THE RADIO AND ELECTRONIC ENGINEER

The Journal of the Institution of Electronic and Radio Engineers

FOUNDED 1925 INCORPORATED BY ROYAL CHARTER 1961

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

VOLUME 29 MARCH 1965 NUMBER 3

INTERNATIONAL EXHIBITIONS

RECENT years have seen a considerable increase in the number of electronics exhibitions held in Europe—a trend which is all part of the general expansion in international trade. These occasions enable the presentation of a company's products to create an impression which counts for far more than advertisements however informative. From the point of view of the prospective customer, the opportunity to carry out at least a preliminary comparison of competing products is well worth while.

But exhibitions, and particularly international exhibitions, provide much more than market places: they are becoming occasions when engineers can widen their horizons, both technical and professional. The opportunities for obtaining a broad picture of developments in one's own and associated fields of work, for attending the conferences and symposia which are held concurrently with many exhibitions and, above all, for meeting engineers of other countries, are all good reasons for urging manufacturing organizations to enable their middle-level as well as senior engineers to travel overseas and visit appropriate exhibitions. In this matter we are not speaking only to British companies—a two-way traffic must be set up so that engineers from overseas are encouraged to visit the various exhibitions held in this country.

In April electronic engineers from all over Europe and indeed the whole Western World will be visiting Paris for the Eighth International Electronic Components Salon at which there will be over 800 exhibitors. (A preview of some of the exhibits by British electronics companies is included with this issue.) In addition to an associated exhibition concerned with Electro-Acoustics, a "Colloque International sur les Techniques des Memoires" is being organized by the Société Française des Eléctroniciens et Radioéléctriciens. It can be confidently predicted that the attendance of 115 000 in 1964, of which 14% were from 65 countries outside France, will be exceeded.

In Great Britain, too, international conferences are being held this year to coincide with major exhibitions. This Institution is collaborating with the Institution of Electrical Engineers and other bodies in organizing conferences on "Components and Materials Used in Electronic Engineering" in May at the time of the London Components Exhibition, and on "U.H.F. Television" during the International Television and Radio Show in September.

The electronic engineering profession is most certainly one which has far more to gain from adopting an international outlook than from a short-sighted policy of trying to be self-sufficient. International exhibitions and conferences are important means of achieving this end. We make no apology for returning to this theme—the Institution is firmly committed to an international outlook and the great advantage of "International Exchange" set out in an editorial article in this Journal last September and underlined by the Institution's President, Colonel G. W. Raby, in his Inaugural Address, can hardly be stressed too often. On the most materialistic level the old saying "Trade follows the flag" can, with every justification, now be paraphrased to read "Trade follows the engineer".

F. W. S.

INSTITUTION NOTICES

Joint I.E.E.-I.E.R.E. Symposium on "Microwave Applications of Semiconductors"

A joint I.E.R.E.-I.E.E. Symposium on "Microwave Applications of Semiconductors" will be held at University College, London, on Wednesday, 30th June to Friday, 2nd July.

Some forty papers will be presented in the course of six sessions, having the following headings:

Microwave Properties of Semiconductor Devices (including design of features and measurements); Generators; P-I-N Diode Circuits; Tunnel Diode Circuits and Mixers; Low Noise Devices; and Systems.

Registration charges for members will be £10, which will include the complete set of preprints, lunch and refreshments on all three days; the registration charge for those who are *not* members of either of the sponsoring Institutions will be £13.

A list of the papers will be published in the April and May issues of *The Radio and Electronic Engineer*.

British Joint Computer Conference

The 1966 British Joint Computer Conference will be held at the Congress Theatre, Eastbourne, Sussex from 3rd to 5th May, 1966. Sponsoring the conference, under the aegis of the United Kingdom Automation Council, are the Institution of Electrical Engineers, the Institution of Electronic and Radio Engineers, the British Computer Society, the British Institute of Management and the Institute of Chartered Accountants.

It is hoped to include the following topics in the programme:

Scientific applications—Combinatorial problems such as scheduling, assignment and timetabling; use of computers for engineering design.

Management planning and decision making—Resource utilization; network analysis; business control at all levels, including production control.

Current problems—The interaction between the design of computer systems (hardware/software) and the needs of the user; how the systems engineer proposes to meet the requirements of the user; how the user proposes to accommodate the limitations of the system.

Engineering of real time systems; and Future Horizons.

Contributions of up to 3000 words are invited for inclusion in the programme. Brief synopses should be sent to the Conference Secretariat, I.E.E., Savoy Place, London, W.C.2, by 31st July, 1965, from whom further details and registration forms will be available in due course.

Conference on U.H.F. Television

An international conference on "U.H.F. Television" will be held in London on Wednesday and Thursday, 1st and 2nd September 1965. Aspects to be covered include receiver and transmitter design, propagation, receiving and transmitting aerials, parametric amplifiers and test equipment.

The sponsoring bodies, the I.E.E. Electronics Division, the I.E.R.E., the I.E.E.E. U.K. and Eire Section and the Television Society, are inviting papers and contributions. Synopses of about 250 words should be submitted by 20th April to: The U.H.F. Television Conference Joint Secretariat, 8–9 Bedford Square, London, W.C.1. Complete papers, and contributions of not more than 2000 words or the equivalent, including diagrams, should arrive by 1st June.

Conference on "Optics in Space"

The Optical Group of The Institute of Physics and The Physical Society is arranging a conference on "Optics in Space" to be held at the University of Southampton from 27th to 30th September, 1965. It is proposed that the main sessions of the conference will cover the following topics: Materials and instrumentation, spectroscopy, telescopes in space, laser techniques, television techniques, optical tracking, optical guidance, photography from space vehicles, environmental testing and wavelength sampling.

Offers of papers of about 15 minutes presentation time should be submitted by 30th June to Dr. H. G. Jerrard, at the Physics Department, University of Southampton. Further details may be obtained from the Meetings Officer, I.P.P.S., 47 Belgrave Square, London, S.W.1.

S.M.P.T.E. Conference

The 98th Technical Conference of the Society of Motion Picture and Television Engineers will be held at the Queen Elizabeth Hotel, Montreal, from 31st October to 5th November, 1965. It is intended to cover the following subjects:

Television and cinema in air transportation, and their use in education; television transmission, automation and test methods; audio-visuals in the developing nations; studio construction and production techniques; and scientific applications.

Offers of papers should be sent to Mr. Gerald Graham, Papers Chairman, S.M.P.T.E., National Film Board of Canada, P.O. Box 6100, Montreal 3.

U.D.C. 621.374.4

A Technique for the Time-transformation of Signals and its Application to Directional Systems

By

W. J. CAPUTI, M.Sc.†

Reprinted from the Proceedings of the Symposium on "Signal Processing in Radar and Sonar Directional Systems", held in Birmingham from 6th–9th July 1964.

Summary: This paper describes a passive time-transformation technique ('Stretch') that permits an exchange of signal time duration for bandwidth. It is not limited to electrical signals or to a particular portion of the spectrum. It is linear in the sense that the principle of superposition is applicable. Good waveform fidelity is preserved and signals that are resolvable at the input remain resolvable at the output, even though the bandwidth may be changed by a large factor. Signal/noise ratios are not affected by the transformation.

The basic components of the system are: (1) the uniformly dispersive delay device in which envelope delay is a linear function of frequency and phase shift is a parabolic function of frequency, and (2) the heterodyne mixer in which the heterodyne signal is linearly modulated in frequency. By proper combinations of these non-linear phase and amplitude elements, systems can be devised to speed up, slow down, or reverse the time scale of the input signal.

In much the same way that a transformer permits matching the impedance properties of a circuit, the technique permits matching of the bandwidth capabilities of the circuit. One of the most obvious applications is to enable the observation or transmission of extremely fast phenomena with economical, low-bandwidth equipment. Other applications include speed-up of signals to permit efficient time-division multiplexing and the generation of accurately controlled fast waveforms.

Application of the technique to wide-bandwidth, wide-baseline directional systems will be discussed. Removal of most of the bandwidth takes place at the individual receiving sites. Because the phase and amplitude characteristics of the signals are preserved, remoting, combining, processing and recording of the information can be performed at practical bandwidths.

1. Introduction

'Stretch' is a general technique for transforming the time scale of signals that are confined within a defined interval of time. Slow-down or speed-up of signals is accompanied by an increase or decrease in the time interval over which they occur. The principles of this technique allow temporal operations to be performed on electrical, electromagnetic, optical or acoustical signals. Although only non-linear passive components are required, the system is linear in the sense that superposition is applicable and good waveform fidelity is attained. This technique is a general solution to the problem of matching data rates of signals to receivers when the bandwidth characteristics of both are fixed.

† Airborne Instruments Laboratory, Cutler-Hammer, Inc., Long Island, New York.

Impedance matching to obtain maximum transfer of power is a well-known concept. The application of this concept to the data rate or information-transmission characteristics may be more often neglected. The principle, a generalization of the 'matched filter' concept,‡ is that the data rate of a signal should be matched to the data-transfer characteristic of a transmission line or receiver. Mismatch results in either loss of signal information or inefficient use of the transmission line or receiver.

For example, consider an experiment in which it is desired to measure the rise-time associated with a non-repetitive millimicrosecond transient by displaying

‡ D. O. North, "An analysis of the factors which determine signal/noise discrimination in pulsed carrier systems", *Proc. Inst. Elect. Electronic Engrs*, **51**, pp. 1016–27, July 1963. (Originally published in 1943.)

the waveform on an oscilloscope. If the signal is applied directly to the input of the oscilloscope, performance and cost inefficiency results because of the expense of obtaining an oscilloscope with a rise-time sufficient to handle the waveform. This is because no advantage has been taken of the fact that the fast rise-time phenomenon occurs over a relatively short interval of time, and only events in the vicinity of the transient are of interest. Thus the basic experiment produces signals with an extremely low duty cycle of information content in spite of the wide bandwidth. If we had a convenient way of reducing the bandwidth of the signal by slowing down the waveform before the signal is displayed on the oscilloscope, we could display the required information on an inexpensive instrument.

The lack of appreciation of the concept of data matching is probably due to the lack of devices that operate in the time domain and that can be used as data-rate transformers. The purpose of this paper is to describe a technique that can be used to provide this function under a wide variety of conditions.

Section 2 explains the general concept of the time-transformation technique by means of a simplified approximate analysis. This discussion is intended to give the reader an understanding of the manner in which the technique operates for a slow-down application and further discusses the approximations used in the analysis of the 'Stretch' technique. Section 3 describes a typical application of the technique to ease the problems of obtaining high directivity in a wide-band receiving system.

2. Analysis of Time-transformation Technique

Figure 1 is a functional block diagram of one implementation of the 'Stretch' technique. The input signal that is to be transformed in time is applied to a dispersive device, mixed against a sweep, and then applied to a second dispersive device that is used as a compressor. For purposes of explanation,

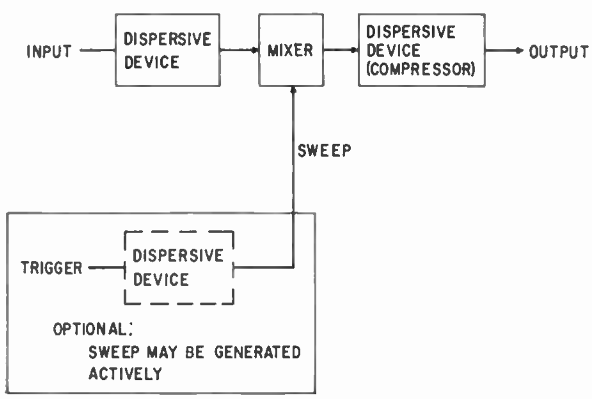
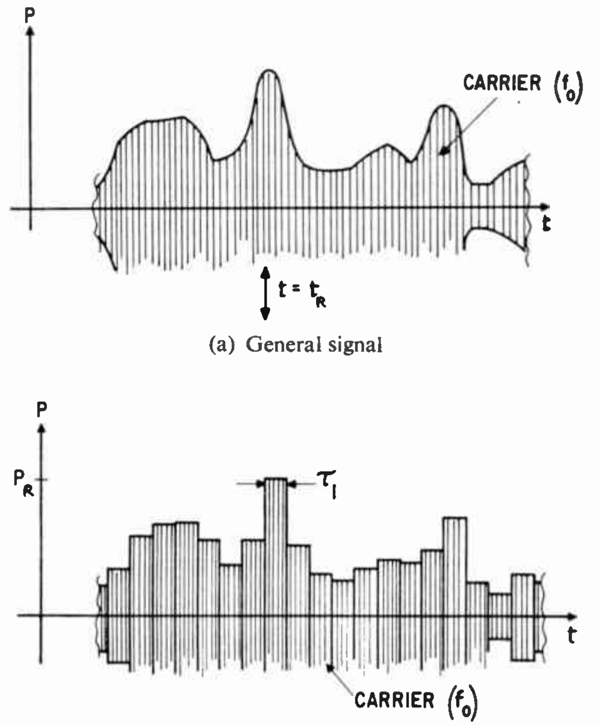


Fig. 1. Block diagram of time-transformation technique.



(a) General signal

(b) Pulse samples

Fig. 2. Input signal.

let us assume that an arbitrary input waveform is to be slowed down in time so as to reduce the bandwidth of the signal.

The assumed arbitrary signal (Fig. 2) is the input to the device. (This signal is shown as modulated on a carrier; however, if the input signal is not of this form, it can easily be modulated on a carrier frequency or the basic principles of the technique can be applied to work with a video waveform.)

Shannon's sampling theorem,† allows us to resolve the general signal of Fig. 2(a) into component pulses. In accordance with the sampling theorem, the elementary pulses used to represent the waveform must be faster than the rise-time of the signal information and have a bandwidth at least equal to the bandwidth of the input signal. If rectangular sampling pulses are assumed, it is easy to see how the complex signal can be approximately represented by the summation of the individual component pulses. Figure 2(b) shows the pulse samples.

Consider first a reference pulse at the centre ($t = t_R$) of the signal described in Fig. 2. This pulse (shown in Fig. 3(a)) has a pulse length τ_1 and an amplitude P_{R1} as determined by the bandwidth of

† C. E. Shannon, "Communication in the presence of noise", *Proc. Inst. Radio Engrs*, 37, pp. 10–21, January 1949.

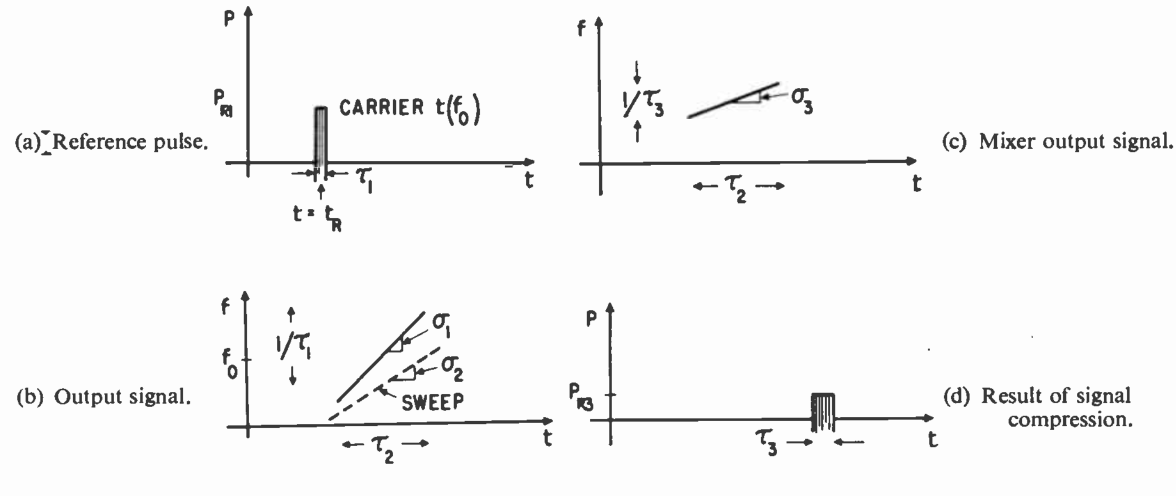


Fig. 3. Sampling pulses.

the sampled waveform and the amplitude of the waveform at the time t_R. If the bandwidth of the input dispersive device is at least equal to that of the signal, a dispersed pulse will be obtained at the output of the dispersive network. The dispersed pulse length and the slope of the frequency modulation under this envelope is determined by the characteristics of the device. Figure 3(b) shows the output signal from the device. The bandwidth over this pulse length is, of course, the same as the bandwidth of the sample, $1/\tau_1$. The dispersed pulse length is

$$\tau_2 = \frac{1}{\sigma_1 \tau_1}$$

where σ_1 is the frequency/delay slope of the device. The amplitude of the pulse is

$$P_{R2} = \frac{\tau_1}{\tau_2} P_{R1}$$

This dispersed reference pulse is now mixed with a wideband sweep. The sweep is a linearly frequency-modulated voltage (as shown in Fig. 3(b)) and is similar to the signal obtained at the output of the dispersive device.

The sweep can be generated either actively if components are available or passively by impulsing a third dispersive device with a trigger.

Since the input signal is to be slowed down, the wide-band sweep frequency/time slope (σ_2) is set slightly below that of the dispersive network (σ_1). The timing of the sweep (shown dotted in Fig. 3(b)) is adjusted to coincide with the dispersed reference pulse.

The difference-frequency mixer output is chosen in this example; hence, the mixer output signal is at a lower carrier frequency and has a reduced slope σ_3, where

$$\sigma_3 = \sigma_1 - \sigma_2$$

This signal (shown in Fig. 3(c)) is simply the difference between the two signals of Fig. 3(b). This mixer output is now applied to the input of the second dispersive device (the compressor).

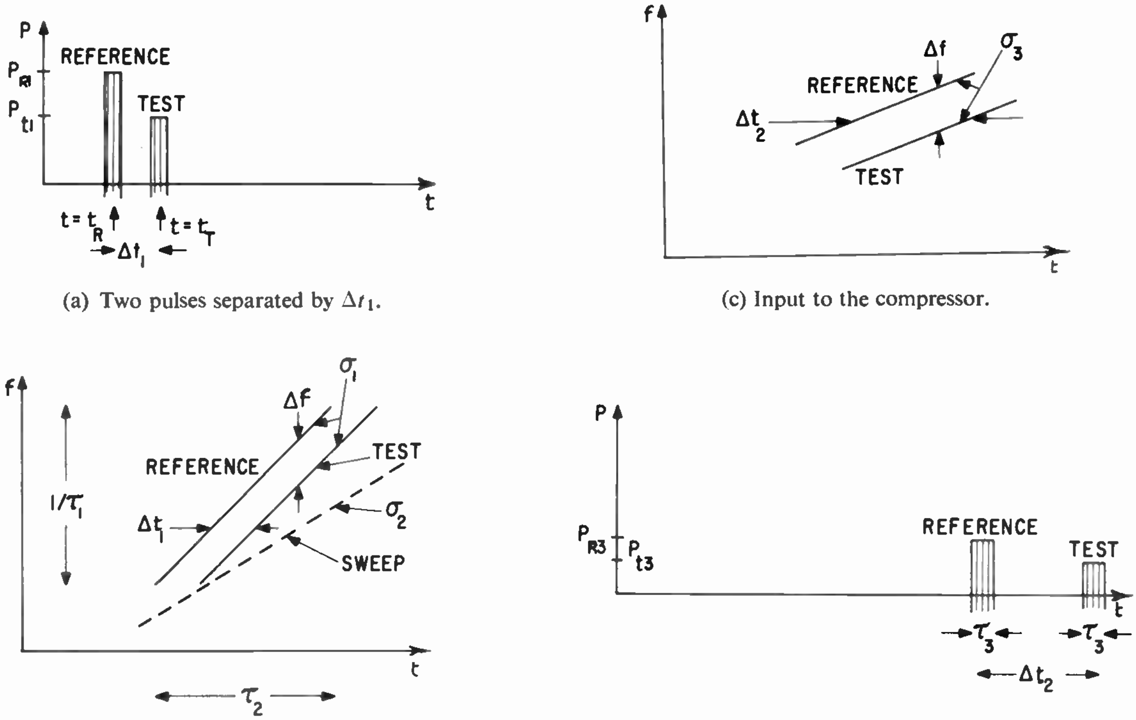
The characteristics of the second dispersive device are chosen to compress the signal that is applied to it. Since the slope of the applied signal is determined by the characteristics of the input dispersive device and of the sweep, this network can be designed independently of the input signal. The result of compressing the signal is shown in Fig. 3(d). The compressed pulse width (τ_3) is the reciprocal of the bandwidth of the applied signal. This bandwidth is determined by the slope (σ_3) and the pulse length (τ_2) of the applied signal:

$$\tau_3 = \frac{1}{\sigma_3 \tau_2} = \frac{\sigma_1}{\sigma_3} \tau_1$$

The amplitude of the output pulse is

$$P_{R3} = \frac{\tau_2}{\tau_3} P_{R2} = \frac{\tau_1}{\tau_3} P_{R1} = \frac{\sigma_3}{\sigma_1} P_{R1}$$

Note that both the output pulse width and the amplitude are proportional to the input pulse width and the amplitude, with a proportionality constant determined by the design of the input and output dispersive devices. Furthermore, since the output-signal energy is always the same as the input-signal energy, no change will result in the signal/noise ratio of a noisy input signal (the noise power is also reduced by the ratio of the input to output bandwidth).



(a) Two pulses separated by Δt_1.

(c) Input to the compressor.

(b) Output of the dispersive network.

(d) Output signal from the compressor separated by Δt_2.

Fig. 4. Relationship between the two signals.

Let us now consider a second 'test' pulse in the train of sampling pulses representing the input waveform. The test pulse, which is similar to the reference pulse, is positioned at an arbitrary time ($t_T = t_R + \Delta t_1$) relative to the reference pulse. Its amplitude (P_{t1}) is proportional to the amplitude of the complex signal at time t_T.

Figure 4 shows the relationships between these two signals as they pass through the transforming apparatus. Figure 4(a) shows the two pulses separated by the time Δt_1. Figure 4(b) shows the output of the dispersive network corresponding to these two pulses; these two dispersed linear f.m. signals are displaced in time by Δt_1. After passing through the mixer, the two signals are mixed against the sweep and, as before, reduced in slope and centre frequency (Fig. 4(c)).

To determine the position of the test mixer output signal, we must consider the fact that at any instant of time the two mixer inputs and the sweep behave like constant-frequency signals. Since both mixer input signals are mixed against a common instantaneous sweep frequency, the two output signals must have a constant frequency separation that is equal to the input frequency separation, as shown in Fig. 4(c). From Fig. 4(b), we can calculate the frequency between the two mixer input signals, namely

$$\Delta f = \sigma_1 \Delta t_1$$

Since the frequency separation is the same at the input and output of the mixer, we can now reverse the procedure to compute the time separation (at any constant frequency) between the two output signals as follows:

$$\Delta t_2 = \frac{\Delta f}{\sigma_3} = \frac{\sigma_1}{\sigma_3} \Delta t_1$$

These two signals (Fig. 4(c)) are now applied to the compressor, and two output pulses are obtained. These two output pulses (Fig. 4(d)) are separated in time by an amount (Δt_2) equal to the separation in time of the equal-frequency components of the two dispersed signals applied to the device. The starting and finishing times of the applied signals are of no significance in determining the time of the output pulses. Thus we see that the separation of the output test and reference pulses in time is increased by a factor equal to that with which each of the pulses is broadened—that is, the ratio of the frequency/time slopes of the input and output dispersive devices. The amplitude of the test pulse is, as before:

$$P_{t3} = \frac{\sigma_3}{\sigma_1} P_{t1}$$

Thus, we have proved that all of the samples of the input waveform are broadened and separated by a controllable factor in passing through the apparatus. In addition, no signal energy has been lost. The output signal is, of course, the reconstituted sum of the output sample pulses. We can conclude that the input signal has been slowed down in time with high waveform fidelity and with no loss of signal/noise ratio.

To determine the maximum time displacement of the test pulse from the reference position, let us assume that the pass bands of the dispersive devices are rectangular in shape, exactly equal to the bandwidth of the signals passing through the devices, and that the signal spectra are rectangular ($(\sin x)/x$ pulse samples). Referring to the previous examples, the bandwidth of the input dispersive network is $1/\tau_1$ and the bandwidth of the output network is $1/\tau_3$. We will assume that the input signal is band-limited with a bandwidth of $1/\tau_1$. If we now review the analysis, we will find that no difficulty arises until we reach the signal that is applied to the output compressive network. Referring to Fig. 4 we see that the time separation of the test signal from the reference signal has been accompanied by a shift in frequency. As we move the test sample pulse away from the reference position (that is, increase Δt), the linear f.m. signal applied to the compressor shifts in frequency (that is, Δf increases) as shown in Figs. 5(a) and 5(b).

The pass band of the output compressor is fixed to accommodate the band defined by the reference pulse; therefore, as the test pulse is moved from the reference position, a loss of spectrum and pulse energy results. This, in turn, results in a gradual degradation of the test-pulse shape and amplitude, as a function of the position. This effect is shown in Fig. 5(c).

We can arbitrarily define an input signal 'window' by assuming that pulses that pass through the system are useful if no more than half of their energy is lost. Although this definition allows a considerable loss of power and bandwidth at the edges of the window, it is useful in many applications. The result of this assumption is that the maximum displacement of the test pulse (at the output) is equal to $\pm\ \Delta t_{2(\max)}$ where

$$\Delta t_{2(\max)} = \frac{1}{2\tau_3\sigma_3(1-\sigma_3/\sigma_1)}$$

The ratio of the output signal maximum displacement divided by the output pulse width (or signal rise-time) is equal to the window width W (expressed in number of resolution elements).

$$W = \frac{2\Delta t_{2(\max)}}{\tau_3} = \frac{1}{\tau_3^2\sigma_3(1-\sigma_3/\sigma_1)}$$

or

$$W = \frac{\tau_2}{\tau_3}\frac{1}{1-\sigma_3/\sigma_1}$$

The term τ_2/τ_3 is the time-bandwidth product of the compressor. In addition, when the parameters are chosen to obtain slow-down of input signals, the term σ_3/σ_1 often becomes negligible. Hence, the window width W becomes equal to the time-bandwidth product of the output dispersive device.

In general, the limits over which good performance can be obtained are related to the engineer's ability to control and predict the performance of the input and output dispersive devices. In a practical implementation, these devices can be constructed to obtain linear dispersion within sufficient tolerance to obtain spurious response and distortion of the order of

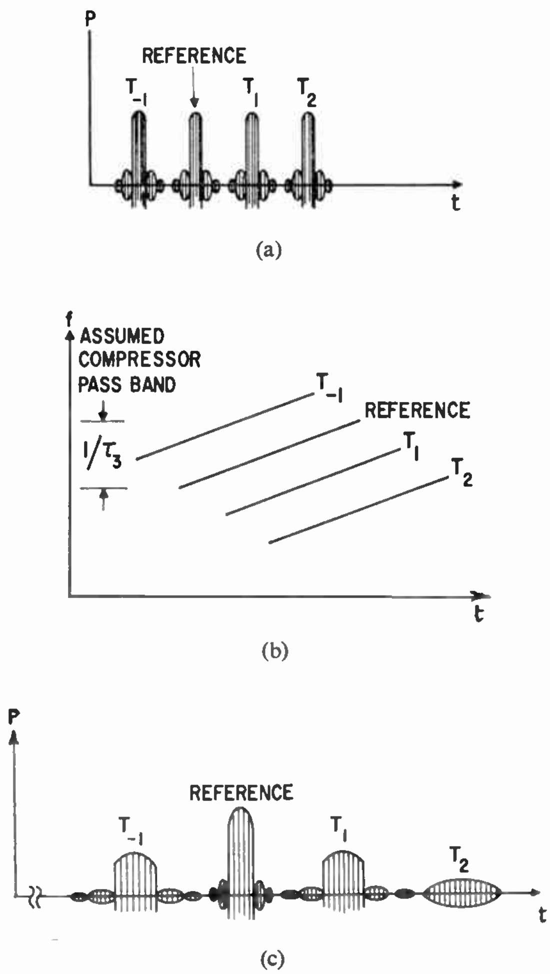


Fig. 5. Effect of increasing Δt.

30 to 50 dB below peak signal. The maximum ratio of input to output time scale is generally determined by the engineer's ability to match the sweep to the input-dispersion characteristic. By using special error-correcting techniques, time slow-down ratios of the order of several hundreds have been achieved.

3. Application of 'Stretch' to Wide-baseline Systems

To illustrate a particular application of the time slow-down technique, we will discuss its application to a wide-baseline radio interferometer.

It is often desirable to obtain high angular resolution in a wide-band receiving system. If the carrier frequency of the signal is relatively low, an attractive means of obtaining high angular resolution is to use an interferometer configuration. Thus, when two receiving aerials are mounted some considerable distance apart, coherent addition of the signals can be performed to generate an equivalent narrow beamwidth. If the separation of the two aerials is large and the received bandwidth is wide (for example, a few hundred megacycles) difficulty may be encountered in feeding the received signals to the processing equipment and in processing and recording the signals. On the other hand, it is usually desirable to avoid narrow-banding of the signal, since the wide bandwidth aids in separating the interferometer lobes.

If (in the case that we are considering) it is desired to receive wide-band signals only over a short time interval, we can make use of the 'Stretch' transformation technique to reduce the bandwidth of the received signals separately at the local and remote sites, and then relay the coherent slowed-down signals to the beam-forming and signal-processing location. This configuration would be most useful when the wide receiver bandwidth is needed for the purpose of correlating one received signal against the other to reduce the interferometer lobe ambiguity.

The configuration is illustrated in Fig. 6. Signals entering the local and remote receivers are immediately dispersed by means of a wide-band dispersion network. Actually, the dispersion inherent in the waveguide run from the aerial feedhorn to the remainder of the equipment can perform this function. After dispersion, the signals are mixed against the output of the wide-band linear swept oscillator. The output signals from the mixer are now in narrow-band form and can be amplified easily and remoted to the processing station via a simple coaxial cable. In addition, the signals are at an intermediate frequency.

Since coherence is desired, a stable locking frequency must be provided to synchronize the phase of the two swept oscillators. This device is shown simply as a stable r.f. source that is gated into the swept oscillator at each site prior to the period during which the oscillator must be swept. This procedure locks the phase of the swept oscillators before sweeping and ensures that the outputs of the two 'Stretch' mixers will be coherent.

It is also necessary to provide the swept oscillator with stable triggering pulses. These need not be wide band, however, as long as the signal/noise ratio is sufficient to ensure accurate triggering of the devices.

The remainder of the equipment shown in Fig. 6 is conventional. Notice, however, that the dispersive network that is normally included after each of the 'Stretch' mixers has been moved to the output of the beam-forming network. Thus, only one network is required.

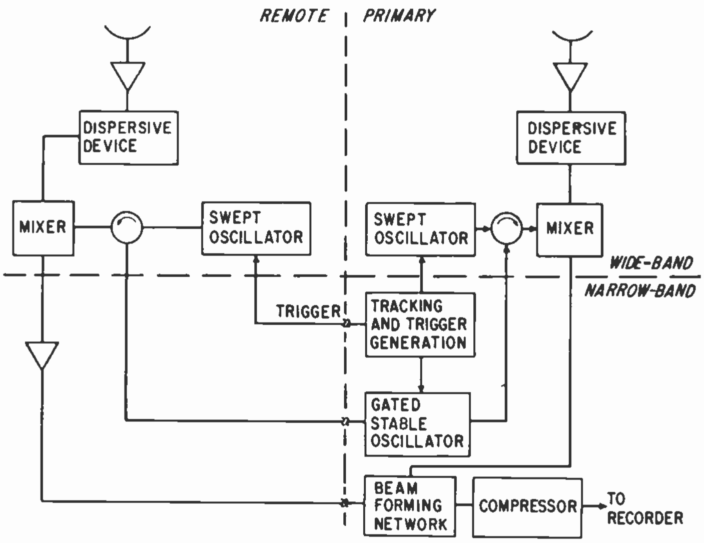


Fig. 6. Interferometer configuration of wide-baseline system.

In summary, we can see that the use of the slow-down technique at the receiving sites has allowed us to simplify greatly the remoting problem of the signals. In essence, only one transmission line is required between the remote site and the local site. This line can carry simultaneously the video trigger, the i.f. received signal and the r.f. coherent phase reference, since all are narrow-band. In addition, the beam-forming circuitry and processing circuitry at the local site is considerably simplified. Also, the narrow-band output signal allows the use of a simple tape recorder or other narrow-band recording means to form a permanent record of the received signals.

4. Conclusions

An approximate analysis has been developed to explain the operation of the 'Stretch' technique. Test results have been obtained that confirm the analysis.

It has been shown that the technique permits convenient linear manipulations of the time and bandwidth coordinates of signals, with good waveform fidelity. By means of this technique, signals can be slowed down, speeded up, or reversed in time with corresponding changes in the spectrum. The basic technique permits passive linear operation on electrical, electromagnetic, or acoustic signals over any portion of the spectrum.

In its simplest configuration, two dispersive elements and one mixer are required. Generation of the sweep from a trigger impulse can be accomplished by a third dispersive element. The use of this technique as applied to any particular type of signal therefore depends only on the availability of dispersive elements and non-linear devices to handle the particular type of energy and the portion of the spectrum involved. For microwave signals, a conventional diode mixer and waveguide operating near cut-off might be used for the essential components. For optical signals, glass is dispersive and photocells operate like mixers. In general, the input bandwidth, the maximum bandwidth reduction, and the amount of information that can be handled by a particular design are determined by the characteristic of the dispersive devices.

The 'Stretch' technique provides a general solution to the problem of matching the data rate obtained from an experiment or other signal source to that which can be handled by an observer, recorder, display or transmission line. Several specific applications are currently being pursued or have been successfully solved over the past three years, using the principles explained in this paper. Time slow-down factors of up to 75 have been achieved with input-signal bandwidths of 200 Mc/s and window widths of 0·5 μs in current operational equipment. Time slow-down of 500 Mc/s bandwidth signals has been achieved in the laboratory, and systems that operate at much higher input bandwidths are in the planning stage.

Manuscript first received by the Institution on 20th April 1964 and in final form on 16th June 1964. (Paper No. 964.)

DISCUSSION

Under the Chairmanship of Mr. R. N. Lord

Dr. R. Benjamin: The operation of this device may be thought of as follows: for each sample the first dispersive network separates co-existing frequencies in time. Next the frequency modulation decreases (or increases) these frequencies by a chosen *proportion*. Then the second dispersive network brings the (scaled) originally co-existing frequencies back into their correct time and phase relation.

However, at the same time, the f.m. produces a frequency difference between originally separated signal samples proportional to their original time separation. Hence, the second dispersive network will produce a proportional increase (or decrease) in that time separation.

Mr. E. D. R. Shearman: If I have understood correctly the system described, it is necessary to know when the event which is to be stretched is going to occur so that the locally generated frequency may be swept at the appropriate instant.

Is any generalization of the method possible which will overcome this necessity for synchronization?

The Author (*in reply*): It is necessary to know when the event is going to occur. The requirement is analogous to the requirement of a conventional oscilloscope for a synchronizing trigger when a small portion of a waveform is to be examined in detail. Of course, as in the conventional oscilloscope, it is always possible to derive the trigger (which starts the sweep generator) from the waveform itself, so that the 'Stretch' circuitry becomes internally triggered.

Mr. R. H. Johnson: Referring to the problem of observing a signal when no *a priori* knowledge of its occurrence is available, has the author considered using a number of transformation and display channels in parallel? In such a system no information would be lost.

The Author (*in reply*): The use of a multiplicity of slow-down channels is indeed one answer to the previous question. It has a disadvantage, however, of requiring a considerable amount of equipment when the time slow-down ratio is large, since the number of transformation and display channels which are needed to insure no loss of information is equal to the time slowdown ratio.

Mr. G. F. Vogt: This stretching technique can be explained by reference to a sampling method. This resembles, however, a method restricted to repetitive processes using as time diversion oscillator a saw-tooth generator† In the example (Fig. 5) the time expansion factor for a portion of a sine wave amounts to approximately 1 to 400. In the extreme case the time can be completely eliminated (ratio 1 to ∞).

The Author (*in reply*): The device you describe is similar in that it will slow down a signal. I feel, however, that your circuit is more analogous to the sampling oscilloscope.

Dr. M. I. Skolnik: How does this compare with a sampling oscilloscope?

The Author (*in reply*): The sampling oscilloscope is another method of slowing down a fast waveform for display. However, the sampling scope requires an exactly repetitive waveform for proper operation. In contrast to this, the 'Stretch' technique will operate properly on a single transient.

Mr. P. Bradsell: Would the author please describe the form and characteristics of the dispersive networks used, particularly in view of the large bandwidths involved?

The Author (*in reply*): In the experimental equipment a simple 150-foot run of RG91U waveguide was used for the wide band dispersion. Of course, the dispersion obtained over the 500 Mc/s band is non-linear in uncompensated waveguide but this non-linearity was corrected by matching the non-linearity with the sweep.

Dr. G. R. Whitfield: What was the width of the 'window' used in the laboratory time transformation system with 500-Mc/s bandwidth?

The Author (*in reply*): The window width was 20 nanoseconds.

Mr. N. S. Nicholls: Is the duration of sweep of the local oscillator selected to be longer than the differential delay of the first dispersive network by less or more than the 'window'?

The Author (*in reply*): The duration of sweep of the local oscillator is equal to the differential delay of the first dispersive network plus the window.

Dr. M. J. Levin: Does the device operate on video signals or does it operate on the envelope only of a band-pass signal?

The Author (*in reply*): In present configurations the device operates on the envelope of an i.f. carrier. Operation on video signals requires the use of video dispersive devices.

Mr. J. Ponsonby: Can one reverse the process to produce very high-speed precision waveforms?

The Author (*in reply*): Time speed-up is obtained by using the sum output of the mixer in Fig. 1. This configuration can be used to produce high speed replicas of narrow-band generated waveforms, subject of course, to the restriction that the waveform to be operated on must fall within a specified window. Thus, for example, a single high speed saw-tooth waveform repeating at a duty cycle equal to the time compression ratio might be generated. Generation of a continuous train of saw-tooth waveforms requires multiple channels.

† G. Vogt, "An analogue polarization follower for measuring the Faraday rotation of satellite signals", *The Radio and Electronic Engineer*, **28**, No. 4, pp. 269–78, October 1964.

Dr. J. M. Blythe: I would like to mention that P. S. Brandon of the Marconi Company has proposed the use of a similar time-compression arrangement with a radar to enable the returns from multiple beams to be displayed on a common p.p.i. or sent over a common link.‡

The Author (*in reply*): I have examined very carefully the referenced patents and I find that no mention is made of time transformation or signal speed-up. These three patents, dated 1953 through 1956, describe the basic operation and two applications of the compressive spectrum analyser (one that uses the conventional mixer and swept oscillator but replaces the usual narrow-band filter with a pulse compression network to allow rapid frequency scanning). In the second patent cited the rapid scanning ability is put to use to scan through the output of several independent c.w. radars in a time duration over which it would normally be possible to scan the output of only one such radar. Mr. Blythe is apparently drawing an analogy between the rapid scanning of the output of the c.w. radar, and the time compression of the output of a pulse radar.

Dr. G. O. Young: Can a time reversal of the waveform be expected? I am concerned about anticipation of signals.

The Author (*in reply*): Time reversal is obtained by using a sweep slope greater than that of the input dispersion network. Reversal with no change in time scale is obtained when $\sigma_2 = 2\sigma_1$. Time reversal of a waveform window is accompanied by a constant delay over the whole window, hence no signals are anticipated.

Mr. G. S. Baldock: On reading the summary, I imagined waveform reversal in time to be related to normal pulse compression. In such a scheme a long pulse, linearly frequency modulated, may be compressed by delaying the lower frequency components with respect to the high (e.g. waveguide near cut-off propagating a signal increasing in frequency with time). With correct filter time-constant, virtually all the components appear at about the end of the long pulse.

It seems that if the filter delay were made twice as long as this, a time-reversed long pulse would be produced.

Is this so, and if not, would it lead to the same result?

The Author (*in reply*): It is possible to derive a passive time invariant filter characteristic which will transform a given waveform into another specific waveform. In your example, the specific input waveform is a linear f.m. chirp pulse and networks may be devised to lengthen or shorten this pulse or as you suggest to time reverse the pulse. However, the network which time reverses one specific input pulse will not necessarily operate in the same manner on another input signal. The technique I have described will transform any input signal.

‡ British Patent Nos. 736,602, 736,607, and 749,588.

U.D.C. 517.512: 621.376.56

The Use of Quantizing Techniques in Real Time Fourier Analysis

By

Professor H. S. HEAPS, B.Sc., M.A.†

AND

P. W. WILLCOCK, M.E.‡

Reprinted from the Proceedings of the Symposium on "Signal Processing in Radar and Sonar Directional Systems", held in Birmingham from 6th–9th July 1964.

Summary: The purpose of the present paper is to present a general theory of the effect of quantizing upon digital computation of Fourier transforms. A similar theory applies in relation to correlation functions. The theory is relevant to spectral analysis and to a number of practical problems in signal detection schemes notably directional systems.

Detection of a weak signal upon a noise background may be dependent upon the ability of a system to distinguish between the correlation functions of the signal and the noise background. The correlation functions are functions of both time and space.

If a receiving array in space is steered by means of time delays between appropriate elements and if the response is processed by digital methods then the computation time determines the time for a complete sweep through the directions of interest. Such time may be reduced by very coarse quantizing before digital processing. A measure of the maximum errors arising from such quantization is provided by the theory.

1. Introduction

Calculations of Fourier transforms and correlation functions arise in many physical studies that involve the analysis of a large amount of numerical data. In some instances, particularly in radio astronomy, geophysics, and acoustics, the amount of data that may be analysed is mainly dependent upon the speed at which the computations may be performed. In other instances, for example when dealing with underwater signals for detection purposes, it is desirable to process data as rapidly as possible. Therefore it is desirable to perform computations by electronic devices; such devices include both analogue and digital ones.

Suppose for example that a receiving array consisting of hydrophone elements in space is to be steered by means of linear processing networks connected to each hydrophone. The outputs of the processing networks are summed and applied to a common total output. The directionality of the array may be changed by changing the response functions of the processing networks. An alternative to the use of electrical networks for the processing is to convert the output of each hydrophone into digital form and to process the signals by digital methods.

Consider the use of a highly directive array to study the dependence of ambient ocean noise upon direction of view. As the array is steered to each direction of interest it is necessary to compute either the autocorrelation function of the noise as a function of a time delay or else to find the Fourier transform of the noise as a function of frequency. In either instance a large amount of computation is required for each angle of view.

Digital methods of computation have a number of advantages over analogue methods, particularly if infinitely clipped functions are used. A function is said to be infinitely clipped if its value is replaced by $+1$ whenever it is positive and by -1 whenever it is negative. Multiplication of clipped functions is equivalent to a test for like polarity while addition is equivalent to a count of unit inputs. The logic of comparing $+1$'s and -1's may be performed very rapidly and so rapidly varying signals may be analysed without time scaling.

Various investigators have considered the effect of complete clipping upon signal processing. J. J. Faran and R. Hills[1] and E. J. Riisnaes[2] have examined the effect of clipping the inputs to a cross-correlation computer in the instance that the original input is Gaussian. J. B. Thomas and T. R. Williams[3] have analysed the polarity coincidence correlator for signals accompanied by Gaussian noise. They concluded that, on the basis of maximizing output signal/noise ratio, the polarity coincidence correlator

† Department of Engineering Mathematics, Nova Scotia Technical College; Mathematical Consultant to Naval Research Establishment, Dartmouth, Nova Scotia, Canada.

‡ Formerly at Nova Scotia Technical College; now with New Brunswick Telephone Company, St. John, N.B., Canada.

could be superior to the multiplier correlator for certain types of noise.

The use of clipped functions in the computation of Fourier transforms has been considered by D. Q. Mayne[4] in the instance in which the function could be expanded in a Taylor series valid over a small interval of time.

J. J. Jones[5] has obtained an expression for the autocorrelation function of the output of a polarity coincidence correlator when the input consists of two sinusoidal signals accompanied by Gaussian noise. He also obtained the autocorrelation function of the output after passage through a bandpass filter. J. L. Lawson and G. E. Uhlenbeck[6] have derived a formula for the autocorrelation function of completely clipped Gaussian noise. A formula for the fourth product moment was obtained by J. A. McFadden.[7]

The subject of coincidence of random pulse trains, which is equivalent to the problem of correlation of completely clipped functions, has been examined in some detail. K. S. Miller and R. J. Schwartz[8] examined the problem of coincidence of two pulse trains with both fixed and randomly varying initial phases, and they determined the coincidence-time fraction for the two cases if only coincidences in excess of a certain specified interval were considered. They extended the results to coincidence of any number of pulse trains. Miller and Schwartz used the theory of linear congruences whereas H. O. Friedman[9] considered similar problems by using probability theory. Friedman also derived an upper bound for the error caused by using probability theory. His solutions were restricted to pulse trains whose fundamental periods were commensurable.

G. M. Weiss[10] obtained an expression for the coincidence fraction of a number of pulse trains by using probability theory. His results were extended by F. Brooks and N. D. Diamantides[11] who obtained an expression for the autocorrelation function of a random pulse train. One of the present authors[12] has derived an upper bound for the error that may result in computation of Fourier transforms and correlation functions by use of clipped functions.

A measure of the error introduced into a Fourier transform by completely clipping both the function and the cosine factor is described in Section 2 and the theoretical limits of this error are stated. Some particular functions are described and shown to correspond to significantly less error than the theoretical maximum. The result of clipping the function but not the cosine factor is described in Section 3. Quantizing is discussed in Section 3 and it is suggested how the choice of quantizing levels should be made.

2. A Measure of the Effect of Infinite Clipping

The Fourier cosine transform of a function $\mathrm{f}(t)$ is the function $\mathrm{F}(\omega)$ defined by

$$\mathrm{F}(\omega) = \int_{-\infty}^{\infty} \mathrm{f}(t) \cos \omega t \, \mathrm{d}t \qquad \ldots\ldots(1)$$

It will be supposed that $\mathrm{f}(t)$ is real, and that the integral defining $\mathrm{F}(\omega)$ exists and may be approximated by a finite sum of the form

$$\mathrm{F}(\omega) = \Delta \sum_n \mathrm{f}_n c_n(\omega) \qquad \ldots\ldots(2)$$

where $c_n = \cos n\omega\Delta$ and $\mathrm{f}_n = \mathrm{f}(n\Delta)$. The time origin, and an integer N, may be chosen so that the terms in the summation (2) are negligible except when $0 < n \angle N$. Let $N\Delta$ be denoted by T.

The infinitely clipped versions f'_n of f_n are defined by the relation

$$\mathrm{f}'_n = \begin{cases} +1 & \text{if } \mathrm{f}_n > 0 \\ -1 & \text{if } \mathrm{f}_n < 0 \\ \pm 1 & \text{if } \mathrm{f}_n = 0 \end{cases} \qquad \ldots\ldots(3)$$

The reason for defining f'_n to be indeterminate as ± 1 when $\mathrm{f}_n = 0$ is that if the computation is performed by an electronic device, or if there are accidental errors in the recording process, then there will be noise present and the noise will determine whether a zero value of f_n is recorded as positive or negative. Thus the definition (3) allows for the presence of an infinitesimal amount of noise.

An approximation to the Fourier transform is given by the quantity

$$\mathrm{F}'(\omega) = \Delta \sum \mathrm{f}'_n c'_n(\omega) \qquad \ldots\ldots(4)$$

in which both the function and the cosine term are infinitely clipped. In computation of (4) each multiplication simplifies to a test for like polarity while the summation is equivalent to a count of unit inputs. For storage of each sample of f_n and c_n just one binary bit is required.

In order to compare the graphs of $\mathrm{F}(\omega)$ and $\mathrm{F}'(\omega)$ it proves convenient to multiply $\mathrm{F}'(\omega)$ by a factor k where

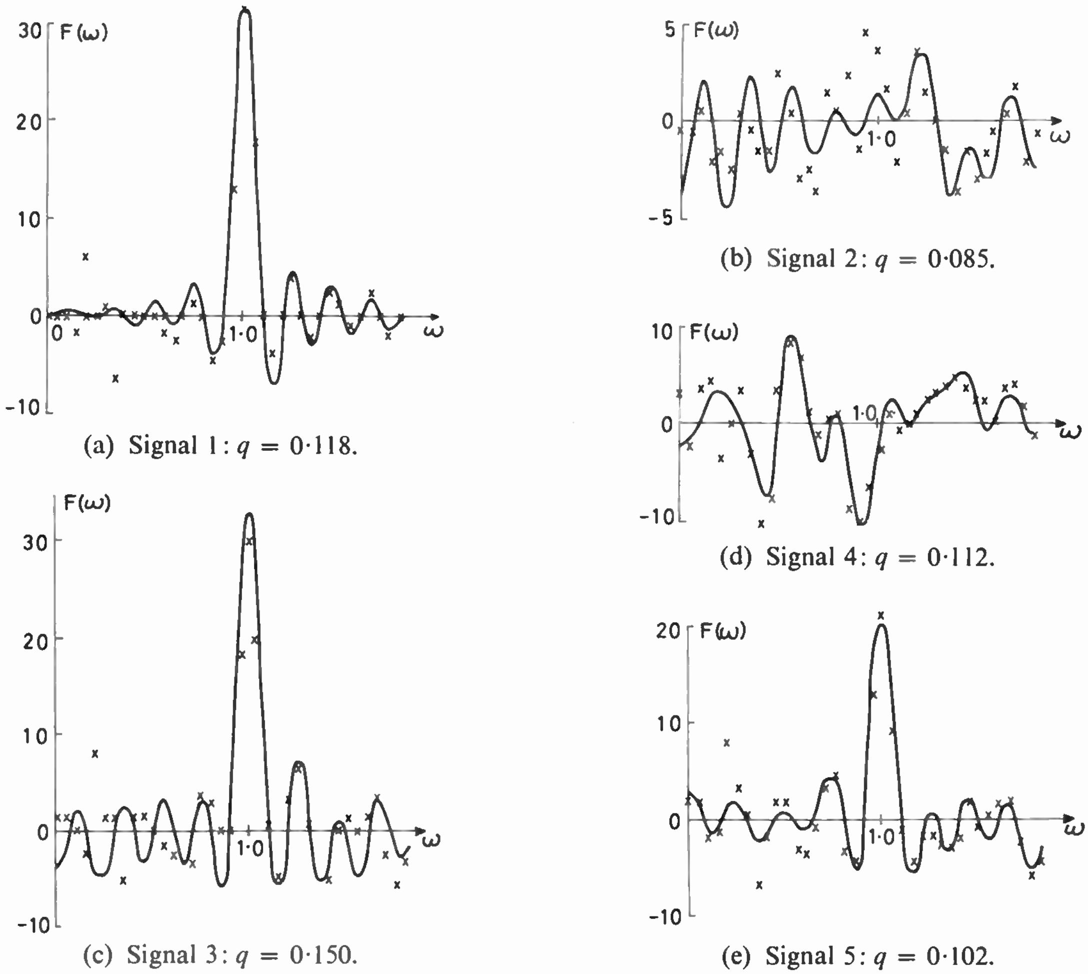
$$k^2 = (1/2N) \sum \mathrm{f}_n^2 \qquad \ldots\ldots(5)$$

Such choice of k makes the areas under the curves $\mathrm{F}(\omega)^2$ and $k^2\mathrm{F}'(\omega)^2$ approximately equal.

For any values of ω the difference between $\mathrm{F}(\omega)$ and $k\mathrm{F}'(\omega)$ is dependent upon the scale of the f_n. However the quantity

$$Q'(\omega) = \frac{\sum \mathrm{f}_n c_n}{(\sum \mathrm{f}_n^2)^{\frac{1}{2}}} - \frac{\sum \mathrm{f}'_n c'_n}{(2N)^{\frac{1}{2}}} \qquad \ldots\ldots(6)$$

is not dependent on the scale of the f_n. The expression $\Delta^{\frac{1}{2}}Q'(\omega)$ is the difference between the two sampled data approximations to



(a) Signal 1: $q = 0{\cdot}118$.

(b) Signal 2: $q = 0{\cdot}085$.

(c) Signal 3: $q = 0{\cdot}150$.

(d) Signal 4: $q = 0{\cdot}112$.

(e) Signal 5: $q = 0{\cdot}102$.

Fig. 1. Graph of $F(\omega)$ for signal 1 computed by eqn. (2). The crosses indicate values of $F'(\omega)$ according to eqn. (4).

$$\frac{\int f(t)\cos\omega t\,dt}{\left[\int f(t)^2\,dt\right]^{\frac{1}{2}}}$$

and provides a convenient measure of the effect of complete clipping.

If $f(t)$ is a piecewise continuous function of t then as N tends to infinity, with T fixed, the functions $\sum f_n c_n$ and $\sum f'_n c'_n$ are asymptotic to N while $(\sum f_n^2)^{\frac{1}{2}}$ is asymptotic to $N^{\frac{1}{2}}$. Thus as N tends to infinity then

$$\Delta^{\frac{1}{2}} Q'(\omega) \to q T^{\frac{1}{2}} \qquad \text{......(7)}$$

where q is a constant which depends upon the function $f(t)$ but not upon N.

It is important to note that $Q'(\omega)$ as defined by eqn. (6) represents the error caused by quantizing but does not indicate the extent of any additional errors introduced by time sampling. The quantity $Q'(\omega)$ forms a convenient measure of the difference between the normalized Fourier transforms (2) and (4) but does not indicate the error caused by using (2) to approximate (1). This latter error may clearly be made as small as desired by sufficiently large choice of N and T. However, in view of the relation (7) it is inadvisable to increase T beyond the minimum required to ensure that (2) is a sufficiently good approximation to (1).

In a previous paper[12] it was shown that for all functions $f(t)$ the value of q satisfies the relation

$$q \leq 0{\cdot}882 \qquad \text{......(8)}$$

and that $q = 0{\cdot}882$ for a particular choice of a function $f(t)$ which is equal to zero over 88% of the range T.

In order to indicate the practical significance of the result (8) the functions $F(\omega)$ and $F'(\omega)$ have been computed for several signals of practical interest and the maximum q has been listed for each example. The signals chosen are defined over the range $-10\pi < t < 10\pi$ as follows:

Signal 1: $f(t) = \cos t$.

Signal 2: $f(t) = n(t)$ = random noise with uniform probability distribution of amplitude between -1 and $+1$.

Signal 3: $f(t) = \cos t + n(t)$.

Signal 4: $f(t) = \cos [t + 2\pi n(t)]$.

Signal 5: $f(t) = \cos [t + \frac{1}{2}\pi n(t)]$.

Throughout the computations the sampling interval Δ was chosen as $\pi/10$ so that $N = 200$.

The graphs of $F(\omega)$ and $F'(\omega)$ for each of the above signals are shown in Fig. 1(a)–(e). Over the range of ω considered the maximum values of q are as shown.

For each signal the maximum value of q is significantly less than the value according to (8). This is not surprising since the function which leads to the maximum value of 0·882 contains many points on the axis and the $F'(\omega)$ is determined by the worst possible choice of the f_n at such points.

3. Effect of Clipping f(t) Only

Instead of the approximation (4) the following may be used:

$$F''(\omega) = \Delta \sum f_n' c_n(\omega) \qquad \ldots\ldots(9)$$

in which $c_n(\omega)$ is not clipped. In computation of (9) each multiplication is equivalent to an addition or subtraction. Since accurate multiplication is generally a slower operation than addition it follows that in many instances computation of (9) instead of (2) involves considerable reduction in computation time. It might be expected that $F''(\omega)$ would approximate $F(\omega)$ more accurately than does $F'(\omega)$.

The function $F''(\omega)$ may be multiplied by a factor k_1 chosen so that the areas under the curves $F(\omega)^2$ and $k_1^2 F''(\omega)^2$ are equal. The required value of k_1^2 is

$$k_1^2 = (1/N) \sum f_n^2 \qquad \ldots\ldots(10)$$

and the quantity corresponding to $Q'(\omega)$ is then

$$Q''(\omega) = \frac{\sum f_n c_n}{(\sum f_n^2)^{\frac{1}{2}}} - \frac{\sum f_n' c_n}{N^{\frac{1}{2}}} \qquad \ldots\ldots(11)$$

In order to simplify the expressions for certain integrals which occur subsequently it will be supposed that T is a whole multiple of $\pi/2\omega$. Such a restriction on T is not a serious one since a necessary condition for existence of the integral (1) is that $f(t)$ tends to zero as $t \to \pm\infty$ and hence provided T is chosen sufficiently large it may be increased to become a multiple of $\pi/2\omega$ while producing negligible change in (2).

For fixed ω in expression (11) $Q''(\omega)$ may be regarded as a function of the N quantities $f_1, f_2, \ldots, f_N$. Then Q'' is a continuous function of $f_1, f_2, \ldots, f_N$ except for discontinuities whenever an f_n is equal to zero since f_n' then has a discontinuity of amount two. As the f_n are varied the greatest value of Q'' occurs either at a discontinuity or else at a point where Q'' is stationary with respect to $f_1, f_2, \ldots, f_N$.

Consider the value of $Q''(\omega)$ at a point $(f_1, f_2, \ldots, f_N)$ at which the f_n satisfy the relations

$$f_r = 0 \text{ for } M \text{ values of } r \qquad \ldots\ldots(12)$$

but at which the remaining f_n's, denoted by f_s, are not zero. Then

$$Q''(\omega) = \frac{\sum f_s c_s}{(\sum f_s^2)^{\frac{1}{2}}} - \frac{\sum f_r' c_r}{N^{\frac{1}{2}}} - \frac{\sum f_s' c_s}{N^{\frac{1}{2}}} \qquad \ldots\ldots(13)$$

For the fixed zero values of f_r, and a fixed value of ω, the maximum value of $Q''(\omega)$ occurs when the f_s are chosen so that for each f_s

$$0 = \frac{\partial Q''}{\partial f_s} = c_s(\sum f_s^2)^{-\frac{1}{2}} - f_s(\sum f_s^2)^{-\frac{3}{2}}(\sum f_s c_s)$$

Thus

$$\frac{f_s}{c_s} = \frac{\sum {}_s^2}{\sum f_s c_s}$$

and hence

$$f_s = \lambda c_s$$

where λ is a constant independent of f_s. The resulting stationary value of Q'' is

$$Q'' = (\sum c_s^2)^{\frac{1}{2}} - \frac{\sum |c_s|}{N^{\frac{1}{2}}} - \frac{\sum f_r' c_r}{N^{\frac{1}{2}}}$$

$$\leq (\sum c_s^2)^{\frac{1}{2}} - \frac{\sum |c_s|}{N^{\frac{1}{2}}} + \frac{\sum |c_r|}{N^{\frac{1}{2}}}$$

$$= (\sum c_s^2)^{\frac{1}{2}} + \frac{2\sum |c_r|}{N^{\frac{1}{2}}} - \frac{\sum |c_n|}{N^{\frac{1}{2}}} \qquad \ldots\ldots(14)$$

Detailed examination of eqn. (14) shows that $Q''(\omega)$ assumes its maximum value of $0{\cdot}761 N^{\frac{1}{2}}$ if $\sum c_s^2$ contains the largest $N-M$ of the $|c_s|$ and if $M/N = 0{\cdot}94$. Thus

$$\Delta^{\frac{1}{2}} Q''(\omega) \leq 0{\cdot}761\, T^{\frac{1}{2}}$$

and the relation (8) is replaced by

$$q \leq 0{\cdot}761 \qquad \ldots\ldots(15)$$

Comparison of the results (8) and (15) suggests that when $f(t)$ is clipped there is little to be gained by preserving the unclipped values of $\cos \omega t$.

4. Two-level Quantization

The process of two-level quantizing may be described as follows. A signal $f(t)$ is examined with respect to an amplitude h. If a sample of $f(t)$ exceeds h then the sample is recorded as having a value a. If the sample lies between zero and h its value is recorded as ca where $c < 1$. Thus the recorded values of positive samples are quantized at two levels. Negative values of f(t) are similarly quantized at $-a$ and $-ca$. Thus the quantized samples f'_n are defined by the relation

$$f'_n = \begin{cases} a & \text{if} \quad f_n > h \\ ca & \text{if} \quad 0 < f_n < h \\ -ca & \text{if} \quad -h < f_n < 0 \\ -a & \text{if} \quad f_n < -h \end{cases} \qquad \ldots\ldots(16)$$

The Fourier transform of f(t) is then approximated by the function

$$F_2(\omega) = \sum f'_n c_n(\omega) \qquad \ldots\ldots(17)$$

The area under the curves $F(\omega)^2$ and $k^2F_2(\omega)^2$ are equal if k is chosen so that

$$k^2 = \frac{\sum f_n^2}{\sum f_n'^2} = \frac{\sum f_n^2}{(N_1c^2+N_2)a^2} \qquad \ldots\ldots(18)$$

where N_1 is the number of samples of f_n that are numerically less than h while N_2 is the number numerically greater than h. A measure of the difference between the normalized transforms is then provided by

$$Q_2(\omega) = \frac{\sum f_n c_n}{(\sum f_n^2)^{\frac{1}{2}}} - \frac{\sum f'_n c_n}{(N_1c^2+N_2)^{\frac{1}{2}}a} \qquad \ldots\ldots(19)$$

The value of $Q_2(\omega)$ depends not only upon the f_n and ω but also upon the values chosen for h and c.

The Fourier transform F(ω) is a measure of the spectral content of f(t) at the frequency ω, and if f(t) contains a strong component of frequency ω it is desirable that $Q_2(\omega)$ should be small. It therefore appears reasonable to choose h and c to minimize $Q_2(\omega)$ in the instance that f(t) = cos ωt.

Now if f(t) = cos ωt, and T is a multiple of $2\pi/\omega$, it may easily be shown that $Q_2(\omega)$ is independent of ω and has the value

$$Q_2 = N^{\frac{1}{2}}\left[2^{-\frac{1}{2}} - \frac{2}{\pi}\frac{c+(1-c)\cos(x\pi/2)}{(c^2x+1-x)^{\frac{1}{2}}}\right] \ldots(20)$$

where x denotes N_1/N. For various values of x the value of c has been chosen to minimize Q_2, the required values of c and the resulting values of Q_2 being as listed in Table 1. The minimum value of Q_2 is 0·015 and occurs when x = 0·402 and c = 0·354.

The results of Table 1 may be applied to a practical computation in the following manner. Suppose a value of h is chosen and N values of f_n are examined with respect to the level h. If xN values are found to be numerically less than h then the appropriate value of c may be read from Table 1 and used to determine the quantizing levels in (16). The value of a is of no importance in computation of a normalized transform. If the value of x differs from 0·4 then a further N samples may be examined with respect to a new value of h chosen to lead a value of x closer to 0·4. Repetition of such a procedure determines a level h for which the quantized transform is most accurate in estimating the extent of strong frequency components.

Table 1

Values of c to minimize q when f(t) is a cosine wave

$x(= N_1/N)$	c	$Q_2/N^{\frac{1}{2}}(= q)$
0	0	0·070
0·1	0·112	·044
·2	·210	·026
·3	·285	·017
·4	·354	·015
·5	·414	·018
·6	·467	·025
·7	·516	·036
·8	·559	·047
·9	·599	·060
1·0	·637	·070

Although x and c may be chosen as above to minimize Q_2 it is of interest to determine the maximum value of $Q_2(\omega)$ that could arise for any signal f_n. However, if $Q_2(\omega)$ is regarded as a function of the f_n and examined as in Section 3 it is found that the f_n may be chosen so that $\Delta^{\frac{1}{2}}Q_2(\omega) = 0{\cdot}761T^{\frac{1}{2}}$ and so there is no improvement over the result expressed in eqn. (15).

5. n-Level Quantization

Further analysis similar to that described in Sections 3 and 4 has led to the following conclusions.

Regardless of the number of quantizing levels chosen in the computation of a Fourier transform by means of eqn. (17) it is always possible to choose a signal f(t) for which q = 0·761. Such a signal however is somewhat artificial in that it is zero over most of its range.

Let $F_n(\omega)$ denote the n-level quantized extension of eqn. (17), the signal f(t) being examined with respect to amplitudes $\pm h_1, \pm h_2, \ldots, \pm h_{n-1}$. Let $N_1, N_2, \ldots$ be the number of samples whose numerical values lie in the ranges 0 to h_1, h_1 to h_2, etc., and let $c_1a, c_2a, \ldots$ be the quantized levels. In order to make $F_n(\omega)$ best approximate F(ω) at the dominant frequencies the levels should be chosen as indicated below:

Infinite clipping: q_1 = 0·092

Two-level: q_2 = 0·015, N_1/N = 0·40, c_1 = 0·35

Three-level: q_3 = 0·0064
N_1/N = 0·25, N_2/N = 0·30
c_1 = 0·21, c_2 = 0·63

Four-level: $q_4 = 0{\cdot}0035$

$N_1/N = 0{\cdot}185, \quad N_2/N = 0{\cdot}194, \quad N_3/N = 0{\cdot}214$

$c_1 = 0{\cdot}16, \quad c_2 = 0{\cdot}46, \quad c_3 = 0{\cdot}74$

Eight-level: $q_8 = 0{\cdot}0010$

$N_i/N = 0{\cdot}10, 0{\cdot}10, 0{\cdot}10, 0{\cdot}10, 0{\cdot}10, 0{\cdot}15, 0{\cdot}15$

$c_i = 0{\cdot}065, \ 0{\cdot}219, \ 0{\cdot}369, \ 0{\cdot}505, \ 0{\cdot}633, \ 0{\cdot}771, \ 0{\cdot}893$

Numerical computation of Fourier transforms of many quantized functions has led to the conclusion that the above values of q_2 to q_8 are, in general, appropriate for estimation of the error. Signals that lead to larger values of q are ones that are zero for most of the range.

6. References

1. J. J. Faran and R. Hills, "Correlators for Signal Reception", Harvard University Research Laboratory Technical Memo No. 27, September 1952.
2. E. J. Riisnaes, "The Effect of Clipping and Sampling on the Performance of Some Signal Detection Systems", Admiralty Research Laboratory, ARL/N30/L, December 1958.
3. J. B. Thomas and T. R. Williams, "On the detection of signals in non-stationary noise by product arrays", *J. Acoust. Soc. Amer.*, **31**, p. 453, 1959.
4. D. Q. Mayne, "The evaluation of the frequency spectrum of a waveform", *Proc. Instn Elect. Engrs*, **110**, p. 390, 1963.
5. J. J. Jones, "Hard limiting of two signals in random noise", *I.R.E. Trans. on Information Theory*, **IT-9**, p. 34, 1963.
6. J. L. Lawson and G. E. Uhlenbeck, "Threshold Signals". (McGraw-Hill, New York, 1950.)
7. J. A. McFadden, "The fourth product moment of infinitely clipped noise", *I.R.E. Trans. on Information Theory*, **IT-4**, p. 159, 1958.
8. K. S. Miller and R. J. Schwartz, "On the interference of pulse trains", *J. Appl. Phys.*, **24**, p. 1032, 1953.
9. H. O. Friedman, "Coincidence of pulse trains", *J. Appl. Phys.*, **25**, p. 1001, 1954.
10. G. M. Weiss, "A note on the coincidence of some random functions", *Quart. Appl. Maths*, **14**, p. 103, 1956.
11. F. Brooks and N. D. Diamantides, "A probability theorem for random two-valued functions, with applications to autocorrelation", *Soc. Industr. Appl. Maths Rev.*, **5**, p. 33, 1963.
12. H. S. Heaps, "Maximum error caused by using completely clipped functions in the computation of Fourier transforms and correlation functions", *Quart. Appl. Maths*, **19**, p. 321, 1962.

Manuscript received by the Institution on 4th June 1964. (Paper No. 965.)

DISCUSSION

Under the chairmanship of Mr. W. K. Grimley, O.B.E.

Mr. C. van Schooneveld: With regard to the clipping of f(t), suppose f(t) = f$_1$(t) + f$_2$(t). Then, in the ideal case, the spectrum of f(t) equals the sum of the spectra of f$_1$(t) and f$_2$(t). This will no longer be true in the proposed non-linear method. Do we still recognize in the computed spectrum the individual contributions? In other words, how severe is the effect of the mutual interference between f$_1$(t) and f$_2$(t)?

The Authors (*in reply*): Since the spectrum of an infinitely clipped function depends only upon the statistics of the axis crossings then the spectra of clipped versions of f$_1$(t) and f$_1$(t) + f$_2$(t) differ significantly only when f$_2$(t) is sufficiently large to affect the statistical distribution of the crossings. In general, if f$_2$(t) is small in amplitude in comparison to f$_1$(t) then its spectrum will have little effect on that of the clipped sum of f$_1$(t) and f_2(t). When f$_1$(t) and f$_2$(t) are of comparable amplitude the spectrum of the clipped sum f$_1$(t) + f$_2$(t) may be very different from the spectrum of the true sum.

U.D.C. 621.391.823

The Mechanism of Interference Pick-up in Cables and Electronic Equipment with special reference to Nuclear Power Stations

By

D. HARRISON†

Presented at a Symposium on "The Operation of Electronic Equipment under Conditions of Severe Electrical Interference" in London on 15th April, 1964.

Summary: The various forms of electrical interference in nucleonic reactor control equipment have been recently studied in the U.K.A.E.A. The results of this work are discussed and an analysis is made of interference due to interconnecting cables. Other causes of interference are mentioned and a method of *in situ* testing is described which can be used to diagnose the sources of interference pick-up and to establish the interference sensitivity of a completed installation.

List of Symbols not Defined in the Text

C capacitance per unit length

i r.m.s. value of disturbing current

L_c that part of a cable conductor inductance which is coupled to its screen

L_E effective earth inductance, per unit length

L_S inductance of cable screen, per unit length

l total cable length

M mutual inductance per unit length of cable

R_c resistance of centre conductor, per unit length

R_0 d.c. resistance of screen, per unit length

R_s resistive part of the screen transfer impedance

t screen thickness

Z_E effective earth impedance in cable/earth system, per unit length

Z_0 characteristic impedance of screen to conductor

Z_{SE} characteristic impedance of screen to earth

Z_{iT} screen transfer impedance, per unit length, of the ith screen

β transmission line wavelength constant

l_c inductance between centre conductor and screen, per unit length

x short section of cable

ρ screen resistivity

μ screen permeability

1. Introduction

Electrical interference is frequently a source of trouble in nucleonic equipment for reactor control or nuclear experiments. This is not surprising since the equipment often has to detect continuous currents of less than 10^{-12} A or pulses less than 100 ns and a few hundred microvolts amplitude on sites where power circuits may carry kilovolts and kiloamperes. Electrical interference can cause spurious instrument readings, alarms or trips and sometimes may inhibit protection circuits.

Electrical interference has been widely studied in the past[1] but nucleonic equipment has not received much special attention. This paper discusses the results of a recent U.K.A.E.A. investigation which was aimed at gaining a better understanding of this difficult subject by experiments on full size models and measurements on reactor sites.

2. Interference Coupling Paths

The electrical interference situation can be divided conveniently into three parts:

(*a*) a disturbing circuit,

(*b*) a coupling path,

(*c*) a disturbed circuit.

It should be noted that remedial measures do not usually affect circuit operation when applied to the coupling path (screening, for example) but may do so when applied to the circuits themselves (spark quenching, for example).

This paper is primarily concerned with the coupling paths between circuits and these may be classified as follows:

(1) Stray coupling—inductive, capacitive and resistive combinations.

(2) Electromagnetic wave coupling.

(3) Real coupling elements—inductor, capacitor and resistor combinations.

† Control and Instrumentation Division, Atomic Energy Establishment, Winfrith, Dorchester, Dorset.

(4) A number of mixed paths and sources of interference such as
thermo-electric contact potentials
electrochemical paths and potentials
common ionization paths
electromechanical paths (piezo-electricity, mechanical charge re-distribution and all forms of microphony, for example).

The distinction between (1), (2) and (3) is one of convenience rather than real, and the treatment which follows is based on lumped or distributed parameters so as to avoid the difficult boundary conditions of the more basic field theory approach. Interference from class (4) is rather special and is not considered further in this paper.

In practice direct stray resistive couplings rarely occur in good system designs and propagated e.m. waves usually give less energy transfer than do the associated local magnetic and electric fields.† In any case, screening techniques for local ('induction') fields are also effective against the propagated e.m. wave of the same frequency.

In relating the coupling forms (1) (2) and (3) above to practical circuits it is convenient to consider the disturbed circuit as enclosed within a surface, outside which the disturbing circuit lies. This surface can most conveniently be taken as the screen which usually surrounds instruments and cables and all interference effects can then be described in terms of electric currents which penetrate this screen. A perfectly conducting screen, in which there are no penetrations, completely protects circuits within from outside electrical disturbances. Practical screens are not perfect in this way.

It is convenient to distinguish between coupling paths which penetrate the screens of electronic units and those which penetrate the screens of interconnecting cables. The former often arise through stray and 'real' coupling elements (power supply conductors, for example) but the problem is local and under control of the circuit designer. In such cases cures are usually simple despite the complex geometry of the situation.

On the other hand, coupling which penetrates the instrument cables is related to the whole installation and is peculiar to it, so that cures are not always obvious or easy although the geometry of the problem is often quite simple.

The overall 'cable induced' forms of interference seem to be the least well understood and are the main subject of this paper. Local interference in electronic units is, however, briefly examined in Section 4.

3. Interference due to Interconnecting Cables

At an early stage in the study it became obvious that most good nucleonic cable installations have very little direct coupling with disturbing circuits. This is because cables which can cause disturbance are nearly always well separated from nucleonic cables, crossings are made at right angles and a great deal of intentional and accidental screening occurs by virtue of interposed earthed plant structures, metal ducting or specially provided screens. On the other hand cables are nearly always tightly coupled to the local ground since they run in close proximity with it, often with a coaxial configuration, as when surrounded by metal ducts or in cable tunnels. Measurements on test models and on reactor sites confirmed that, except in cases where long close parallel runs occur between disturbed and disturbing cables, the measured interference is consistent with coupling to local earth disturbances.‡

The problem of analysis may therefore be divided into the following parts:

(1) The relationship between disturbing circuit currents and the resulting earth disturbance.

(2) The relationship between earth structure disturbances and the resulting cable screen currents.

(3) The relationship between cable screen currents and interference developed at the cable entry to the instrument.

(4) The relationship between the interference developed at instrument cable entries and the resulting unwanted effects in the equipment.

If these relationships are expressed in terms of frequency (the transfer function in ω, for example) the overall behaviour is then the product of the separate parts. The analysis may of course, be extended to include a multiplicity of screens between (3) and (4) above.

3.1. *The Relationship between Interference developed at Instrument Cable Entries and the Resulting Unwanted Effects in the Equipment*

These relationships are a function of the individual items of equipment, and the user must be familiar with them before he can attempt analysis of interference troubles. Although instruments will obviously respond to interference at input terminals if it falls in the signal pass band, there may also be responses well outside this band. These may be due to over-

† In a practical situation, if the energy level is sufficient to cause interference the distance from the source will nearly always be less than $\lambda/2\pi$ (λ is the wavelength of the interfering signals) which, for free space, makes the local field the greater.

‡ Where this is not true, the principles which follow still apply but the analysis has to be performed between each disturbing system with which direct coupling occurs.

loading, rectification (particularly in logarithmic channels), or simply because the instrument has a number of spurious pass bands (as, for example, in chopper amplifiers).

Connections other than signal inputs may have unintended internal coupling to signal circuits and can produce interference induced by the cables to which they are connected so that the designer should know the response of all points where connections are made to the instrument. External connections to meters, recorders, relays and power supplies cannot be ignored. It is usually a simple matter to calculate or measure response to interference at the instrument cable entries in terms of the acceptable input level over the frequency range of interest.

3.2. *The Relationship between Cable Screen Currents and the Interference developed at the Instrument Cable Entry*

We are here concerned with the mechanism of interference resulting from screen currents in interconnecting cables; a later section will discuss how these currents originate. For simplicity we consider coaxial cables although the conclusions may readily be extended to other cable types, including multi-core or multi-screen cables.

Figure 1 is a simplified lumped parameter representation of a short section of coaxial cable. It can be assumed, with negligible error, that currents induced in the centre conductor are small compared with the screen currents which induce them, so that the conductor/screen circuit does not load the screen/earth circuit. Since, for a coaxial cable, $M = L_s = L_c$ the only potential difference tending to produce currents in the centre conductor is that due to $iR_s\,\delta x$.† As a result a current is induced in the centre conductor of value

$$\frac{iR_s\,\delta x}{(Z_1+Z_2)}$$

since the effects of $C/2.\delta x$ and $R_c\,\delta x$ may be ignored in relation to the impedances Z_1 and Z_2.‡ When the cable as a whole is considered the effect of the current induced at dx has to be transferred to the cable terminations ($x = 0$) and the superposition theorem permits us to integrate for all x between $x = 0$ and $x = l$.

When the cable length approaches or exceeds a quarter wavelength the screen current distribution along the cable length is not uniform and resonances can occur in both the screen/earth system and the centre conductor/screen system. The wavelength constant, β, for both screen/earth and centre conductor/screen has about the same value so that the lower resonances occur at about the same frequencies; at these frequencies sensitivity to interference may be raised by 30 dB or more.

In this section we are concerned only with the centre conductor/screen system.

Assuming constant screen current and one value for β we can perform the integration over the length of cable for two extreme cases of cable termination—the case of the properly terminated cable and the case of open circuit termination at one end and short circuit termination at the other.

In the first case we find the received interfering current is given by

$$i_{\text{int}} = \frac{iR_s l}{2Z_0}\left[\frac{\sin(2\beta l+\delta\phi)-\sin\phi}{4\beta l}+\frac{\cos\phi}{2}+ + j\left(\frac{\cos(2\beta l+\phi)-\cos\phi}{4\beta l}+\frac{\sin\phi}{2}\right)\right] \quad \ldots\ldots(1)$$

where ϕ defines the position of the screen standing wave at the equipment and is commonly, but not always, of value zero. $\phi = \beta l'$ in Fig. 1. In the second case we find

$$i_{\text{int}} = \frac{iR_s l}{2Z_0}\left[\tan\beta l\cos\phi + \sin\phi - \frac{\tan\beta l\sin\phi}{\beta l}\right] \quad \ldots\ldots(2)$$

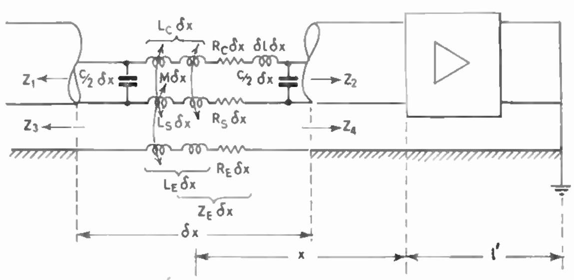


Fig. 1. Representation of a section δx of a cable.

The bracketed parts of these relationships are plotted in Fig. 2, for $\phi = 0$, as a function of βl (which, for constant l, is proportional to frequency). The expressions (1) and (2) describe conditions for a lossless line and in Fig. 2 the measured differences in practical cases are indicated. The bracketed part of these expressions may be regarded as a 'frequency sensitive function' whilst the first part is an amplitude term. The average level of interference may be adjusted by manipulation of the first term and the size or position of the peak responses may be adjusted by manipulation of the bracketed terms.

† In a cable in which $M \neq L_s \neq L_c$ the impedance responsible for this potential difference would include the additional uncoupled impedances of the screen.

‡ This is true except under certain conditions of resonance when this approximation accounts for the differences between the 'lossy' and 'lossless' cables referred to later.

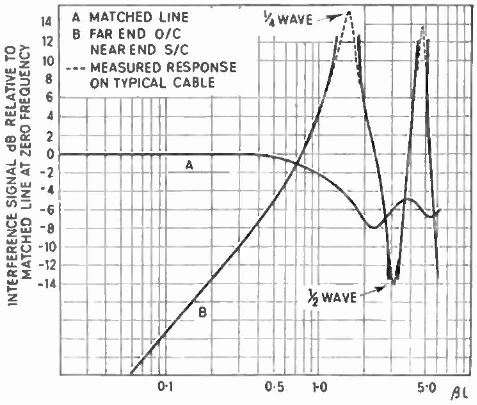


Fig. 2. Response of cable conductor/screen to screen current.

The following points of interest may be noted:

1. Even if the screen currents are of constant amplitude a badly terminated line will become very sensitive to interference at the resonant frequencies; properly terminated lines are not so affected.
2. For the unmatched case considered here the first resonance occurs near the line quarter wavelength frequency. Where the line is short circuit at both ends the first resonance turns out to be at the half wavelength frequency as does the first voltage resonance in a line open circuit at both ends.
3. Interference is reduced by reduction of both R_s and l. Changes in Z_0 will usually result in a proportional change in signal level so that the nuisance value of the interference will not be altered.
4. Reduction of l also has the effect, with badly terminated lines, of reducing the number of resonant peaks in a given frequency band.

In practice the properly terminated line is usually found in connections between units, the open circuit/short circuit line often occurs when virtual earth amplifiers are used with ionization chambers and the open circuit/open circuit line when valve head amplifiers are connected to pulse ionization chambers. When matched input transistor head amplifiers are used in pulse channels an intermediate case arises (open circuit/matched line) which has the characteristics of curve B below $\beta l = 0{\cdot}5$ and is similar to curve A above this.

All these combinations can occur for common mode interference in thermocouple amplifiers depending upon whether the hot junction is grounded or not and on the common mode to ground impedance of the thermocouple amplifier.

In the preceding discussion it has been assumed that R_s is simply the resistive component of the screen impedance. Because of skin effect the situation is more complex than this but it is possible to specify a transfer impedance which may replace R_s in the above expressions. This transfer impedance is the ratio of e.m.f. developed in the cable conductor/screen circuit to the external screen current per unit length of cable.[2, 3, 4] The finite thickness of the screen can be divided into shells, as in Fig. 3, to form an attenuating transmission line through the screen thickness, which, after allowing for the energy reflection at the inner surface of the screen, and letting the shell thickness tend to zero, yields a transfer impedance Z_T given by

$$|Z_T| = \frac{R_0 x}{\sqrt{\cosh x - \cos x}} \qquad \ldots\ldots(3)$$

where $x = 2t\sqrt{\dfrac{\pi\mu_0\mu_r f}{\rho}}$

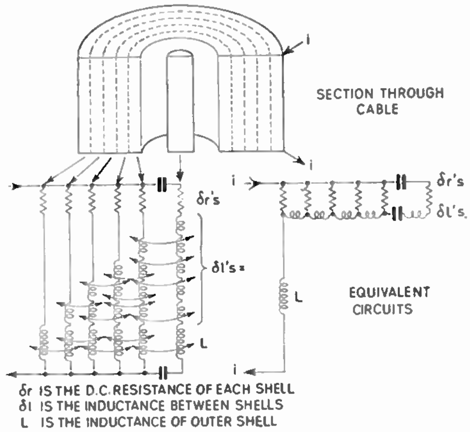


Fig. 3. Current penetration in a cable screen.

In Fig. 4 Z_T is plotted against frequency for a number of typical cables. Desirable features for high screening efficiency can be deduced by observing that R_0 should be low and x high.

We may note particularly the way in which screening efficiency increases with frequency for all but braided cables. In the case of braided cables the effect of the separate screen strands and their lay is to introduce a small inductance on the screen which is not mutually coupled with the centre conductor.[3] At frequencies above those at which this inductance presents an impedance comparable to Z_T for an equivalent solid screen cable the screening efficiency falls since the replacement for R_s in eqns. (1) and (2) is not now Z_T but something of the form $Z_T + j\omega L_x$ where L_x is the additional 'uncoupled' inductance.

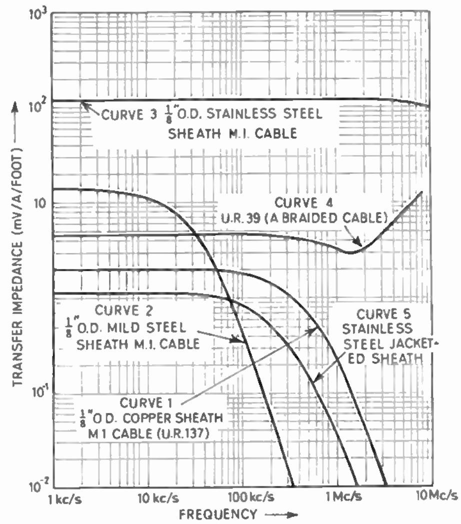


Fig. 4. Interference rejection of coaxial cables (experimental curves).

From the foregoing, if the screen currents are known, it becomes possible to calculate the interference developed in terms of the cable (its length, geometry and materials of construction), the circuit terminations and the instrument response. Precision is rarely justified in such a calculation but considerable accuracy is possible especially if cable losses are included in the calculation.

It is now necessary to examine the causes of interfering currents which traverse the cable screens.

3.3. *The Relationship between Earth Structure Disturbances and the Resulting Cable Screen Currents*

This relationship is similar in principle to that of the cable screen and conductor but is more complex since the geometry is rarely coaxial nor is the spacing uniform along the length of the cable. Furthermore the earth currents which induce the cable screen currents are distributed about the earth system in a complex way which defies detailed analysis.

However, the non-coaxial configuration only means that R_s in Fig. 1 becomes a complex impedance, and the variability of spacing does not, in practice, have the big effect that might be expected unless marked discontinuities occur. These can often be allowed for since they should be obvious from an inspection of the installation. As an example, Fig. 5 shows the measured impedance/frequency characteristic of a screen/earth system at Trawsfynydd Power Station between the reactor and control room in which the resonant frequencies due to a far end short circuit are within a few percent of calculations based on cable length and from which the characteristic impedance can be calculated to lie between 100 and 250 Ω over the whole frequency range. At other sites impedances between about 75 and 250 Ω for screen/earth systems have consistently been measured despite changes in spacing throughout the length of the cable run.

With regard to the unpredictability of the earth currents themselves the greatest coupling will occur when these currents follow the route of the cable screen and, on the basis of this worst case, the simplified lumped parameter network of Fig. 1 may be used.

With regard to the screen termination there are normally only two cases to consider—the case with both ends grounded and the case with one end grounded and the other open circuit.

If we integrate over the cable length for the contribution of the sections, x, we find the interfering current appearing at the equipment input, when matched, is given by

$$i_{\text{int}} = \frac{i_E Z_E R_s l^2}{8 Z_{SE} Z_0}\left[\frac{l^2\beta^2+1}{\beta^2 l^2}\tan\beta l - \frac{1}{\beta l}\right] \quad \ldots\ldots(4)$$

for the grounded/open circuit screen and

$$i_{\text{int}} = \frac{i_E Z_E R_s l^2}{8 Z_{SE} Z_0}\left[\frac{3}{\beta l} + \cos\beta l\right] \quad \ldots\ldots(5)$$

for the grounded/grounded screen.

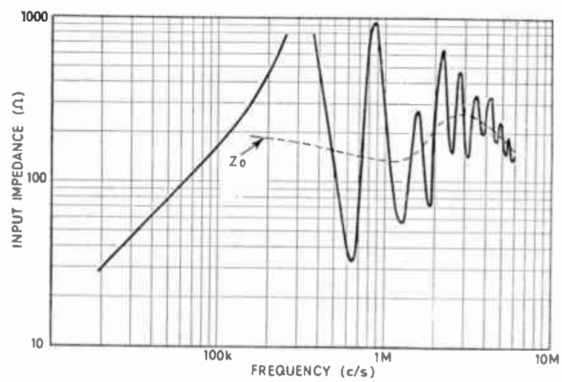


Fig. 5. Input impedance of instrument cable screen (far end s/c) at Trawsfynydd Nuclear Power Station (150 yd).

These expressions assume lossless lines, no loading by screen circuit on the earth circuit or the centre conductor circuit on the screen and equal values of β for screen/earth and conductor/screen transmission lines. A travelling wave is assumed in the earth but an expression of essentially the same form is obtained for standing waves and, in practice, most earth currents have a high standing wave ratio. The expressions in the brackets are plotted in Fig. 6 against βl and a measured result on a cable in the *Zenith* reactor is also given for comparison. Again we have the separation into an 'amplitude term' and a 'frequency function'.

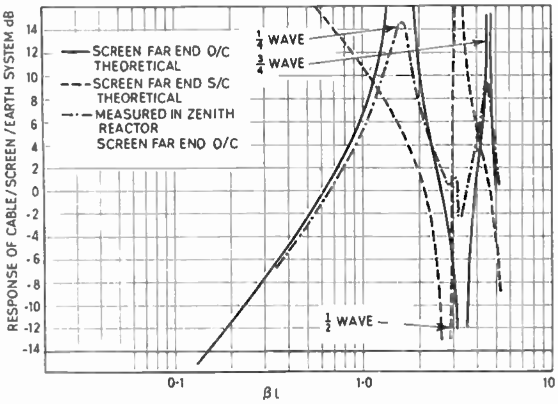


Fig. 6. Response of cable/screen/earth systems.

The following points of interest occur in these expressions:

1. The grounded/grounded screen resonances occur at multiples of the half wavelength with nodes at approximate multiples of the quarter wavelength.
2. The grounded/open circuit screen resonates at multiples of the quarter wavelength with nodes at approximate multiples of the half-wavelength.
3. The grounded/grounded screen is very much more sensitive to low-frequency pick-up than is the other configuration and this emphasizes the importance of single point earthing of screens where there is sensitivity to the lower frequencies—frequencies below which βl is less than unity. $\beta = \omega LC$ and LC is typically about 5×10^{-9} seconds per metre.
4. The effect of increasing cable length is first to cause more resonant peaks to occur in a given frequency bandwidth and also to raise the general level of sensitivity by l^2.
5. The expression shows the desirability of having low values for Z_E and Z_T. One factor which makes Z_E low is high mutual inductive coupling between screen and earth currents, the extreme case of which is the coaxial configuration and this is sometimes approached in duct systems. A high value of Z_{SE} is desirable providing this is consistent with low Z_E. Partly for this reason, a centrally-spaced cable in a duct produces less interference than one which is asymmetrically spaced.

A situation sometimes arises in which, say, a radiation detector forms a significant capacitance to ground so that the cable screen at this end cannot be regarded as open circuit. In this case the first resonance is reduced to the frequency at which the equation $\frac{1}{\omega C'} = Z_0 \tan \beta l$ is satisfied (approximately). C' is the chamber capacitance to ground.

This modification only becomes significant when C' exceeds about one-quarter the total cable capacitance. At higher frequencies the resonances occur at frequencies closer to those of the grounded/grounded screens. At very low frequencies the interference is increased according to the factor

$$\frac{3C' + C_1}{C_1}$$

and both these effects can be important with short cable lengths such as occur on pulse channels with valve head amplifiers.

The above principles may be elaborated to include multiple screens and ducts and these are briefly mentioned in Section 4.

3.4. *The Relationship between the Disturbing Circuit Currents and the Resultant Earth Disturbances*

This relationship is the most difficult to analyse because of the wide variation in the form of the disturbing current and circuit layout. Interference studies at A.E.E. have been mainly concerned with the disturbed circuit; the behaviour of disturbing circuits has not been studied in much detail. Further work is justified on this part of the problem.

We may note, however;

1. The way in which the disturbing circuit affects the earth system is the same in principle as the way in which the disturbed circuit is affected by it, but the direction of energy transfer is reversed.
2. Much more coupling exists between most disturbing circuits and earth than between earth and the instrument channels. This is because disturbing circuits, being relatively insensitive

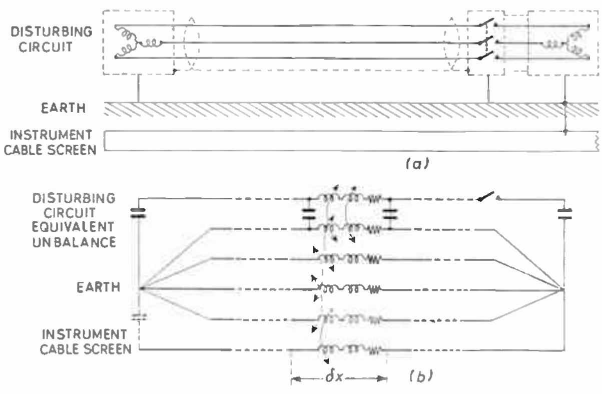


Fig. 7. Representation of a disturbing circuit.

to outside interference, do not force designers to take suppression measures. In particular balancing and screening is usually poor.

3. Much high frequency interference comes from switching or switch-like operations. (Contactors, commutation or welding, for example.)

Figure 7(a) depicts a machine supply on a reactor system. Unbalanced currents will flow on switching on or off because:

(*a*) The breaker contacts do not make or break together.

(*b*) Stray capacitances to ground may be large and highly unbalanced.

(*c*) Screening of cables may be poor, especially where stranded steel armour is used without an inner lead sheath.†

The mechanism of high-frequency earth current disturbance seems to be due to the switching pulse which produces resonances determined by the separation distance between major electrical discontinuities to ground. These resonances appear to fall in the same frequency band as those of the instrument cables, which is to be expected since the lengths involved are often similar. Resonances over long lengths do not seem to occur, probably because of poor energy transfer through the discontinuities.

The equivalent earth current of equations (4) and (5) is only a proportion of this resonant current, but the proportion is difficult to define. The simplified lumped parameter representation of Fig. 7(b) illustrates the problem (for the case where 'direct' coupling is negligible). In a case where the earth return route for the disturbing current follows that of the signal cables, this proportion must still be well below unity and measurements on models suggest that practical worst values do not exceed about 0·1.

The model of Fig. 7 is useful in indicating where improvements can be effected and the calculations for determination of resonances which can be applied to it can be useful from a diagnostic point of view.

4. Interference Suppression Techniques

This paper has been concerned primarily with interference induced by the connection of units together by interconnecting cables. Many other causes of interference have been found during the study, most of which are due to local conditions in or around electronic units or racks. In evaluating a system for interference susceptibility it is most important to correct these local sources of trouble before studying the system as a whole. In this section, therefore, the opportunity is taken, in describing suppression techniques, of mentioning all those other sources of interference which the author believes to be important. The situation is dealt with in the order that might be adopted in the course of analysing a real system.

4.1. *Electronic Units*

A number of conductors, in addition to signal inputs, usually penetrate the chassis screen. The circuit designer must remember that all incoming connections routed in cables will bring in interference and some, such as mains supplies, will even have direct connection to interfering sources. Coupling to the signal circuits may be fairly direct, in the case of signal or test inputs, or may be by means of internal strays. Many of these causes of interference can easily be eliminated by screening or filtering although sometimes it is necessary to mount the filter outside the electronic unit, if the interfering current can be large, as in the case of mains filtering. It is particularly important to filter power supplies at head amplifiers, even if the power source itself is quiet, because of cable induced disturbances. Signal pass-bands should, of course, be no greater than the circuit performance demands and designed signal levels should be as high as possible.

Currents which run in the screen or chassis of an electronic unit can cause trouble if the screening is inefficient. The author considers that all units should be interference free with currents of about 50 mA anywhere on their outside surfaces. Some present day designs are unable to tolerate 500 μA.

Most of the unwanted currents arrive by way of cable screens so it is good practice to ensure that these are well bonded together and can find the rack earth with a minimum of travel through the unit chassis. Louvres and holes in the unit screens can cause considerable trouble as can poor contacts between screening surfaces. In one pulse head amplifier, pick-up from current arriving by way of an input cable screen was reduced by 30 dB merely by cleaning two mating surfaces of the unit housing.

The screen portion of all cable connectors and units should be tight and make good contact. For example, a 100 mΩ contact is equivalent to about 30 ft of braided cable (UR 39) or 300 ft of an equivalent solid copper sheathed cable in terms of interference pick-up at about 1 Mc/s.

A big improvement can be obtained by designing for balanced circuits in the signal paths. Twin cables can be obtained with capacitance balance of about 3% and transformers can be designed for use in the pulse channel frequency range (20 kc/s to about 6 Mc/s)

† In one measurement a 2-in dia. p.v.c. sheathed power cable without a lead sheath gave a transfer impedance at 1 Mc/s 10^5 times greater than an equivalent cable with a lead sheath.

with balance factors of the same order or better. Generated interference, as determined, say, by eqns. (4) and (5) is reduced by the effective circuit balance factor.

When transformers are used with pulse counting radiation detectors a difficulty arises for it becomes almost impossible to achieve a balanced system without the use of a capacitor at the detector end.[5]

4.2. *Electronic Racks*

The racks offer an opportunity of passing cable screen currents to ground before they reach the electronic units. In the author's experience this is a most worthwhile thing to do and requires only the bonding of all incoming cable screens to ground at entry. It is important that all incoming cables are so treated and desirable, too, to provide power supply input filters at the same point so as to avoid large circulating currents within the rack volume.

Some units, such as head amplifiers, may be isolated from ground and in these cases cable screens may be bonded together so that screen currents flow through the system without reaching the amplifier chassis. Small units, such as head amplifiers, may often be mounted so as to avoid the passage of earth currents through their chassis or may be designed to force such currents to follow a coaxial configuration.

4.3. *Instrument Cables*

Cables should have a low value of screen transfer impedance and the conditions for this are, briefly:

Low resistivity screen materials.

Maximum practical screen thickness.

High screen permeability.

Braided cables should be used with particular caution and lumped capacitance to ground at the far end should be kept to a minimum, especially for short cables.

Double-screen cables can have very high screening efficiency and may be a solution to interference problems in bad cases. These cables may, however, be awkward to handle and are certainly more difficult to terminate. In high temperature systems it may be convenient to remember that leakage resistance across the insulation between the two screens can be as low as a few hundred ohms without increasing interference pick-up.

Double-screened cables may be analysed by the methods of Sections 2 and 3. The expression for interference pick-up becomes rather complex, involving multiple resonances, but the multiplier includes the term $Z_{1T}\ Z_{2T}/Z_{SE}$ which can have a very small value since Z_{SE} may be about 20 ohms and Z_{1T} and Z_{2T} may be less than 1 mΩ. The author considers, however, that the complexity and cost of ensuring double screening throughout, including the electronic chassis, is not worth while and double screened electronic designs have not been adopted in this country. Double screening can, however, be very effective on cables alone where severe interference is encountered.

4.4. *Cable Systems*

Although the correct choice of cable can do much to reduce interference pick-up a great deal can also be done in the arrangement of the installation. Apart from reducing the energy transfer to the cable there is also the possibility of reducing the amplitude of resonant peaks (from which most interference comes) or shifting them outside the equipment pass band.

We may note the following:

All cables should be as short as possible.

Instrument cables should be well separated from disturbing system cables, parallel runs should not occur and every opportunity should be taken to use earth structure to provide screening. Crossings of the cable route by disturbing cables should be made at right angles. Where it can be arranged, the location of switch gear and machinery should avoid giving unbalanced earth current along the route of the instrument cables.

Earthing of systems at one point only makes a very great improvement at the lower frequencies but the effect at higher frequencies is small, altering the frequency but not the amplitude of the resonant peaks.

When a remote head amplifier is employed the resonances at the sensitive end of the system, the head amplifier input cable, may often be shifted outside the pass band by earthing the head amplifier locally, since the resonance now involves only the input cable. With this arrangement very heavy low-frequency disturbances occur in the head to main amplifier connection and balanced transformers or double-screened connections are usually essential in this link. The head amplifier local earth may be chosen to reduce front-end pick-up still further by making connection to the earth which surrounds the detector. Interference has been reduced by over 40 dB on one U.K.A.E.A. reactor by this means.

Metal ducting may be used to provide a measure of double screening but, to be effective, the longitudinal conductance must be high. System designers must beware of arranging duct systems in the form of a main trunk with branches spreading to various parts of the reactor site since this arrangement will encourage earth currents from all over the site to flow in the main trunk. Ducting of this sort may do more harm than good. Because only part of the installation can be enclosed in ducts the analysis of ducted systems is complex and outside the scope of this paper. Where the local earth forms a coaxial configuration with the cables, as with ducting, the cable screen impedance to ground, Z, can be increased without increasing Z_E, by running the cables centrally in the duct. Equations (4) and (5) indicate the improvement this gives since Z_{SE} may, typically, be increased by a factor of 4 in this way.

In the case of rigid extensions attached to radiation pulse counting detectors it is often practicable to raise Z_{SE} by increasing the inductance of the screen/earth system. This may be done by adding ferrite rings along the length of the extension, so dimensioned that the increase in capacitance does not outweigh the increase in inductance. Wrapping of cables with high permeability material is also effective by serving to raise Z_T.

One way of eliminating, or reducing, cable screen resonances is to provide proper terminations for the cable screen system. A straight resistive termination is impracticable since although the termination might reduce the resonant peaks by 10 : 1, it would increase the injected signal by least a factor of 10^4. An effective termination can be obtained, however, by winding the entire cable through ferrite rings to give an inductance greater than $4\,L_s l/\pi$ and by connecting a terminating resistor across the cable screen between the inductor so formed. A 15 dB improvement on a 2000 ft cable has been achieved by this means, the limitation being due to the non-uniform nature of the screen/earth impedance, which cannot be properly matched by a simple resistor.

Another method of reducing resonant peaks is to provide a resistor to ground at the far end of the screen, if this is normally open circuit. Only the first resonant peak is greatly affected and this occurs when the resistor provides the best possible termination for the line. The low-frequency pick-up is raised as a result and a compromise value may be preferred depending, of course, upon the pass band of the instrument channel. In a low-frequency channel the technique may be positively harmful.

When a cable has inadequate screening a considerable improvement may be achieved by taping a low resistance braid to it, along its length. The resulting mutual coupling may be regarded as reflecting the low resistance of the braid into the cable screen and the method has been successfully employed on detector connections in some U.K.A.E.A. low power reactors.

Another similar and very effective technique is to line detector holes with a tube of low resistivity material. The earthing arrangement must depend on the arrangement of screen earthing and whether the tube itself can be isolated from the reactor structure. It is most successful when the tube is isolated and connected to a local, grounded, head amplifier chassis.

4.5. *Reactor Earth System*

From eqns. (4) and (5) it is obvious that the earth structure impedance should be kept as low as possible, especially along the line of the instrument cable runs. This is not just a matter of increasing conductor cross-sections since even a heavy earth braid will have considerable inductance, the reactance of which may equal the effective resistance in the audio-frequency range. Above this frequency, increase in cross-section is relatively ineffective. The inductance may be greatly reduced by providing a number of parallel-spaced earth paths which follow a straight line. A net or mesh covering the area of interest ought to be a favourable arrangement.

The main nucleonic earth should be derived near the control room or main equipment location and should provide earthing for nothing else. A separate, remote 'quiet' earth for the electronic equipment is most undesirable since it encourages the flow of circulating currents between the instrument screen system and the local 'noisy' earth.

4.6. *Disturbing Circuit System*

Suppression techniques on disturbing circuits are not a subject for this paper. With regard to minimization of coupling most of the preceding principles should apply but translation into practical terms is not always possible especially on heavy power plant.

The author believes that most reactor installations could benefit from more careful consideration of the location and routing of heavy power plant and cables in relation to the instrument systems at the design stage. The choice and amount of cable screening required should also be related to the expected magnitude of the current disturbances and, in this connection, there seems to be a pressing need to determine values of Z_T for typical power cables.

5. An Interference Measuring Technique

The tracking down of causes of interference after instruments have been commissioned and when the reactor is in operation is an extremely difficult business. It is however possible to obtain a measure of the sensitivity of instruments to electrical interference before this stage is reached by applying artificial disturbances to the site earth structure. The author has adopted a sinusoidal disturbing signal because of the ease with which the results may be analysed, especially for diagnostic purposes.

The disturbance is made from a power signal source located near the main equipment and feeding some convenient point on the reactor structure via a cable which is well separated from the instrument cables and from the earth, as far as is practicable. This simulates a worst case of earth disturbance since earth currents will follow the general line of the instrument cables, are not tightly coupled to their own cable and all the sent current is 'unbalanced'.

The resulting interference in each instrument channel is measured and related to the disturbing circuit current as it enters the earth structure at the far end.

By comparing the instrument interference with the disturbing current reaching earth at the far end of the cable the effect of disturbing circuit resonances is eliminated. This is because the disturbing current distribution is a standing wave at nearly all frequencies and the far end current is always the peak value of this wave; the near-end current may have any value.

Such a test determines the efficiency of the reactor earth system and the instrument installation as a whole and experience so far has been that instruments should tolerate disturbances of 100 mA to ground at any frequency above about 10 kc/s. This current corresponds to the expected disturbance from an unbalanced 440-V system of about 50Ω characteristic impedance, when resonating as a result of switching on, in which 10% of the disturbance reaches the reactor structure to which the instrument cable is coupled. There is less certainty about what standard disturbance to apply at lower frequencies.

The diagnostic value of the method may, perhaps, be seen from the following list of standard procedures:

1. The main equipment is operated with all cables disconnected except power supplies. Local interference is examined with the disturbing system operating and should be negligible. At this stage disturbing currents may be fed into the equipment rack or on to power supply lines as well as to the reactor structure.
2. Each cable screen is then connected in turn, without connection of the cable conductors.
3. Each cable is connected properly, in turn, but remote equipment (head amplifier, for example) is not switched on.
4. Remote equipment is switched on, but connected only to its power supplies and output cable.
5. The remaining remote equipment cables are connected in turn, first through cable screens only, until the entire channel is operating.

The amplitude/frequency response at each stage indicates where the interference is most severe and serves to give a measure of the effectiveness of any remedial action taken. No advance from one step to another can be taken until interference has been reduced to insignificant levels, otherwise the results which follow will be too complex for analysis.

Experiments are in hand at the present time to speed up the measurement process by applying a swept frequency disturbance. In parallel with this we are examining the possibility of achieving the same result by application of a pulse disturbance, in which the conclusions are drawn from examination of the pulse response in the time domain.

6. Conclusions

The forms of interference described in this paper seem to be the main sources of interference in nucleonic equipment. The principles of analysis which are described are not new, nor are many of the remedial techniques which are discussed, but the author has tried to show that instrument systems can be designed to operate in a reactor environment without interference troubles and that most of the criteria of design are calculable.

The method of instrument interference sensitivity measurements has had considerable success in the U.K.A.E.A. and the author believes that this, or some equivalent technique, should be adopted as a standard commissioning procedure.

7. Acknowledgments

The author would like to thank those members of Design and Manufacturing Division, A. E. E. Winfrith, and especially Mr. W. R. White, who assisted in the study by making many thousands of measurements on models of instrument systems, and those members of Control and Instrumentation Division who undertook similar work on the various reactor sites.

He would also like to thank the United Power Company and the C.E.G.B. for permission to make measurements in the Trawsfynydd Nuclear Power Station and for their co-operation during the work.

8. References

1. "Bibliography on Radio Interference", *I.R.E. Trans. on Radio Frequency Interference*, **RFI-4**, No. 1, February 1962.
2. L. Krugel, "The screening effect of the outer conductors of flexible coaxial cables", *Telefunken Z.*, **29**, No. 114, pp. 256–64, 1956.
3. L. Krugel, "Multiple screening of flexible coaxial cables", *Telefunken Z.*, **30**, No. 117, pp. 207–14, September 1957.
4. R. W. Levell, "A Preliminary Study of the Characteristics of Mineral Insulated Cable for Reactor Control Pulse Counting Channels." AEEW–M.369.
5. E. P. Fowler, "Design Criteria for Pulse Transformers used in Neutron Detector Pulse Counting Channels." AEEW–R.306, October, 1963.

Manuscript first received by the Institution on 6th April 1964 and in final form on 11th December 1964. (Paper No. 966.)

Accelerating Medical Progress through Medical Engineering

By

V. K. ZWORYKIN, D.Sc. (*Honorary Member*)†

We are all aware of and impressed with the rapid progress of science in many fields. The time between the conception and the widespread utilization of an idea can be very short under appropriate circumstances. For example, the transistor was conceived of only a few years ago, yet transistor devices are now in universal use.

There is, however, one field in which progress, though sorely needed, is not as rapid as might be desired. This is in the general field of the application of modern technology, which is largely electronic in nature, to the problems of biology and medicine. This field has been given many names. It has been called medical electronics, medical engineering, biophysics, bio-medical technology. In the United States, for example, bio-medical engineering is now becoming established as a scientific discipline. The National Institutes of Health annually support a large number of fellowships in Biological and Medical Engineering in order more adequately to satisfy the great demand for personnel trained in this manner.

In many countries there are now societies with international affiliation. These groups are in close contact so that their meetings whether on a local or an international level, are filled with the ferment of the communication of new ideas, techniques, experiments and instrumentation. The scientific journals are filled with information on new instruments, many of them developed on a 'one-off' basis with government support.

Although in clinical medicine and pharmacology there is little problem in achieving the widespread adoption of new medical and surgical therapies, and in making them available to the individual patient at relatively low cost, this is largely not the case in bio-medical engineering. There are a number of reasons which contribute to this state of affairs. The first of these is the natural and proper conservatism of the medical practitioner. However, of even greater importance is the lack of economic incentive for a producer to manufacture a small number of devices at high cost, with no firm market established and ready.[1] Many fine and possibly life-saving ideas in medicine therefore languish not for lack of government support to make the first one, which is of course prohibitively expensive, but for lack of the necessary pump-priming to ensure the widespread distribution of the necessary devices and techniques, making them available not only at the largest medical centres, but in the local hospitals and at the general practitioner's level.

Although this is a generally acknowledged state of affairs, not much has been done on either side of the Atlantic to alleviate the situation. However, I have just returned from an extended trip into a number of Eastern European countries and found there an already existing working mechanism to accelerate the application of bio-medical technology to human problems. This mechanism is the 'Scientific Medical Engineering Research Institute', a relatively mission-oriented organization devoted not so much to theoretical investigations as to the production of specific badly needed medical or technical equipment for the solution of pressing problems. This type of institute does not appear to be correspondingly duplicated in the western countries.[2]

In Czechoslovakia, Russia and East Germany, I saw or know of complete institutes established to conceive, do research upon, design, develop, produce and market pilot models of promising new instruments. These institutes have generally followed an organizational pattern consisting of a director, a staff of engineers, medical scientists and supporting personnel supplied with workshops, laboratories, hospitals, clinics and production facilities. The Scientific Research Institute for Experimental Surgical Instruments and Apparatus in Moscow has largely devoted itself to the instrumentation of surgical procedures such as vascular anastomis, and electronic anaesthesia.[3] Another institute in Moscow is the Institute for Scientific Instruments, which is developing a whole range of electronic and other equipment ranging from operating room lamps to magnetic devices to measure the thickness of electroplating on surgical instruments, as well as such specialized diagnostic equipment as automatic blood-cell-counters. Also in Russia is the Vishnevsky Institute, which is pioneering in the immediate application of computers to case history storage and diagnosis.[4] A new institute was recently established for clinical and experimental surgery under the direction of the Medical Academician, B. V. Petrovskii. The Russian government has made far-reaching plans to expand its support of medical engineering to a level approximating twelve times that in the United States.[5]

† Chairman of the Board of Governors of the International Institute for Medical Electronics and Biological Engineering; Honorary Vice President of the Radio Corporation of America.

In Eastern Germany the PREMA Institute under the direction of Kosumplik is developing a number of novel diagnostic electronic instruments including such devices as foetal monitors, cardiotachometer and recorders. A recently-created institute is the Research Institute of Medical Electronics and Modelling. East Germany has its Carl Carus Institute and the Manfred von Ardenne Institute in Dresden.[6] Each of these has a list of impressive accomplishments in bio-medical engineering to its credit. In Prague there is a similar institute under Dr. Peleska.

How can we in the United Kingdom, France and similar countries hope to compete, within our democratic society, with the massive support which a socialist or communist state can bring to bear on the problems it decides to attack?

Certainly very large-scale projects cannot be undertaken by private resources alone, and even government facilities may sometimes be inadequate. Already the United Kingdom has become aware of the realities of limited research and development budgets and an important project such as the development of the supersonic transport came under review.

Questions have been raised in the United States Senate on the wisdom of our space budget expenditures. In some sections of the electronic industries the pinch of contract cancellations has been felt.

The budgets which might be required for medical technology are likely to represent much more moderate investments than space and military programmes, but may well pay much greater dividends. Furthermore, in this area international competition might more readily be replaced by international co-operation.

The pattern for accelerating the practical utilization of research achievements in bio-medical engineering in the West would necessarily deviate from that followed in Eastern countries. It might well take the form of the establishment, by private initiative, of a series of relatively small bio-medical engineering research institutes, eventually having the sponsorship of professional associations. The International Institute for Medical Electronics and Biological Engineering in Paris, closely linked with the International Federation for Medical Electronics and Biological Engineering, may serve as an example. Such institutes might then seek support for their projects from the appropriate agencies of their respective governments.

Medical electronics and biological engineering can contribute materially to the advancement of medicine and human welfare. By removing a major roadblock in the way of the achievement of its goals, the encouragement and support of the research institutes here envisioned would help to bring us a step nearer to the "Great Society" which we all seek to realize. A rededication of a small portion of our present efforts in this direction would, I am sure, be amply repaid. There will be no dearth of volunteers ready and able to assist in such work.

References

1. *I.E.E. Trans. on Bio-medical Engineering*, **BME-11**, No. 4, p. 122 (footnote), October 1964.
2. R. P. Grant, C. P. Huttrer and C. G. Metzner, "Biomedical science in Europe", *Science*, **146**, pp. 493–501, 23rd October 1964.
3. M. G. Anan'yev, "The Russian Institute", *Amer. J. Medical Electronics*, **2**, No. 1, pp. 70–8, January–March 1963.
4. G. A. Zedgenidze, "Soviet medical engineering", *Amer. J. Medical Electronics*, **2**, No. 3, pp. 255–68, July–September 1963; **3**, No. 4, pp. 271–80, October–December 1964.
5. "Soviet medical electronic plans", *Amer. J. Medical Electronics*, **3**, No. 2, p. 140, April–June 1964.
6. M. von Ardenne and P. G. Reitnauer, "The medical section of the Manfred von Ardenne Research Institute", *Amer. J. Medical Electronics*, **3**, No. 1, pp. 50–56, January–March 1964.

U.D.C. 621.318.132: 681.32

Large Capacity Magnetic Film Stores— A Design Approach

By

P. I. BONYHARD, Ph.D.†

AND

W. S. CARTER, Ph.D.‡

Summary: The organization of large capacity magnetic film stores is considered. It is shown that the best compromise is obtained when all the strip transmission lines are extended in length until the limit imposed by attenuation and delay is reached. The number of continuous sections of both word and digit lines, and hence the number of auxiliary electronic circuits required for a store of given capacity and speed, is expressed in a novel manner as a function of the cross-sectional dimensions of strip lines. This expression is optimized to give the optimum store organization. The demands made on constructional techniques and electronic circuit elements by such a store are considered and the limits within which magnetic film stores are likely to be economical are suggested.

List of Symbols

Z_0	characteristic impedance
Z	impedance of free space $\simeq 377\,\Omega$
v	signal velocity
w	strip line width
p	strip line spacing or pitch
d	strip line thickness
h	separation between strip line and ground plane
l	length of line
ρ	resistivity
R	resistance
A, A'	voltage or current attenuation along the line
Δ	equivalent free space thickness of magnetic film
T, T'	time delay
l_A	attenuation limit of line length
l_T	delay limit of line length
W	index referring to word lines
D	index referring to digit lines
G	shunt conductivity of lines per unit length
$\tan\delta$	loss tangent
N	number of bits (or crossing lines) corresponding to a line
δ	skin depth
l'	corrected line length
C_{WD}	line to line capacitance
C_{WO}, C_{DO}	line to ground capacitance

† International Computers and Tabulators (Engineering) Ltd., Gunnels Wood Road, Stevenage, Hertfordshire.

‡ Formerly of I.C.T. (Engineering) Ltd.; now with Fairchild Corporation, California, U.S.A.

1. Introduction

The exploitation of the potential advantages of ferromagnetic films (i.e. intrinsic switching speeds in the order of a nanosecond, low fabrication cost and planar drive line construction) is a possible way of making large and fast random access computer stores economically.

The successful achievement of such a large storage system depends as much on optimizing the drive line characteristics and using a selection scheme which is compatible with the high density construction, as it does on using films with the desirable switching characteristics.

In a very large store it is advantageous to make all conducting strip lines as long as the limit imposed by transmission properties and to place the largest possible number of elements under each strip, so that the cost per bit of auxiliary electronic circuitry associated with each line is minimized. This leads to the use of multi-words,[1,2,3] i.e. the number of digits under a word line is a multiple of the computer word length. The limitation of the length of strip lines and the number of elements which may be placed under them are dependent on the transmission properties of the lines which in turn depend on the cross-sectional geometry of the lines. The pitch at which the lines may be packed depends on the line cross-section because of the stray field coupling which causes interaction between storage elements. It is possible therefore to express the number of continuous sections of both word and digit lines required for a store of given size as a function of the cross-sectional dimensions. This expression can then be optimized so that the total cost of electronic circuitry is minimized.

Because treating the cycle time as a continuous variable would be rather difficult, only two values of the mean cycle time, 100 ns and 1 µs are considered. It is believed that the mean cycle time of most practical magnetic thin film stores will be between these extremes, so that the design parameters corresponding to the intermediate values can be found by interpolation. A magnetic thin film store much faster than 100 ns is unlikely, owing partly to the difficulties encountered in the design of selection and amplification circuits with the large power gains required and partly to the increasing competition from using tunnel diode and similar systems. A magnetic thin film store much slower than 1 μs is similarly unlikely owing to the existence of well established ferrite core storage systems at these speeds.

The development of a suitable non-destructive read scheme, which has an effect on the overall cost of the store, is assumed in the first instance. The effects of eliminating it are then considered separately.

The properties of the magnetic films are not considered in detail except where our experience suggests they impose a limit on packing density below the optimum derived from the transmission line properties. Otherwise it will be assumed that films having the required properties can be produced. It has to be mentioned, however, that possible improvements in the properties of films used would contribute greatly to the reduction of costs, if they permitted the use of lower drive currents. This could partly be achieved by producing films with reduced anisotropy constant and dispersion as both parameters contribute to the magnitude of the word and digit drive pulses required.

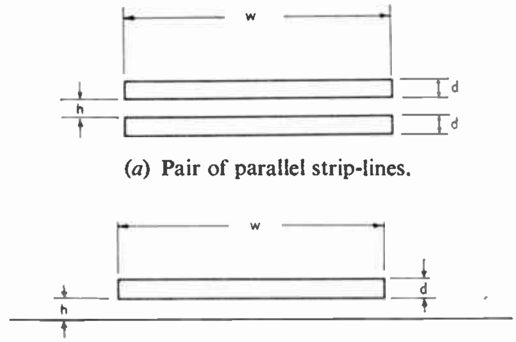
2. Transmission Properties of Parallel Strip Lines

The theoretical treatment of the transmission properties of parallel plate systems is extremely complex. A good review of the many different approaches and results is given by Harvey.[5] The cross-sections of the two types of interest here are shown in Fig. 1. These two may be considered equivalent to the first approximation. For our discussion an accuracy of ± 20% is adequate; therefore, as comparison with the more rigorous expressions given by Harvey shows, the simplest approximation suffice. Treating the lines as parallel plates of infinite extension one gets for the characteristic impedance Z_0 and the propagation velocity v

$$Z_0 = \sqrt{\frac{L}{C}} = \sqrt{\frac{\mu_0(h+\Delta)/w}{\varepsilon_0\varepsilon_r w/h}} = \sqrt{\frac{\mu_0}{\varepsilon_0\varepsilon_r}}\frac{\sqrt{h(h+\Delta)}}{w}$$

$$v = \frac{1}{\sqrt{LC}} = \frac{1}{\sqrt{\mu_0\varepsilon_0\varepsilon_r}}\sqrt{\frac{h}{h+\Delta}} = \frac{c}{\sqrt{\varepsilon_r}}\sqrt{\frac{h}{h+\Delta}}$$

where, using the notation of Fig. 1, the width of striplines is w, their separation h, their thickness d. L and C are the inductance and capacitance per unit length, μ_0 and ε_0 the permeability and dielectric constant of free space, and ε_r the mean relative dielectric constant of the medium between the strips.



(a) Pair of parallel strip-lines.

(b) Strip-line and substrate.

Fig. 1. Strip transmission line configurations.

$$\sqrt{\frac{\mu_0}{\varepsilon_0}} = Z \simeq 377\ \Omega$$

and

$$\frac{1}{\sqrt{\mu_0\varepsilon_0}} = c \simeq 3\times 10^8 \text{ metres/second}$$

Δ is the thickness of free space that would give the same induction as the magnetic film between the strip lines. For digit lines $\Delta \simeq 0$. It may appear that for L the internal inductance should also be considered and $(h+d)$ should be put in place of h. This, however, according to Harvey, would give a worse approximation.

The resistance R of a pair of lines of length l and strip line width and thickness w and d respectively is $R = \frac{2\rho l}{wd}$. Hence the total attenuation factor is

$$A = \frac{R}{2Z_0} = \frac{\rho l\sqrt{\varepsilon_r}}{Zd\sqrt{h(h+\Delta)}}$$

corresponding to a ratio of voltage or current at the two ends of the line of

$$\frac{V_2}{V_1} = \frac{I_2}{I_1} = l^{-A} \simeq 1-A$$

The delay time along the line,

$$T = \frac{l}{v} = \frac{l}{c}\sqrt{\varepsilon_r}\sqrt{\frac{h+\Delta}{h}}$$

For our purpose it is most suitable to express the length of lines as limited by attenuation and delay. These are, respectively,

$$l_A = \frac{AZ}{\rho\sqrt{\varepsilon_r}}d\sqrt{h(h+\Delta)}$$

$$l_T = \frac{cT}{\sqrt{\varepsilon_r}}\sqrt{\frac{h}{h+\Delta}}$$

for the word line, whereas for the digit line these expressions reduce to

$$l_A = \frac{AZ}{\rho\sqrt{\varepsilon_r}}\,dh$$

$$T = \frac{cT}{\sqrt{\varepsilon_r}}$$

If $l_A = l_T$ is required, as it would be, to make the best use of the lines, the equation

$$d(h+\Delta) = \frac{cT\rho}{AZ}$$

$$\left(dh = \frac{cT\rho}{AZ} \text{ for the digit line}\right)$$

has to be satisfied.

So far only the attenuation due to resistive losses has been considered. The attenuation due to dielectric loss may be expressed as

$$A' = \frac{GZ_0 l}{2} = \frac{l}{\lambda}\,\pi \tan\delta$$

where G is the shunt conductivity per unit length, λ the signal wave-length and δ the loss angle. A' is negligible, provided that $\tan\delta$ is small.

Lastly, the effect of one set of parallel strip lines upon any one of the other, perpendicular set of parallel strip lines has to be considered. A set of strip lines can be thought of as a heavily slotted plane conductor. Owing to the slotting, the effect of this plane upon the inductance of the strip is insignificant, but the capacitance may be affected considerably. The interline capacitance is connected to earth via half the characteristic impedance (Fig. 2), which is equivalent to another capacitance and another resistance shunting to earth in parallel. A check of the typical magnitude shows clearly that the shunt resistance (about 0·1 MΩ for 10^4 crossing lines) is too high to cause any significant attenuation, whereas the shunt capacitance is very nearly equal to the inter-line capacitance. The effect of this extra capacitance is to decrease both the characteristic impedance and the signal velocity and thereby decreases both l_A and l_T by the same factor. This factor can be calculated from geometrical considerations alone, giving

$$\sqrt{1 - \frac{h_W}{2h_D - h_W - 2d_W}}$$

for the word line

and $$\sqrt{1 - \frac{h_W + d_W}{2h_D - h_W - d_W}}$$

for the digit line,

where indices W and D refer to the word and digit lines, respectively.

3. Read and Write Cycle Times

The read and write cycles have to be considered in detail to determine the delay that can be tolerated, and the rise-time and skin depth, corresponding to a given mean cycle time, before considering details of store organization. A mean cycle time of 100 ns will be considered, assuming that non-destructive read is available and assuming an average of 2 read operations for 1 write operation to make up the mean cycle time.

It is possible to terminate the far end of the line, i.e. the end remote from the pulse generator or output amplifier, either in the characteristic impedance of the line or in a short circuit. In the former arrangement, the Z_0 termination, the rise-time and delay-time can be decided independently. In the latter arrangement, shorted termination,[1,4] the delay time must be less than, or equal to half the rise-time.

The read and write timing diagram for both types of termination are illustrated in Fig. 3. The continuous lines correspond to delay along the digit lines only, whereas the dotted lines represent the cycle when there

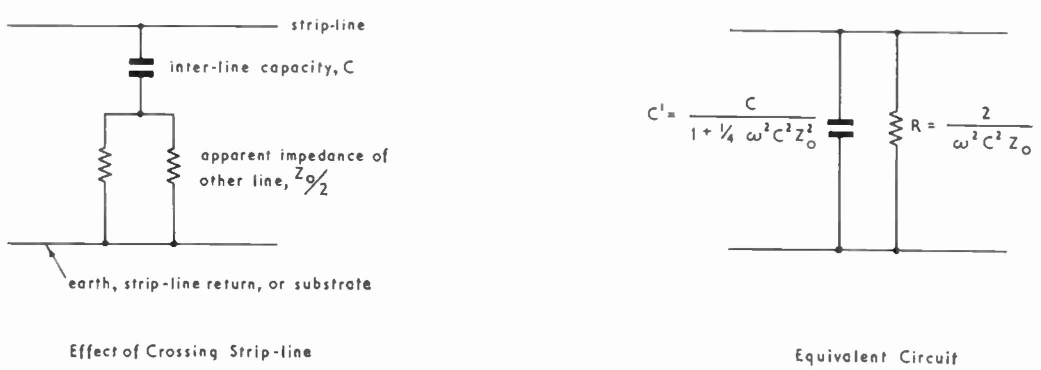


Fig. 2. Equivalent circuit of the effect of inter-line capacitance.

Typical values: $C < 0{\cdot}1$ pF, $\omega C < 10^{-6}\ \Omega^{-1}$, $Z_0 \simeq 20\Omega$. Thus $C' \simeq C$ and the total conductance due to 10^4 lines is less than $10^{-5}\ \Omega^{-1}$.

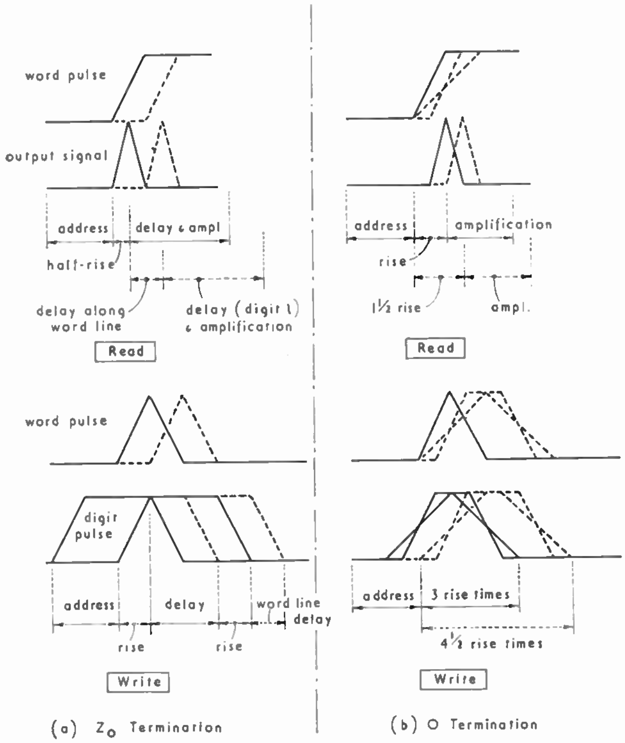


Fig. 3. Timing of read and write cycles.

is the same delay along the word lines as well. In the case of shorted terminations there are more possible output waveforms than indicated on the diagram, where only signals corresponding to maximum delay are shown. In the write cycles some tolerance must be allowed for in the overlap between the word and digit pulses, which is not shown on the diagram. No extra time is allowed for recovery, as this can take place during the next addressing time.

The total cycle time may now be expressed as follows:

(*a*) Z_0 termination:

read time = address time + $\frac{1}{2}$ rise-time + digit delay time + amplification time + word delay time.

write time = address time + 2 rise-time + digit delay time + word delay time + tolerance time.

mean cycle time = $\frac{1}{3}$ (2 read time + 1 write time) + address time + $\frac{1}{3}$ (2 amplification times + tolerance time) + rise-time + digit delay time + word delay time.

For a 100 ns mean cycle time it is reasonable to assume that the electronics time is 50 ns (address time 20 ns, amplification and tolerance time 15 ns each, for instance) so that 50 ns is allowed for the rise-time plus the total delay time.

(*b*) Zero termination:

read time = 1 rise-time + address time + amplification time + ($1\frac{1}{2}$ rise-times).

write time = 3 rise-times + address time + tolerance time + ($1\frac{1}{2}$ rise-times).

mean cycle time = $\frac{5}{3}$ rise-times + electronics time as before + ($\frac{5}{6}$ rise-times).

Thus $\frac{5}{3}$ rise-times + ($\frac{5}{6}$ rise-times) = 50 ns, where the figures in brackets are to be added only if the delay along the word line is not negligible.

The Z_0 termination allows a degree of freedom insomuch that only the sum of the rise and delay times is given, whereas for the zero termination the rise-time and hence the delay time (half the rise-time) and the skin depth are given as

$$T_{delay} = 15 \text{ ns } (10 \text{ ns})$$
$$\delta = 23 \ \mu\text{m } (18{\cdot}7 \ \mu\text{m})$$

where the figures in brackets again refer to significant delay along the word line. The skin depth, δ was calculated for copper at a frequency whose cycle time equals 4 rise-times.

4. Optimum Store Organization

The optimum organization of a large capacity store (i.e. the most favourable number of bits corresponding to each continuous word and digit line) will now be determined from the relationships derived above.

The treatment is simplified by using a target capacity of 10^8 bits in preference to retaining the capacity as a variable, and predicts the number of matrix units required to accommodate the number of bits. A matrix unit, in this context, refers to a single two-dimensional array of continuous strip lines; and consequently has a bit capacity of the product of the number of bits under a word line times the number of bits under a continuous section of digit line.

The maximum economic capacity for an individual store (i.e. the capacity beyond which the cost per bit does not reduce) can then be determined from these results and is independent of the chosen target capacity, provided that this is larger than the largest value obtained for a unit. It will be seen that 10^8 bits meets this requirement and is also the size considered by many authors.[2,3,6]

The choice of 100 bits per word was made partly for simplicity and partly to cope with the expected increase in word size. The store, chosen as an example may then consist of either 10^6 single words or a smaller number of multi-words. In the multi-word organized store results will not be much affected by the computer word size assumed. It transpires that, if it is constructionally feasible, there is a large saving in electronics for multi-word organization. First, however, the single word organization will be considered.

4.1. *Single Word Organization*

In the case of the single word organization the word lines are short and the delay along them probably negligible. As the number of word lines is fixed, the object here is to minimize the number of digit lines or, put the other way, maximize the number of bits (or words) corresponding to each digit line. This number, N_W will be expressed now in terms of the transmission properties of the strip lines. Indices D and W will refer to digit and word lines, respectively, throughout. From the expressions derived in Section 2,

$$l_D \leqslant \begin{cases} \dfrac{AZ}{\rho\sqrt{\varepsilon_r}}\, d_D h_D & \text{for } h_D \geqslant \\[2ex] \dfrac{cT}{\sqrt{\varepsilon_r}} & \text{for } h_D \leqslant \end{cases} \left.\vphantom{\begin{matrix}a\\a\\a\\a\end{matrix}}\right\} \frac{cT\rho}{AZd_D} \qquad \ldots\ldots(1)$$

In order to keep stray fields coupling at a reasonable level the pitch of the lines, p, must be proportional to $(h+d)$.

Because there are insufficient experimental results to determine the constant of proportionality (K), it will be treated as a variable for the present but given the value 10 later.

$$p_D = K(h_D+d_D), \qquad l_W = 10^2K(h_D+d_D)$$

and

$$l_W \leqslant \begin{cases} \dfrac{AZ}{\rho\sqrt{\varepsilon_r}}\, d_W\sqrt{h_W(h_W+\Delta)} \\[2ex] \dfrac{cT}{\sqrt{\varepsilon_r}}\sqrt{\dfrac{h_W}{h_W+\Delta}} \end{cases}$$

As the word line will be necessarily very short, the delay inequality is easily satisfied, so that d_W and h_W can be reduced until the attenuation limit is reached. Hence maximum word packing density is obtained by minimizing (h_W+d_W), subject to the condition that

$$d_W\sqrt{h_W(h_W+d_W)} = \frac{\rho\sqrt{\varepsilon_r}\,l_W}{AZ} = \frac{10^2K\rho\sqrt{\varepsilon_r}(h_D+d_D)}{AZ}$$

The result will be anticipated by assuming $h_W \ll \Delta$ so that

$$d_W\sqrt{h_W} = \frac{10^2K\rho\sqrt{\varepsilon_r}(h_D+d_D)}{AZ\sqrt{\Delta}}$$

The minimum of (h_W+d_W) occurs at

$$d_W = 2^{\frac{1}{3}}\left[\frac{10^2K\rho\sqrt{\varepsilon_r}(h_D+d_D)}{AZ\sqrt{\Delta}}\right]^{\frac{2}{3}} \qquad \ldots\ldots(2a)$$

and

$$h_W = 2^{-\frac{2}{3}}\left[\frac{10^2K\rho\sqrt{\varepsilon_r}(h_D+d_D)}{AZ\sqrt{\Delta}}\right]^{\frac{2}{3}} \qquad \ldots\ldots(2b)$$

so that, finally,

$$N_W = \frac{l_D}{p_W} = \frac{l_D}{K(h_W+d_W)} = \frac{1}{200K} \times$$
$$\times\left(\frac{AZ}{\rho\sqrt{\varepsilon_r}}\right)^{\frac{5}{3}}\Delta^{\frac{1}{3}}\,\frac{d_D h_D}{(h_D+d_D)^{\frac{2}{3}}}$$

where

$$d_D h_D = \frac{cT\rho}{AZ}$$

The maximum of N_W is at $h_D = d_D = \left(\frac{cT\rho}{AZ}\right)^{\frac{1}{2}}$. If, however, this gives $d_D > \delta$, it cannot be achieved, and $d_D = \frac{cT\rho}{AZ\delta}$ should be taken.

N_W can now be calculated both for Z_0 and zero terminations. In the case of zero termination the signal is made up as the sum of a direct and a reflected wave so that the effective attenuation is lower. A total attenuation of 37% down the line causes only 10% difference in signal level under these conditions. It appears, therefore, that a larger value of N_W is obtained by using the zero in preference to the Z_0 termination, because the larger value of A this gives in the above expression for N_W more than compensates for the corresponding smaller value of T.

Numerical values of all geometrical parameters taking $K = 10$, are worked out below. Using the values from those given in Section 3 of $T = 15$ ns, and $\delta = 23$ µm together with the following values:

$$A = 0{\cdot}3$$
$$c = 3\times10^{10}\ \text{cm s}^{-1}$$
$$\varepsilon_r = 4$$
$$Z = 377\ \Omega$$
$$\Delta = 50\ \mu\text{m}$$
$$\rho = 1{\cdot}7\times10^{6}\ \Omega\,\text{cm}$$

$$\left(\frac{cT\rho}{AZ}\right)^{\frac{1}{2}} = 26\ \mu\text{m} > \delta\text{,thus} \begin{cases} d_D = \delta = 23\ \mu\text{m} \\[1ex] h_D = \dfrac{cT\rho}{AZ\delta} = 29{\cdot}5\ \mu\text{m} \end{cases}$$

$$l_D = \frac{cT}{\sqrt{\varepsilon_r}} = 225\ \text{cm}$$
$$p_D = 525\ \mu\text{m}$$
$$l_W = 5{\cdot}25\ \text{cm}$$

gives

$$\left(\frac{l_W\rho\sqrt{\varepsilon_r}}{AZ\sqrt{\Delta}}\right)^{\frac{2}{3}} = 1{\cdot}71\ \mu\text{m}$$

From this we obtain $d_W = 2{\cdot}16$ µm, $h_W = 1{\cdot}08$ µm.

These results give a word pitch of 32·4 μm which is unlikely to be realized in practice. Taking, still rather optimistically, $p_W = 100\ \mu\text{m}$ and $l_D = 2$ m, we get $N_W = 20\,000$ as an upper limit of what may be feasible at this speed.

The delay along the word line,

$$T' = \frac{l_W\sqrt{\varepsilon_r}}{c}\sqrt{\frac{\Delta}{h_W}} = 3 \text{ ns}$$

is negligible, as expected.

If each continuous section of digit line accommodates 20 000 bits, then 5000 separate sections of digit lines are required in the 10^8 bit store, which has been chosen as an example, so that each of 100 complete digit lines has to be broken into 50 separate sections.

4.2. *Multi-Word Organization*

In the single word store considered above, the cost of auxiliary circuits is almost entirely governed by the cost of the selection circuit element associated with each word line, all other circuits being much smaller in number. It is apparent therefore that the cost of the electronic circuitry may be reduced by reducing the number of word lines. This can be done by increasing the length of word lines, which, in the single word construction are very short. The use of non-destructive read would allow selection at the end of digit lines, which leads to the requirements of only one output signal selector and one digit pulse selector per digit line.

If there are N_W word lines corresponding to a digit line and N_D digit lines corresponding to a word line, the total number of word lines is C/N_D and the total number of digit lines C/N_W, where C is the number of bits in the store (10^8 in the case considered here). Thus the total cost of circuitry is proportional to $(1/N_D) + (v/N_W)$, where v is the ratio of the cost of the selector elements at the end of each digit line to the cost of selection element at the end of each word line. If the word selection element is a transistor, as is generally considered necessary at fast cycle times,[1,7] v is likely to be 1 to 2. In the single-word store $N_W = 20\,000$, $N_D = 100$, as deduced above, so that the $1/N_D$ term dominates the cost. This term would still be the largest if N_D were increased 10 times. It will be shown below that such an increase is about the maximum feasible. The multi-word organization is therefore advantageous if a satisfactory construction can be found and the cost of electronic circuitry may be reduced by nearly a factor of 10.

The optimum choice of parameters for the multi-word organization, assuming that the delays along the word and digit lines are equal, is obtained by starting with the optimized single-word store discussed above. However, $T = 10$ ns and $\delta = 18{\cdot}7\ \mu\text{m}$ must now be used in accordance with the considerations of Section 3.

The object now is to have more bits under each word line. This can be achieved by:

(*a*) reducing the length of the digit lines, l_D so that the digit pitch can be reduced; or

(*b*) altering the thickness and separation of the word lines, d_W and h_W so that the word lines can be made longer.

(*a*) is uneconomical, as the length of the digit line is nearly proportional to the square of the pitch according to eqn. (1). (*b*) however can be effected, as it will be shown, with only the slight disadvantage of a moderate increase in the word pitch. Taking h_W as the independent variable, the other parameters may be expressed as follows:

$$d_W = \frac{cT\rho}{AZ(h_W+\Delta)} \quad \left(\begin{array}{l}\text{to equalize the attenuation}\\ \text{and delay limits of the word}\\ \text{line lengths}\end{array}\right),$$

$$l_W = \frac{cT}{\sqrt{\varepsilon_r}}\sqrt{\frac{h_W}{h_W+\Delta}} = \frac{AZ}{\rho\sqrt{\varepsilon_r}}\, d_W\sqrt{h_W(h_W+d_W)}$$

$$d_D = \delta = 18{\cdot}7\ \mu\text{m (constant)},$$

$$h_D = \frac{cT\rho}{AZ\delta} = 24\ \mu\text{m (constant)},$$

$$l_D = \frac{cT}{\sqrt{\varepsilon_r}} = 150 \text{ cm (constant)}.$$

Also

$$N_W = \frac{l_D}{p_W} = \frac{l_D}{K(h_W+d_W)}, \quad N_D = \frac{l_W}{p_D} = \frac{l_W}{K(h_D+d_D)}.$$

For the single word case ($N_D = 100$) the condition $(h_W+d_W) \ll h_D$ is satisfied. Now this is no longer necessarily the case and therefore one of the following steps must be taken:

(*a*) Increase h_D. In this case the digit pitch would increase more than the word line length and so N_D would decrease rather than increase;

(*b*) Put the word lines outside the digit lines. This results in either extremely short digit lines or extremely large word pitch causing an intolerable decrease of N_W;

(*c*) Allow for the effects of inter-line capacitance as shown in Section 2. This is the only satisfactory alternative, giving:

$$l'_W = l_W\sqrt{1-\frac{h_W}{2h_D-h_W-2d_W}}$$

$$l'_D = l_D\sqrt{1-\frac{h_W+d_W}{2h_D-h_W-d_W}}$$

p_W, l'_W, l'_D, N_D and $(1/N_D)+(2/N_W)$ are plotted as a function of h_W in Fig. 4. Again K is taken as 10.

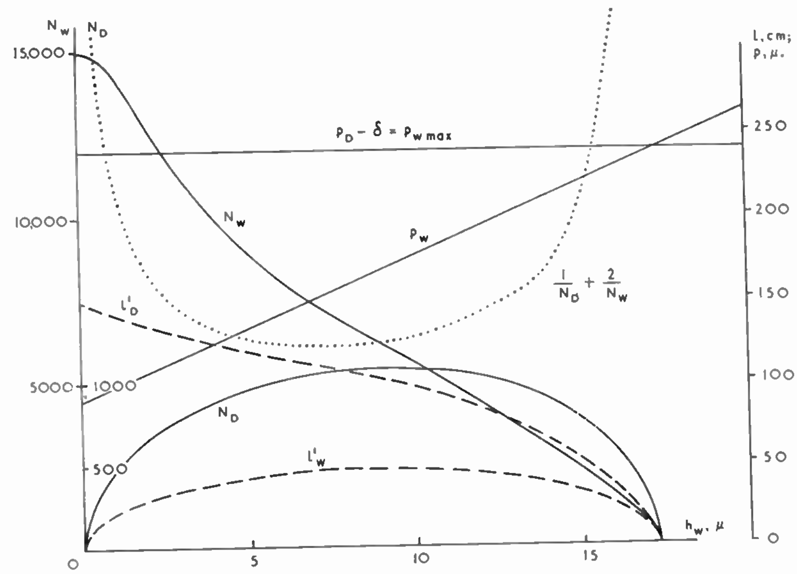


Fig. 4. The variation of store parameters as a function of h_W, 100 ns cycle time.

As anticipated, $\dfrac{1}{N_D}$ predominates over $\dfrac{2}{N_W}$ and, therefore, the minimum electronics cost nearly corresponds to the point where N_D is maximum. Practical values, as suggested by the graph, would be $N_D = 1000$, $N_W = 7150$ and, correspondingly, 14 sections for the example of 10^8 bits. This gives a total of 100 000 word lines and 14 000 digit lines, as opposed to 1 000 000 word lines and 5000 digit lines in the single word organized store. Figure 5 shows schematically how the multi-word store is organized, using the values derived here.

5. Destructive Read

The use of non-destructive read was assumed in Section 4. The effects of only having destructive read available will be considered now. Obviously the need to re-write information, in order to preserve it after interrogation, will increase the mean cycle time by some factor between 1 and 2 depending on the way the store is used. It will be assumed here that such an increase in the cycle time is accepted, so that all the calculation so far need not be repeated for a 100 ns mean destructive read cycle time. Another effect of destructive read is that the multi-word organization becomes less attractive, as a large number of complete digit circuits have to be provided. Using the numerical values derived in Section 4, the following circuits are required, when the read is destructive:

(*a*) Single word organization;

1 out of 10^6 word selection;

5000 digit-sense gating elements;

100 complete digit-sense circuits;

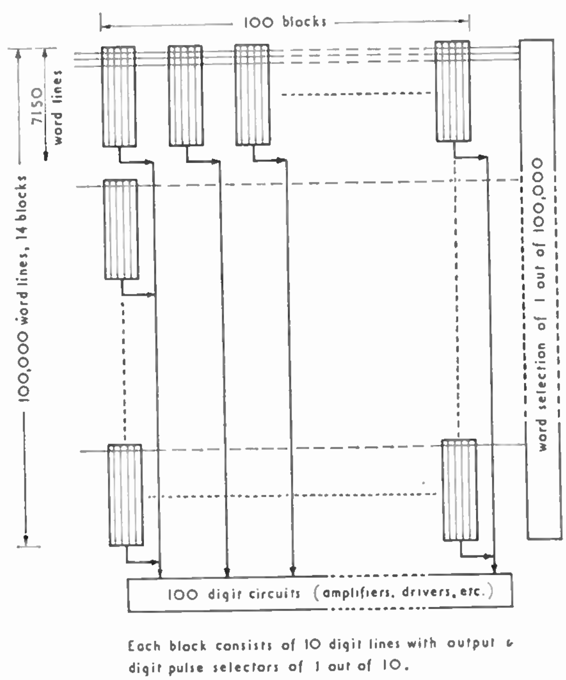


Fig. 5. Organization of the 10^8-bits, 100-ns cycle-time multi-word store.

(*b*) Multi-word organization;

1 out of 10^5 word selection;

14 000 digit-sense selector elements;

1000 complete digit-sense circuits.

Assuming that the digit-sense selectors and digit-sense gates are identical and, as before, 2 times as expensive as the word selection elements, it can be deduced from the above figures that there is a real saving in going to multi-word organization, provided that the complete digit-sense circuit costs less than 980 times as much as the word selection element. If the latter is a transistor, this factor is about 20 rather than 980, and the saving is still about an order of magnitude. Consequently the fact that the read is destructive has no significant effect on the economical balance between single and multi-word organizations.

6. One-microsecond Cycle Time

It is quite simple to repeat the calculations of Sections 3, 4 and 5 with the cycle time increased by one order of magnitude and all other quantities altered accordingly. The results are given below:

(*a*) Single word organization.

$T = 150\ \mu\text{s}$	$\delta = 73\ \mu\text{m} = d_D$
$h_D = 93\ \mu\text{m}$	$p_D = 1{\cdot}66\ \text{mm}$
$l_D = 22{\cdot}5\ \text{m}$	$l_W = 16{\cdot}6\ \text{cm}$
$d_W = 5{\cdot}75\ \mu\text{m}$	$h_W = 1{\cdot}45\ \mu\text{m}$

p_W is still unrealistically low, and will be taken as 100 μm.

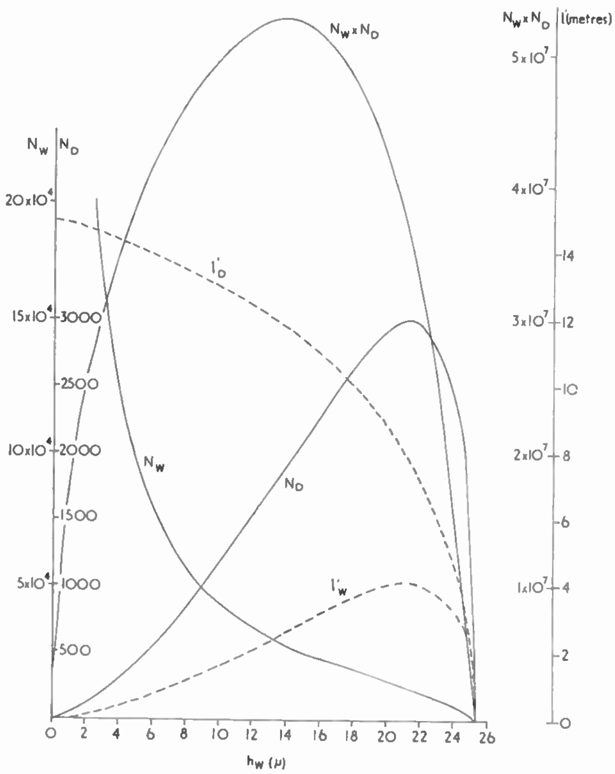


Fig. 6. Variation of store parameters as a function of h_W at 1 μs cycle-time.

Hence N_W can be taken as 200 000 and only 5 sections of digit lines are required.

(*b*) Multi-word organization.

The treatment of Section 4 cannot be applied here directly. If the equation $d_W = \dfrac{cT\rho}{AZ\,(h_W+\Delta)}$ was used, (d_W+h_W) would always be nearly as large or larger than h_D, which is not permissible. This means that the equalization of the attenuation and delay limits of the word line length is not practical. The word line will be attenuation limited so that $d_W = 2h_W$ relationship, used in a single word case in order to minimize (h_W+d_D), as it was derived in Section 4.1 on single-word organization and expressed in eqns. (2), is still the best compromise. T', the delay along the word line will then be smaller than the maximum permissible, T. As T' may not be negligible in comparison with T, the cycle time corresponding to significant delay along the word line must be used here. Taking again h_W as the independent variable, all other parameters may be expressed as follows:

$$T = 100\ \mu\text{s}$$

$$\delta = 59\ \mu\text{m} = d_D$$

$$h_D = 76\ \mu\text{m}$$

$$p_D = 1{\cdot}35\ \text{mm}$$

$$d_W = 2h_W$$

$$l'_D = \frac{cT}{\sqrt{\varepsilon_r}}\sqrt{1-\frac{3h_W}{2h_D-3h_W}}$$

$$l'_W = \frac{AZ}{\rho\sqrt{\varepsilon_r}}\,2h_W\sqrt{h_W(h_W+\Delta)\left(1-\frac{h_W}{2h_D-5_W}\right)}$$

$$T' = \frac{l'_W\sqrt{\varepsilon_r}}{c}\sqrt{\frac{h_W+\Delta}{h_W}}$$

$$p_W = 3Kh_W = 30h_W$$

l'_W, l'_D, $N_W = l'_D/p_W$, $N_D = l'_W/p_D$ and $(N_W \times N_D)$ are plotted as functions of h_W in Fig. 6, T', which is not plotted in the figure, never rises over 50 ns (i.e. $T/2$). The remarkable result is that over 40 million bits can be accommodated in a single unit for quite a wide range of all other parameters. Having two sections for each digit line, as it is usual for balancing, a $0{\cdot}8 \times 10^8$ bit store appears to be feasible. The word pitch need not be below 0·5 mm, which makes construction much more attractive.

At a cycle time of 1 μs, non-linear ferromagnetic materials (ferrite cores in particular), present a feasible word selection element as an alternative to transistors or diodes. The use of ferrite cores for word selection reduces the cost of electronics considerably, possible by more than one order of magnitude. It may also simplify construction by substituting a core threading operation for that of making electrical connections. Switching currents into long word lines, however, is only possible if rather large cores are used. Assuming that the flux of the core has to be twice the maximum flux of the magnetic film, the necessary core cross sectional area, (S) may be expressed as $S = 20 l_W D$, where the saturation of the core is assumed to be 10 times less than that of the thin magnetic film, whose thickness, D will be taken as 500 angstroms. Thus 1 mm^2 of core cross-section is required for every metre of word line length. It seems from these figures that a compromise could be found using cores of a few millimetres in height and diameter to drive word lines of the required length. The use of core selection of course changes the relative economics of single and multi-word organizations. Multi-word organization may still be advantageous since, for instance, by doubling the number of bits under a word line, 500 000 cores would be saved, at the expense 200 digit-sense selection elements in the 10^6 words, 10^2 bits per word store.

7. Constructional Considerations

Before the physical dimensions of a large store of the type considered can be given, the inter-plane spacing is required. The limit to this is set by the minimum spacing between substrates that can be tolerated. Bringing a conducting plane up to a magnetic film storage plane reduces the word and digit fields as well as reducing the magnitude of the output signal. The acceptable limit occurs approximately when the digit line is half-way between the planes, which has the effects of:

(*a*) Halving the digit line field, but not affecting the relative value of its stray field.

(*b*) Halving the output signal, if a common digit-sense line is used.

(*c*) Slightly reducing the word line field, but decreasing its relative stray field, because the word line will be nearer to the film plane.

The required increase in digit drive current is offset by a corresponding decrease in the digit line impedance.

For the 10^8 bit store of a 100 ns cycle time, the above considerations give a limit of about 100 μm spacing between substrates. This combined with an estimated minimum substrate thickness of about 200–400 μm gives an inter-plane pitch of about 0·5 mm.

The store will have to be divided into planar sub-units. This will be done most conveniently if the one dimension of each unit will just contain all the digits of a store-word, and the other dimensions lie between the value which will contain the square root of the total number of word lines and the value that will make the word selection matrix geometrically square. Using the former value we get a store made from 300 planes of 50 cm (in the word line direction) by 4·5 cm, which will be stacked into a height of 15 cm, for the multi-word store. The single word store, on the other hand, would have 10 000 planes of 5 cm × 10 cm, the stacked height of the planes being 450 cm. The latter dimensions would give a delay along the lines of the word selection matrix which would adversely affect the cycle time. This, coupled with the increased cost and complexity of the word selection matrix, makes the single word arrangement much less attractive. In the case of the multi-word store, however, the 15 cm height is very reasonable so that the stacking pitch of 0·5 mm could well be increased if this is found to be advantageous.

These considerations give rise to three main problems of construction:

(1) Mounting the substrates so that they provide a continuous ground plane for the drive lines.

(2) Fabricating the drive lines so that they remain continuous and are spaced from the substrate by a constant amount.

(3) Making electrical connexions between the word lines and the word selection matrix.

It seems reasonable that present techniques can be extended to coat and test substrates with the dimensions of 5 cm × 50 cm, and that these can be mounted at the required pitch for the multi-word store. The continuous substrate would help in making the drive lines continuous. At the proposed spacing from the film the insulation of the word lines could be achieved by depositing on material, such as cellulose, from a solvent. A continuous conductor could then be laid down by electroplating on an initial deposit of metal, which could then be etched into strip lines. The digit conductor spacing is sufficient to allow the use of the conventional technique of making lines on a flexible insulator.

The problem of making connexions to such a selection matrix appears so formidable that it seems unlikely that transistors will be used, so that the realization of a store of this specification will also depend on the successful development of a selection component which is more compatible with the planar structure of the films. Thin-film transistors,[8] thin-film tunnelling diodes[9] or integrated circuits are possible contenders here.

By comparison a 10^8-bit 1-μs cycle store will be physically much larger, requiring 250 magnetic film plates of 240 cm × 8 cm which probably cannot be made economically in one piece. Because the digit line spacing is larger, the minimum spacing between plates will be three times that given for the 100 ns store. However, a much larger spacing which will almost certainly be required for constructional reasons can be used as the larger cycle time does not limit the selection line length so severely.

The problem of fabricating the drive lines is basically the same as for the 100 ns store, the main difference being the extra length required.

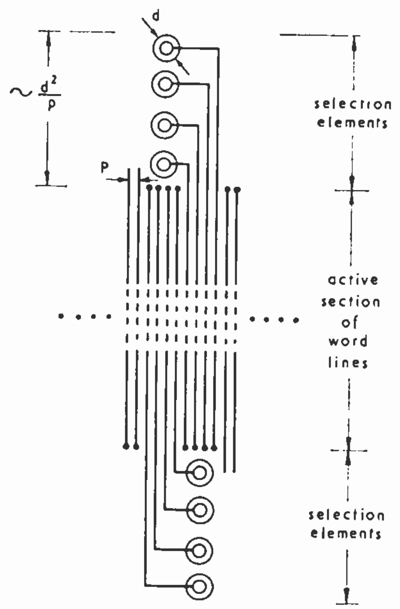


Fig. 7. Arrangement of selection elements driving word lines at high packing density.

The low cost of ferrite core selection will almost certainly preclude any other means of selection at this cycle time. The disparity between core diameter and the word line pitch can probably be made small by increasing the substrate thickness and the height of the cores. Figure 7 gives a geometrical arrangement to overcome any disparity and shows that the increase in word line length need be little more than n times the core diameter, where n is the ratio of the core diameter to the word line pitch.

8. Conclusions

The novelty in the approach to the design of large capacity thin-film stores presented here lies in expressing the number of bits which can be associated with unbroken sections of word and digit lines as functions of the cross-sectional dimensions of the strip lines. The expressions obtained can then be optimized accordingly to some criterion, such as the maximum number of bits in a single unit or the minimization of the auxiliary electronics.

The results give an estimation of the limitations, in terms of size and speed, of the economically feasible thin-film stores, and also indicate the best dimensions of strip lines, packing densities, plate size, etc., for such a store.

It has been found that, when all parameters are optimized, a total of 100 000 word lines and 14 000 digit lines are required in a 10^8 bits capacity 100 ns non-destructive read thin-film store. This corresponds to word and digit line pitches of about 150 μm and 430 μm, respectively. The problem of making connections between the word lines and the word selection elements, which, (relying on present day techniques) would almost certainly be transistors, appears formidable, if not impossible. Therefore, the word line pitch would probably have to be increased, and this would also increase the number of digit lines. It may be concluded from these considerations that the economical size for a 100 ns mean cycle time non-destructive read store would be between 10^6 and 10^7 bits. Above this size the cost per bit cannot be expected to reduce significantly as the size is increased.

In the case of a slower, 1 μs mean cycle time store, a capacity only slightly below 10^8 bits appears economically attractive, from the point of view of the construction as well as the electronics, if a ferrite core matrix is used for word selection. Here optimization yields the values of 50 000 word lines and 3200 digit lines giving a total capacity of $0{\cdot}8 \times 10^8$ bits. The optimized word and digit line pitches are 500 μm and 1350 μm, respectively, which simplifies construction.

It must be emphasized, however, that the numerical results given here for the number of bits corresponding to a line is inversely proportional to the value of K, as defined in Section 4.1, which has been taken rather arbitrarily as 10. The total capacity of a multi-word store or store section is therefore inversely proportional to K^2.

Non-destructive read facilities have been shown to be advantageous. If such facilities are not available, the effective cycle time increases by some factor between 1 and 2, provided that the store is driven at the same speed. The increase of complexity in the digit-sense circuits necessitated by destructive read, however, affects the costs per bit only marginally in very large stores.

Strip lines terminated in short circuits at their remote ends appear somewhat preferable to strip lines terminated in their characteristic impedances, owing to the higher attenuation that may be tolerated when the former type of line is used.

The use of multi-words, i.e. placing the maximum possible number of bits under each word line, not only reduces the cost of electronic circuitry per bit by an order of magnitude, but also simplifies construction as the optimum pitches are higher. Multi-word organization appears to be preferable for all forms of large capacity magnetic thin-film store.

9. References

1. A. V. Pohm, R. J. Zingg, T. A. Smay, G. A. Watson and R. M. Stewart, Jr., "Large, high speed DRO film memories", 1963 INTERMAG, paper 9.5, (I.E.E.E. Publication T.149), April 1963.
2. J. I. Raffel, "Future developments in large magnetic film memories", *J. Appl. Phys.*, **35**, No. 3, p. 748, March 1963.
3. A. V. Pohm, R. J. Zingg, J. H. Hoper and R. M. Stewart, Jr., "Analysis of 10^8 element magnetic film memory systems", 1964 INTERMAG, paper 5.3, (I.E.E.E. Publication T.159), April 1964.
4. E. E. Bittmann, "Thin film memories, some problems, limitations and profits", 1963 INTERMAG, paper 9.1, (I.E.E.E. Publication T. 149), April 1963.
5. A. F. Harvey, "Parallel-plate transmission systems for microwave frequencies", *Proc. Instn Elect. Engrs*, **106**, Part B, No. 26, pp. 129–40, (I.E.E. paper No. 2770E, March 1959.)
6. H. W. Fuller, T. L. McCormack and C. P. Batterel, "System and fabrication techniques for a solid state random access mass memory", INTERMAG, paper 5.5, (I.E.E.E. Publication T.159), April 1964.
7. W. Dietrich, L. J. Dunlop, H. P. Louis and W. E. Proebster, "The design of a one-million bit 100 nanosecond magnetic film memory", 1963 INTERMAG, paper 9.2, (I.E.E.E. Publication T.149), April 1963.
8. P. K. Weimer, "The TFT—a new thin film transistor", *Proc. Inst. Radio Engrs*, **50**, No. 6, pp. 1462–9, June 1962.
9. M. Silverberg, "Integrated computer techniques", R.C.A. Report EM-6220, December 1962.

Manuscript first received by the Institution on 2nd September 1964 and in final form on 9th November 1964. (Paper No. 967/C76.)

STANDARD FREQUENCY TRANSMISSIONS

(Communication from the National Physical Laboratory)

Deviations, in parts in 10^{10}, from nominal frequency for **February 1965**

February 1965	GBR 16kc/s 24-hour mean centred on 0300 U.T.	MSF 60 kc/s 1430–1530 U.T.	Droitwich 200 kc/s 1000–1100 U.T.	February 1965	GBR 16 kc/s 24-hour mean centred on 0300 U.T.	MSF 60 kc/s 1430–1530 U.T.	Droitwich 200 kc/s 1000–1100 U.T.
1	− 149·8	− 151·6	− 4	15	—	− 151·1	0
2	− 151·1	—	− 3	16	− 151·7	− 150·7	− 1
3	− 150·7	− 150·2	− 3	17	− 150·3	− 151·3	0
4	− 150·8	− 151·1	− 3	18	—	− 151·3	− 1
5	− 150·3	− 150·8	− 2	19	—	− 149·6	—
6	− 150·7	—	− 2	20	− 149·3	− 150·8	—
7	− 151·0	—	− 2	21	− 150·9	− 150·0	+ 2
8	− 151·1	− 151·1	− 2	22	− 149·8	− 151·1	+ 2
9	− 150·5	− 151·3	− 2	23	− 150·9	− 150·8	+ 2
10	− 150·6	− 150·1	− 2	24	− 150·1	− 150·9	+ 2
11	− 151·0	− 152·0	− 2	25	− 150·8	− 151·0	+ 3
12	− 150·3	− 150·8	− 2	26	− 150·4	− 150·6	+ 1
13	− 148·6	− 147·8	− 1	27	− 150·4	− 150·6	+ 2
14	− 148·1	− 148·9	0	28	− 150·5	− 150·5	+ 2

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

Note: *The phase of the GBR and MSF time signals was retarded by 100 milliseconds at 00 00 U.T. on 1st March 1965. The frequency offset for 1965 will be* -150×10^{-10}.

The Lincompex System for High Frequency Radio Telephony

By applying modern electronic techniques, the Research Branch of the British Post Office has developed a new type of terminal equipment for high frequency radio telephone circuits, which gives improved performance, more particularly when reception conditions are poor on account of noise and fading. Using this new equipment, known as LINCOMPEX, satisfactory service is obtained in some conditions when two-way communication is impossible with the conventional equipment. The system is at present undergoing a field trial on a radio circuit between London and New Delhi.

In spite of the introduction of submarine telephone cables with submerged repeaters, long distance overseas telephone circuits to many parts of the world will continue to be provided by h.f. radio, for a considerable time to come. The variable performance of these circuits with changing radio propagation conditions makes them somewhat unreliable links, and the new system is therefore an important contribution to the improvement of world communications. The Post Office hopes that it will eventually be adopted by all telephone administrations operating h.f. radio links.

Considering one direction only of a two-way radio link, at the sending end, speech is applied to a 'compressor' which gives a constant output signal level whatever the speech level at the input. Even the syllable to syllable changes in speech loudness are smoothed out. Thus the radio transmitter is always fully modulated and the best possible signal/noise ratio is obtained at all times. The compressor also produces a separate signal, called the control signal, which is a measure of the instantaneous speech level at the compressor input. This control signal is transmitted over the radio link, with the compressed speech, but by a separate channel.

At the receiving end, the compressed speech and control signal are separated and the speech applied to an 'expander', a variable-gain device controlled by the control signal, which faithfully restores the original variations of level or loudness of the speech. Thus the compressor and expander perform complementary functions because of the common control signal—hence the name LINCOMPEX. The control signal is transmitted over the radio link by frequency modulation of a sub-carrier, because this type signal is relatively immune to the effects of fading. The control signal is derived at the sending end by rectification of the input speech. It is applied to the compressor and causes it to give a constant output level. The control signal is also applied to a frequency modulator the output of which is combined by means of filters with the compressed speech, and passed to the radio transmitter.

At the receiver, the two signals are separated by filters. The effect of variations of signal strength due to fading will have been considerably reduced by the normal automatic gain control of the radio receiver, but any residual variations of level of the speech are removed by a constant volume amplifier. The control signal is demodulated, in the conventional manner for frequency modulated signals, by a discriminator. The speech is then applied to an expander, which, under the control of the recovered control signal, restores the original speech level variations.

Since the speech and control channels are of widely differing bandwidth, their times of transmission from input at the sending end to output at the receiver necessarily differ appreciably. It has been found that an improvement in performance is obtained by equalizing these times. This is achieved by inserting delay networks in the speech channel.

In order to appreciate how the new system improves performance, it is necessary to consider the previously existing, or 'conventional' system. At the sending end of an existing-type system there is a compressing device or constant volume amplifier (c.v.a.) which maintains a constant mean output level independent of the level at the input, so that the radio transmitter is operated at near full modulation. There is, however, no control signal, and in order to preserve the character of the speech, the c.v.a. is relatively slow acting and does not smooth out the syllable to syllable variations of level. These variations are therefore correctly reproduced at the receiver output. Since, however, no signal is sent to the receiver to indicate how much compression has been performed, there is no fixed relationship between the mean level at the input to the transmitter c.v.a. and the mean level at the receiver output.

For a low level input signal there would be a considerable net gain round the loop formed by the go and return paths, which would make the circuit unstable and cause oscillation. To prevent this, devices known as 'singing suppressors' are fitted. These consist of switches operated under the control of the speech signals in the two directions which ensure that only one direction of transmission can be connected through at any one time. The voice-operated control of the switches is differential. Thus if Party B is talking, Party A can break in and cause the switches to change over by talking sufficiently loudly. When radio conditions are very bad, however, it is possible for 'lock-out' to occur, i.e. one talker is unable to break in, because even when talking loudly, he cannot override the effect of the noise received on the radio receiver, which has the same switching effect as speech received from the distant end.

The great advance of the LINCOMPEX system over the present system is that there is a fixed relationship between the levels at the compressor input and expander output. We have, in fact, a circuit of constant loss, like a line circuit, and singing suppressors are unnecessary.

Furthermore, in the intervals between words the control signal conveys to the expander the instruction that its output level should be very low, so the expander inserts a high attenuation or loss. Thus in the silent intervals—when it would be most noticeable—the noise from the radio link is muted.

These are the two main advantages of LINCOMPEX, the constant loss feature which makes singing suppressors unnecessary and thus avoids the 'lock-out' trouble, and the muting of noise between utterances.

U.D.C. 534.1: 621.372.542

Progress in Electromechanical Filters

By

M. BÖRNER, Dr.rer.nat.†

Presented at a meeting of the Electro-Acoustics Group in London on 11th November 1964.

Summary: The development of techniques for manufacturing electromechanical filters is discussed and the change from single rods with turned-in necks to built-up construction described. Single-sideband filters and miniature types are then described and methods of eliminating spurious modes are shown.

1. Introduction

Fifteen years ago, first in the United States, later on also in Europe, research and development work began in the field of new types of electromechanical filters for frequencies up to 500 kc/s. Quite a number of research institutes in industry and in universities participated, but of the many variations of possible filter constructions which were at first proposed, only a small number have proved to be practicable. Initially, these were those manufactured by the Collins Company, who have been active in this field since its inception. An early filter, still in use, had bending-plate resonators and coupling rods, attached by means of spot-welding. Later on further practicable filter types were developed which will be discussed in this paper.

The filters used in the early stages of development, and to some extent until now, were manufactured out of a single metal rod by turning in necks (RCA, Marconi, Telefunken). In newer types, however, a technique is used which consists of first turning the individual resonators piece by piece, then tuning them to the right frequency, and afterwards assembling the complete structure by spot-welding the coupling members on to the resonator surfaces. This spot-welding process can now be performed so precisely that no post-tuning is necessary. In perfecting this technique, however, it has been found that all possible requirements for mechanical filters cannot be met by one type only. On the one hand filters must be made smaller and smaller, following the present trend of miniaturization in electronics. On the other hand some types of filters require ever-increasing accuracy with respect to their tolerances for frequency and attenuation behaviour. This is in keeping with the demand for more and more transmission channels to be established within very limited frequency ranges.

Such requirements led to the development of special single-sideband filters (15–20 circuits, carrier frequency approximately 200 kc/s) and to the development of miniaturized filters (6–10 circuits, passband 455 kc/s). These two trends are dealt with in this paper. Filters are described with regard to their electrical and mechanical properties as well as their production techniques.

Finally, possible future developments are discussed. All the known electromechanical bandpass filters are minimal-phase-shift networks, in which the transmission properties depend strongly on the phase properties. In the case of fast data-transmission, filters are needed with a frequency-independent delay-time for transmitted signals. This is achieved with additional all-pass elements. It seems possible to achieve integrated mechanical bandpass-all-pass structures for electromechanical filters in a very simple way, using a construction similar to that employed in the single-sideband filters. Not very different from such phase-compensated filters are filters in which attenuation poles have been achieved in the neighbourhood of the passband.

2. Single Sideband Filters

The development of electromechanical filters will be discussed in chronological order, starting with the single-sideband filter. It will be seen how the knowledge gathered with respect to tuning and production techniques for this type of filter assisted in further development. In a technique described by Roberts and Burns,[1] an attempt was made to produce single-sideband filters,[2] by turning in necks into a single metal rod (Fig. 1). Using longitudinally vibrating magnetostrictive transducers, the first cylindrical metal resonator is brought to torsional resonance by means of tangential forces acting on its surface. The torsional vibrations are coupled to the second resonator by means of a thin torsional $\lambda/4$ line. This process is then continued along a series of resonators and coupling elements until at the output end the mechanical energy is again converted to electrical energy by two longitudinal transducers vibrating in antiphase, as at the input end. Figure 2 shows one reason why such filters have not been successful. In addition to the proper transmission range for torsional vibrations, further transmission ranges exist, which are caused by

† Telefunken Research Institute, Ulm, Western Germany.

Fig. 1. Torsional-mode single-sideband filter made by turning necks in a rod (with ferrite transducers).

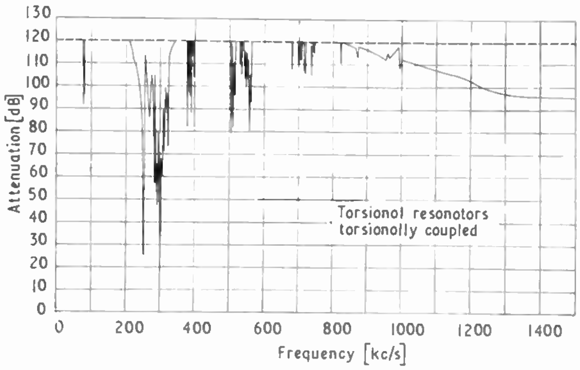


Fig. 2. Spurious responses in a filter of the type shown in Fig. 1.

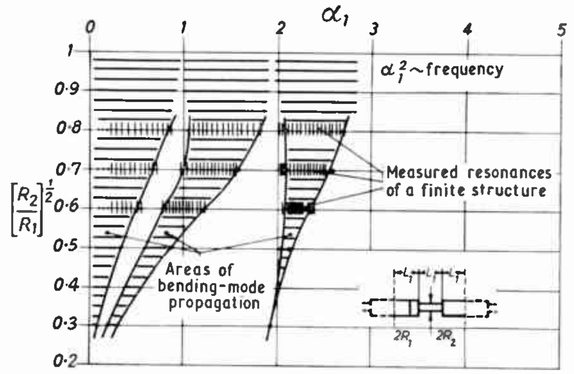


Fig. 3. Calculated bending-mode propagation in an infinite periodic structure.

unwanted resonances. The bending modes are particularly disturbing due to their large bandwidth. They cannot be suppressed by changing the cross-sectional dimensions, due to the following fact:

For torsional vibrations, the coupling between neighbouring resonators is given by $(R_c/R_r)^4$ where R_c is the radius of the coupling elements and R_r that of the resonators. This fourth-power dependence of coupling factor upon diameter ratio calls for relatively small diameter differences at the usual bandwidth for single-sideband filters. In complete contrast to this is the case where longitudinal vibrations are used. There the coupling only depends on the square of the diameter ratios, and for the same relative bandwidth greater diameter differences are needed. The coupling for bending modes in particular is then also much weaker.

Figure 3 shows the theoretical bandwidth of periodic bending-mode-lines as a function of diameter ratio.[6] With larger diameter differences the bending-mode propagation disappears, and the few remaining bending-mode resonances could be suppressed for example by using different diameters for the filter resonators.[7, 8]

Another difficulty with such one-piece filters still remains: Each individual resonator has to be tuned after turning, to a few cycles per second. This can be done for a particular resonator only when the neighbouring resonators are clamped (or alternatively very strongly detuned). This clamping method, however, no longer works at ultrasonic wavelengths of the order of 1 cm, because the required mechanical short circuits can no longer be realized.

Figure 4 shows a single-sideband filter, assembled with individual resonators.[7, 9] The resonators oscillate torsionally, while the coupling wires, which are spot-welded on the surfaces of the resonators and are a quarter of a wavelength long, vibrate longitudinally. Using this construction it is possible to manu-

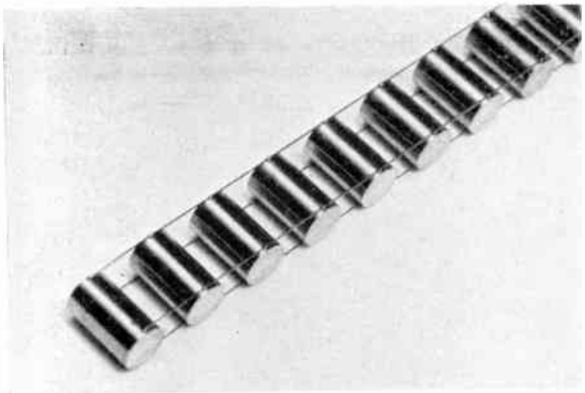


Fig. 4. Single-sideband filter structure using torsional $\lambda/2$-mode resonators coupled with longitudinally vibrating $\lambda/4$-wires.

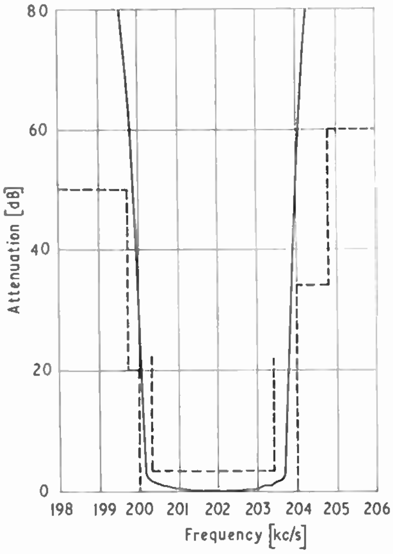


Fig. 5. Transmission curve of a FE 21 type single-sideband filter.

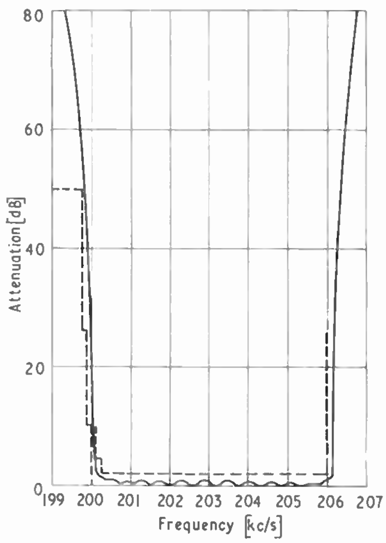


Fig. 6. Transmission curve of a FE 25 type single-sideband filter.

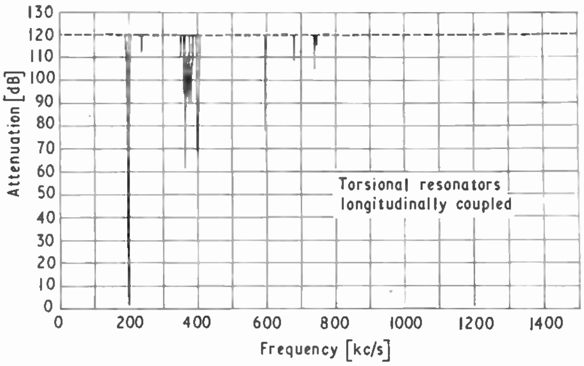


Fig. 7. Spurious responses in a FE 22 type single-sideband filter.

facture single-sideband filters with high precision and very few subsidiary modes. Figures 5 and 6 show transmission curves for two different types, first a filter for one single-sideband channel and then one for two single-sideband channels. Figure 7 shows the subsidiary mode characteristic. Because the driving forces for torsional vibrations are applied to the resonator surfaces, and not as moments in the middle of the end surface—as in the filter shown in Fig. 1—only small wire diameters are needed for the coupling elements. Bending modes are therefore practically not propagated. The remaining subsidiary modes come from λ-vibration of the coupling wires and harmonic overtones of the $\lambda/2$-torsional resonators. It is sufficient to drive this filter with an asymmetrically

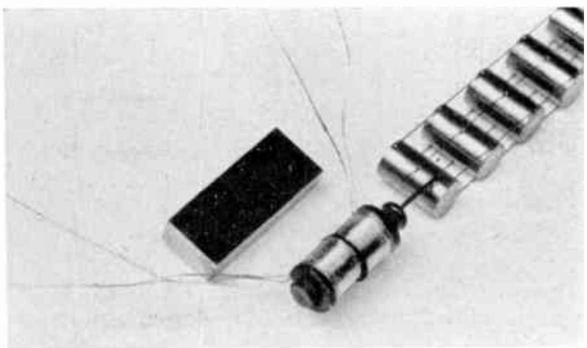


Fig. 8. The electromechanical transducer of a FE 21 type single-sideband filter.

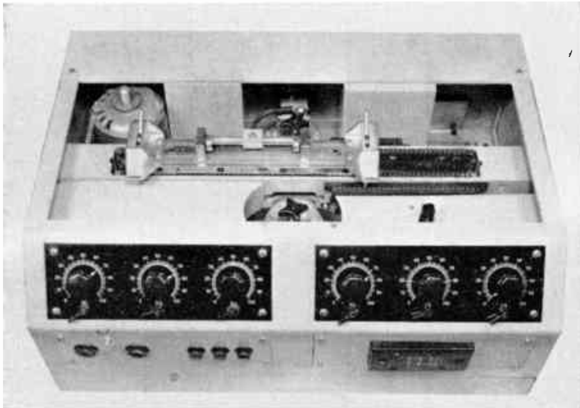


Fig. 9. Automatic spot-welding machine for single-sideband filters.

attached, longitudinally vibrating electromechanical transducer which consists of a ferrite rod, $\lambda/2$ long, biased by a permanent magnet (Fig. 8). The resonators are tuned in a largely automatized apparatus. The tuning accuracy is of the order of ± 3 c/s. The diameter tolerance of the coupling wires is small and is of the order of $\pm$ 1 μm. Figure 9 shows an automatic spot-welder, by means of which the coupling wires are welded to the resonators. The resonators are suspended in a plastic comb and are kept in position by additional wires.

A major problem in the construction of good filters which should also be mentioned is that the behaviour of the filter at different temperatures and the long-term constancy of its characteristics are almost entirely determined by the corresponding behaviour of the metal resonators and the transducer materials. Nickel-iron alloys are considered as resonator materials and it is possible, by making use of an adjustable interaction between magnetostrictive characteristics and internal stresses, to hold the resonance frequency constant to $\pm 1 \times 10^{-6}$ deg C in the working temperature range of about -40°C to $+80$°C. Unfortunately these materials mostly exhibit an ageing in their characteristics, and only by means of age-hardening using further additives (for example, beryllium, titanium, molybdenum) in the alloys, can usable materials be achieved. The internal stresses are adjusted by means of a correspondingly extensive cold-working, when the rod-shaped material is drawn. The hardening, which also has an influence on the temperature coefficient, takes place by means of a fairly precise annealing process at about 600° C. This guarantees, with good materials, a good temperature coefficient as well as a small ageing characteristic ($< 1 \times 10^{-7}$/week) and a high quality factor (between 15 000 and 25 000).

Figure 11 shows the behaviour with temperature of the resonance frequency of a torsional resonator (frequency 200 kc/s).

Figure 12 shows the behaviour of the resonance frequency shift with temperature for a cobalt-substituted ferrite which was specially developed for use in electromechanical filters.[10, 11, 12] The characteristics of the electromechanical transducer just discussed may deteriorate by about one order of magnitude, without influencing the filter characteristics. In addition to these characteristics, the coupling factor of the electromechanical transducer is of interest. It is important that this should be constant, as it determines the transformation of the electrical terminating resistance into a mechanical termination which has to be matched to the mechanical impedance of the filter.

Figure 13 shows the behaviour of the electromechanical coupling factor with temperature.

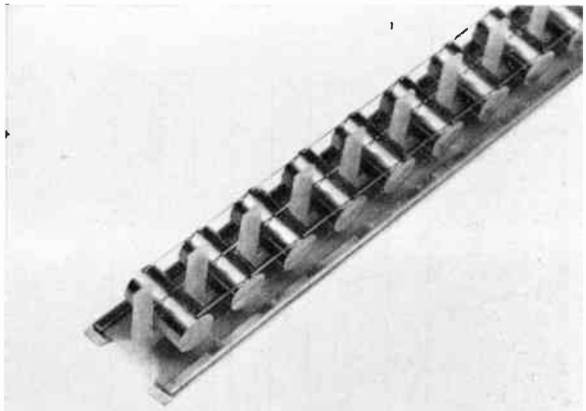


Fig. 10. Suspension of a single-sideband filter.

3. Miniaturized Filters

When the first electromechanical band filters were developed, no one foresaw the technique of miniaturization which has since become more and more successful in the field of electronics. Filter sizes were then in keeping with the dimensions of miniature tubes. It will be shown in this section how it has proved possible to miniaturize electromechanical filters. Some cautionary remarks with respect to the possibilities of miniaturization will also be made.

It may be seen that the filter quality depends materially on the quality of the resonators, which is in turn strongly dependent on the volume to surface ratio of these resonators. This is apparent from the consideration that disturbances in the resonator quality (ageing, temperature coefficient of the resonance frequency, Q-factor) emanate principally from

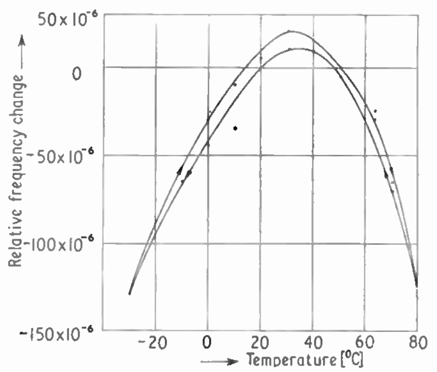


Fig. 11. Relative change of the resonance frequency of 200 kc/s torsional resonators with temperature.

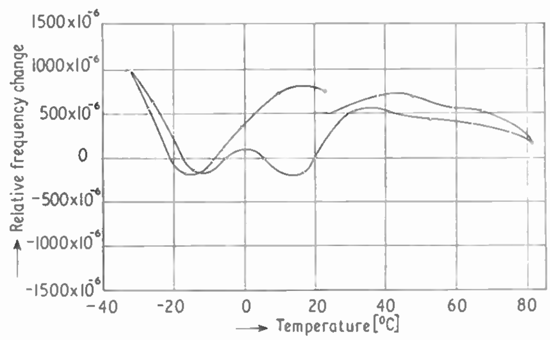


Fig. 12. Relative change of the resonance frequency of a 200 kc/s longitudinally vibrating ferrite transducer with temperature.

the surface in the form of changes in the crystal lattice and wandering of dislocations. Here one has to consider primarily the surface damage during the essential machining and tuning processes.

This then has two consequences:

(1) When geometrically similar resonators are compared, the lower the frequency the better the resonator.

(2) At the same frequency, the resonator whose shape comes closest to the spherical is best.

On account of spurious mode suppression and manufacturing considerations, as well as requirements for the filter structure as a whole (of which the resonator is always only a part), it is necessary to deviate from these somewhat oversimplified geometrical standards. Yet the fact remains that, for example, a resonator for a good single-sideband filter at 200 kc/s has better characteristics than one for 300 kc/s. The fact that, for equal *relative* quality (equal $T_k(f)$, equal relative ageing rate, equal Q-factor) the filter with the better *absolute* characteristics will always be achieved at the lower frequency has not even been taken into consideration here. With the resonator materials available at present, single-sideband filters with commercial quality can only be achieved at frequencies up to about 200 to 250 kc/s. A further substantial miniaturization of this type of filter is not contemplated at the moment.

The position is somewhat different in the case of normal i.f. filters for a.m. and f.m. receivers, which are also needed in large quantities, as well as filters for single-sideband transmission with simple requirements.

For these filters a midband frequency of about 455 kc/s is prescribed. At this frequency the greatest reduction in size can be achieved by selecting a filter structure in which both resonators and coupling elements are driven in flexure. However, apart from the quality deterioration already mentioned which takes place when the resonators become too small, and the fact that very small bending-mode transducers are hard to make, considerations with respect to spurious resonances in particular call for a different approach.

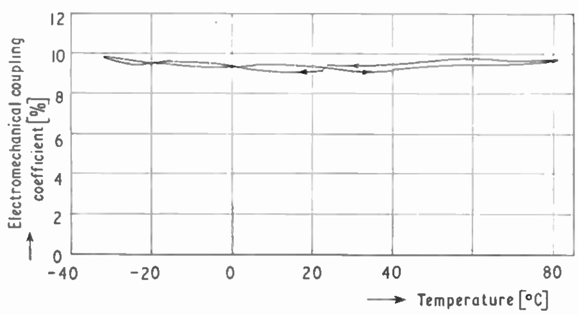


Fig. 13. Electromechanical coupling coefficient of a longitudinally vibrating ferrite transducer vs. temperature.

We therefore prepared to use $\lambda/2$-longitudinal resonators, which are easy to make and tune. Such resonators, coupled by means of short bending-mode leads, were used to build the miniature filter structure shown in Fig. 14. Several coupling wires fixed to the cylindrical resonator and transducer surfaces by spot-welding may be seen.[13] Once again the resonators are made out of a stabilized nickel-iron alloy and this offers the additional advantage that, using a coil and slight magnetic bias, the resonators can be checked by means of the magneto-strictive effect, so that for production tuning the resonance frequency can be determined accurately. Instead of the magneto-strictive ferrite transducers used previously, piezo-electrical transducers were used for this filter with obvious advantages.

Since the transducer now needs no driving coil, it fits very snugly into the filter structure. It consists of a small tube made of piezoelectric ceramic, with silver coatings on the inner and outer surfaces. The outer coating is strengthened galvanically to form a silver-nickel layer which can be welded. These coatings are also used as electrodes to polarize the transducer.

Figure 15 shows this transducer. The driving voltage is connected to the inner coating by means of a wire making contact at the nodal point for $\lambda/2$ longitudinal vibrations. The outer coating, which carries the coupling wires, is also the earthed counter-electrode. When an alternating voltage with the frequency of the $\lambda/2$ longitudinal resonance is applied, this resonance is excited in a roundabout way via the cross-contraction.

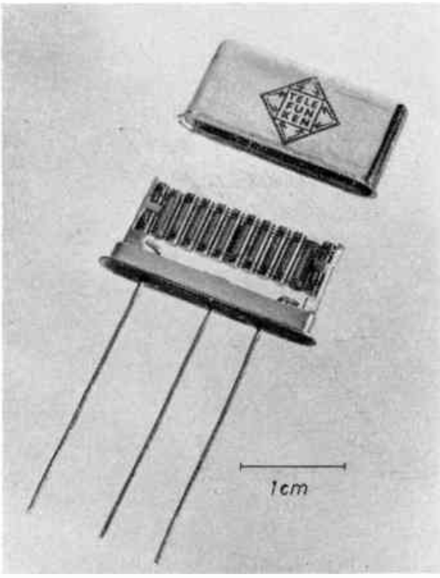


Fig. 14. Miniaturized electromechanical filter using $\lambda/2$ longitudinal resonators and bending-mode coupling wires.

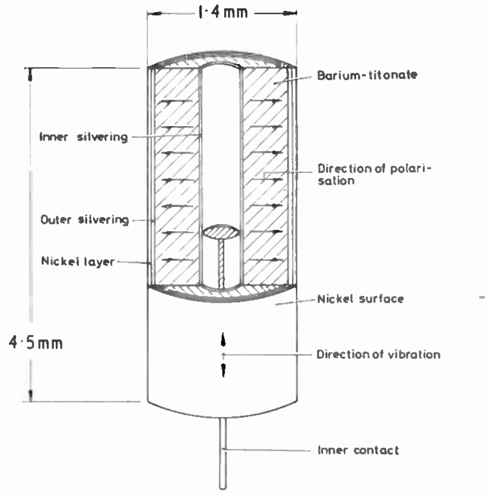


Fig. 15. The electromechanical transducer of a miniaturized filter.

The operation of the bending-mode coupling wires will be described in greater detail. It is well known that the bending wave has a propagation velocity which depends on the cross-section of the lead as well as on the frequency (where the lowest mode with respect to the cross-sectional dimension is always understood). Thus it is possible to make a bending-mode vibrator as short as desired at a given frequency. The same applies to a coupling wire which nevertheless should not show any resonances in the relevant frequency range (here approximately 455 kc/s).

The diameter of the coupling leads may be chosen in such a way that the resonators are packed together as tightly as possible. The length of the coupling leads is judiciously chosen as a quarter wavelength. This has the advantage that all the resonators can be pre-tuned to the midband frequency, even when the coupling between neighbouring resonators varies, corresponding to optimal dimensioning of the filter. The reason why the dimensioning becomes complicated is because the propagation of waves in a bending-mode vibrator is characterized by four components instead of two, as with longitudinal or torsional vibrators. These components vary along the length of the lead corresponding to the wave propagation.[14]

Figure 16 shows the basic element of a miniature filter. At the connecting point between bending lead and resonator the following bending values apply:

Displacement y

Displacement slope $\dfrac{\partial y}{\partial x} = y'$.

Total moment of the normal forces

$$N = EJ\frac{\partial^2 y}{\partial x^2}$$

(E = Young's modulus, J = second moment of area of the rod cross-section around an axis through the neutral bending plane and perpendicular to the vibration plane).

Shear force $S = EJ\dfrac{\partial^3 y}{\partial x^3}$ on the rod.

If the diameters of the longitudinal resonators and of the bending rods were optional, then all four values would have to be considered with respect to their influence on the vibrational behaviour. However, for relatively narrow-band filters we try to achieve very weak coupling only, so that we can assume that the bending leads must have a much smaller diameter than that of the resonators. Apart from this we wish to attach the thin coupling wires by electrical spot-welding. Both stipulations immediately lead to the simplification $y' = 0$, and this simplification further leads to the possibility of mathematically treating the bending coupling lead in complete analogy to the well-known torsional or longitudinal coupling leads. The bending coupling lead, which furnishes $\lambda/4$ coupling in particular, is then described by a chain-matrix with vanishing elements in the main diagonal.

For the length l_{K1} of the $\lambda/4$ coupling lead (the corresponding 'free' wave propagation of course does not exist!) one obtains[14]

$$l_{K1} = 0{\cdot}667\sqrt{R_B \cdot \lambda_L} \qquad \ldots\ldots(1)$$

and for the $3\lambda/4$ lead

$$l_{K2} = 1{\cdot}552\sqrt{R_B \cdot \lambda_L} \qquad \ldots\ldots(2)$$

R_B is the radius of the coupling lead and λ_L the wavelength for longitudinal vibrations in the bending lead at the filter's midband frequency.

The coupling provided between two $\lambda/2$ longitudinal resonators with radius R_L, denoted as $(K_{12})_1$ for $\lambda/4$ coupling or $(K_{12})_2$ for $3\lambda/4$ coupling, is given by[14]

$$(K_{12})_1 = 1{\cdot}396\,\frac{R_B^{\frac{5}{2}}}{R_L^2\lambda_L^{\frac{1}{2}}} \qquad \ldots\ldots(3)$$

or

$$(K_{12})_2 = 1{\cdot}605\,\frac{R_B^{\frac{5}{2}}}{R_L^2\lambda_L^{\frac{1}{2}}} \qquad \ldots\ldots(4)$$

The method of calculation has proved successful. A filter with midband frequency of 455 kc/s was built, for example, with a calculated bandwidth of 5·2 kc/s. The measured bandwidth was 4·9 kc/s. The validity of these relationships, however, reaches a limit when plate-shaped resonators are considered.

The transmission characteristics of this filter are of course not only determined by the vibration types used for the resonators and coupling leads, which only come into play in the actual transmission range. This already emerged clearly from the remarks about single-sideband techniques. The bending-mode resonances of the longitudinal resonators, as well as bending- and longitudinal coupling of these bending vibrations to one another, have also proved to be important for the behaviour of miniature filters in the neighbourhood of the passband.[15]

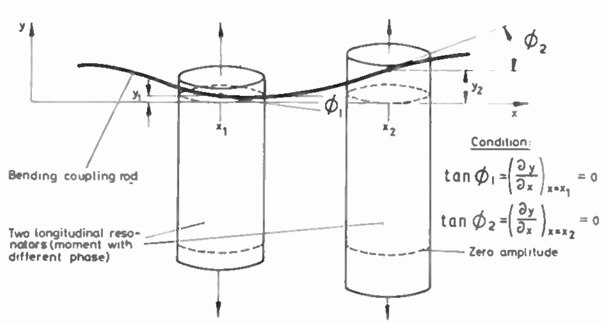


Fig. 16. Two longitudinal resonators with bending-mode coupling.

A most important way of suppressing these spurious resonances, which could make a filter completely useless, consists in making the electromechanical transducer sensitive to the longitudinal mode only. However, it is not possible to achieve more than 20 to 30 dB spurious mode suppression with this method due to small residual unsymmetry, which cannot be completely avoided in miniature filters. The only remaining way out is to look for arrangements whereby the bending modes of the resonators are only coupled very weakly. If, for example, one attaches the bending coupling wires to those positions on the longitudinal resonators where the bending modes of the resonators have their nodes, then the coupling approaches zero. In Fig. 17, R_s denotes a value which is proportional

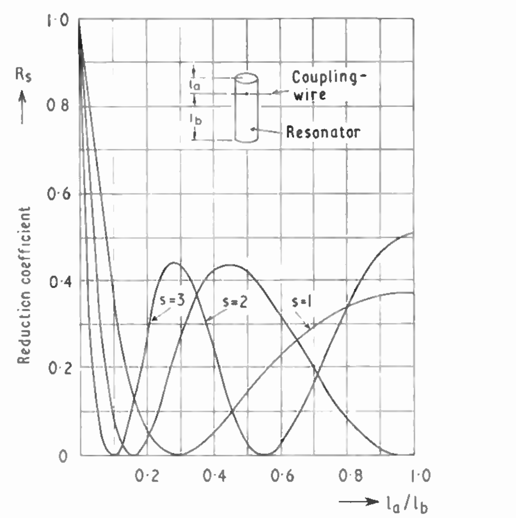


Fig. 17. Reduction of the coupling of undesirable bending-modes by changing the point of connection between coupling wires and longitudinal resonators.

to the coupling of bending resonances (with $s+1$ nodes) for the longitudinal vibrator. For the position of attachment of the coupling wires given by $l_a/_bl = 0{\cdot}13$, optimum decoupling is achieved for the second ($s = 2$) and the third bending harmonic ($s = 3$).[13] This is also the case which is of most interest in practice. With arbitrary increase of the diameter of the longitudinal resonators (at constant longitudinal resonance and therefore constant length except for the Rayleigh correction factor), the first bending mode always lies far below the $\lambda/2$ longitudinal resonance, because, according to Timoshenko's theory, the bending wave velocity soon becomes independent of the diameter. The second and third bending harmonics ($s = 2$ and $s = 3$) however, can already lie in the immediate vicinity of the actual transmission band, so that the adjustment of $l_a/l_b = 0{\cdot}13$ becomes technically very interesting. The use of so thin a longitudinal resonator such that very high bending modes ($s > 3$) occur in the neighbourhood of the transmission band, is in any case not to be recommended, as with the higher orders the relative frequency separation of bending resonances, and therefore also the spurious mode separation, becomes ever smaller.

To sum up, one can say that by this method, neighbouring resonators can be decoupled, with regard to the second and third bending mode, to the order of 1 : 40.

The question of how great the coupling of bending modes actually is, and how great this is in comparison with the coupling of the main mode in the actual transmission band of the filter, is also interesting. For the sake of comparing the bandwidth for spurious resonances to the main transmission bandwidth, we define, as ratio of the two bandwidths the spurious mode ratio N. This is at the same time a measure of the suppressibility of spurious resonances. If this ratio N is small in relation to 1, then a further method of improving the spurious mode suppression may be used. This method does not work when $N > 1$, (as has already been noticed with single-sideband filters). By means of small changes in geometry, which do not disturb the main transmission characteristics (for example, small changes in resonator diameters, or using small tubes instead of solid cylinders), it is possible to achieve a condition where the bending resonances of the few remaining spurious modes are detuned more strongly than they are coupled. If $N \ll 1$, the production spread usually suffices to let the spurious modes vanish completely.

The most disturbing spurious mode spectrum in our case is caused by the second and third bending modes of the longitudinal resonators, where the bending vibrations are polarized in the plane of the filter, and the energy is transmitted by the thin coupling leads through longitudinal drive. In this case the spurious mode ratio N is calculated as

$$N = VF_r \frac{1{\cdot}41}{m} \sqrt{\frac{\lambda_L}{R_L}} \sqrt{\frac{\lambda'_L}{R_B}} \qquad \ldots\ldots(5)$$

In this equation m has the values $m = 7{\cdot}85$ (for $s = 2$) and $m = 11{\cdot}00$ (for $s = 3$) respectively. λ_L is the wavelength for longitudinal waves in the middle of the main transmission band and λ'_L the wavelength for longitudinal vibrations in the middle of the relevant spurious mode spectrum (for example at $s = 2$ or $s = 3$). The quantity V is found from

$$V = \frac{1}{\sin(2\pi l_{kr}/\lambda'_L)} \qquad \ldots\ldots(6)$$

and F_r equals $1{\cdot}14$ or $0{\cdot}99$ according to whether respectively $\lambda/4$ or $3\lambda/4$ bending coupling is used for the main transmission band.

For the filter shown in the following example, we have

$$l_k = 1{\cdot}7\,\text{mm}\,(3\lambda/4\,\text{coupling}),$$
$$\lambda_L \simeq \lambda'_L = 10\,\text{mm}$$
$$2R_L = 1{\cdot}5\,\text{mm}, \qquad 2R_B = 0{\cdot}1\,\text{mm},$$
$$V \simeq 1{\cdot}25 \quad \text{and} \quad F_r \simeq 1.$$

For $m = 7{\cdot}85$ N becomes 6 and for $m = 11{\cdot}00$, N becomes 4. The spurious mode ratio is therefore > 1 in both cases, and only by reducing the coupling by choosing the point of attachment of the coupling wires to the resonator so that $l_a/l_b = 0{\cdot}13$ (optimum for $s = 2$ and $s = 3$) can a point be reached where $R_s = 1/40$, so that the effective spurious mode ratio $N' = R_sN$, becomes appreciably smaller than 1. We did in fact find, experimentally, that only a few spurious modes with suppression > 60 dB occurred, while without these measures to suppress them, very wide bands of spurious resonances readily occurred. Figure 18 shows the transmission curve of an I.f. filter, (type FZ 65) with a bandwidth of 6 kc/s.

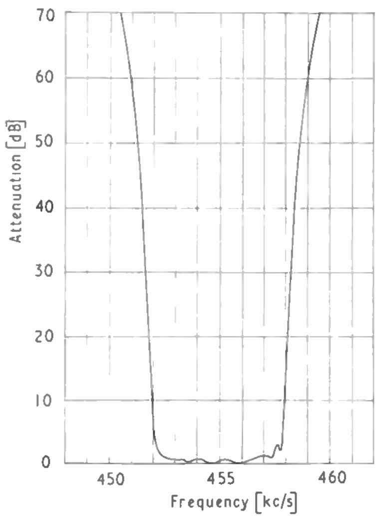


Fig. 18. Transmission curve of a FZ 65 type miniaturized i.f. filter.

4. Further Developments

The newer developments shown thus far are essentially complete. What further improvements can still be expected in this field? A comparison with conventional filter techniques shows that a broad category of

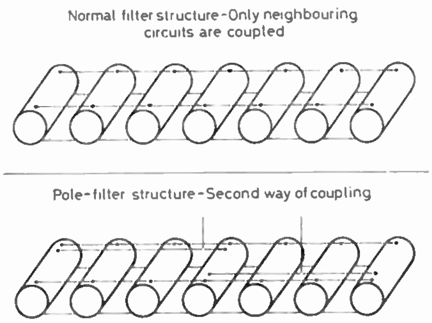


Fig. 19. Realization of pole-filters by two-way coupling.

filters have up to the present not yet been realized mechanically. The mechanical filter types realized up to now correspond to simple coupling filters. How can one realize filters with attenuation poles and non-minimal-phase filters? There has been a whole series of interesting experiments in this field, dating back several years. Mason[16, 17] was the first to try such experiments. Unfortunately all these approaches failed in their practical realization, principally on account of a large amount of uncontrollable spurious modes. Lukas[18] describes how a second resonance of the same resonator may be used to achieve definite attenuation poles. We found a simple way to build filters with attenuation poles by not only coupling neighbouring circuits to one another, but also to non-neighbouring circuits by secondary coupling leads.[19] (Fig. 19.)

It is surprising that the relatively long coupling leads needed for several apparently possible structures (Fig. 19 shows only one example of many) are completely free of spurious modes, although one would expect them to have a multitude of bending resonances and therefore closely spaced spurious modes. It is true that with the excitation of the longitudinal mode in long thin wires (as in the example of Fig. 19) a large number of bending resonances exist. The characteristic impedance, however, that can be assigned to these resonances, differs so strongly from that of the substantially more massive resonators that have to be coupled that these resonances are practically not coupled at all to the actual energy exchange in the filter. Figure 20 shows a picture of a realized pole filter, and Fig. 21 shows the measured attenuation curve of such a filter. Trceva treats filters with secondary coupling as pure electrical coupling filters.[20] Fairly concrete concepts about the realization of all-pass filters, with which the delay curve may be adjusted independently of the attenuation curve,

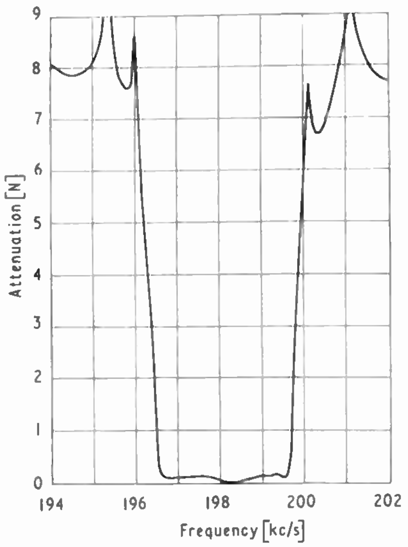


Fig. 21. Transmission curve of a pole-filter using two-way coupling.

already exist. Figure 22 shows such an element with all-pass characteristics.

Brief mention may be made in passing of filters, which are obtained by making notches in plate-shear mode quartz sheets. It is likely that such filters

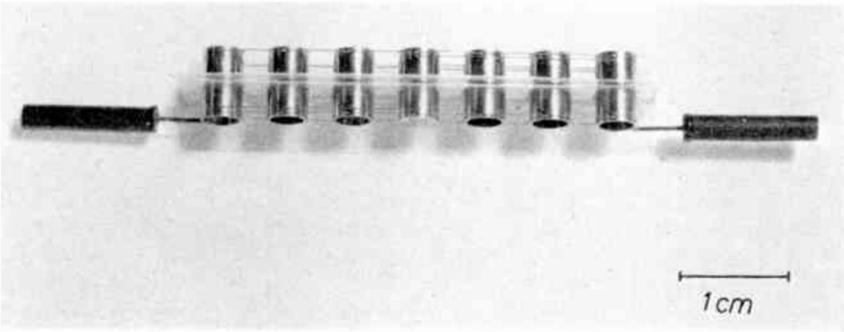


Fig. 20. A pole-filter using two-way coupling.

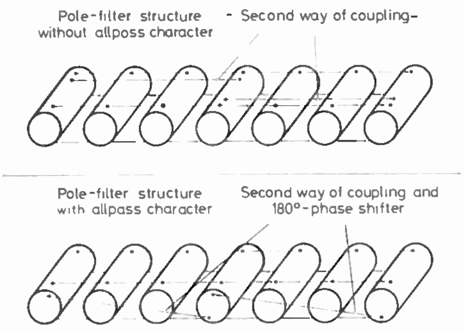


Fig. 22. Realization of non-minimal-phase mechanical filters with all-pass portion.

can only be manufactured with a small number of circuits.

A further demand is for filters of the present quality, but for frequencies above one megacycle. The requirements set, however, regarding their behaviour with temperature and above all their Q-factors cannot be fulfilled with present materials. A few interesting trial experiments which have been undertaken with a view to mastering the possible vibrational structures at such high frequencies[22] make little reference to these requirements in fact and therefore do not point the way towards realizing the present quality at still higher frequencies. Nevertheless it is possible that new applications will be found for such mechanical resonators in electrical filters.

5. References

1. W. van B. Roberts and L. L. Burns, "Mechanical filters for radio frequencies", *R.C.A. Review*, **10**, p. 348, September 1949.
2. R. W. George, "Electro-mechanical filters for 100 kc carrier and sideband selection", *Proc. Inst. Radio Engrs*, **44**, p. 14, January 1956.
3. G. L. Grisdale, "Electro-mechanical filters for use in telecommunication equipment", *Brit. Commun. Electronics*, **6**, p. 768, November 1959.
4. H. Bache, "A practical electromechanical filter", *Marconi Review*, **22**, No. 134, p. 144, 1959.
5. W. Struszynski, "A theoretical analysis of the torsional electromechanical filters", *Marconi Review*, **22**, No. 134, p. 119, 1959.
6. M. Börner, "Biegeschwingungen in mechanischen Filtern", *Telefunken Z.*, **31**, p. 115 and p. 188, 1958.
7. M. Börner, E. Kettel and H. Ohnsorge, "Mechanische Filter für die Nachrichtentechnik", *Telefunken Z.*, **31**, p. 105, 1958.
8. H. D. Pieper, "Mechanische filter", *Telefunken Z.*, **32**, p. 279, 1959.
9. M. Börner, E. Dürre and H. Schüssler, "Mechanische Einseitenbandfilter", *Telefunken Z.*, **36**, p. 272, 1963.
10. C. M. van der Burgt, "Ferrites for magnetic and piezomagnetic filter elements with temperature independent permeability and elasticity", *Proc. Instn Elect. Engrs*, **104B**, Supplement No. 7, p. 550, 1957. (I.E.E. Paper No. 2287R, 1957.)
11. C. M. van der Burgt, "Controlled crystal anisotropy and controlled temperature dependence of the permeability and elasticity of various cobalt-substituted ferrites," *Philips Res. Rep.*, **12**, p. 97, 1957.
12. S. Schweitzerhof, "Über Ferrite für magnetostriktive Schwinger in Filterkreisen", *Nachrichtentechn. Z.*, **11**, p. 179, 1958.
13. M. Börner and H. Schüssler, "Miniaturisierung mechanischer Filter", *Telefunken Z.* (To be published.)
14. M. Börner, "Mechanische Filter mit Biegekopplung", *Archiv Elektr. Übertrag.*, **15**, p. 175, 1961.
15. M. Börner, "Zum Nebenwellenverhalten von Mikromodulfiltern (Biegenebenwellen in biegegekoppelten Longitudinalfiltern)", *Archiv Electr. Übertrag.*, **16**, p. 532, 1962.
16. W. P. Mason, U.S. Patent 2,345,491 (1944).
17. W. P. Mason, "Physical Acoustics", I Part A, (Academic Press, New York, 1964).
18. J. Lukas, "Plattenförmige elektromechanische Filter mit Dämpfungspolen", *Archiv Elektr. Übertrag.*, **17**, p. 230, 1963.
19. M. Börner, "Mechanische Filter mit Dämpfungspolen", *Archiv Elektr. Übertrag.*, **17**, p. 103, 1963.
20. E. Trceva, "Hochfrequenzbandfilter mit Dämpfungspolen bei endlichen Frequenzen und geebnetem Betrag im Durchlassbereich", *Nachrichtentechnik*, **12**, No. 12, p. 450, 1962 and **13**, No. 1, p. 36, 1963.
21. W. Poschenreider and F. Schöfer, "Das elektromechanische Quarzfilter—ein neues Bauelement für die Nachrichtentechnik", *Frequenz*, **17**, No. 3, p. 88, 1963.
22. E. Gikow, A. Rand and J. Gianotto, "Functional circuits through acoustic devices", *I.R.E. Trans. on Military Electronics*, **MIL-4**, No. 4, p. 469, 1960.

Manuscript received by the Institution on 12th October 1964. (Paper No. 968/EA18.)

DISCUSSION

Under the chairmanship of Mr. C. T. Chapman

Mr. S. Kelly: In view of the extremely close limits for resonant frequency of the individual resonators which in turn call for comparable mechanical tolerances (i.e., ± 1 part in 100 000) I would be grateful if the author could indicate how he obtains these mechanical limits in practice. Or, if adjustment is subsequently applied, how is this performed?

Dr. M. Börner (*in reply*): The tolerance of ± 3 c/s at a frequency of 200 kc/s is obtained by an individual grinding process. In successive steps the resonator's frequency is measured, it is ground, measured and ground again and so on until the frequency lies within the guaranteed values. For this process we use special machinery.

Mr. G. C. W. Harvey: We have heard a great deal from Dr. Börner about the fact that the lower the frequency the easier the design problem of electromechanical filters. At what lower frequency limit does this type of filter cease to be attractive compared with conventional components and what are the design problems at these lower frequencies?

Dr. Börner (*in reply*): We have built mechanical filters down to 30 kc/s. I think this is the frequency range at which filters made with inductors have the same absolute stability as mechanical filters. Filters using inductors and capacitors have the advantage that they can be designed and built in small numbers much more cheaply than mechanical filters, which only have advantages if great numbers of the same type are required. The main design problem is how to get a good resistance against the creation of an electrical output caused by mechanical shock. The other problem is how to match the relative very broad transmission bands electrically by electromechanical transducers. PZT-ceramics with a high electromechanical coupling coefficient seem to have this possibility down to some kc/s.

Mr. H. Zakaria: The paper states that in perfecting the spot welding process all requirements of mechanical filters cannot be met. The impression I get is that the spot welding operation is unsuited for a filter with a tight attenuation-frequency characteristic requirements. Does the above statement imply this? The paper also states that automatic equipment achieves resonators tuned to $\pm$ 3 c/s. I would like to know what sort of accuracy is achieved and the tolerances allowed in the spot welding technique to ensure that the accuracy of machining is not upset. It would also be interesting to know how much the spot welding process detunes the resonators and if anything is done about this detuning.

I am pleased to note that Dr. Börner gave some thoughts to the problem of spurious responses. The bending mode was mentioned as being the main cause of spurious responses on the one-piece torsional filter body with push-pull drive. I understand by a bending mode that the body under consideration is excited in flexural vibrations. I would like to bring to the attention of Dr. Börner that Mr. W. Struszynski† attributes the difficulties that were encountered with spurious responses on a very similar torsional filter design mainly to the transverse shear mode and the differential transverse shear mode. There is clearly a need for more thought to be given to the nature of the mode of vibration causing these spurious responses.

In Fig. 2 Dr. Börner showed a level of 120 dB discrimination in the measurement of spurious responses, and it would be interesting to know the input at which the filter was driven.

It is apparent from Fig. 13 relating the variation of the electromechanical coupling with temperature, that the ferrites used in Dr. Börner's transducers are very stable at about 9% or 10% coupling. I am rather curious to know more about the stabilization process used on the ferrite sample. Has the graph, for instance, been drawn after any particular heat treatment or demagnetization of any sort? I would also like to know if there is a simple ratio between the stable value of electromechanical coupling and its initial value when the ferrite is fully saturated.

Mention has been made of the dependence of the filter quality on the surface quality of resonators. This would obviously upset the advantages of the heat treatment if the resonators were to be tuned. I would like to know if the order of this effect during the tuning operation of the resonators is known.

In mentioning plate-shaped resonators Dr. Börner referred to quartz sheets. Whilst his remarks may be true of these materials, I wonder if they also apply to metallic plate-shaped resonators. The possibility of manufacturing filter bodies in the shape of a punched metallic strip has been suggested recently by A. K. Losev.‡ Mr. Losev also suggested that by using higher order electromechanical iterative filters, the working frequency range can be increased to higher than one Mc/s. Would Dr. Börner like to comment on this type of filter in view of its relatively easier mass production?

Dr. Börner (*in reply*): I think the spot welding process is the best known one for assembling resonators to a filter structure. The tolerances of the tools used to assemble resonators before spot welding are smaller than 1/200 mm. The detuning of a resonator for 200 kc/s single sideband filters by one welding spot is 10 c/s $\pm$ 1 or 2 c/s.

Spurious frequencies in the filter described by Mr. W. Struszynski occur by bending vibrations of the whole structure. The explanation that these spurious modes are caused by shearing is not right, for this type of vibration does not exist as a free vibration, because it breaks the law of conservation of angular momentum. But the normal theory of bending vibrations of periodic bending structures is only valid for mechanical filters with ratios of diameter to length of resonators smaller than about 1 : 20. In the case of Marconi filters this theory only gives the rough properties of the bandwidth of the spurious modes (and their nature) but not the right frequencies. The calculation in this case becomes more complicated, because Timoshenko's theory of bending is more complicated than Lord Rayleigh's theory. But we have measured the continuous transition of bending-mode spurious frequencies in structures with small diameter (they are in good accordance with our theory) to structures with the same length (the same torsional frequency) but with higher diameter as used in the Marconi filter.[6] So we are sure that the spurious modes belong to bending vibrations.

The input level for measuring the 120 dB level in the stop band was 3 V at 600 ohms.

There is no simple ratio between the stable value of the electromechanical coupling coefficient and its initial value when the ferrite is fully saturated. The temperature stabilization is effected by addition of cobalt and not as in metallic resonators by additional strain treatment with a specific temperature dependence.

Heat treatment has a great influence not only on the temperature coefficient, long term stability and the Q of the

† W .Struszynski, "A theoretical analysis of the torsional electromechanical filters", *Marconi Review*, **22**, No. 134, pp. 119–43, 1959.

‡ A. K. Losev, "Multi-element couplers in filters", *Elektrosuriaz*, No. 1, pp. 1–8, January 1964.

resonator but also on its frequency. Because this effect is not very reproducible no heat is allowed after tuning.

It is possible to make plate-shaped metallic resonators at frequencies higher than 1 Mc/s. But the Q of metallic resonators, the temperature coefficient, the stability of frequency and the reproducible behaviour of spurious modes are not adequate to make very good filters in this way at present.

Mr. J. Birch: Dr. Börner has said that mechanical filters are easier to construct at the lower frequencies but limitations are imposed by the ability of such filters to withstand mechanical shock. Considering the use of some of these filters to replace quartz crystal filters in telephone communication systems, are not the problems of frequency ageing and unwanted resonances of greater significance than mechanical shock in such an application? Would Dr. Börner care to amplify his remarks concerning the ageing properties of the mechanical filter and also explain the passivation technique used to lower the ageing rate?

Dr. Börner (*in reply*): Part of your question has already been covered in my previous answers. Of course at frequencies of 60 kc/s and more, the problem of suppressing the spurious modes is very interesting. But I think this problem has been solved for some special types of mechanical filters with bandwidths not greater than 4 kc/s.

The passivation is achieved by preventing the dislocations from travelling. The tendency to travel is higher with higher inner strains in the metal. By the heat treatment these strains are lowered but unfortunately, because of their dependence on temperature coefficient, not down to zero. This would be the state with highest entropy and thus the most stable one. The more unstable state with remaining strain is stabilized by means of a precipitation hardening (stability better than 1 in 10^7 per week).

Mr. R. K. T. Galpin: What is the effect of spot-welding the support wires in the middle of the resonator? I should have thought that this was undesirable.

Dr. Börner (*in reply*): Indeed spot welding of supporting wires in the middle of the resonators is undesirable. But this technique has so great advantages with regard to supporting the filter that nevertheless we have used this method. Its effect on the stability is tolerable.

Mr. S. J. Partridge: Has Dr. Börner considered using a high remanence ferrite for the transducer resonator thus removing the need for a separate polarizing magnet?

With regard to the 200 kc/s s.s.b. filter described, what is the approximate level of third-order distortion products arising from frequencies in the pass band? How does this compare with the same distortion arising in the pass band from intermodulating stop band frequencies?

Dr. Börner states that after heat treatment and strain relief, machining of the resonator surfaces near to vibration anti-nodes to tune the resonator results in considerable loss of long term and temperature stability. Can this effect be overcome by a further strain relief treatment after tuning, the detuning of the resonator due to this treatment being allowed for at the tuning stage?

Would Dr. Börner like to comment on the stability of the piezo-electric coupling and the resonant frequency with temperature of the barium-titanate transducer used in the 455-kc/s i.f. filter described?

Dr. Börner (*in reply*): Ferrite transducers with a high enough coercive-force and a good electromechanical coupling coefficient do not exist.

Only passband-passband intermodulation is of interest, if two speech channels are to be transmitted in the same filter (AIII/b-modulation). The cross-talk attenuation between two channels within the same passband of a filter (Telefunken type FE 25 and FE 26), is approximately 60 dB at a level of $-22{\cdot}5$ dB.[9]

The long term stability is effected by machining the resonator's surface near the vibration-nodes. To get a good resonator, the frequency has to be the right one, the inner strain must have the right level (not zero!) for a good temperature coefficient and by precipitation hardening, using additives like beryllium, titanium or molybdenum, a good Q must be established. I think the best way to get this is to carry out first the heat treatment for a low temperature coefficient and a high Q (and a low ageing rate) and then the tuning process without damaging the sensitive part of the surface of the resonator. Any other succession of processes do not seem to be reproducible.

We now use PZT-ceramic in the transducers of our miniature-filters with an ageing rate of about 5% for the coupling coefficient and of $0{\cdot}1$% for the frequency during the first year of ageing (later on it is smaller).

U.D.C. 621.37/9.002

The Design Principles and Rôle of a Comprehensive Unit System of Electronic Equipment

with particular reference to the Harwell 2000 System

By

H. BISBY, B.Sc., A.K.C.†

Summary: The development of a unit system of electronics for a large research organization has evolved design principles which are thought to apply to other comprehensive unit systems. These principles, and the benefits which attend such systems, are reviewed.

1. Introduction

Electronic instruments which measure, analyse and record nuclear events are being used in increasing numbers and are employing techniques of increasing complexity. This trend is common in other fields of instrumentation, and there is now a growing awareness that the progress of much research is limited by the capability, flexibility in application, and the reliability of instrumentation. The pressures, therefore, in the discipline of electronics to contrive improvements in these and other respects (compactness, costs, heat generation) are probably greater now than ever before.

The achievement of such improvements is, however, not without its contending factors. On the one hand, complexity in design and expertise in the use of new electronic devices and techniques is implied together with sound development, prolonged trials and effective control over production. On the other hand, there is an acute shortage of electronic designers and technicians and a demand from the scientist for immediate availability of contemporary instrumentation.

The electronics team of any large research organization, which depends extensively on instrumentation, must come ultimately face to face with these problems and realize that to continue the practice of designing instruments specific to each requirement is not only inefficient in terms of time and effort of the team, but also prohibitive in terms of capital and maintenance costs.

This situation became evident some few years ago in the Atomic Energy Research Establishment, Harwell, and the expedient adopted by the Electronics Division took the form of a comprehensive unit system of electronic equipment.

The major instrument requirement in the Establishment is to detect, measure and analyse radiations from nuclear events and the task of the unit system is to provide a range of basic units from which can be derived general purpose, and also specific, electronic complexes having advanced technical features and a degree of flexibility in use, not hitherto available with laboratory type instrumentation.

The completion of this task is now within sight and three years of operational experience have been obtained. It is therefore appropriate to review the design principles, which have developed with this growing experience, the extent to which these have been applied and the effects which the unit system has had both on the work of the Establishment and on the Electronics Division itself.

2. Design Principles of a Comprehensive Unit System

An electronic complex can usually be resolved into an assembly of simpler parts, each having an assigned function, and it is common practice to design and construct these parts as separate physical units. In this way the complex can be built up from units which have a size and weight suitable for handling. These units, however, must be capable of being interconnected in respect of electrical power supplies and signals, i.e. have some measure of electrical compatibility.

The term 'unitized' is universally applied to instrument systems which extend this breakdown principle by stipulating that units should have only one function and that the physical form of each unit should follow a modular pattern in one, or more than one, of its linear dimensions, for example front panel height or width. This introduces the concept of mechanical compatibility of units within an overall system.

Many varieties of such unit systems exist because modular unit construction enables complexes of varying size to be assembled within a pre-designed mechanical framework which may be extended or contracted to suit the size of the complex. Furthermore, the diagnosis of a failure is made easy and replacement of a faulty unit rapidly effected, thereby contributing to a substantial reduction of 'down-time' of the complex and the load which a 'first-line' maintenance engineer would otherwise have to bear. The modern computer is a typical example which exploits these advantages, the modular units being separate, or groups of, logic functions arranged

† U.K. Atomic Energy Authority, Electronics Division, Atomic Energy Research Establishment, Harwell.

physically in a specific order to provide a complex with a specific task.

Some unit systems have the added feature that units having the same function, but with different technical characteristics, are readily interchangeable one with the other thus providing a measure of flexibility in the operational role of a complex.

Unit systems with only the simplest forms of mechanical and electrical compatibility between units can be justified for the reasons outlined above when large and expensive complexes are involved, whether these are permanent or semi-permanent installations. In addition, the acceptance of a module frame reduces the engineering design effort which is needed for non-modular units and their assembly in groups.

For smaller complexes, the case for a unit system is less convincing on these grounds since unit frames and frameworks, needed for assembly, introduce extra costs, and compatibility rules (i.e. standardization) may impose restrictions on design freedom. However, in a research organization, where both a large number and variety of large and small complexes are required, with technical and operational characteristics which are subject to frequent changes, and where a large team of designers is involved, an even more comprehensive approach to unitization can be justified.

The justification is in proportion to the number of functions which the various complexes have in common. One such universally common feature, for example, is a.c. to d.c. power conversion. In nuclear instrumentation there are many others such as direct-voltage generation for polarizing detectors, digital pulse generation for timing purposes, analogue pulse amplification and analysis. It is therefore a class of instrumentation which is amenable to unitization and likely to show most benefit since the variants of a particular function can be considerable. A unit having one of these common functions will be referred to as a 'basic' unit. There will, however, be 'special' units designed specifically to extend the application of a complex of 'basic' units.

A comprehensive system of 'basic' and 'special' units will be one possessing a high degree of flexibility in both technical performance and operational usage, and the following design principles have been evolved for such a system:

(i) The engineering aspect of the system, i.e. the unit-frame and the framework for assemblies, should be designed on a modular dimensional pattern.

(ii) Any unit having one or any number of module dimensions should be capable of fitting into any one station of appropriate dimensions in the framework.

(iii) Any unit should be easily and rapidly replaceable by one of like physical dimensions although the latter may or may not have a similar function.

These principles demand mechanical standardization of unit frames and framework with rigid adherence to dimensional tolerances for all time so that a production batch of units can be guaranteed to fit into frameworks produced many years earlier, or vice versa. They provide complete flexibility in the spatial disposition of one unit relative to others—this often being an operational requirement of some importance—and the advantages already mentioned of complete interchangeability.

(iv) A.c. and a full range of d.c. supplies should be available at each station and only be connected to or disconnected from a unit when it is plugged into or withdrawn from a station.

(v) Each 'basic' unit should be unifunctional.

(vi) Each 'basic' unit should be self-contained and therefore require no intermediate components, apart from cable assemblies, to connect it to the power lines or other units.

These principles imply electrical standardization of the a.c. and d.c. power lines, in respect of voltage, ripple, stability, regulation and noise, and also pin allocations for inter-connections. The former will ensure a supply system which can be guaranteed to designers of units and the latter give rise to electrical compatibility of all units of the system. From the operational point of view, (iv) removes any possible hazard which could arise when interchanging units and (vi) should make it a simple operation for a user, with little knowledge of electronics, to assemble complexes or interchange units.

(vii) Ventilation of any unit, for the purpose of heat removal, should be automatically incorporated in the system engineering and not rely on being provided at the discretion of unit designers, nor made the responsibility of users.

(viii) All facilities, to which a user may have need of access for normal operation, calibration or standardization, should be available without having to withdraw a unit from its station.

Inadequate ventilation of electronic equipment is a large factor contributing to unreliability, yet inadequate ventilation still occurs because designers are preoccupied with problems on other design aspects, or do not make due allowance for conditions which can be generated, in all innocence by users, when units are associated in a complex. The solution therefore is to design the modular system, and impose design principles on each unit, such that any unit can operate within specification no matter with what units it is associated or how these are disposed relative to it. Thereby, again, a user has freedom to choose the disposition which suits best his operational requirement for controls, indicators and connectors.

(ix) Signals, required for passing control or data information between units, should be standardized to ensure information compatibility between units and avoid interposing external buffers, convertors or translators between units in the system.

For the user, this principle again simplifies the problems of assembling a complex. It also makes possible the design of one unit without knowing, in great detail, the units with which it might be required to operate within a complex. This feature will often simplify a design problem.

(x) The controls, indicators and connectors should be designated according to a system code of practice.

This principle covers an operational factor which is seldom given the consideration which it deserves. A completely flexible unit system should be designed such that any user, assembling the 'basic' units to form a complex to his requirement, should only need to know what the unit does and not necessarily how it does it. The provision of a fully detailed literature specification is not the complete answer, although it is obviously a very necessary requirement. The solution rests in providing adequate information near the controls, indicators and connectors. It is important therefore that symbols or terms have a clear significance throughout. This principle can be applied by rigid adherence to a glossary of symbols and terms, based on international or national standards; where no such standards exist in a special terminology, the glossary should clearly define the symbols or terms. The designation code should also include general rules, to be applied wherever possible, on the disposition of controls, indicators and connectors, and even the type and style of lettering to be used. This may appear to the casual observer to be taking standardization too far, yet it is a feature which gives a complex the appearance of uniformity and neatness and demonstrates that all units have been designed both thoroughly and with a sense of purpose.

3. Design Features of the Harwell 2000 Series Unit System

The design principles have been evolved as experience with this unit system has grown over the past seven years, particularly during the last three years. They now apply fully to the Harwell 2000 system which provides general-purpose, and also specific complexes of laboratory type nucleonic equipment. The 2000 series units are modules which plug into a small-assembly framework, termed a shelf. This shelf, in turn, can be fitted into a large-assembly framework, termed a rack, so that a complex can range from a few units in a single shelf to many units and shelves in several racks.

The mechanical design of the 2000 system is based on the standard Post Office rack, or its proprietary variants, since this rack is internationally accepted for electronic complexes. The rack is modular, in that the vertical spacing of fixing holes allows panels of 19 in (48·3 cm) width and height $N \times 1{\cdot}75$ in (4·45 cm) to be accommodated over its frontal area. Prior to the 2000 system, all equipment designed by Electronics Division at Harwell had 19 in panels, whose heights were chosen appropriate to the size and number of components used in the unit. These units, compared with the 2000 series units, although modular in respect of front panel height, compatible with a particular framework for assembly, and of reasonable size and weight cannot be considered as part of a comprehensive unit system.

Before going further into the mechanical features of the system it is necessary to consider the purpose of the plug-in facility of the unit system. Obviously, it allows a rapid interchange of units and ideally should provide complete electrical integration of a unit with other units in terms of power supplies and signals. This can be done by allocating pins on the plug/socket connectors to specific voltage lines and specific signals which are common over a fair range of functions. In nucleonic complexes, such common signals are very few so that it is desirable to route only voltage supplies via the plug-in connector because the connector would otherwise require many pins, be large and expensive. Therefore all signals must enter, or leave, via the front, rear, or both faces of each unit. Side entry is excluded on other grounds. For small complexes, the choice has little or no significance, but for large complexes, signal connections at the rear have a nuisance value, when replacing units, and imply access between the rear of a rack and an adjacent wall of 2·5 ft at least. Front interconnections have the disadvantage that they are subject to the mischievous person and cables may obscure indicators or present an untidy appearance. A mixed system of rear and front interconnectors is likely to have the advantages of neither and the disadvantages of both. Thus as one must be chosen, operational convenience makes front mounting the obvious choice and in the 2000 series, all signal connectors are on the front panel of a unit. This feature also applies to all controls and indicators to which a user has need of access.

3.1. *Mechanical Features*

The choice of dimensions for plug-in units is not easy. They should impose no restriction on the use of typical components, nor limit the number of components that can be reasonably assumed necessary for the most complicated single function. The front panel area in particular should be chosen to allow for connectors, indicators and controls appropriate to such a function. On the other hand, a large basic dimension may cause the majority of units to be uneconomical in volume or front panel utilization. A full-depth unit having a 19-in panel of least modular height (1·75 in), for example, would not accommodate many types of normal components and if twice this height (3·5 in) would have a volume too large for the majority of single function units. The alternative is to accept a larger modular height (8·75 in) and divide into equal parts the front aperture width (17·4 in) of a rack. This alternative has been adopted in most unit systems, the divisors ranging from 4 to 12 for systems which have controls and indicators on front panels, or larger where this does not apply. The choice in the 2000 system was governed by the dimensions of valves, when mounted with their long axis horizontal to give good heat dissipation, and meters of reasonable readability, i.e. 3·0 in diameter. Thus the divisor 5 is used and the smallest module width is 3·375 in.

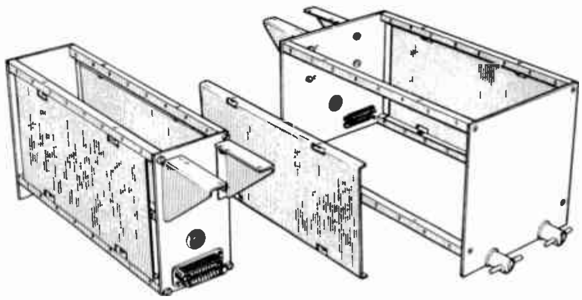


Fig. 1. Perspective of unit-frame.

The modular height of 8·75 in includes the shelf framewoik. The choice of the unit depth presents no problem if a particular rack is specified. At Harwell, there is a large capital investment in racks of all varieties, and to avoid excluding a major part of this investment for the 2000 series, the overall depth (15 in) of the unit was chosen to suit the most commonly used rack having the smallest back to front dimension. The depth available for circuits is about 12 in since a volume is reserved for a.c. transformers outside the rear panel to isolate them from heat generating components and remove them as far as reasonably possible from components sensitive to alternating fields (Fig. 1).

The mechanical construction of the unit-frame must be such that tolerances on linear dimensions are maintained in production, when circuit components are included and also during operational use. The frame should have also an open form to give through ventilation. Its cost must be low since although it will be required in large numbers, it contributes to the cost of unitization. The unit-frame of the 2000 series has strong rear and front panels (0·125 in thick) joined by four nickel-plated brass tie-bars which are machined square at their ends to hold parallel the two panels. Screws, through holes in the panels can engage the ends of the tie-bars, and when tightened up give the frame remarkable rigidity and a dimensional tolerance within $\pm 0{\cdot}015$ in on any dimension. This arrangement also allows the sides to be opened up for servicing as in Fig. 5.

The plug-in facility on each unit is provided by a socket mounted in a specified location on the rear panel. Figure 1 gives a perspective view of a unit frame. The tie-bars are pre-drilled and tapped to accommodate side covers and component boards—see Figs. 5, 6 and 7. Frames can be built up from kits to form single, double or triple width units using identical tie-bars.

The shelf has to accept any modular width of unit or combination of widths up to 5; provide a connector, switch, indicator and fuses for an a.c. mains supply; provide a means of interconnecting power lines between units and between shelves; and incorporate ventilation of units. Figure 2 shows the shelf design features. Each of the five stations has a shrouded plug on the rear framework. To avoid close tolerancing on the positions of the shelf plugs and the unit sockets, the sockets are loosely fitted. A double or triple width unit has only one socket to engage one of the shelf plugs, in which case the other one or two plugs are not in use, although services are fed to them. This may appear to be an unnecessary expense, but in fact it permits one design of shelf to be used throughout, no matter what complex is desired. Inter-shelf connection of power supplies is at the rear of the shelves. This is permissible on the grounds that complexes are much less frequently re-arıanged than are units. Ventilation of a unit or complex is obtained by means of a full width, full depth deflector plate fitted into a space at the base of the shelf. This plate can occupy one of two positions, illustrated in Fig. 3, to give forward or rearward ejection of warmed air. The flow direction indicated by the dotted line is used when the rack is fitted with an extractor fan, otherwise and normally, the other flow direction is used. Each shelf provides an air current in parallel with other shelves to prevent the accumulation of hot air at the top of a rack. The rails, which guide the units during insertion or extraction, can be rotated out of position and stowed away to convert two single width stations to one of double width and so on. The unit-frames do not bear on the guide rails when fully inserted and locked in, to avoid

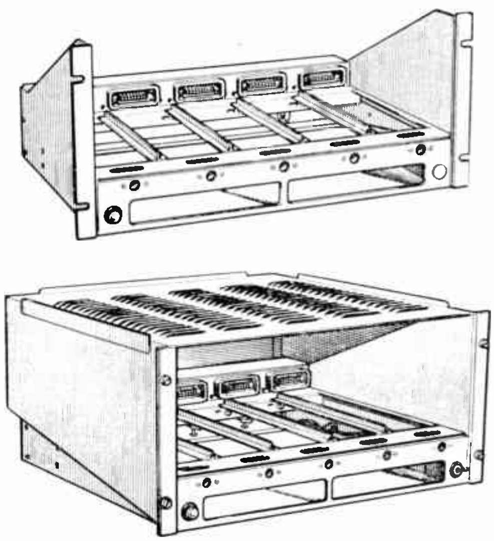


Fig. 2. Perspective of shelf.

close tolerancing. This is done by mounting the shelf-plug in such a way that the rear part of the unit rises when the lock is engaged. The lock is an adaption of a Dzus fastener with a toggle bar. It is loosely fitted to the front panel of a unit and, with the unit fully inserted, rotation through 90 deg engages diametrically opposite claws on the barrel of the fastener with a strong spring-clip on the shelf. This action holds the unit-socket firmly in the shelf-plug and gives a good earth connection between unit-frame and shelf. It is also a rapid means of engaging or disengaging a unit from a shelf.

3.2. *Electrical Features*

The unit/shelf and the shelf/unit connectors have each pin allocated to a specified voltage line and all connectors on one shelf are in parallel via a cable-form inside the rear cover (Figs. 2 and 3) for protection and safety. The supplies on these connectors comprise all the standard direct voltages and two 240 V, 50 c/s mains voltages, one delayed, after switching on the shelf, by 60 seconds on the other. This permits heaters of valves to be warmed up before h.t. is applied but has otherwise no significance.

A design feature of the system is that the d.c. power units are themselves in unit form so that such power units as necessary, inserted into the shelf complex, are energized from the mains and supply their outputs via the shelf cable-form. An alternative approach, on some unit systems, is to build a power unit at the back of each shelf so that it does not occupy front panel space. To provide the full range of power at all voltages, this power unit could be large and expensive and yet possibly not be fully used in most shelves. The 2000 series feature allows a power unit of capability appropriate to a given combination of units to be chosen from a range of simpler power units and is therefore more flexible and more economic. In addition the power units can be exchanged rapidly and access is available to fuses, trips and 'on line' indicators from the front panel.

For valve circuits the standard preferred voltages are +300 V and −150 V, although −300 and +150 V are also available, from a selection of three units, one providing +300 V and −150 V at a smaller output rating than the two others which separately give ±300 V and ±150 V. To supplement these units and for use in large complexes, 19-in panel bulk power units are available which have a large output capability and control all shelves to which they supply power. For transistor circuits the standard preferred voltages are +15 V, −15 V with other standard voltages of −5 V, −10 V and −30 V additionally available. These again are obtainable from a range of plug-in units, each providing all the voltages but with different output current capabilities. A further pin allocation is made for a 24-V unstabilized supply primarily intended for a battery supply in transportable assemblies of 2000 series equipment.

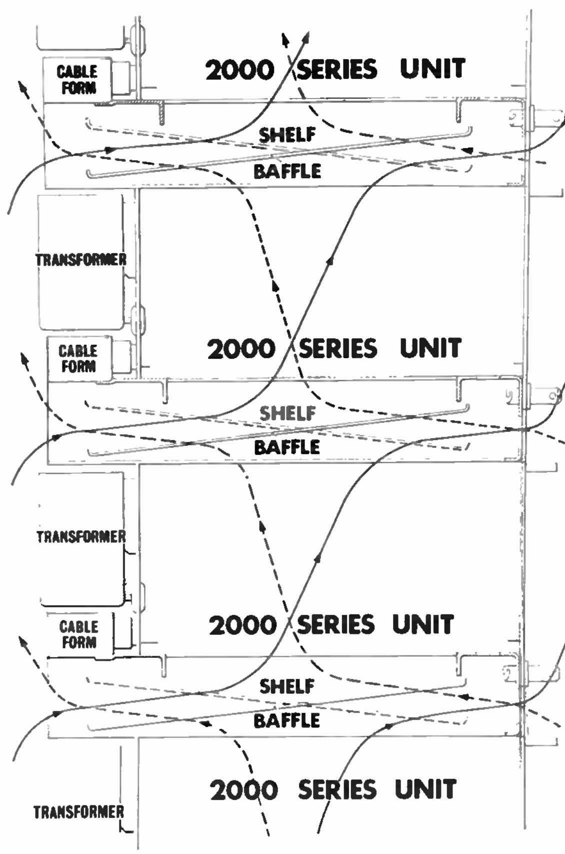


Fig. 3. Sectional view of shelf assembly demonstrating ventilation.

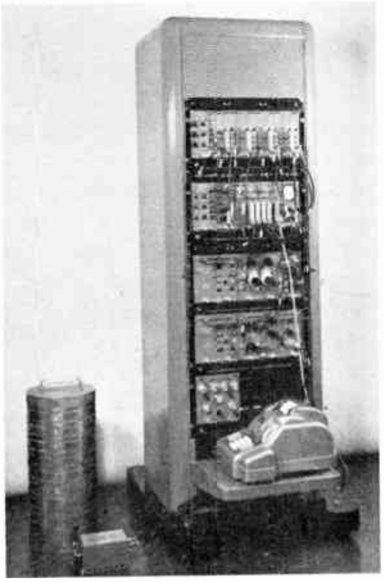


Fig. 4. A typical complex of 2000 Series units.

All these standard supplies are completely specified and designers employ the preferred voltages wherever possible, although no restrictions are imposed on the use of the other voltages. A specification is also laid down for the maximum tolerable voltage surge which any unit is allowed to put out on to the common supply lines. Designers are requested to arrange that single (double) width units do not demand more than 0·25 (0·5) of the output capability of one commonly used l.t. power unit, so that one such power unit can always supply units in a fully occupied shelf. Occasionally this condition cannot be met and then the units concerned may be powered direct from the mains supply with the one proviso that inclusion of an a.c./d.c. converter in the unit does not increase the unit width (say from single to double) above what would otherwise be required.

The above approach to routing and supplying power to plug-in units allows any size of complex to be assembled with one design of shelf-unit, which can be mass produced at reduced prices, again necessary to bring down the cost of unitization. In fact two types of shelf are used, both physically identical in their metalwork and cabling, to reduce the cost still further. They are the 'control' and 'slave' shelves; the latter does not provide switching, fusing and delay of mains supply. The control shelf has two fixed inter-shelf connectors, but the slave shelf has one only and a plug on a flying lead. This allows a control shelf to be positioned in between two or more slave shelves and is an important feature in reducing voltage losses and surges on the routing wires when units demanding heavy currents cannot be situated, for operational reasons, in a station adjacent to a power unit.

3.3. *Standard Coupling Signals*

The design principle has already been discussed which requires that a comprehensive system should have a recognized form of signals for passing information between units. This information can be contained in an analogue or digital signal. The analogue signal is a current or voltage waveform whose value at any time is related to some parameter(s) under observation. The digital signal, in the 2000 series system, is a voltage signal which has two recognized states, i.e. 'present' and 'absent', and can therefore be described by the term 'binary signal'. Ternary signals, i.e. having three recognized states, have been discarded because their use created considerable difficulty with 'fan-in' circuits.

The binary signals are classified into binary coupling levels (b.c.l.) and digital coupling pulses (d.c.p.). The former is a voltage signal which can have one of two recognized states 'present' or 'absent', this conveying one bit of information (digit). The voltage amplitude in these states must be within prescribed limits, but has no significance otherwise. The duration of the signal in either state may have significance in relation to the time scale of interest, thus conveying a second bit of information. A digital coupling pulse is a voltage signal which also can have one of two recognized states 'present' or 'absent', this conveying one bit of information. Its amplitude in these states must be within prescribed limits, but otherwise has no significance. Its duration in the 'present' state, however, although confined within prescribed limits, has no significance in relation to the time scale of interest.

With these definitions, a family of levels and pulses has been derived suitable for operation in so called 'slow' (< 1 Mc/s), 'medium' (> 1 Mc/s < 10 Mc/s) and 'fast' (> 10 Mc/s < 100 Mc/s) operational fields. The derivation had one particular principle in mind; that the voltage limits defining the two states, and the transit times between the states, had to be compatible between pulses and levels so that units could be designed, if appropriate, to accept both, either as alternative inputs to a connector or as simultaneous inputs via separate connectors to a common circuit element. The preferred binary signals are specified as in Tables 1 and 2. Some hybrid signals between the three major classifications are omitted for the sake of clarity. It will be observed that in each classification a b.c.l. and a d.c.p. can have identical waveforms under certain conditions.

Some conventions are essential in using these signals and a few typical examples are of interest:

(i) The voltages set out are measured with respect to a recognized earth such as the earthed outer screen of a cable or a core connected to the zero volt terminal of a power supply.

(ii) In the 'absent' state, the signal waveform should be as near to zero volts as possible.

(iii) When no external connection is made to an input, this has the same significance as applying a signal in the 'absent' state. The actual voltage relative to earth of such a 'free' terminal should normally comply with the specified limits of the signal 'absent' state.

(iv) No convention is preferred for the binary significance of the two states, i.e. either 'present' or 'absent' can represent binary '0' or '1' or an 'enable' or 'inhibit' function.

(v) The functioning of output circuits should be normal after the output terminal has been connected to earth through zero impedance for an indefinite period.

Table 1

Specification of standard coupling signals

(a) Binary Levels

Characteristic	⩽ 1 Mc/s		⩽ 10 Mc/s	⩽ 100 Mc/s
Designation	Type 0	Type 1	Type 2	Type 3
Symbol	L0	L1	L2	L3
'signal absent'	± 0·75 V	± 0·75 V	± 0·15 V	± 0·075 V
transients	± 1·0 V	± 1·0 V	± 0·3 V	± 0·15 V
'signal present'	− 5, − 10 V	− 5, − 10 V	− 2, − 5 V	− 1, − 2·5 V
changeover thresholds	− 1, − 5 V	− 1, − 5 V	− 0·5, − 2 V	− 0·25, − 1 V
changeover times	Stated	< 300 ns	< 20 ns	< 2 ns
minimum duration (either state)	200 ns	200 ns	20 ns	2 ns
input impedance	> 2 kΩ	> 2 kΩ	100 ± 4 Ω	50 ± 2 Ω
load impedance	> 500 Ω	> 500 Ω	100 ± 4 Ω	50 ± 2 Ω
input current	± 0·75 mA	± 0·75 mA	± 3·0 mA	± 3·0 mA
input circuit protection	− 30 V	− 30 V	− 5 V	− 2·5 V
coupling	—	—	—	—

(b) Digital Pulses

Characteristic	⩽ 1 Mc/s	⩽ 10 Mc/s	⩽ 100 Mc/s
Designation	Type 1	Type 2	Type 3
Symbol	P1	P2	P3
'signal absent'	± 0·75 V	± 0·15 V	± 0·075 V
transients	± 1·0 V	± 0·3 V	± 0·15 V
'signal present'	− 5, − 10 V	− 2, − 5 V	− 1, − 2·5 V
dv/dt thresholds	− 1, − 5 V	− 0·5, − 2 V	− 0·25, − 2 V
maximum duration at least negative dv/dt threshold	600 ns	60 ns	6 ns
minimum duration ('signal present')	200 ns	20 ns	2 ns
input impedance	> 2 kΩ	100 ± 4 Ω	50 ± 2 Ω
load impedance	> 500 Ω	100 ± 4 Ω	50 ± 2 Ω
input current	± 0·75 mA	± 3·0 mA	± 3·0 mA
input circuit protection	− 30 V	− 5 V	− 2·5 V
coupling	a.c. or d.c.	d.c.	d.c.

(vi) The preferred method of distributing an output signal to several units is by feeding the signal into parallel electrical paths (fan-out) which terminate in the separate units, rather than by transmission (i.e. directly or by regeneration) through the units in series. Thus for Type 0,1 levels and Type 1 pulses each output circuit is capable of 1:4 fan-out via a passive system. For Type 2,3 levels and pulses, 1:2 passive fan-out is possible, but higher fan-out ratios must normally be via an active fan-out system. In all cases the outputs from fan-outs must satisfy the specifications in Table 1, as appropriate.

Analogue signals may again be pulses or levels. The significant feature of a level is its amplitude whereas for pulses both amplitude and time integral may be significant. Analogue pulses are standardized in the 2000 system but only in respect of their maximum amplitude (± 5 V w.r.t. earth) and no classification has yet been attempted.

3.4. *Designation*

Designation is concerned with the terms and symbols used on front panels to indicate the purpose or function of connectors, controls and indicators. A code for the 2000 series has been set out to establish a degree of uniformity of designation between units and to reduce the possibility of misinterpretation. To be ideally successful, the code should guide designers on the choice of terms and symbols which will indicate the correct use to be made of front panel devices without having recourse to descriptive literature. The code therefore includes the principles of designation, and lists of approved terms, abbreviations and symbols.

Certain areas of a front panel are allocated to basic information, such as descriptive title of unit, catalogue number, serial number, manufacturer's name and voltage and current demands of the unit. The code also defines the size and type of lettering, colour of front panel, filling for engraved letters and lamp signals. However, there are some principles which are of general interest:

(i) Abbreviations are not allowed unless there are good reasons for them, such as lack of space, improved appearance or significantly decreased engraving costs. Symbols, on the other hand, must be used at all times.

(ii) The designation of a unipole connector, accepting or delivering a binary signal, should give the condition associated with the signal 'present' state and the class of signal.

(iii) Where ambiguity is likely to arise, because of close packing, terms are to be linked to the appropriate controls etc., by a white line.

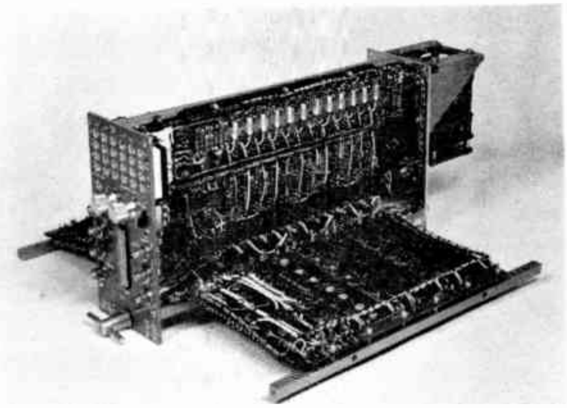


Fig. 5. A scaler, opened for servicing.

Table 2

Characteristics of standard coupling signals

Designation to be used when reference is made to levels, or pulses, except where letter symbols are used.

Letter symbols to be used when abbreviation is required in any context and must be used in the designation of front panels.

'Signal absent.' A voltage signal between these limits, at an input or output terminal will be associated with a 'signal absent' state.

'Signal present.' A voltage signal, between these limits, at an input or output terminal will be associated with a 'signal present' state.

Transients. The 'signal absent' state is extended to these limits for transient excursions of the signal.

Change-over thresholds are voltage values through which the signal voltage must pass in changing between the two states.

Change-over times (levels only). The time period during which the signal voltage exists between the change-over thresholds.

Minimum duration. A 'signal present' or 'signal absent' state for levels, a 'signal present' state for pulses will not be recognized unless the signal exists in the recognized state for a duration equal to or greater than the period specified.

Maximum duration (pulses only). The time, for which the signal voltage extends more negatively than the least negative dv/dt threshold, shall not be greater than the values quoted.

(dv/dt) *thresholds (pulses only).* Rates of change of voltage (dv/dt) will be measured between the voltage values quoted.

Load impedance. The voltages and times specified to be measured across resistors values quoted, when connected across an output terminal and earth.

Input impedance. The values of resistance quoted to be those which an input terminal presents to the signal.

Input current. The current flowing through an input terminal, when connected to an output terminal, shall be within the limits quoted when the signal is in the 'Signal Absent' state.

Input circuit protection. The functioning of an input circuit must be normal after the voltages quoted have been applied across the terminal and earth for an indefinite period.

(iv) Units quoted on a meter scale should not differ from those given on a panel.

Most of the terms included in the approved list are taken from British Standard Glossaries on Nuclear Science (B.S. 3455:1962) and Automatic Data Processing (B.S. 3527:1962), and spelling conforms with the "Concise Oxford Dictionary". There are nevertheless some terms which require some further resolution, to avoid ambiguity. A demonstrable example occurs in the indiscriminate use of terms such as 'clear' and 'reset', 'expand' and 'stretch', 'trip' and 'trigger', 'dead-time' and 'paralysis'. Where terms can be assigned a clear function this is done, i.e. 'clear'—set to zero, 'reset'—set to original state, 'expand'—multiplication, 'stretch'—multiplication to a fixed limit; however, where terms have no separate meaning, only one is approved. Thus 'trip' and 'paralysis' are omitted. Similar selection is done on possible alternatives of abbreviations.

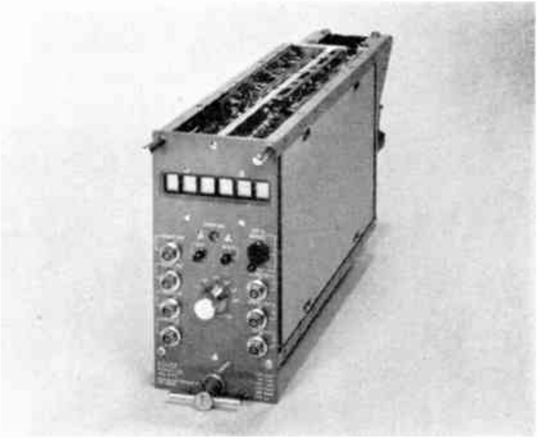


Fig. 6. Typical single-width unit (Scaler 2142).

The symbols are those universally adopted for such things as ohms (Ω), pulse recurrence frequency (p.r.f.), volts (V) etc., but some, which are peculiar to the system have been invented. For example, L1 and P2 for binary level Type 1 and digital pulse Type 2, BP ($\pm$) for bipolar pulses having analogue information in the positive portion of the waveform and BP ($\mp$) when the information is in the negative portion.

Figures 6 and 7 show examples of typical single- and double-width units.

4. Facts about the Harwell 2000 Series System

The system takes its name from the practice of allocating type numbers, in the number series 2000–2999, to units which are compatible with the system. HARWELL is a registered mark to ensure that units bearing it satisfy formal specifications. As implied earlier, valve equipments predominate in the serial numbers 2000–2050, but thereafter units are almost exclusively transistored. Of the 180 serial numbers issued, an analysis shows that 55% are allocated to 'basic' units of wide application, 25% to 'special' units and the remainder are reserved or discarded numbers.

Among the 'basic' units are unit-frames (3 types), shelves (2), and shelf-extension units (3) into which single, double and treble width units can be plugged to be held in front and clear of neighbouring units in the normal shelf. This provides access to the side of the unit for observation whilst in an operational condition. The complement of 'basic' units, required for nuclear measurements, will not be described in detail. It includes a large variety of pulse amplifiers (12), discriminators (5), single-channel pulse amplitude selectors (3), nine scalers, four of which have access for read-out on to types of recording machines, coincidence units (2), clock-pulse generators and timers (5), ratemeters (3), amplitude/digital and time/amplitude converters, a random pulse generator, a meter recorder, e.h.t. generators (3) and power units for valve (4), and transistor (2) circuits. In addition, the series is accompanied by a range of head-units suitable for proportional, scintillation and semi-conductor detectors, a standard range of interconnecting cables and other accessories.

The full range of units has been originated by more than 30 designers mainly in Electronics Division but some more recently in the National Institute for Research in Nuclear Science. The numbers of unit-frames and shelves used in the past four years are 15 000 and 3600 respectively and about 11 000 other units have been issued during the same period, representing a total capital expenditure of about £900,000. The design, development and production of this equipment have been done in collaboration with some 20 firms in the British scientific instrument industry.

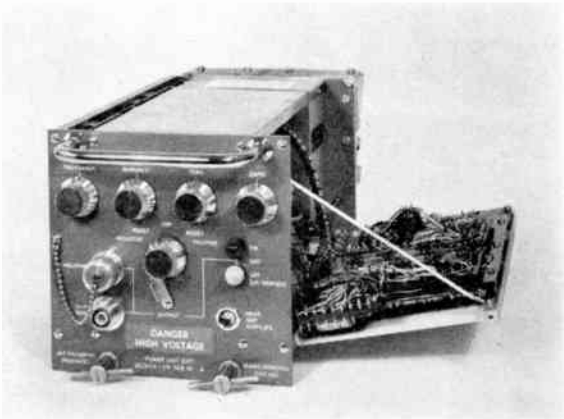


Fig. 7. Typical double-width unit (E.H.T. Power Unit 2124).

5. The Effects of a Comprehensive Unit System

Although it is true that the extent of the task of attaining a full complement of units in a comprehensive system can be underestimated, it is also true to say that all the benefits derived from such a system, in all their various ramifications, cannot be predicted, nor their full value realized, when the task is first undertaken. Some benefits, obvious in themselves, were sufficient justification for attempting the task, but many others have become apparent as experience has grown with the 2000 system. Those which could be applicable to any similar system are worth some discussion.

Unitization alone was foreseen as a means of introducing flexibility in operation by virtue of the combination of basic units to form many complexes having specific roles, but which could be changed readily by the substitution or addition of other basic units.

In practice a much greater degree of flexibility than this implies has evolved. Designers, faced with the prospect that 'basic' units must be capable of wide application, as part of a system, have come to think more carefully about the essential features which such units should possess. This has caused operational features to be introduced which would otherwise have been omitted, and at very little extra cost. The outcome has been that units are fulfilling operational roles which were not foreseen at the design stage.

The use was expected of 'special' units to extend the facilities available from a complex of 'basic' units and thereby make available these special complexes at short notice. There is now clear evidence that, in addition, a unit system can take immediate advantage of advances in electronic techniques and devices, as these occur, by partial or re-development of a few 'basic' units to keep the facilities available in touch with the demands of the scientist, again with only a small deployment of effort.

As larger machines, such as accelerators, reactors and analysers, have become available, the rapid diagnosis and replacement of faulty units made possible by a unit system, assumes greater importance in reducing the 'down-time' of the many instrument complexes surrounding these machines, thus reducing wastage of expensive machine time and also the scientific effort geared to the machine programme.

Increased reliability can contribute to this problem and is also essential if the maintenance and repair effort are to be kept within reasonable grounds. The contribution to improved reliability, by the sole factor of unitization, is confused in the case of the 2000 series by the changeover from valves to transistors during its evolution, and is therefore difficult to assess directly. Some indirect conclusions can be drawn from a recent tentative survey which shows that an improvement of the order of 3:1 has been obtained on complexes using transistor units, as compared with the nearest equivalent valve type complexes. This may be considered as a surprising result in view of the reputedly higher reliability of solid-state devices over valves. That higher values have not materialized is due to the increased complexity and the large component packing density attainable in transistor units—the six-decade decimal-indicated scaler in Fig. 6, for example, is a single-width module containing well over 1000 components of which less than 20% are transistors and diodes. Thus, apart from the use of greater numbers of active components, the association of these with a larger number of passive components, whose reliability has not been improved by the same factor as between valves and transistors, largely cancels out the higher overall improvements on unit reliability which might be expected. The reduction in reliability has been kept smaller because of three factors. The first is the built-in ventilation of the 2000 series; secondly because all units must be compatible with the system, a high uniform standard of design has to be maintained on all units; finally, the technical and test specifications for unifunctional units have been found to be simpler than for comparable multifunctional units, thereby permitting better control over production and acceptance testing.

The prospect of standardization and the degree of enforcement essential to a comprehensive system were received with some apprehension by designers who feared that their freedom of design might be restricted. Experience has shown that this apprehension has not, and should not, be realized if standardization is based on, and develops with, experience. Standardization has, in fact, relieved designers of the commonplace in design, for example mechanical engineering aspects, power units and input and output circuits, and allowed them to concentrate on those design aspects of a unit of greater interest and more appropriate to their specialist ability and knowledge.

Dealing further with the engineering development associated with units, the unit-frame was devised as a method of establishing dimensional compatibility between units. Again, it has proved to be more than this. Designers and those involved in building prototype equipment have become so familiar with the frame that they can progress rapidly, and in one step, from the 'bird's nest' state of a circuit design to a respectably engineered model suitable for copying on a production scale. This contrasts with the 'model-after-model' progress of development in other equipments and obviously makes a large contribution to shortening the period between the design stage and the time when large numbers are in use. The unit-frame has come to be used on a large scale for minor or even major experimental circuits since the shelves and power units are often already available in the laboratory concerned, or if not can be obtained readily from stock.

The adoption of the unit system has had completely unpredictable effects on the organization and methods of operation of the team adopting it. In design and development, there appears a greater sense of purpose since each 'basic' unit is part of the one system; the processes of development, design approval, the organization of component lists, kits, and circuit diagrams for production and maintenance has been streamlined and reduced to almost routine procedures, and finally, the distribution, or in the commercial world sales, staff have become familiar more readily than ever before with the technical capability of each unit and some standard complexes of 'basic' units. The reasons for these effects are diverse, but fundamentally it is because the complicated problems associated with processes of designing, developing, producing and distributing equipment are, like the electronic complexes themselves, broken down into smaller bits which occur at different times and which are therefore simpler to analyse and overcome.

6. Conclusions

The design principles appertaining to any comprehensive unit system have been demonstrated by reference to the Harwell 2000 Series system. There is no doubt that this system is achieving its primary aim of providing nucleonic instrument complexes which have both a high degree of operational flexibility and reliability and are contemporary with advances in electronic techniques. Justification for the enormous amount of effort put into the system is borne out in terms of making complexes more readily available to scientific programmes and in breaking down the problems of design, development and production of electronic equipment.

Perhaps the system's most significant role, for an electronic engineer, apart from proving that standardization is no bar to advancement in instrumentation, is that a unit system can be an effective influence in bringing together, by its sense of purpose, individual designers, thereby stimulating ideas, achieving an easy exchange of experience and reducing the duplication of effort. In the author's opinion this role could be equally effective on a national or even international scale.†

7. Acknowledgments

The Harwell 2000 Series Unit System could not have achieved its primary aims, nor this paper have been presented, without the unfailing patience, advice and co-operation of many colleagues in the Electronics Division at A.E.R.E. Harwell. In particular, the author would like to acknowledge the work of Messrs. K. Kandiah and E. A. Sayle during the formative period of the system, the interest and advice throughout of Mr. E. H. Cooke-Yarborough and the invaluable part played by Mr. F. H. Hale in the past three years.

Manuscript received by the Institution on 26th October 1964. (Paper No. 969)

† "European Standards of Nucleonic Equipment (ESONE)", Euratom Report EUR 1831e.

Radio Engineering Overseas . . .

The following abstracts are taken from Commonwealth, European and Asian journals received by the Institution's Library. Abstracts of papers published in American journals are not included because they are available in many other publications. Members who wish to consult any of the papers quoted should apply to the Librarian, giving full bibliographical details, i.e. title, author, journal and date, of the paper required. All papers are in the language of the country of origin of the journal unless otherwise stated. Translations cannot be supplied.

ELECTRONIC MICROVOLTMETER

A relatively simple and sensitive microvoltmeter has been developed in Czechoslovakia, using a special circuit with a switching contact, working as a modulator and demodulator with corresponding action. The basic range for full deflection of the indicating meter is 50 microvolts approximately, the time constant of the amplifier is 0·2 seconds, noise and reading shift is less than 2 microvolts per hour; input impedance is 15 kΩ to 1 MΩ per millivolt according to sensitivity adjustment by feedback.

"An electronic microvoltmeter with a contact modulator", Miroslav Pacák, *Slaboproudý Obzor*, (*Prague*), **25**, No. 10, pp. 571–76, October 1964.

HIGH-SPEED MICROWAVE SWITCHES

A Japanese paper describes some experimental results on microwave nanosecond switches using silver-bonded diodes in the 11 Gc/s region. Although these diodes were originally developed as variable capacitance elements for use in parametric amplifiers, it was concluded from the experiments described that they also have an excellent performance for nanosecond switching of microwaves.

Their main advantages are low insertion loss (as low as 1 dB), large switching ratio (greater than 20 dB), fast rise and fall times (as small as 0·5 ns) and moderate power handling capacity (as high as 100 mW).

Semiconductor diode switches have a non-linear characteristic inherently. This characteristic can be used for pulse shaping in microwave nanosecond switching. But, for pulse shaping, the transient response of the switching diode is important, and from experimental work it was concluded that the silicon silver-bonded diode SiSBR has an excellent transient response.

"High speed microwave switches using silver-bonded diodes in the 11-Gc region", Kazuhiro Miyauchi and Osamu Ueda. *Review of the Electrical Communications Laboratory*, N.T.T., **12**, No. 5–6, pp. 281–94, May–June 1964.

S.S.B. TRANSMISSION FOR BROADCASTING

The constant increase in the power and number of broadcasting transmitters makes it desirable to reduce their bandwidths within the channels allocated to them. Single-sideband modulation is one means to this end. If signals so modulated can be received by sets currently in use, the single-sideband modulation will be 'compatible' with the conventional double-sideband system, and s.s.b. can be introduced gradually. A 'squaring method' has been developed in Holland for obtaining the s.s.b. signal required for this purpose. The underlying idea is that amplitude modulation sufficiently free from distortion can be achieved by adding a corrective side component, lying within the transmitted sideband, to a non-corrected s.s.b. signal. The circuitry needed to modify a standard a.m. transmitter for this method of s.s.b. modulation can be accommodated in a simple switching unit.

Tests show that compatible s.s.b. modulation offers better reproduction of the high audio frequencies without any troublesome increase in non-linear distortion. Greater freedom from interference is also possible, enabling the transmitter to cover a wider area.

" 'Compatible' single-sideband modulation", T. J. van Kessel. *Philips Technical Review*, **25**, No. 11/12, pp. 311–19, 1963–64. (In English.)

FREQUENCY STORAGE

French engineers have developed a device to store an incident frequency of some unknown value within the widest possible frequency band. On reception of a short input signal characterized by its carrier frequency the device acts as a responder, i.e. it supplies an output signal at the same carrier frequency as the incident signal but of longer duration.

The responder used is a wide-band oscillator whose reaction loop includes essentially a travelling-wave amplifier tube and a delay line. First the basic principle of the ideal loop is considered. Account is then taken of the effect of noise, of loop gain irregularities and of saturation effects.

The experimental work mentions a number of measurements made on the output signal characteristics and gives some idea of the possibilities of the responder device.

"Design of a frequency storage device", V. Biggi and J. Dardenne. *Annales de Radioélectricité*, 19, No. 76, pp. 97–109, April 1964.

PULSE FREQUENCY DISCRIMINATOR

A Czech engineer has given details of a new method for obtaining the difference frequency of a pulse waveform from two input signals of pulsed waveforms. A pulse frequency demodulator with tunnel diodes is described and its properties are stated in detail. The demodulator uses signal summation and its output is a two-phase voltage or the difference frequency and a two-value signal carrying the information about the positive or negative value of the difference frequency.

"A pulse frequency demodulator", L. Beneš. *Slaboproudý Obzor*, **25**, pp. 528–33. September 1964.