for $n = 82$ and 96, the number of operations is nearly double those required for $n = 86$. Also the number of operations is not proportional to n^2 but nearly to n for high values of n. This makes the calculation of a great number of harmonics, such as 120 or 160, possible and even easy, since with the schedules prepared by the writer and a calculating machine, 120 harmonics can be obtained in three hours. This makes the method a practical proposition.

Schedules for 12, 16, 24, 36, 48, 60, 72, 84, 96, 120 and 160 harmonics have been prepared.

Once this was done, the writer compared this method with all mechanical and other methods suggested. He found that the labour required with mechanical methods is generally greater, sometimes much greater, whilst other electric or electronic methods either require very complicated apparatus, which can only be found in few laboratories, or require apparatus which are not available commercially. In all cases the method of calculation compares favourably with others and requires no apparatus.

The practical importance of these facts has been considerably extended by further works of the writer, who showed that the methods of harmonic analysis can be applied to the spectrum analysis of non-periodic curves.

The calculations give exactly the same results as if a selective amplifier having a selectivity curve of the type $\left(\frac{\sin x}{x}\right)^2$ was used, the frequency of selection being varied n times along the spectrum. He applied his method to the analysis of recorded aircraft and telephone noise, and got the spectrum in each case with an amount of labour not exceeding ten hours. The methods have been used in many laboratories with great success. The writer has lost his manuscript and calculations during the German invasion of France, but hopes to publish a new study when completed.

III. References

(1) *General Treatise.*

1. W. E. Byerly. " An Elementary Treatise on Fourier's Series and Spherical, Cylindrical and Ellipsoidal Harmonics."
2. A. Eagle. " A Practical Treatise on Fourier Theorem and Harmonic Analysis." Longmans, Green & Co., London, 1925.
3. E. A. Guillemin. " Communication Networks." John Wiley & Sons, New York, 1931.
4. M. Levy. " Series de Fourier et Analyse Harmonique."

(2) *Papers.*

(*a*) **Fourier Series**

5. Lagrange. "Oeuvres completes." T.I. p. 39 and p. 553. (Original studies published in 1759 and 1762 to 1765).
6. A. Clayton. *Jour. I.E.E.*, 1920-21, p. 491. "Corrective Terms for Lagrange's Formulæ."
7. G. Koehler and A. Walther. "Fouriershe Analyse von Funktionen mit Sprüngen Ecken und ahnlichen Besonderheiten, Archiv für Elektrotechnik, 12 Nov. 1931, pp. 747-758 (Formulæ of discontinuity and examples of application).

(*b*) *Applications.*

(8) W. L. Sullivan. "Analysis of Systems with Known Transmission-frequency Characteristics by Fourier Integrals." *Electrical Engineering,* May 1942, pp. 248-56.

9. A. V. Bedford and G. L. Fredendall. "Transient Response of Multi-Stage Radio Frequency Amplifiers." *P.I.R.E.*, April 1939, pp. 277-84.
10. H. Nyquist and K. W. Peleger. " Effect of the Quadrature Component in Single-side-band Transmission." *Bell System Technical Journal*, Jan., 1940, pp. 63-73.
11. L. B. Arguimbau. "Transient Response of a Broadcast System." *General Radio Experimenter.* April 1940.
12. M. Levy. (1) "Transformations selectives. Application à l'analyse des mélanges de sinusoides." *Comptes Rendus de l'Academie des Sciences*— t. 198, p. 2222, 25 juin 1934.

    (2) " The Impulse Response of Electrical Networks, with Special Reference to the Use of Artificial Lines in Network Design." *Journal I.E.E.*, Dec. 1943, pp. 153-164.
13. Hamburger. Letter to the Editor. *Wireless Engineer*, February 1944.

SOME NOTES ON THE OPERATIONAL CALCULUS

by

L. Jofeh, A.M.I.E.E.*

1. Introduction

Until fairly recently a radio engineer could gather an adequate understanding of the behaviour of the circuits and elements with which he worked by studying their steady state responses. Within the last decade, however, this has rapidly become inadequate, and the study of the transient responses of networks has become increasingly more important and necessary. The truth of this remark is more or less self evident if one considers such topics as the design of television transmitting and receiving apparatus, delay lines of lumped elements, pulse and scanning transformers, and automatic regulators and servo mechanisms.

It is the object of this short paper to draw attention to the operational calculus, one of the most effective

*A. C. Cossor Ltd.

methods for the analytical study of transient phenomena. However, in order to place the operational calculus in something approaching true perspective, and to give at least a partial answer to the question why a busy engineer should go to the pains of learning a new branch of mathematics, it is desirable to consider briefly some alternative analytical processes.

2. Alternative Methods

Pride of place amongst the alternative processes for the study of transient phenomena must be given to the classical method of solution of differential equations, a process in no small measure responsible for the many triumphs of engineering. As is well known, the solution of a differential equation by classical methods involves finding both a particular integral (giving the steady state solution) and a complementary function, the latter having as many arbitrary constants as the order of the differential equation; in order to satisfy the boundary conditions of the problem it is necessary to determine the values of these constants, and this is precisely the great weakness of the classical method. So long as the differential equation is of fairly low order the work involved is not very great, but it quite rapidly becomes prohibitive as the order of the differential equation increases.

A second and very powerful method lies in the use of Fourier analysis and synthesis. As applied to the study of transient phenomena, the work may be divided into four stages, namely (*a*) calculation of the steady state amplitude and phase responses of the system to be examined, (*b*) Fourier analysis into harmonic components of the signal to which the system is to respond, (c) modification of the relative amplitudes and phases of the harmonic components in accordance with the steady state responses of the system and (*d*) summation of the modified components. While the first three stages are usually relatively straightforward, the summation of the harmonics tends to be an extremely laborious process, at least in the absence of mechanical synthesisers. It is probably true to say that it is largely as a result of this excessive demand on time and labour that so many fallacious short cuts are attempted: of these, the most common is to assume that the network has a constant time delay for all harmonics and a constant amplitude for the first n harmonics and zero reponse for the $(n + 1)$th and higher harmonics, a set of assumptions completely at variance with Bode's[1] well-known laws.

Despite this disadvantage, and the further disadvantage that the results obtained do not permit of the ready appraisal of the effects on the transient response of the various circuit elements forming the network, the Fourier method is extremely valuable: indeed, when considering the transient response of a multistage I.F. amplifier, for example, this method is probably the most useful of those available in the absence of a differential analyser,[2,3]

The operational calculus provides a method having the advantage of giving solutions (where they exist) in closed forms, as does the classical method of solution of differential equations: it has the further very great advantage that the boundary conditions are satisfied automatically, thereby avoiding the principal disadvantage of the classical method. Yet another attraction of the operational calculus is that its operations are for the most part algebraic and arithmetical rather than analytic. This list of advantages of the operational calculus over the alternative processes for the study of transient phenomena, although formidable, is by no means exhaustive; indeed, the feature most likely to appeal to the engineer, and one not so far mentioned, is that it is possible to construct tables of operational forms, or transforms, by means of which the work of analysis is rendered almost as easy as, for example, the use of logarithm tables.

It will not be possible in a paper of this length to offer anything like adequate substantiation of these claims: however, so many excellent texts are available that a skimped exposition would be superfluous as well as undesirable.

3. Systems of Operational Calculus

Broadly speaking, there are three points of view from which the operational calculus may be approached.

The first of these is due to Heaviside,[4,5] but is studied much more conveniently by reference to the works of more recent authors,[6,7,8]. Heaviside's own approach to the subject was both intuitive and deductive: his results are scattered widely through his many engineering papers, and to say the least, he was not concerned to spend much time upon proving the validity of his methods. However, the work of a great number of very able mathematicians has shown clearly the scope and limitations of the Heaviside operational calculus and it is to-day a flexible, easily handled method of extremely wide applicability.

A second line of approach is through Fourier Integral Analysis and Mellin's Inversion Theorem[9]. This process has a somewhat greater range of applicability than Heaviside's calculus, but makes rather greater demands upon the user's mathematical ability.

The third method, and probably the most powerful and widely applicable, is that of the Laplace Transform,[10,11]. A proper appreciation of the Laplace Transform calls for a fairly large amount of mathematical equipment, including especially a more than nodding acquaintance with the ideas of functions of a complex variable.[12,13] However, if the user is content to obtain a general idea of the process, and thereafter to use the results given in Tables of Transforms, the Laplace Transform method need have no terrors.

An excellent comparison of the three methods of approach to the operational calculus with one another and with the classical method of differential equations is given in the table on page 9 of Gardner and Barne's book.[11]

I. Boundary Conditions

As originally expounded, the Heaviside calculus was supposed only to be capable of handling the boundary conditions of the problem if the initial state of the system was one of complete quiescence; that is to say, if the problem to be studied was the transient behaviour of an electrical network, all the mesh currents, charges and potentials are assumed to be zero initially. This point of view imposed some restrictions on the applicability of the calculus, though generally these could be overcome by the use of a variety of artifices, in particular by invoking Thevenin's theorem. This procedure was unsatisfactory, even when adequate, as it is undesirable to have to adopt special measures to deal with particular cases.

However, as Jeffreys[7] and others have pointed out, such an *ad hoc* procedure is quite unnecessary, as the process of construction of the operational form, or transform, of a function is inherently capable of handling other initial conditions than those of quiescence in a systematic manner. Although eighteen years have elapsed since the original publication of Jeffreys' book, it is still quite common to find some confusion existing about this matter: the old fallacy is long dying (see, e.g. page 10 of Ref. 11), so that it is felt that some time might usefully be spent on a brief study of this topic.

We may begin by defining the operational form $\overline{\phi}(p)$ of a function of time $\phi(t)$. In the notation adopted by Carson[6] and Bush[8] the operational form is defined by the infinite integral

$$\overline{\phi}(p) = p\int_0^\infty \epsilon^{-pt}\,\phi(t)\,dt \quad \ldots\ldots\ldots\ldots\ldots\ldots(1)$$

and the existence of this integral, for a sufficiently large positive value of p, is the necessary and sufficient condition for the existence of the operational form. To clarify this before proceeding further, we may consider a few simple examples.

Example 1. $\phi(t) = \epsilon^{-at} \quad t \geqslant 0$

Then, if $R(p) > R(a)$, i.e. if the real part of p is greater than the real part of a, the integral (1) exists, and

$$\overline{\phi}(p) = p\int_0^\infty \epsilon^{-pt}\,\epsilon^{-at}\,dt$$

$$= p\left[\frac{-\epsilon^{-(p+a)t}}{p+\alpha}\right]_0^\infty$$

$$= \frac{p}{p+a} \quad \ldots\ldots\ldots\ldots\ldots\ldots(2)$$

Example 2. $\phi(t) = \sin\omega t \qquad t \geqslant \theta$

$$\phi(p) = p\int_0^\infty \epsilon^{-pt}\sin\omega t\,dt = \frac{\omega p}{p^2+\omega^2} \quad (3)$$

Suppose now that we are considering the transient response of a network, and that the differential equation connecting the response, $x(t)$ with the impressed signal $F(t)$ is

$$\{a_nD^n + a_{n-1}D^{n-1} + \ldots\ldots + a_1D + a_0\}\,x(t) = F(t) \quad \ldots(4)$$

where, as usual, D^r stands for $\dfrac{d^r}{dt^r}$

Let us denote the initial values of x, $Dx \ldots D^rx$ by $x_0, x_1 \ldots x_r$ respectively.

To transform equation (4) into its operational form, we perform the operation indicated by (1), i.e. we multiply both sides of the equation by $p\epsilon^{-pt}$ and integrate the product with respect to t over range 0 to ∞. The result of performing this operation on the R.H.S. of (4) gives

$$p\int_0^\infty \epsilon^{-pt}\,F(t)\,dt = \overline{F}(p) \quad \ldots\ldots\ldots\ldots\ldots(5)$$

Taking the expression on the L.H.S. term by term

$$a_0\,p\int_0^\infty \epsilon^{-pt}\,x\,dt = a_0\,\overline{x}(p) \quad \text{(by definition)}\ldots(6)$$

$$a_1\,p\int_0^\infty \epsilon^{-pt}\,Dx\,dt = a_1\,p\Big[\epsilon^{-pt}\,x\Big]_0^\infty + a_1p^2\int_0^\infty \epsilon^{-pt}\,x\,dt$$

$$= -a_1\,p\,x_0 + a_1\,p\,\overline{x}(p) \ldots\ldots(7)$$

$$a_2\,p\int_0^\infty \epsilon^{-pt}\,D^2x\,dt = a_2p\Big[\epsilon^{-pt}\,Dx\Big]_0^\infty + a^2\,p^2\int_0^\infty \epsilon^{-pt}\,Dx\,dt$$

$$= -a_2\,p\,x_1 - a_2\,p^2\,x_0 + a_2\,p^2\,\overline{x}(p) \ldots(8)$$

and so on, on the assumptions that the limits

$$\lim_{t\to\infty}(\epsilon^{-pt}\,D^r\,x) = 0 \qquad r \leqslant n$$

and that the integral $p\int_0^\infty \epsilon^{-pt}\,x(t)\,dt$ exists, if p is greater than some finite positive number.

Proceeding in this way, and adding, we have finally

$$\{a_n\,p^n + a_{n-1}\,p^{n-1} + \ldots\ldots + a_1\,p + a_0\}\,\overline{x}(p)$$
$$= \overline{F}(p) + a_n\,\{p^nx_0 + p^{n-1}x_1 + \ldots + p^2x_{n-2} + px_{n-1}\}$$
$$+ a_{n-1}\{p^{n-1}x_0 + p^{n-2}x_1 + \ldots + p^2x_{n-3} + px_{n-2}\}$$
$$+ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$
$$+ a_2\{p^2x_0 + px_1\} + a_1px_0 \qquad (9)$$

By the use of similar methods, we also see that had the original equation (4) included on the L.H.S. a term such as $a_{-1}\int x\,dt$, and if x_{-1} denoted the initial value of the integral, we should have had

$$a_{-1}\,p\int_0^\infty \epsilon^{-pt}\Big[\int x\,dt\Big]dt = \frac{a_{-1}\,\overline{x}(p)}{p} + a_{-1}\,x_{-1} \ldots\ldots (10)$$

The solution of equation (9) for $\overline{x}(p)$, which may then be interpreted to give $x(t)$, the unknown of the original equation (4), thus includes the boundary conditions x_{-1}, x_0, $x_1 \ldots\ldots x_n$, as required.

We see from (9) that the process to be adopted in setting up the operational form from the original differential equation is as follows :—

(*a*) replace Ft by $\bar{F}(p)$

(*b*) replace $\int x\,dt$ by $\bar{x}(p)/p + x_{-1}$

(*c*) replace $x(t)$ by $\bar{x}(p)$

(*d*) replace Dx by $p\,\bar{x}(p) - px_0$

$D^2 x$ by $p^2\,\bar{x}(p) - (p^2 x_0 + p\,x_1)$

. . . .

$D^r x$ by $p^r\,\bar{x}(p) - (p^r x_0 + p^{r-1} + p^{r-2} x_2 + \ldots + p^2 x_{r-2} + p\,x_{r-1})$

(*e*) Collect terms in $\bar{x}(p)$ and divide the coefficient of $\bar{x}(p)$ into the sum of all other terms, thus obtaining a polynomial fraction in *p*.

The preceding rules, derived quite straightforwardly from the infinite integral definition of the operational form, give the justification for Heaviside's procedure of identifying *p* with d/dt and $1/p$ with $\int dt$ when the initial conditions are those of quiesence.

5. Methods of Interpretation

The interpretation of the operational formulæ may be performed in a variety of ways : for details of these methods it is necessary to refer to one of the standard texts, since an adequate treatment is quite impossible in the space of a short paper ; the following remarks, however, may not be out of place.

Probably the most straightforward method is to expand the expression into partial fractions and to gather the interpretation of the various terms by reference to a Table of Transforms.

A simple example may help to illustrate this process. Suppose that a circuit consisting of a resistance R and an inductance L in series is connected across a battery of EMF E_0, and that at the instant when the current in the circuit is I_0 an additional battery of E.M.F. E is switched into the circuit : it is required to determine the current in the circuit.

The differential equation for the system is

$$\left.\begin{array}{ll}(LD+R)\,I = E + E_0 & t \geqslant 0 \\ I = I_0 \ \ at\ \ t = 0 & \end{array}\right\} \qquad \ldots\ldots\ldots\ldots (11)$$

Then, following the method described in Section (4), the operational equation is

$$\bar{I}(p) = \frac{E + E_0 + LpI_0}{Lp + R} \qquad \ldots\ldots\ldots\ldots (12)$$

On looking up a Table of Transforms, e.g. Bush, *loc. cit.*, p. 380, formulæ (2) and (3), we find

$$\frac{1}{p+b} = \frac{1}{b}(1 - \epsilon^{-bt})$$

$$\frac{p}{p+b} = \epsilon^{-bt}$$

where the = sign is to be read as " has the following interpretation for $t \geqslant 0$." Hence, from (12)

$$\bar{I}(p) = \frac{1}{p + R/L} \cdot \frac{E + E_0}{L} + \frac{p}{p + R/L} \cdot I_0$$

$$\text{and so}\ \ I(t) = \frac{E + E_0}{R}\left(1 - \epsilon^{\frac{-Rt}{L}}\right) + I_0 \epsilon^{\frac{-Rt}{L}} \qquad \ldots\ldots (13)$$

At the other extreme of the scale one may place the interpretation obtained by utilisation of the inversion transformation and contour integration. Formally, if

$$\bar{x}(p) = p \int_0^\infty \epsilon^{pt}\, x(t)\, dt \qquad \ldots\ldots\ldots\ldots (14)$$

$$\text{then}\ \ x(t) = \frac{1}{2\pi j} \int_{c - j\infty}^{c + j\infty} \epsilon^{-pt}\, \frac{\bar{x}(p)}{p} \,.\, dp \qquad \ldots\ldots (15)$$

where *c* is greater than the real part of any pole of $\bar{x}(p)$. This method is unnecessarily difficult in many instances, even when the calculus of residues is utilised for the evaluation of the contour integral, but in more advanced work, especially that arising from partial differential equations, it is extremely useful and powerful.

In certain circumstances, it is not possible to obtain a solution in closed form ; this state of affairs occurs most frequently in the study of systems with distributed parameters, e.g. cable and radiation problems. In such instances it is often possible to obtain two series solutions, one a power series, convergent but most laborious to compute except for small values of the independent variable, because of its slow convergence, and the other a divergent, asymptotic solution which is nevertheless fairly readily computed at large values of the independent variable. An excellent account of the theory of the asymptotic solution is given by Carson (*loc. cit.*, Ch. V).

6. Conclusion

It may be that, quite unwittingly, the impression will have been gathered that the operational calculus is difficult to understand and use : if so, the present writer is to blame, because in fact the true position is quite the reverse of this. Like all other branches of mathematics, the operational calculus takes some time to learn, and a fair amount of practice is needed to give one confidence and speed of working, but in the writer's experience the effort is more than amply repaid. In their book on applied mathematics, von Kármán and Biot[14] use as the heading to the chapter on the operational calculus a quotation from Heaviside which might aptly serve as a motto for those approaching the subject for the first time ; it is as follows :—

" Shall I refuse my dinner because I do not fully understand the process of digestion ? "

7. Bibliography

1. Bode, H. W. "Relation between Phase-lag and Attenuation." *Bell Telephone Journal* 19 (1940), 421.
2. Bush, V. "The Differential Analyser. A New Machine for Solving Differential Equations." *Franklin Inst. Journal* 212 (1931) 447-488.
3. Hartree, D. R. 34th Kelvin Lecture. "Mechanical Integration in Electrical Problems." *Journal Inst. Elec. Eng.* 90 (1943), 435.
4. Heaviside, O. "Electrical Papers" (2 vols). *Macmillan.*
5. Heaviside, O. "Electromagnetic Theory" (3 vols.). *Van Nostrand.*
6. Carson, J. R. "Electric Circuit Theory and Operational Calculus." *McGraw Hill.*
7. Jeffreys, H. "Operational Methods in Mathematical Physics. *Cambridge University Press.*
8. Bush, V. "Operational Circuit Analysis." *Wiley.*
9. McLachlan, N. W. "Complex Variable and Operational Calculus. *Cambridge University Press.*
10. Carslaw, H. S., and Jaeger, J. C. "Operational Methods in Applied Mathematics. *Oxford University Press.*
11. Gardner, M. F., and Barnes, J. L. "Transients in Linear Systems," Vol. 1. *Wiley.*
12. Phillips, E. G. "Functions of a Complex Variable." *Oliver & Boyd.*
13. Whittaker, E. T., and Watson, G. N. "A Course of Modern Analysis." *Cambridge University Press.*
14. von Kármán, T., and Biot, M. A. " Mathematical Methods in Engineering. *McGraw-Hill.*

DISCUSSION

Mr. G. L. Hamburger: I should like to make a contribution to the operational calculus.

In the days of television—which, I hope, will speedily be revived—one of the problems was to make an amplifier not only pass frequencies up to the megacycle region, but also faithfully transmit rectangular pulses of 50 c/s repetition frequency. Let us draw one stage of the actual amplifier and its equivalent circuit with the assumption that we use pentodes with stabilised screen supply.

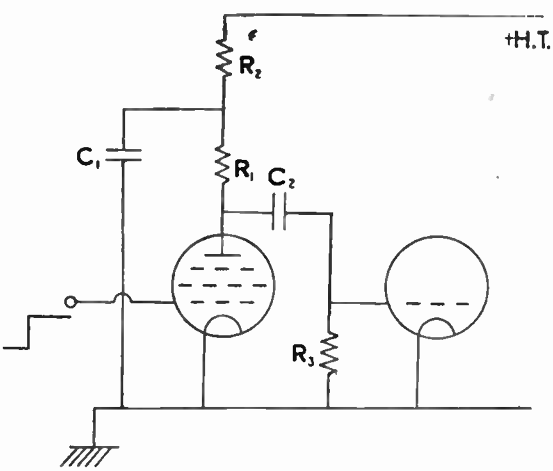


Fig. 1(*a*).—*Actual amplifier circuit.*

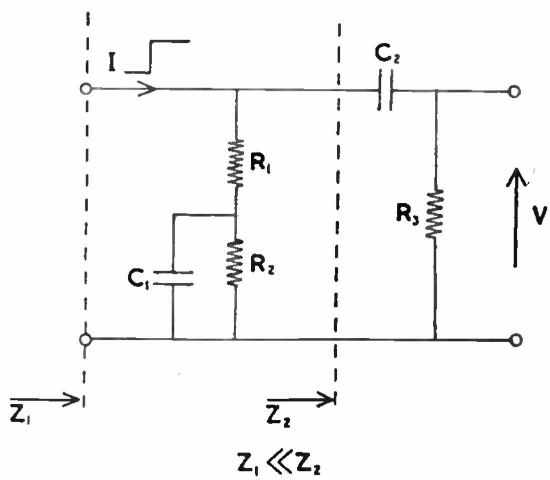


Fig. 1(*b*).—*Equivalent circuit of the insterstage coupling.*

The supreme test for our amplifier under the defined operating conditions will be its response to unitstep. We realise, of course, that any A.C.-coupled amplifier is necessarily "D.C.-losing." in other words, a unit step impressed upon the first grid must give a response on the grid of the second valve, which inevitably must go down to the quiescent zero value after some time, as indicated in Fig. 2b. Amongst the three principally possible responses the one starting with a horizontal tangent is obviously the best, and we will now try to find out under what conditions this can be achieved, and also which factors will draw out the response as long as possible.

The response voltage V can be worked out in terms of the input pentode current I under the assumption of a high grid leak value R_3 as

$$(1)\ldots\ldots V = I1\left[R_1 + \frac{R_2\,\frac{1}{pC_1}}{R_2 + \frac{1}{pC_1}}\right]\left[\frac{R_3}{R_3 + \frac{1}{pC_2}}\right]$$

$$= I1\,\frac{[R_1(1 + pC_1R_2) + R_2]\,pC_2R_3}{(1 + pC_1R_2)(1 + pC_2R_3)}$$

and arranged in powers of p,

$$(2)\ldots\ldots V = I1\,\frac{p^2[C_1C_2R_1R_2R_3] + p[C_2R_3(R_1 + R_2)]}{p^2[C_1C_2R_2R_3] + p[C_2R_3 + C_1R_2] + 1}$$

$$= I1\,\frac{p^2a + pb}{p^2c + pd + 1}$$

This result does not yet convey much to us.

(*a*) (*b*)

Fig. 2.—Unit step impressed on the first grid gives these fundamentally possible responses. They all tend to zero.

However, bearing in mind that we desire a horizontal tangent of the response it appears that we can achieve the result by developing the response into a power series of p, and using one of those formulæ mentioned in Bush's book interpreting p in terms of time t, namely

$$(3)\ldots\ldots\ldots\ldots \frac{1}{p^n} = \frac{t^n}{n!}$$

The power series in p can be obtained by simple division:

$$(4)\ldots(p^2a + pb) : (p^2c + pd + 1) = \frac{a}{c} + \frac{1}{p}\left[\frac{bc - da}{c^2}\right]$$

$$+ \frac{1}{p^2}\left\{\frac{1}{c^2}\left[\frac{d}{c}(da - bc) - a\right]\right\}\ldots\ldots =$$

$$= k_0 + \frac{k_1}{p} + \frac{k_2}{p^2} + \frac{k_3}{p^3} + \ldots\ldots$$

By substituting the conversion formula we obtain

$$(5)\ldots \frac{V}{I} = k_0 + \frac{k_1}{1!}t + \frac{k_2}{2!}t^2 + \frac{k_3}{3!}t^3 + \ldots\ldots,$$

the response as function of time, actually as a power series of t. Now let us discuss this result term by term.

The first term, k_0, gives the step voltage produced at the grid of the second valve. By successively substituting from equations (4) and (2) we find that

$$(6)\ldots k_0 = \frac{a}{c} = \frac{C_1C_2R_1R_2R_3}{C_1C_2R_2R_3} = R_1$$

a result which can immediately be verified from the equivalent circuit of Fig. 2a. The capacitors cannot suddenly alter their charges so that the sudden voltage jump is—in this circuit—faithfully transmitted as the voltage drop developed across R_1.

The next term, the first order term, can similarly be calculated as

$$\frac{k_1}{1!} = \frac{bc - da}{c^2} = \frac{1}{c^2}\Big[C_1C_2R_2R_3C_2R_3(R_1 + R_2)$$

$$- C_1C_2R_1R_2R_3(C_2R_3 + C_1R_2)\Big] =$$

$$= \frac{C_1C_2R_2R_3}{C_1^2C_2^2R_2^2R_3^2}\Big[C_2R_3R_1 + C_2R_3R$$

$$- C_2R_1R_3 - R_1C_1R_2\Big]$$

$$(7)\ldots \frac{k_1}{1!} = \frac{C_2R_3 - R_1C_1}{C_1C_2R_3}$$

This coefficient determines the tangent of the response immediately after the step and evidently it can be positive, zero, or negative according to the choice of the circuit elements. This can also be seen by inspection of Fig. 3.

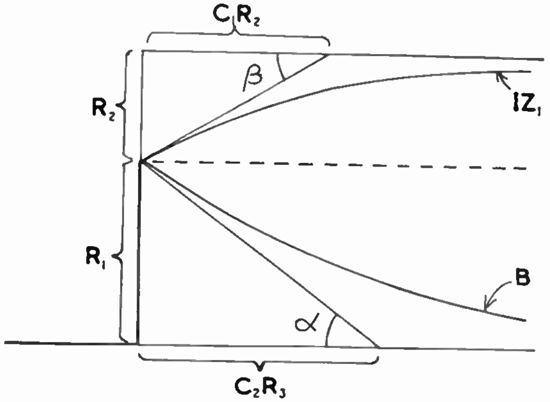


Fig. 3.—Tan α = tan β

$$\frac{R_1}{C_2R_3} = \frac{R_2}{C_1R_2} \therefore R_1C_1 = R_2C_2$$

The step current flowing into R1, and R_2 with C_1 in parallel produces an immediate step voltage proportional to R_1, and as C_1 is being charged up with the time constant C_1R_2 the voltage reached finally is proportional to $R_1 + R_2$. This voltage trend marked as IZ_1 is shown in Fig. 3 as the top curve. It acts as a source e.m.f. for the C_2R_3 series combination, since the impedance of the latter is very much greater than Z_1. If C_2R_3 were fed by the simple unit step proportional to R_1 it would lose the D.C., at a rate inversely proportional to its time constant C_2R_3 (curve B). However, by impressing upon Z_2 the e.m.f. IZ_1 instead of the single unit step, the rise of IZ_1 will at first prevent the fall of B if both tangents have the same absolute value, namely, if $\alpha = \beta$. From this follows that

$$(8)\ldots R_1C_1 = R_3C_2$$

a result immediately obtainable by equating (7) to zero.

Equation (8) answers our first question, namely, it defines the condition when the response will start with a horizontal tangent.

The second point, namely, how to prolong the nearly horizontal part of the response as much as possible, can be answered from the second order coefficient.

$$\frac{k_2}{2!} = \frac{1}{2!\,c^2}\left[\frac{d}{c}\,(da-bc)-a\right]$$

and by using equations (2) and (8) this becomes :

$$(9)\ldots\ldots\ldots \frac{k_2}{2!} = -\frac{1}{2\,R_2C_1^2}$$

Since the second order term can be considered as a measure of the deviation from constancy it appears that prolonging the response would necessitate a reduction of this second order coefficient, or an increase of $R_2C_1^2$. It is valuable to know that it pays better to increase the decoupling capacitor C_1 instead of decoupling resistance R_2, since the former improves matters with its square.

" I could have told you without mathematics that you want a fat condenser," I can almost hear some of my friends say. I know it is obvious that the D.C.-losing process can be slowed down by increasing the time constants, but it is not obvious at a first glance of the mere circuit diagram that and where there is an optimum of operating conditions. This can only be found by proper analysis.

This example may serve as a demonstration to show that the processes involved in the evaluation of transient responses can really be of the simplest type : Ohm's law extended for alternating currents, algebraic manipulations which any successful matric student should be able to carry out, and at last, looking up of the right conversion formula in the right book. This, I admit, may be the trickiest part.

Dr. H. Moss : I congratulate Mr. Hamburger on his efforts, but I want to put forward a case which is more suitable for beginners. It is the problem of the network consisting of two resistors, R_1 and R_2 in series ; each resistor is shunted by a condenser C_1 and C_2 respectively, and the network is to operate as a voltage divider with a division ratio constant at all frequencies.

The working is a little easier than in the case put forward by Mr. Hamburger. The equations are more straightforward, and the result can be verified by much the same sort of argument as Mr. Hamburger used in his typical analysis.

Mr. L. Grinstead (Chairman) : The speakers have shown that problems which at first appear quite insoluble can be attacked in a reasonably quick manner.

The idea of operational calculus is really to simplify the type of mathematical apparatus by reducing it from differential equations with boundary conditions to algebraic operations. The latter may, in cases, be cumbersome, nevertheless they are straightforward.

Mr. Jofeh has given us a very useful bibliography which will be available to us all.

I was rather amazed by Mr. Levy's contribution because some of the things, for which he wrote down the answers immediately, would be quite impossible for the average person, and difficult even for a man versed in advanced methods.

Time does not permit of a very long and continuous discussion, but the object of the meeting has been accomplished and we are very grateful to Mr. Jofeh and Mr. Levy.

TRANSFERS AND ELECTIONS TO MEMBERSHIP

The following elections and transfers were recommended by the Membership Committee at their meeting held on May 3rd, 1946. At this meeting, the Committee considered a total of 27 proposals for transfer or election to Graduateship or higher grade membership.

Transferred from Associate to Associate Member

MORRISON, Donald Finlay	Maldon
PROUDLOVE, Charles Bertram, B.Sc.	Bristol

Transferred from Student to Associate

FEHER, George	Haifa
HAMILTON, James Richard	Greenford

Elected to Associate Member

GOODWIN, Herbert Frederick	London, S.W.15

Elected to Associate

ARSCOTT, Ronald Harry	Sittingbourne
GRENFELL, William Edward	Uxbridge
MALLOY, John James	Harrow
MARCHANT, Albert Charles	Little Singleton
MERRIDEW, James Norton	Buxton
NICHOLSON, Kelvin, B.Sc.	Southampton
STEVENS, Brian Ronald, B.Sc.	Malvern Link
TYERS, Arthur	Eastleigh
WHITEHOUSE, Martin Henry Leigh, B.A.(Hons).	Edenbridge.

STUDENTSHIP REGISTRATIONS

The General Council have also approved the following 25 registrations as Students, wite effect from May 3rd, 1946.

BARCLAY, William Edward	Dublin
DEUTSCHBERGER, Wladyslan E.	Glasgow
DOHERTY, James Andrew	Dublin
DOWLING, Patrick Joseph	Co. Kildare
FOULDS, William H. H.	Glasgow
GREEN, Arthur Eric	Liverpool
GROSVENOR, Kenneth D. James	Walsall
HALAHAN, Henry Crosby	Liss
HIPPLE, Henry	Liverpool
LEE, Charles Het Hien	Singapore
LEVIS, George Charles	Eire
LONG, William Joseph	Sidmouth
MASON, Denis Connell	Wolverhampton
MONAHAN, Leo Joseph	Dublin
PORTMAN, Reginald Leonard	Halesowen
PRADHAN, Keshar B., B.Sc.	Bombay
RAGHUNATH RAO, A. L. N., B.Sc.	Bezwada
REDMOND, James	Finchley, N.3
STEVENSON, Andrew	Norwich
THORNHILL, Sidney	Liverpool
TSAN, Ling, B.Sc.	London, W.1
WATSON, George Ernest	Ipswich
WILDE, George Thornton W.	Prescot, Lancs
WIRRELL, Gordon Wilson	Manly, Australia
YORKE, Albert Kilpatrick	Co. Westmeath