

Ph.D. THESIS

JAN. 1960

" STRATEGY AND DESIGN OF SELF OPTIMISING FILTERS "

William Peter Lanham WILBY

B.Sc.(Eng.), A.C.G.I.

Electrical Engineering Department,
Imperial College of Science and Technology,
South Kensington, London, S.W.7.

Supervisor: Professor D.Gabor F.R.S.

SUMMARY

The thesis is concerned mainly with the design and construction of a machine which can act as a general non-linear filter, predictor, or simulator, and which automatically optimizes its performance by a learning process. Several special techniques which were developed, including a new system of analogue multiplication, are discussed in detail. Typical results obtained using the completed machine are given which show that the technical difficulties involved in constructing such a machine have been successfully overcome. Suggestions for future research and development are made, and the potential usefulness of the machine is discussed.

Note

Mr.R.Woodcock has also worked on some of the technical problems involved in constructing the machine, and his work is described in a Ph.D. thesis entitled " Non-linear Parallel Filters and Circuits for their Optimization ", Imperial College, 1960. Where, for completeness, his work has been included in this thesis, it is explicitly referred to as his, to avoid confusion.

1	INTRODUCTION	Page
1.1.	General	5
1.2.	Linear systems	5
1.3.	Non-linear systems	6
1.4.	Realization of the optimum non-linear filter	7
2	THEORETICAL BASIS OF THE FILTER	
2.1.	The filter operator	8
2.2.	Limitation of the number of coefficients	10
2.3.	Adjustment of the coefficients	11
2.4.	The teaching signal	16
3	GENERAL DESIGN CONSIDERATIONS	
3.1.	Requirements	17
3.2.	Realization of the filter	20
4	STRATEGY	23
5	DESIGN AND PERFORMANCE OF SYSTEM PARTS	
5.1.	Tape recorder	26
5.2.	Pulse rate modulation system	31
5.3.	Tape control unit	43
5.4.	Coefficient store	47
5.5.	Analogue multiplication	50
5.6.	Piezo-magnetic multipliers	58
5.7.	Teaching apparatus	74
5.8.	Sequence control and switching circuits	85
5.9.	Teleprinter output	105
5.10.	Auxiliary output	124
5.11.	Power supplies	126
6	INTERCONNECTION OF LEARNING FILTER PARTS	127

	Page
7 RESULTS OBTAINED WITH THE LEARNING FILTER	
7.1. Scaling of a sine wave	131
7.2. Phase shifting and scaling	133
7.3. Filtering of a sine wave, from a sine wave plus noise	136
7.4. Simulation	141
7.5. Survey of experimental work and present state of the machine	143
8 DISCUSSION	
8.1. Design	144
8.2. System performance and limitations	144
8.3. Some alternative methods	146
8.4. Future development of the existing machine	150
8.5. The use of learning filters	152
8.6. Status of the learning filter as an "intelligent machine"	157
9 REFERENCES	160

1 INTRODUCTION

1.1. General

A filter can be defined as a device which operates on an input signal, to produce an output signal uniquely determined by the present and/or past values of the input signal. The mathematical operation which the filter performs on the input signal may be linear or non-linear.

Methods of designing linear filters for separating signals of different frequencies have been extensively studied, and are well established, but this is a relatively simple case of the general filtering problem. In a more complex situation, for example where the true message $T(t)$ has an appreciable amount of noise $N(t)$ added to it, the design of a filter to extract the true message from the incoming signal $S(t)$ is a formidable task. The time series $S(t) = T(t) + N(t)$ is known up to the present instant, and efforts to extract $T(t)$ are limited by the operations available, and by what is known of the statistical structure of the two time series $T(t)$ and $N(t)$.

1.2. Linear Systems

N. Wiener has successfully attacked the problem of obtaining the optimum linear filter in the above situation¹. The criterion of Wiener (first used by Kolmogoroff²) to define the optimum filter, was that the mean square difference between the output of the filter and the true message, should be a minimum. The papers of Wiener and Kolmogoroff are now generally regarded as classics. In their treatments of the problem they

assumed stationary time series with observations available for all past times, and an infinite frequency band. The great mathematical effort which was required to obtain the optimum linear filter is discouraging, since the linear filter is an especially simple special case of the more general non-linear filter. In many situations a non-linear filter may be expected to improve upon the performance of a linear filter.

1.3. Non-Linear Systems

For system design and analysis, the theory of linear systems, is well developed, but there is a definite lack of an adequate non-linear theory. Thus the linear theory often has to be used for systems which are 'almost linear'. Alternatively it is assumed that a device is linear over certain ranges, which is useful for very non-linear elements such as diodes. The great class of systems which lies between these extremes is often tactfully ignored. Yet it is worth noting that very many useful devices in communication and control are not 'nearly linear'.

In filtering pulses from random noise, one method used is to amplify and clip, thus re-creating clean, flat topped pulses. This is very non-linear filtering and is an extreme case. The filter only recognises amplitude information and is completely confused if the signal/noise ratio is too low.

Sometimes it is essential to use a non-linear device, e.g. the mixer valve in a superhet. For other applications a non-linear device may not be essential but may give good results, e.g. the use of non-

linear resistor discs in switching circuits for spark suppression. A non-linear device can make possible a very simple and cheap solution to an awkward problem. For example in television receivers the E.H.T. supply often has poor regulation, which since the C.R.T. beam current varies widely with time, can spoil picture quality. The regulation can be improved by draining some current from the supply through a non-linear resistor which takes more current when the tube takes less and vice-versa, thus tending to keep the E.H.T. voltage constant.

In such simple cases as the above the effect of the non-linear element can be computed using graphical methods, but with more complex systems this fails. It is plain that non-linear devices can be very useful, but often in practice non-linearity is avoided, not because it would offer no improvement, but because its effect cannot be calculated. This state of affairs imposes annoying restrictions on the versatility and performance of the systems and filters in current use, and has prompted a great deal of interest and effort towards the development of analysis and design methods for non-linear systems³. So far the results, though often theoretically interesting, have proved of very little use to the practising engineer.

1.4. Realisation of the Optimum Non-Linear Filter

In 1954 D.Gabor published his important suggestions for bypassing this severe stumbling block⁴. His main idea was to "take a hint from nature" and make a completely general non-linear filter with many

adjustable parameters, and to make the filter capable of learning how to do its job. In other words the non-linear filter is to be self-optimising and it is to adjust its own parameters; completely freeing the engineer from the arduous task of trying to calculate what they should be.

It was this basic idea which led in 1956 to the commencement of the work described in this thesis, an attempt to design and construct a self-optimising non-linear filter.

In principle such a machine is of great generality. It can perform optimum filtering and prediction in the presence of noise. It can simulate systems of which only the input and output records are known. It can also be used to optimize the performance of a control system.

2 THEORETICAL BASIS OF THE FILTER

2.1. The Filter Operator

In any practical case the signal $S(t)$, on which the filter is to operate, will be received in a finite frequency band of width F . In such a case the signal is completely determined by its amplitude samples taken at intervals $1/2F$ ⁵ (Shannon's Sampling Theorem).

Gabor⁴ has pointed out that acceptance of this practical limitation greatly simplifies the task of obtaining a realizable formulation of the optimum non-linear filter.

We therefore depict the signal as in fig.1, and consider only operations on the amplitude samples $S_0, S_1, \dots, S_n$ of the time

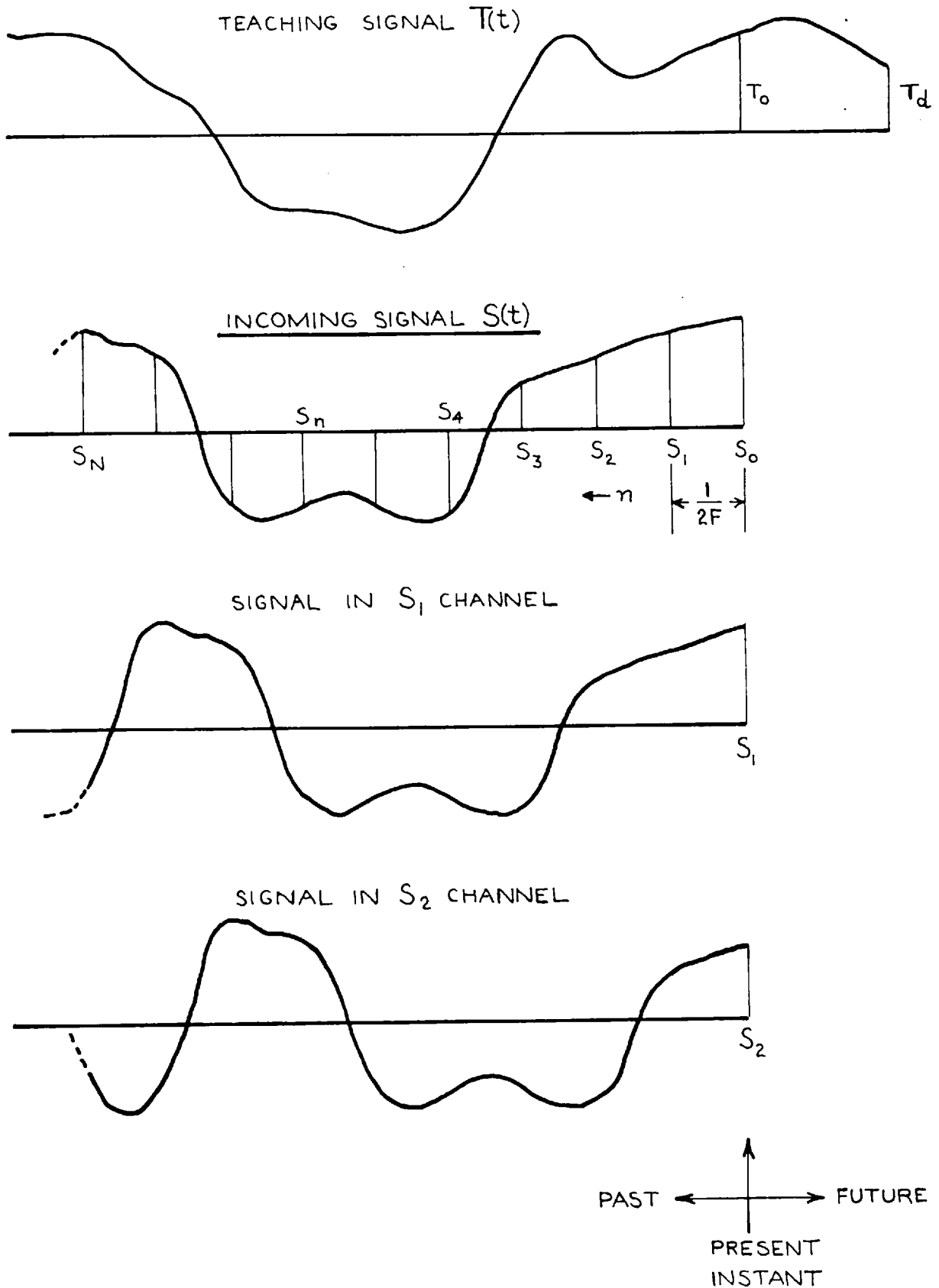


FIG. 1 SAMPLING OF INPUT SIGNAL

series $S(t)$.

If we now associate a coefficient C_n with each signal sample, and form the output of the filter at the present instant:-

$$S_o' = \sum_0^N C_n S_n \quad (1)$$

we have a filter which (given enough terms) can imitate any linear filter. Filters of this type have been studied by Kallman, Espley and Reswick.

To avoid having an infinite series for equation (1), we have to assume that samples of $S(t)$ beyond a certain time $N/2F$ in the past, have a negligible effect on S_o' . It is reasonable to suppose that in practice the most recent samples are likely to be the most important. Equation (1) may be extended to give:-

$$\begin{aligned} S_o' = & \sum_0^N C_n S_n + \sum_0^N \sum_0^N C_{n_1 n_2} S_{n_1} S_{n_2} \\ & + \sum_0^N \sum_0^N \sum_0^N C_{n_1 n_2 n_3} S_{n_1} S_{n_2} S_{n_3} + \dots \end{aligned} \quad (2)$$

The filter is now a general non-linear filter, and it is to be optimized by adjusting the many coefficients $C_n, C_{n_1 n_2}$, etc.

2.2. Limitation of the Number of Coefficients

The number of coefficients increases very steeply with the order of equation (2) because of the large number of possible products of the signal samples in groups of two, groups of three etc. Thus if

the significant past extends over N sampling intervals there will be $(N + 1)$ terms of first order, $\frac{1}{2} (N + 1)(N + 2)$ terms of second order, $\frac{1}{6} (N + 1)(N + 2)(N + 3)$ terms of third order, and so on.

Thus it is apparent that even a filter operator with many terms will only be able to deal adequately with processes which have a strong perspective, such that the significant past of non-linear effects steeply decreases with their order.

One reasonable method of limiting the number of coefficients is as follows. The signal samples are weighted in inverse proportion to their delays and the significance of the $(N + 2)^{\text{nd}}$ sample is made zero.

The rule for the product terms is that they are omitted if their significance is less than that of the $(N + 1)^{\text{th}}$ sample. It is stressed that the above rule is an arbitrary one. The most suitable weighting function will depend on the process being handled, and must be chosen in the light of experience. In certain (mainly artificial) cases the correct weighting function might be specified. In general the exact form of the weighting function used to cut down the number of coefficients will be unimportant, provided there are enough product terms remaining for the process in question.

2.3. Adjustment of the Coefficients

The first step in setting up a systematic procedure for adjusting the coefficients to their optimum values, is the adoption of

some success criterion. The one adopted for the Learning Filter is the minimum mean square error criterion already mentioned, this criterion confers one very important advantage which will emerge later.

Suppose now that we are aware of the desired filter output or target function T_d , where d is an integer, positive for delays, negative for prediction. Writing down the mean square error, Y , we have:-

$$Y = \overline{[S_o' - T_d]^2}$$

$$\text{i.e. } Y = \overline{\left[\sum_0^N C_n S_n + \sum_0^N \sum_0^N C_{n_1 n_2} S_{n_1} S_{n_2} + \dots - T_d \right]^2} \quad (3)$$

This on expansion yields a positive definite quadratic form of the variables C_n , $C_{n_1 n_2}$, etc., consequently there must be one unique minimum. Carrying out the minimization of the mean square error one is led to a set of simultaneous linear equations with the filter coefficients as the unknowns. Solving these equations (say 100 of them) is such a formidable task that Gabor⁴ has proposed the adoption of Southwell's Relaxation routine, where one coefficient is adjusted at a time, and the optimum setting is approached by a series of approximations.

If we adopt this method we can apply the minimum mean square error criterion in a simpler way, which eliminates the tedious algebra resulting from the expansion of equation (3), and gives a good intuitive picture of what the Learning Filter is doing, in terms of the signals present.

Suppose we are interested in adjusting one particular coef-

efficient C_n , associated with a sample or sample product S . We rewrite the expression for the mean square error as:-

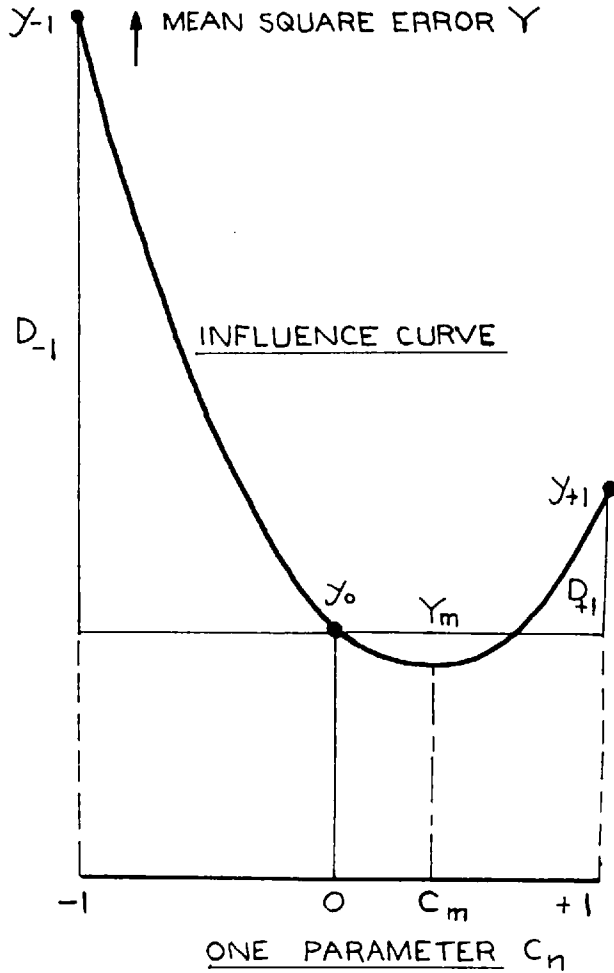
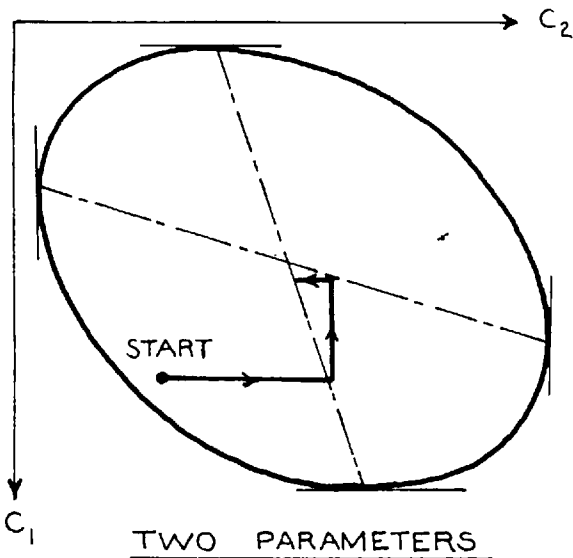
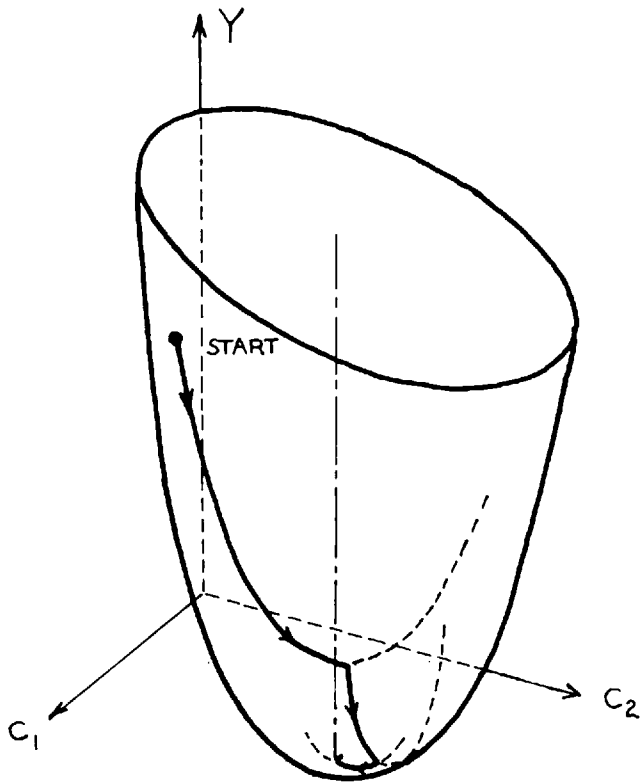
$$Y = \overline{[\phi + C_n S - T_d]^2}$$

The signal ϕ is the output of the Filter with one particular coefficient $C_n = 0$. We can form the signal $(\phi - T_d) = e$ in the Filter and this is the error between the Filter output with $C_s = 0$, and the desired output. Hence:-

$$Y = \overline{[C_s S + e]^2}$$

$$\text{i.e. } Y = \overline{S^2} C_s^2 + \overline{2 e S} C_s + \overline{e^2} \quad (4)$$

This is the equation of a parabola, and since $\overline{S^2}$ and $\overline{e^2}$ must be positive there is a value of C_s which makes Y a minimum (fig.2). The mean square error as a function of the set of N coefficients is a paraboloid in $N + 1$ dimensions. It is shown for two coefficients in fig.2. It will be clear that the value of C_m obtained from equation (4) depends on the values of all the other coefficients, but that as we set the coefficients in turn to their optima we are certain to proceed towards the minimum value of Y . Also, provided we deal with all the coefficients enough times it does not matter what order we take them in; since we always move downwards on the paraboloid we are certain to get to the bottom eventually. This is the particular advantage conferred by the minimum mean square error criterion as applied to the Learning Filter; we are



AS A FUNCTION OF ANY ONE OF THE PARAMETERS c_n THE MEAN SQUARE ERROR IS ALWAYS A PARABOLA, BUT THIS PARABOLA DEPENDS ON THE VALUES OF ALL THE OTHER PARAMETERS

FIG. 2 DESCENT TO MINIMUM ERROR.

certain to reach a unique optimum solution if we are patient enough.

The parabola of equation (4) is termed the "influence curve" of the coefficient C_s . If we know three values of Y corresponding to three values of C_s , it is completely defined, and we can calculate C_m , the coefficient value corresponding to a minimum mean square error Y_m . It was originally intended to do this, taking $C_s = 0, -1, +1$, but equation (4), as written above, suggests a much better method. (This derivation of equation (4) arose when considering the design of the influence curve computer, Section 5.7.). We can immediately write, by differentiating equation (4):-

$$dY/dC_s = 2 \overline{S^2} C_s + \overline{eS}$$

and for a minimum:-

$$C_m = - \overline{eS} / \overline{S^2} \quad (5)$$

It is interesting to note that $(\overline{eS})$ is the cross-correlation between one sample signal (or sample product) and the error in the output of the Filter, and that $\overline{S^2}$ is the mean power of the sample signal. The coefficient C_m can therefore be thought of as a normalized cross-correlation coefficient. We may note from equation (5) that the coefficient C_m is only zero if there is absolutely no correlation between the sample signal and the error signal. In all other cases therefore adjustment of the coefficient to its calculated value C_m will produce some reduction in the mean square error.

2.4. The Teaching Signal

In section 2.3. we used the 'true message' or 'desired filter output' in forming the filter error. If the Filter is arranged so as to compute automatically values of C_m , and to adjust its own coefficients to the calculated values, then it is progressively 'learning' the correct settings. By inference the desired filter output is called the 'teaching' signal.

Suppose we wish to train the Filter as an optimum simulator. We have some 'black box' with an input signal and an output signal. If the output is used as the teaching signal $T(t)$ and the input as the input signal $S(t)$ to the Learning Filter then the Filter will do its best to produce $T(t)$ from $S(t)$, and will convert itself into a model of the system using only the normal input and output records of the system.

Suppose we wish to produce an optimum predictor. In this case we use only one signal during the training period, and the teaching track is obtained by sampling the signal at an advanced point.

To produce an optimum filter the input is a signal plus noise, and a record of the pure signal is used as the teaching signal.

The teaching signal is only used during the training period, once the filter has optimized itself the teaching signal is dispensed with. This is best clarified by an example. If we have a filter that is sufficiently complicated to handle music we can teach it what is music, and what is not, by feeding it with a large variety of pieces of music, pure on the teaching track, noisy on the input. After the filter has

optimized itself we feed it with a fresh piece of music plus noise (not used during the training period) and the filter will be able to filter the music from the noise. It will recognize what frequencies are permissible and reject any noise at other frequencies (together with off-pitch notes). It will also be aware of the probabilities and conditional probabilities of different notes and this will aid in the discrimination.

The Filter is adapted to a stationary time series, that is a series whose statistical properties do not change with time, consequently it will be rather conservative in its tastes and will attenuate music which is too progressive.

The Filter can adapt to a non-stationary time series if it is allowed to go on learning and if the learning time is considerably shorter than the time taken for the time series to change its properties.

Once the Filter has learned its job it can act as a prototype for other simpler filters, which are functional copies of it, but which do not themselves learn.

3 GENERAL DESIGN CONSIDERATIONS

3.1. Requirements

The basic requirements for making the Learning Filter are:-

- 1) A storage medium containing $S(t)$ and $T(t)$ from which the two time series can be read out many times during the training period.

- 2) An arrangement whereby delayed versions of the waveform $S(t)$ can be obtained (S_1, S_2, S_3 etc.).
- 3) Apparatus for multiplying waveforms together to form terms such as $S_{n_1} S_{n_2}, S_{n_1} S_{n_2} S_{n_3}$, etc.
- 4) Apparatus for multiplying the waveforms $S_n, S_{n_1} S_{n_2}$, etc. by variable coefficients to form terms such as $C_n S_n$, etc.
- 5) Apparatus for adding up the many terms of the operator polynomial to give:-

$$S_o' = \sum_0^N C_n S_n + \sum_0^N \sum_0^N C_{n_1 n_2} S_{n_1} S_{n_2} + \text{etc.}$$

- 6) Apparatus which computes Y over one run of the stored record for three values of a given coefficient.

If in addition the optimization is to be fully automatic with no human intervention then we must also have:-

- 7) A minimum computer, which from the three values of Y obtained at the end of a run, computes C_m and Y_m .
- 8) Apparatus which automatically adjusts the coefficients to the calculated values of C_m .
- 9) Sequence control of the whole process which arranges read out of $T(t)$ and $S(t)$ at the correct time and initiates computation and adjustment.

The above is summarized in the simplified block diagram, fig.3.

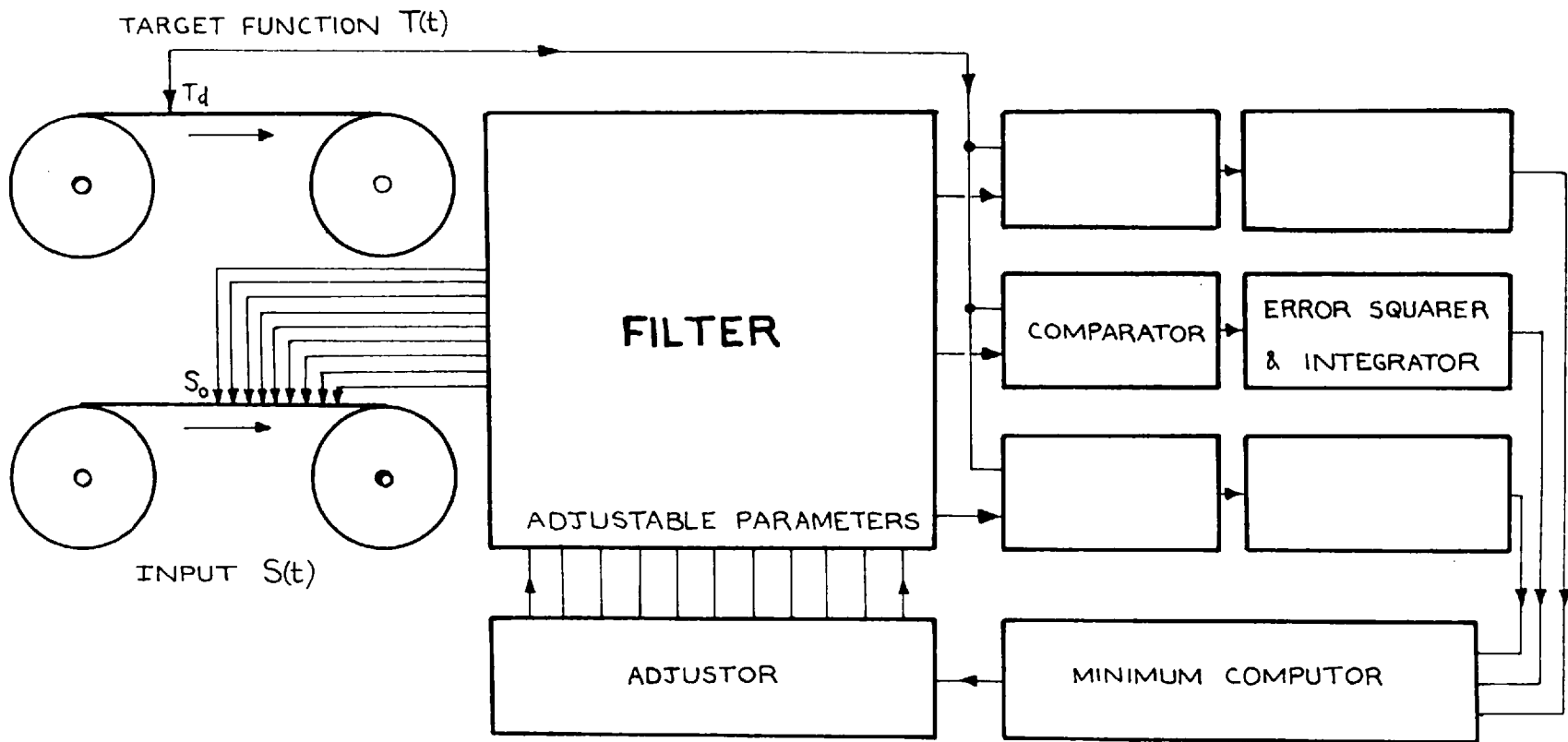


FIG. 3. LEARNING FILTER - SIMPLIFIED BLOCK DIAGRAM

It is necessary to specify the main features of the machine and these were chosen to be:-

- (a) Maximum input frequency = 200 c/s
- (b) Number of delayed samples = 18
- (c) Number of coefficients in the polynomial (eq.2.) = 100
- (d) Teaching records to comprise of the order of 10,000 amplitude samples known with an error of less than 1 %.

In deciding the above figures the aim was to specify a machine which would be sufficiently complicated to be of some real use and interest, and not just a pilot model for larger machines.

3.2. Realization of the Filter

The most convenient storage medium is magnetic tape. With normal amplitude recording, wear of the tape and 'dropouts' give errors much greater than 1 %, but at the standard tape speed of $7\frac{1}{2}$ "/sec. and using a maximum signal frequency of 200 c/s it is possible to use some type of frequency or pulse width modulation. Pulse rate modulation was finally chosen. With magnetic tape the necessary delayed versions of the signal can be produced by recording the signal simultaneously on 18 tracks on the tape and by using staggered pickup heads for playback. A further track is needed for the teaching signal, and a spare track for control purposes, making 20 in all.

It was originally considered that the Hall Effect multipliers described by Kuhrt & Hartel ⁶ would prove to be the most suitable method

for multiplying the delayed signals together. Extensive development was done by Mr.R.Woodcock who produced several successful units, but the Hall Effect multiplier was eventually discarded in favour of a completely new analogue multiplier (section 5.6.) which turned out to be better in this application.

The obvious choice for the 100 variable coefficients is 100 potentiometers. By applying a signal to one end of a potentiometer, and minus the same signal to the other end, then any coefficient between +1 and -1 can be obtained as the wiper is moved from one end to the other. The 100 potentiometers have to be servo-controlled in a fully automatic machine.

The simplest method of adding up the many terms of the operator polynomial is to use a conventional high gain analogue adding amplifier.

The computation of Y is a straightforward problem, and at first it was thought that digital circuits would be necessary to attain the required accuracy. It was later realised that a simple modification would enable analogue circuits to do the job (section 5.7.). This also changed the minimum computer from a digital to an analogue device.

Sequence control of the whole process in an automatic machine can be realised using relay and uniselector switching circuits.

The above account is summarised in the block diagram of the complete machine (fig.4.).

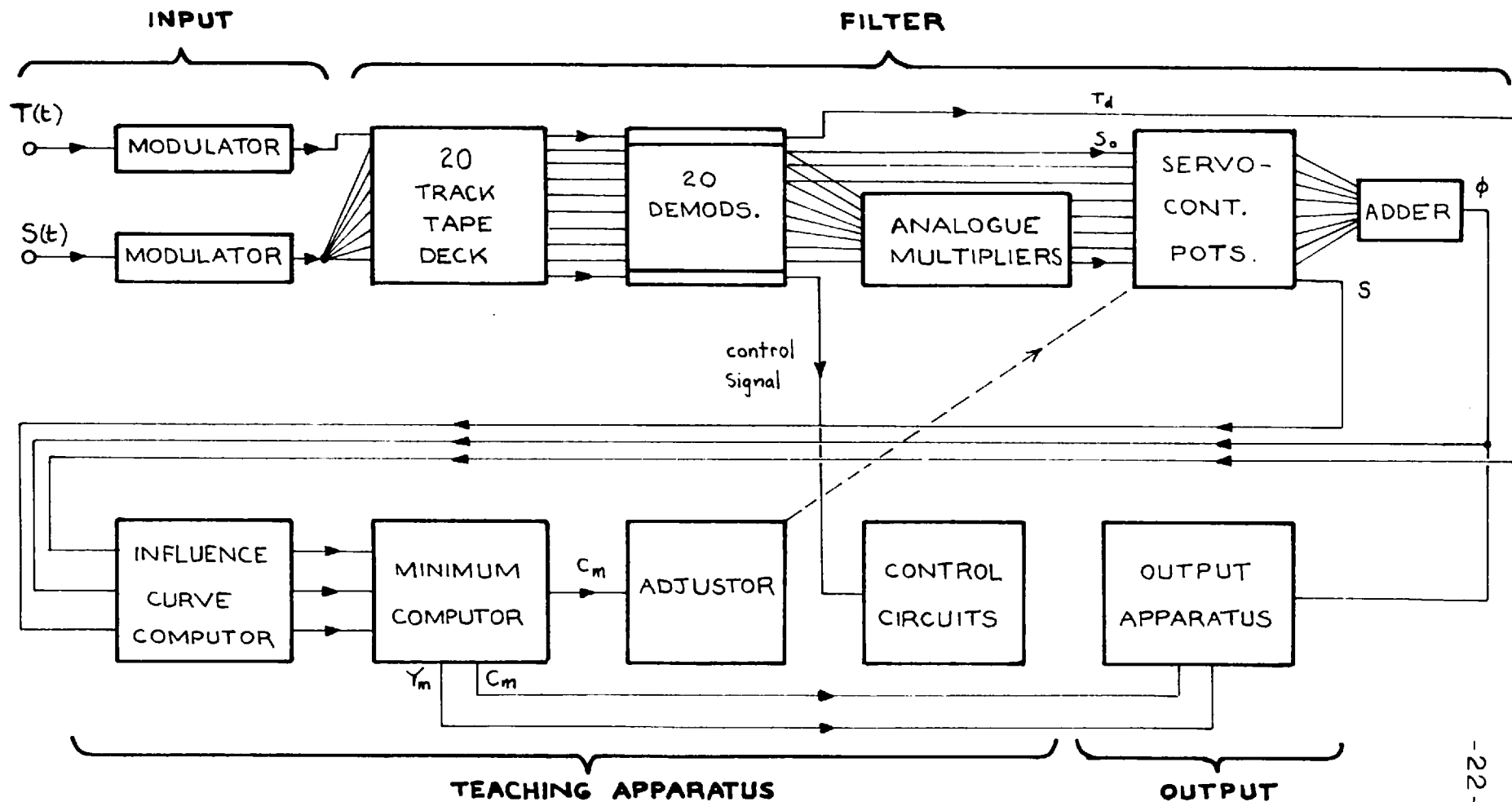


FIG. 4.

REALISATION OF THE LEARNING FILTER

4 STRATEGY

It was pointed out in Section 2 that the learning process converges to a unique solution, with all coefficients set to definite values, in whatever order we proceed with the adjustment of the coefficients to their respective optima. This is inevitable since each adjustment produces a decrease in the mean square error. With so many possible routes to the correct answer, we are still left with the task of deciding the best strategy of descent.

Strategy I

The first intention was to follow the "Southwell" relaxation method ⁷. That is, to adjust always the coefficient which gives the greatest reduction in the mean square error. Suppose there are 100 coefficients. We must first obtain the values of Y_m for every coefficient and store these values, together with the corresponding values of C_m . This requires 100 runs of the tape record and consumes quite a long time. If the record is 100 secs. long it takes 3 hours. We also need a store for 200 numbers each known to at least 0.1 %. Once we have the 100 values of Y_m we must search through them to find the smallest. Having found the smallest the corresponding coefficient must be set to the appropriate value of C_m . This is a straightforward computing problem, and equipment was designed which would do the job. It became clear however that to build the equipment would be both very costly and very time consuming. Also since each coefficient affects the C_m for every other coefficient (section 2) when one C_m has been

set the other 99 stored values are, strictly speaking, incorrect. We must therefore perform another 100 runs before setting another coefficient. At this rate there would only be one adjustment every 3 hours and the machine may be completely worn out before ever reaching a solution. We might improve slightly upon the above by assuming that we can set, say, the 10 most important coefficients using the stored values, and by assuming that the order of importance does not change after they are set. We would then only rank the coefficients in order of importance infrequently.

Strategy II

Realization of the above difficulties led to a much simpler method. The really significant fact is that although the machine takes a long time (100 secs.) to obtain Y_m and C_m for a particular coefficient, it takes only a short time (1 sec.) to adjust the coefficient. At the end of each run in Strategy I we obtain C_m for a particular coefficient. Since it takes only one second to adjust the coefficient we may as well do so. At the end of 100 runs we have now made 100 adjustments, including the coefficient which would have been adjusted by Strategy I. Furthermore at each adjustment the effect of all previous adjustments is taken into account, and so each step brings us safely towards the solution. It is clear that Strategy II must converge far more quickly than Strategy I. Also we have done away with the storage for 200 numbers, and with the apparatus required to arrange the Y_m 's in order of importance; a very large item in the original plans.

Mechanization of the Learning Process

Strategy II is readily mechanized since the machine can deal blindly with one coefficient after another. However almost certainly some coefficients will have only a small effect on the mean square error and it is largely a waste of time to keep on adjusting them just as often as the more important ones. If we are to endow the machine with some discrimination we must do it by a method which requires little extra equipment and computing time. One method is to note the reduction in Y_m which occurs when a coefficient is adjusted to its new setting C_m , and to use this as a measure of the significance of the coefficient. If the reduction in Y_m is less than a certain pre-set amount we can instruct the machine to miss the coefficient on the next cycle of the learning process, arranging to check its significance again later on in the learning process.

The machine can therefore be fully automatic, but even so it is desirable that manual control should be possible to enable the effect of different strategies and measures of significance to be tried out in practice, without changing the design of the machine. This was kept in mind during the design on the control circuits.

5 DESIGN AND PERFORMANCE OF SYSTEM PARTS

5.1.1. Tape Recorder

The recorder used for the magnetic tape record is shown in fig.5. The design was based on a standard 16 track tape deck manufactured



by Epsylon Ltd. The deck accommodates 1200' reels of tape and has two speeds, $7\frac{1}{2}$ "/sec. and 15"/sec. The tape width was increased from 1" to $1\frac{1}{4}$ " to allow for 20 tracks. A standard Epsylon power supply, and a standard control unit for manual control of play, record, spool, and rewind, were obtained. All the operations are controlled

fig.5. TAPE DECK

by relays and solenoids, thus making the deck suitable for inclusion in an automatic system.

5.1.2. Delay Unit

The standard stack of 16 record/replay magnetic heads, was



replaced by the specially designed delay unit (fig.6.). The working sketches were prepared by Professor D.Gabor before the research was commenced, so that progress would not be held up, and the detailed drawings were done by Mr.R.Woodcock. The design was modified slightly by Epsylon Ltd. during manufacture. The essential

fig.6. DELAY UNIT

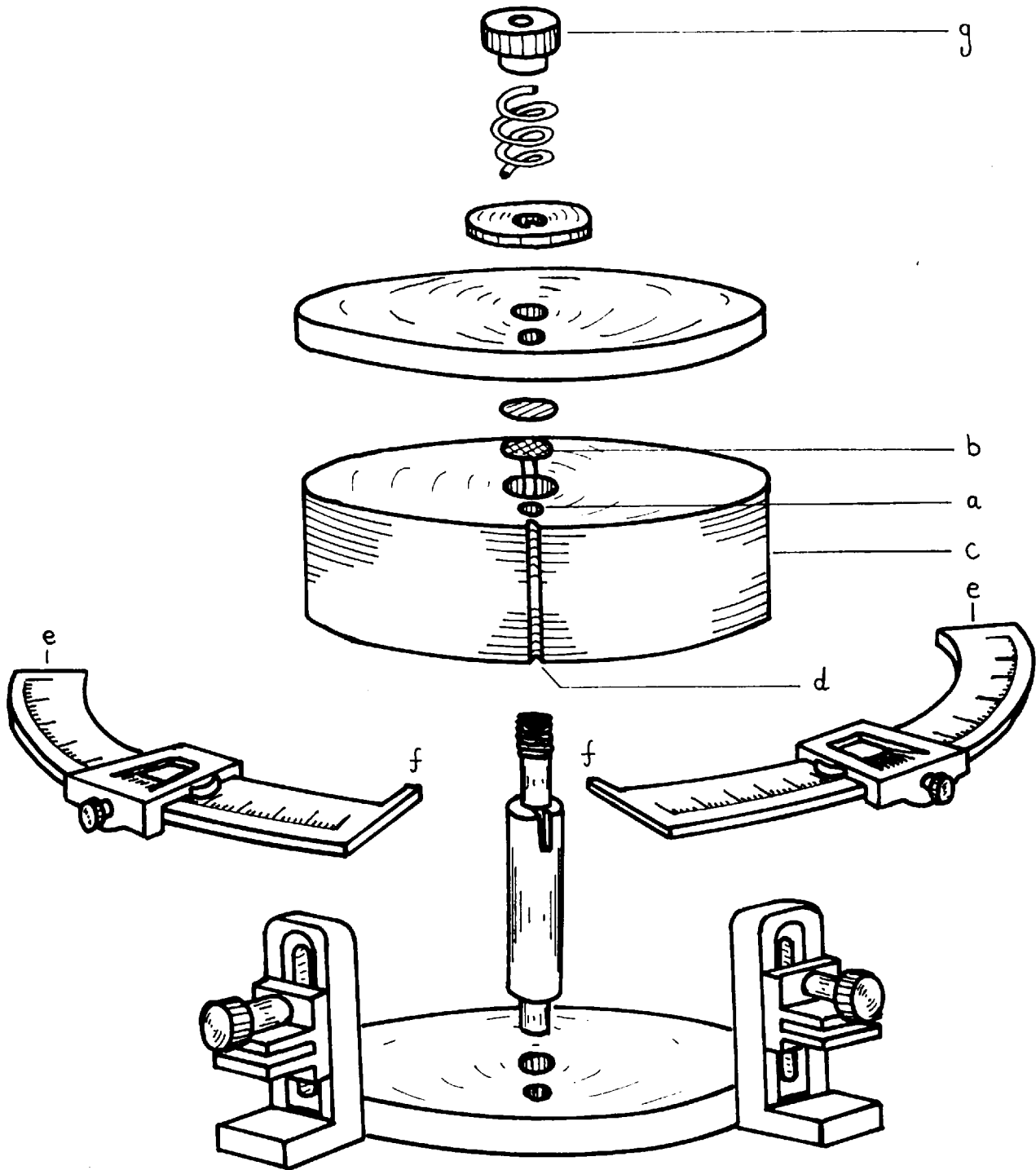


FIG. 7.

DELAY UNIT - EXPLODED SKETCH

part of the design is a cylindrical stack of 20 flat discs. A standard Epsilon record/replay head is inset in each disc with a circular mu-metal magnetic screen above and below (fig.7.). A locating hole (a) is provided in each disc so that the 20 heads (b) can be lined up at the front of the cylinder (c). The tape runs around the front of the cylinder over a sector of 70° . At the back of each disc is a slot (d). Two vernier scales (e), each with a finger (f) which fits the slots at the rear of the discs, slide in horizontal guides which can be moved up and down vertically. Signal delays of up to 0.2 secs. are obtained by recording with the heads in line and by staggering the heads on playback. Stagger is obtained by locking a disc in position using one vernier scale. The disc above is then rotated to the required position using the other vernier scale. This is then locked and the first vernier is removed and used to adjust the third disc, and so on. After all adjustments have been made the stack is clamped down using the lock-nut (g).

Satisfactory performance has been achieved from this unit, though experience has shown that a few improvements could be made.

5.1.3. Tape Deck Console

A console was designed to house the tape deck, and its control and power supply units (fig.8.). Provision was also made for mounting 20 preamplifiers (section 5.2.2.) and a record/replay switch unit. (section 5.1.4.). An air blower unit was installed to cool the capstan motor, since it was anticipated that this motor would have to run



continuously for many hours when operating automatically.

fig.8. TAPE DECK CONSOLE

5.1.4. Switch Unit

The circuit of the record/playback switch unit is given in fig.9. Switching is done by means of a 40 pole 2 way wafer switch. On record, 18 of the magnetic heads are connected in series, the other two being brought out separately for recording the teaching and control tracks. The group-divide switch enables the 18 signal heads to be divided into two series groups of 9 each, an extra modulator then being necessary. On playback the outputs from the 20 heads go to the inputs of the 20 pulse pre-amplifiers. So that the tape recorder can be used for other purposes if required, the 50 way plug carrying the head leads can be removed from the switch unit and plugged into a socket which brings the head connections out to a 50 way socket on the back panel of the tape deck console.

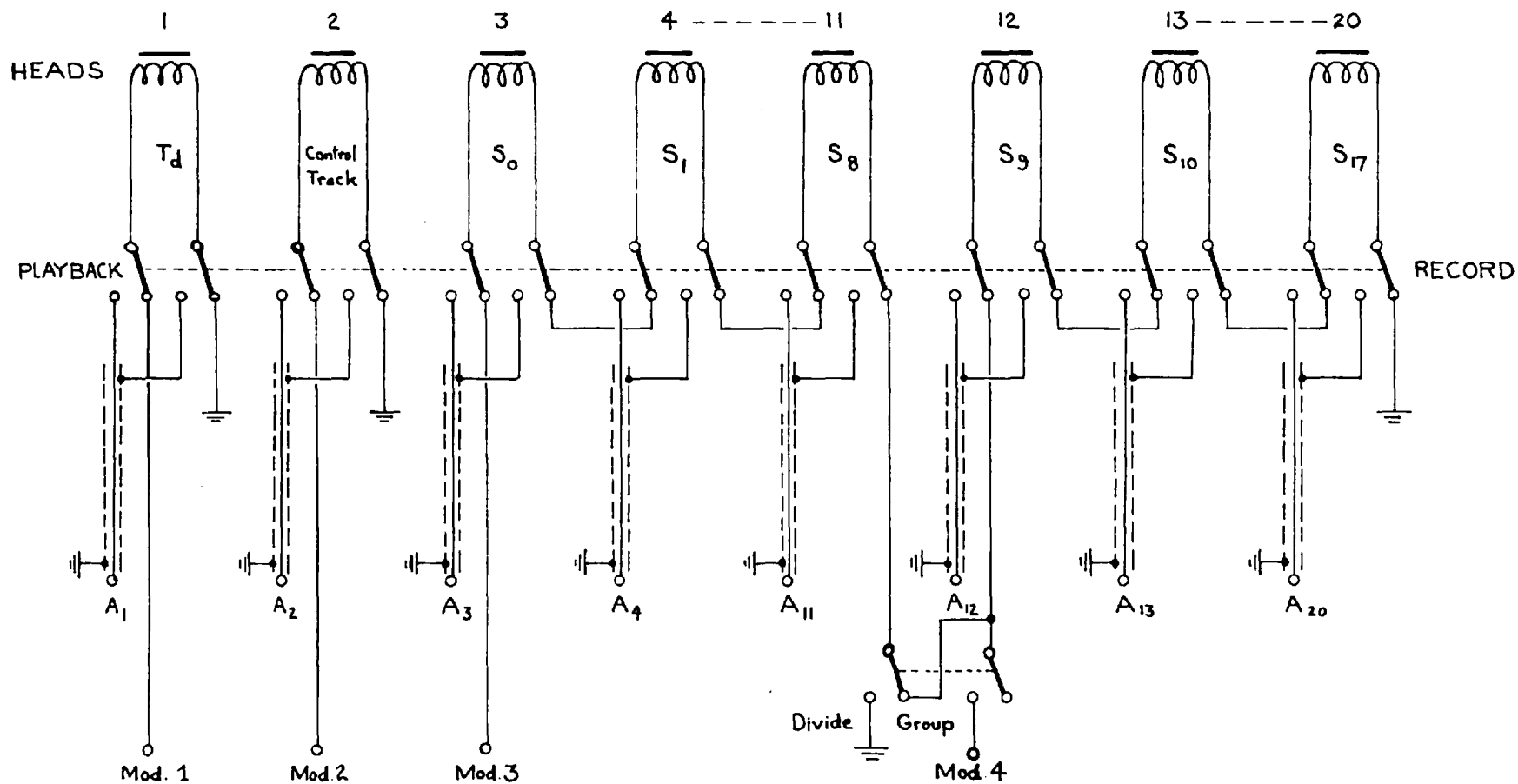


FIG. 9. SWITCH UNIT

5.2. Pulse Rate Modulation System

If a sine wave is amplitude recorded on magnetic tape, then on playback the envelope of the wave appears extremely ragged, and drops in signal level of 50 % and more are quite frequent. These changes in signal level make some form of pulse recording essential. This has the advantage that only 'pulse or no pulse' must be recognized at any point on the tape, and amplitude fluctuations are unimportant. This method also guards against wear of the tape. If a signal is recorded by varying the pulse frequency, then variations in the tape speed due to imperfect tape transport mechanism will cause a spurious output signal, and introduce noise into the system. For this reason it is desirable to vary the pulse frequency over as wide a range as possible.

5.2.1. Modulator

The modulator used (fig.10.) is based on a circuit due to Erath & Smith⁸, and called by them a "slope modulator". In the circuit of the complete modulator (fig.11.) the slope modulator is made up of V_1 , V_2 , and V_3 . The circuit makes use of the fact that in a 6AM6 pentode (V_1) the current is virtually independent of anode voltage over a wide range. If V_1 is used to dis-

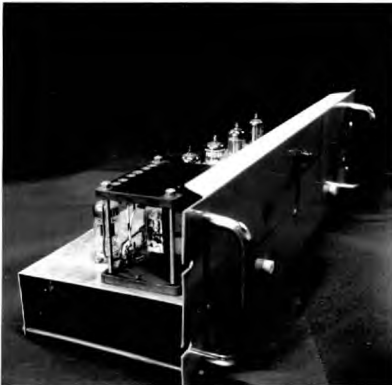


fig.10. MODULATOR

charge the capacitor C_1 , then since a constant current flows out of

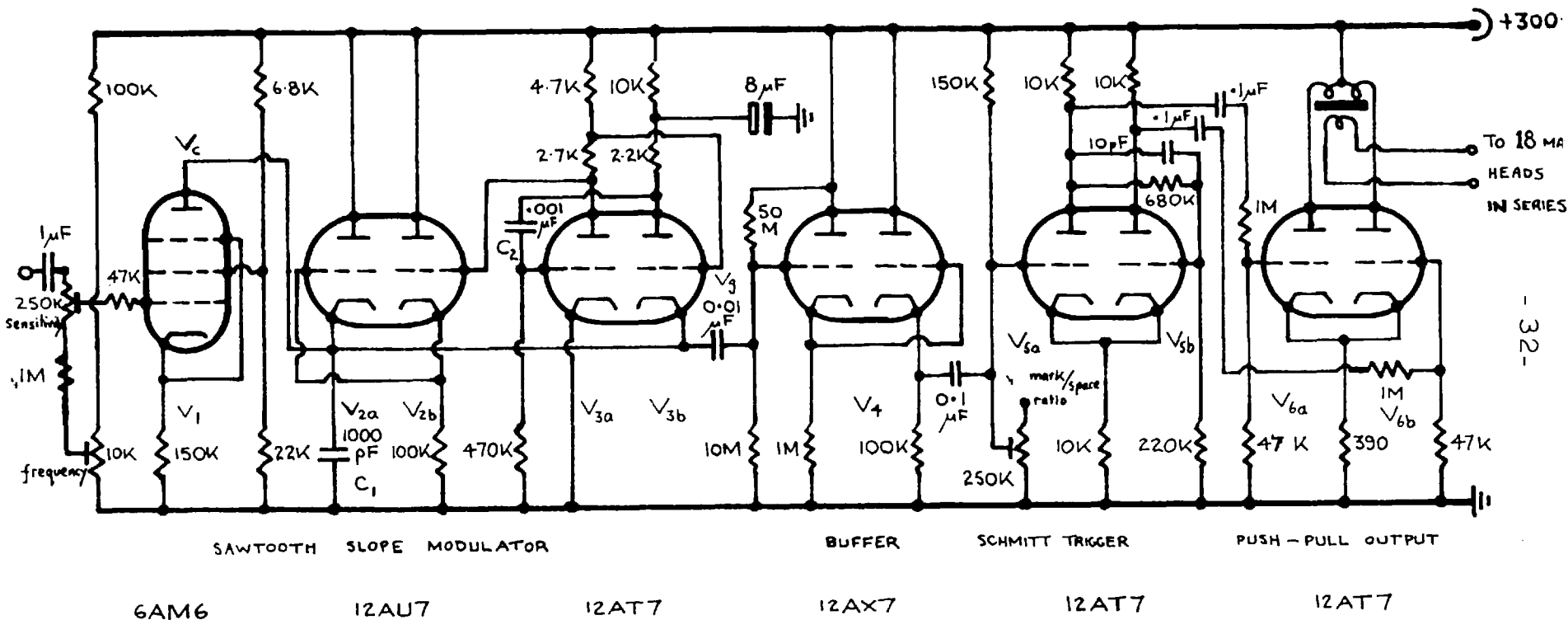


FIG. 11.

MODULATOR FOR USE WITH 20-TRACK TAPE RECORDER
(PULSE RATE MODULATION)

the capacitor, the voltage across the capacitor, V_c , falls linearly with time (fig.12.).

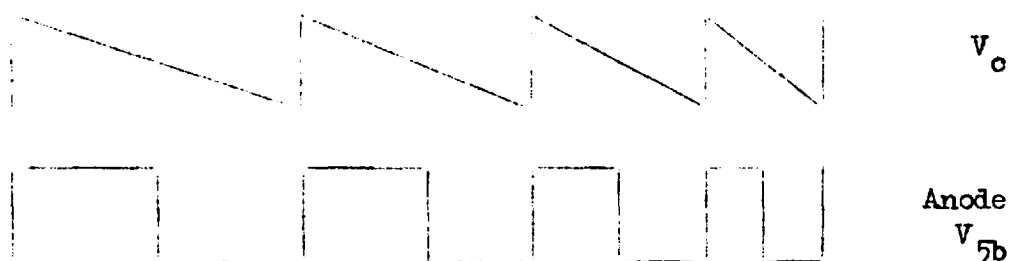


fig.12. MODULATOR WAVEFORMS

The valves V_{3a} and V_{3b} form a trigger circuit. V_{3a} is normally conducting since its grid-cathode voltage is zero and V_{3b} is turned off since its grid voltage V_g is kept below V_c by the current in V_{3a} . When V_c has fallen sufficiently for $(V_c - V_g)$ to equal the grid base of V_{3b} , current starts to flow in V_{3b} and its anode voltage starts to fall. This fall is coupled to the grid of V_{3a} and V_{3a} starts to turn off, increasing the grid voltage and anode current of V_{3b} . The process is regenerative and the anode voltage of V_{3a} rises rapidly to + 300 v. This is coupled by the cathode follower V_{2b} to the grid of the condenser recharge valve V_{2a} . V_{2a} is turned hard on in series with the condenser C_1 , recharging it very quickly. When the charge on C_2 has leaked away the trigger circuit $V_{3a} - V_{3b}$ reverts to its initial state and a new linear rundown of V_c commences. The rate at which current flows out of the condenser C_1 can be varied linearly by varying the grid voltage of V_1 . This varies the slope of the sawtooth wave produced (fig.12.), and hence the repetition frequency. The circuit was designed

to have a centre frequency of 4 Kc/s with no input voltage to the pentode grid. An input voltage of +5 volts gives a frequency of 7 Kc/s and -5 volts gives 1 Kc/s. These figures are compatible with a tape speed of $7\frac{1}{2}$ "/sec. The circuit is linear over the range 1-7 Kc/s and its performance is shown in fig.13. The pentode cathode resistor provides negative feedback which improves the linearity and also makes the circuit insensitive to changes in V_1 . The sawtooth wave from the slope modulator is coupled to a Schmitt trigger circuit $V_{5a} - V_{5b}$ via the cathode follower stage V_4 . Square waves equal in amplitude but opposite in phase are obtained from the anodes of V_{5a} and V_{5b} fig.12. These square waves are fed to the grids of the push-pull output stage $V_{6a} - V_{6b}$. The transformer used in the output stage was designed specially for this application. The transformer is wound in 6 sections, and a shell construction using 4 C-type cores made from grain orientated silicon iron strip is employed, to reduce leakage inductance. The transformer frequency response is 3 dB down at 23 cycles/sec and 110 Kc/s. The transformer feeds up to 18 recording heads connected in series, with a current square wave of 50 mA peak. This current is sufficient to saturate the magnetic tape. With 18 heads the power delivered is 1 watt. Two modulators have been built for single input problems, the teaching track modulator can be used at a different time to record the control track. Provision has been made for the third modulator, required for two input problems.

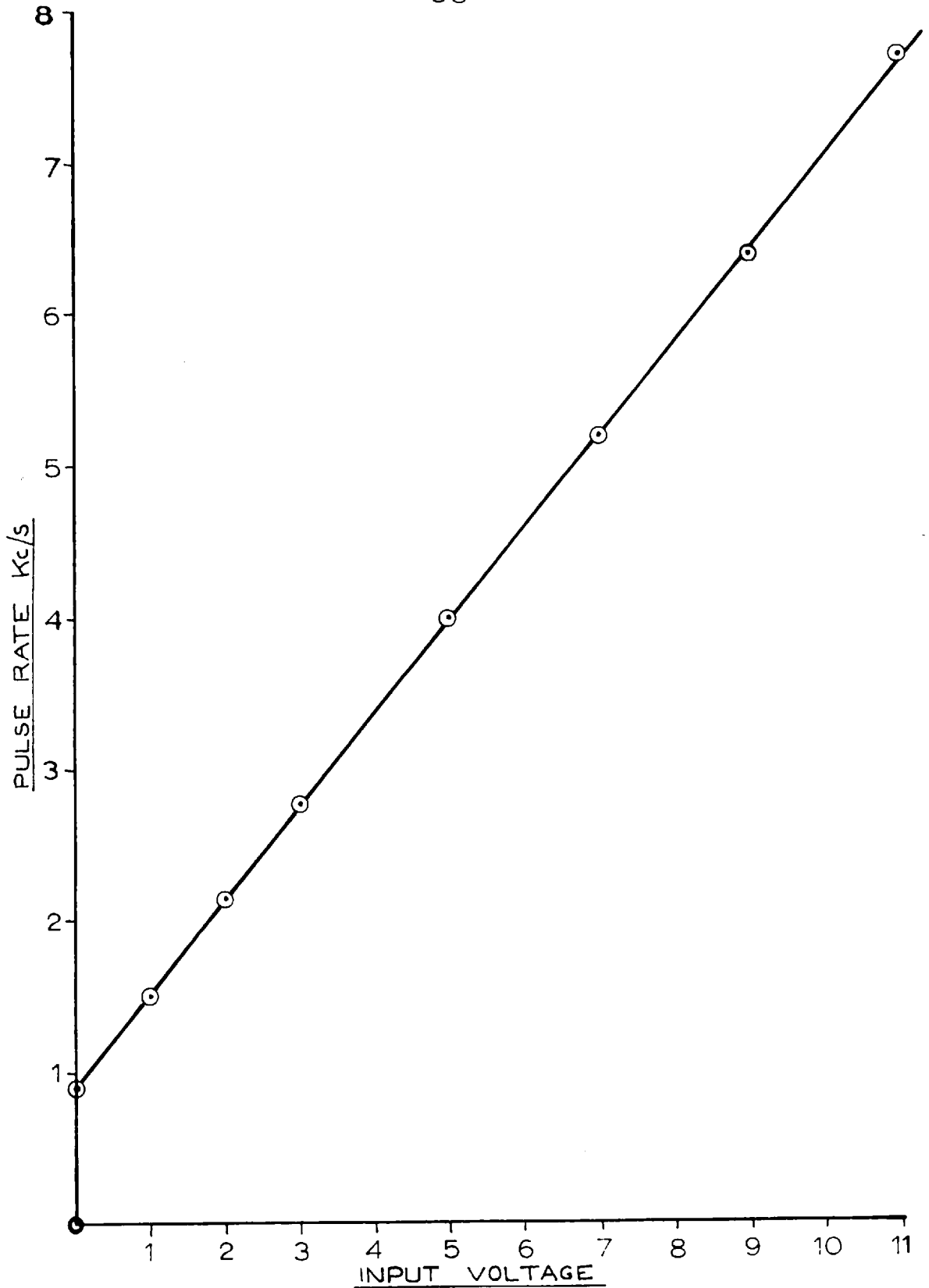


FIG. 13.

PULSE RATE MODULATOR - LINEARITY

5.2.2. Pulse Pre-Amplifier

The signal picked up from the magnetic tape on playback is proportional to the rate of change of flux in the magnetic head. The head output signal is made up of positive and negative pulses corresponding to the vertical edges of the recorded square wave. The size of the pulses is increased at the input of the pulse pre-amplifier (fig.14.)

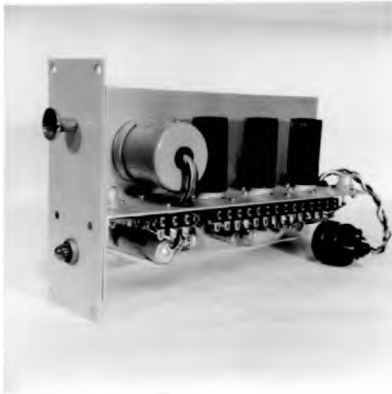
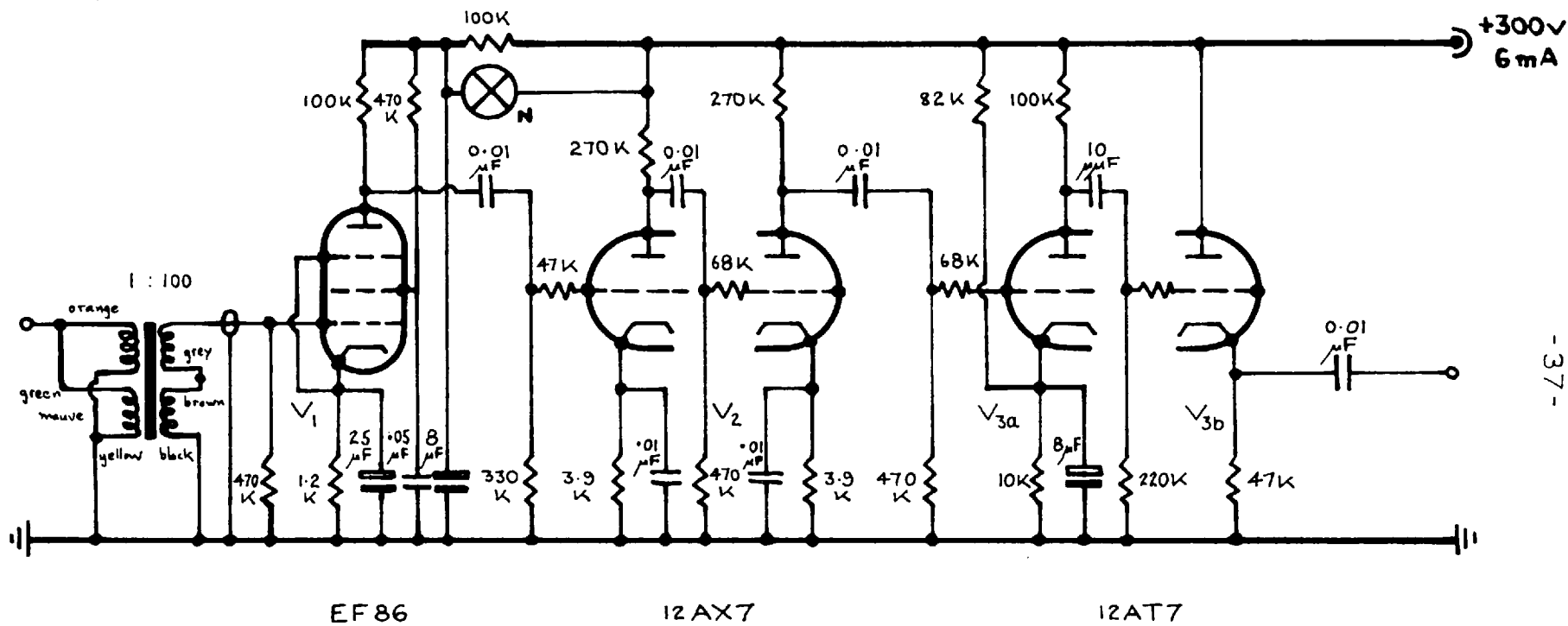


fig.14. PRE-AMPLIFIER

by means of a 100:1 step-up transformer. They are then amplified by a low noise EF86 gain stage. (V_1 , fig.15.) The pulses are further amplified by V_2 , a two stage 12AX7 clipper amplifier. The positive pulses correspond to the vertical edge of the original sawtooth wave, and all the recorded information is contained in the rate of occurrence of these pulses. The negative pulses can therefore be discarded. They are removed together with the system background noise by V_{3a} . The cathode of this valve is held at +30 volts and consequently only positive pulses greater than +25 volts are amplified. The pulses appearing at the anode of V_{3a} are differentiated and yield an output of sharp positive spikes uniform in size and shape. These spikes act as trigger pulses at the input of a pulse rate demodulator.



**FIG. 15. PREAMPLIFIER FOR USE WITH 20-TRACK TAPE RECORDER
(PULSE RATE MODULATION)**

5.2.3. Pulse Rate Demodulator

The first stage of the demodulator (fig.16.) is a monostable cathode coupled trigger circuit. Normally V_{lb} (fig.17.) is conducting

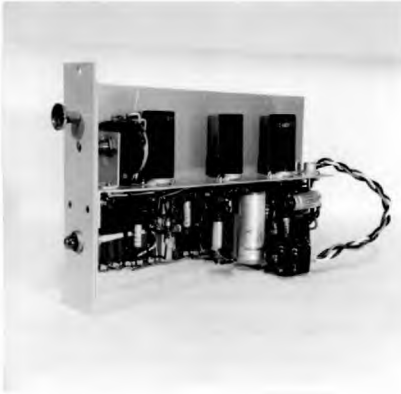
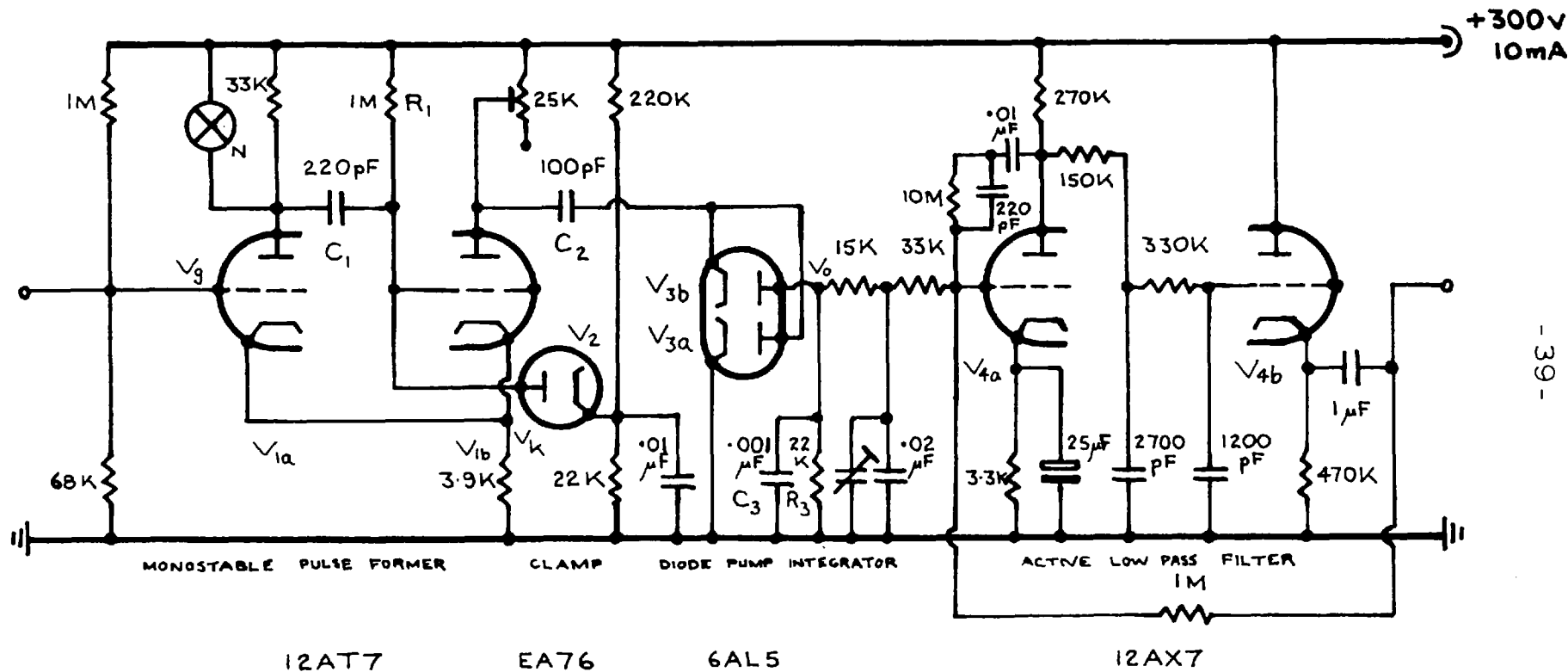


fig.16. DEMULATOR

since its grid resistor is returned to +300 volts and V_{la} is off since V_k is greater than V_g . The diode V_2 , clamps the grid of V_{lb} at +30 volts and ensures that the resting state of V_{lb} is always the same. If a positive pulse arrives at the input, V_{la} turns on and V_{lb} turns off, the change being regenerative. At

a fixed time later, determined mainly by the time constant $C_1 R_1$, V_1 reverts to its original state. Pulses of fixed length appear at the anode of V_{lb} . Their rate of occurrence depends on the recorded signal. The anode resistor of V_{lb} can be reduced to zero without affecting the operation of the circuit and forms the demodulator gain control. The neon indicator across the anode resistor of V_{la} is normally off, and only lights when a signal is passing through the demodulator. This feature has proved very useful.

The train of standard pulses from the first stage is integrated by the diode pump integrator $V_{3a} - V_{3b}$. When the input pulse arrives, current flows through V_{3a} and condenser C_2 obtains a charge equal to EC_2 . At the end of the pulse V_{3b} conducts, and if C_3 is much greater than C_2 , C_2 is discharged, the voltage across C_3 (V_o) being



- 30 -

FIG. 17.

DEMODULATOR FOR USE WITH 20-TRACK TAPE RECORDER
(PULSE RATE MODULATION)

increased by EC_2 / C_3 . The presence of R_3 means that charge is continually lost from C_3 , and since for a steady state the charge lost must equal the charge gained.

$$\begin{aligned} V_o / R_3 &= nEC_2 \quad \text{where } n \text{ is the number of pulses/sec.} \\ V_o &= nEC_2 R_3 \end{aligned}$$

That is, there is a linear relationship between output voltage and pulse rate. The output of the diode pump integrator has a large ripple at the pulse rate. A low pass filter is required which allows the signal frequencies to pass, but which attenuates strongly at the carrier pulse frequencies. This was realized by means of the specially designed active R-C filter ¹⁰, $V_{4a} - V_{4b}$. The filter response is 2 % down at 250 c/s and 40 dB down at 1 Kc/s. The filter also provides gain and a low impedance output. Since frequencies below 5 c/s were not required the output is A.C. coupled. A D.C. coupled output stage would permit the recording and playback of signals extending in frequency down to D.C.

A trigger pulse rate of 7 Kc/s gives +5 volts out and 1 Kc/s gives -5 volts. The overall gain of the record/playback system is therefore unity.

5.2.4. Overall Performance of Record/Playback System

The overall linearity of the record/playback system is shown in fig.18. The overall frequency response is given in fig.19.

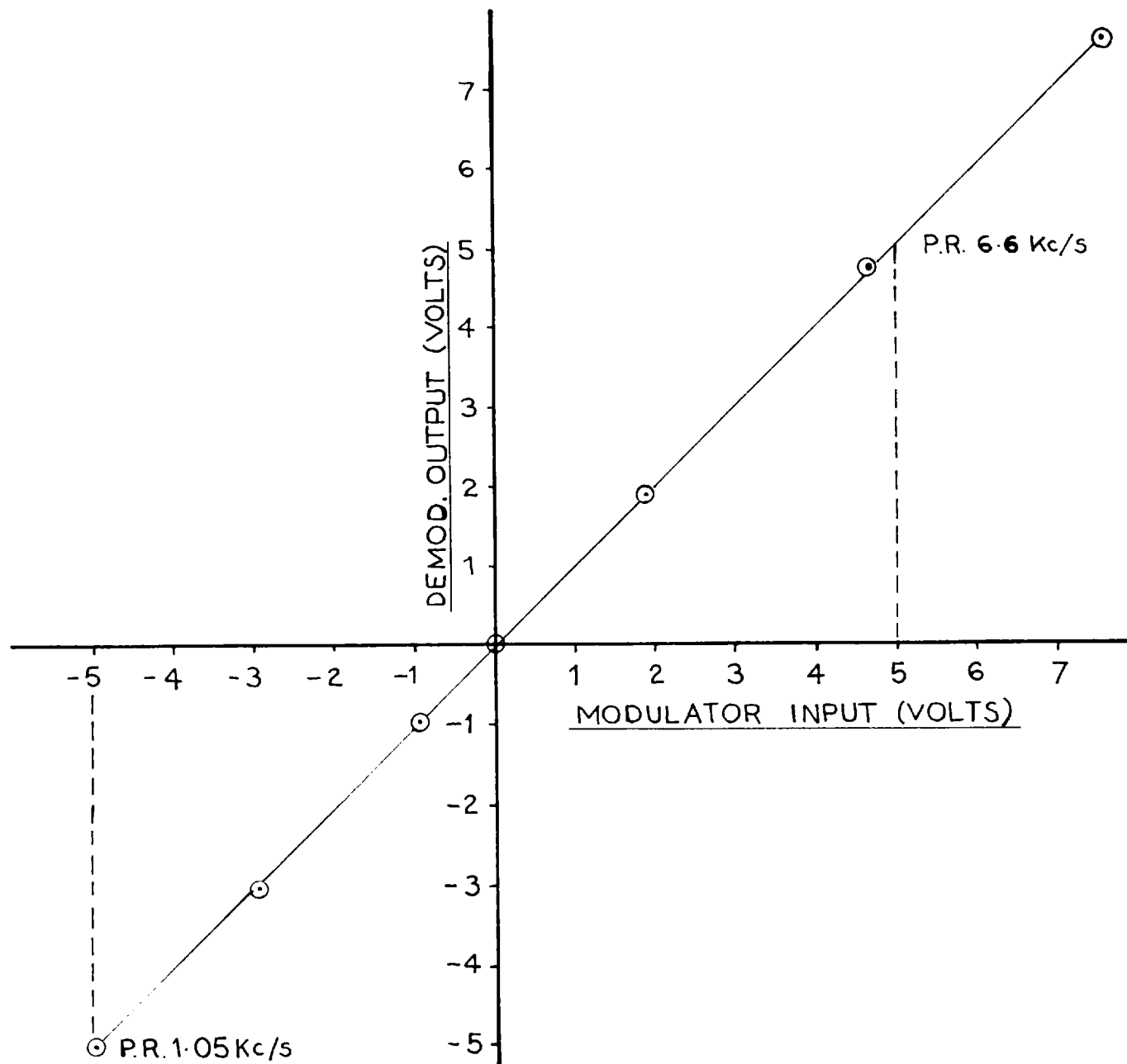


FIG. 18. PULSE RATE MODULATION SYSTEM - OVERALL LINEARITY.

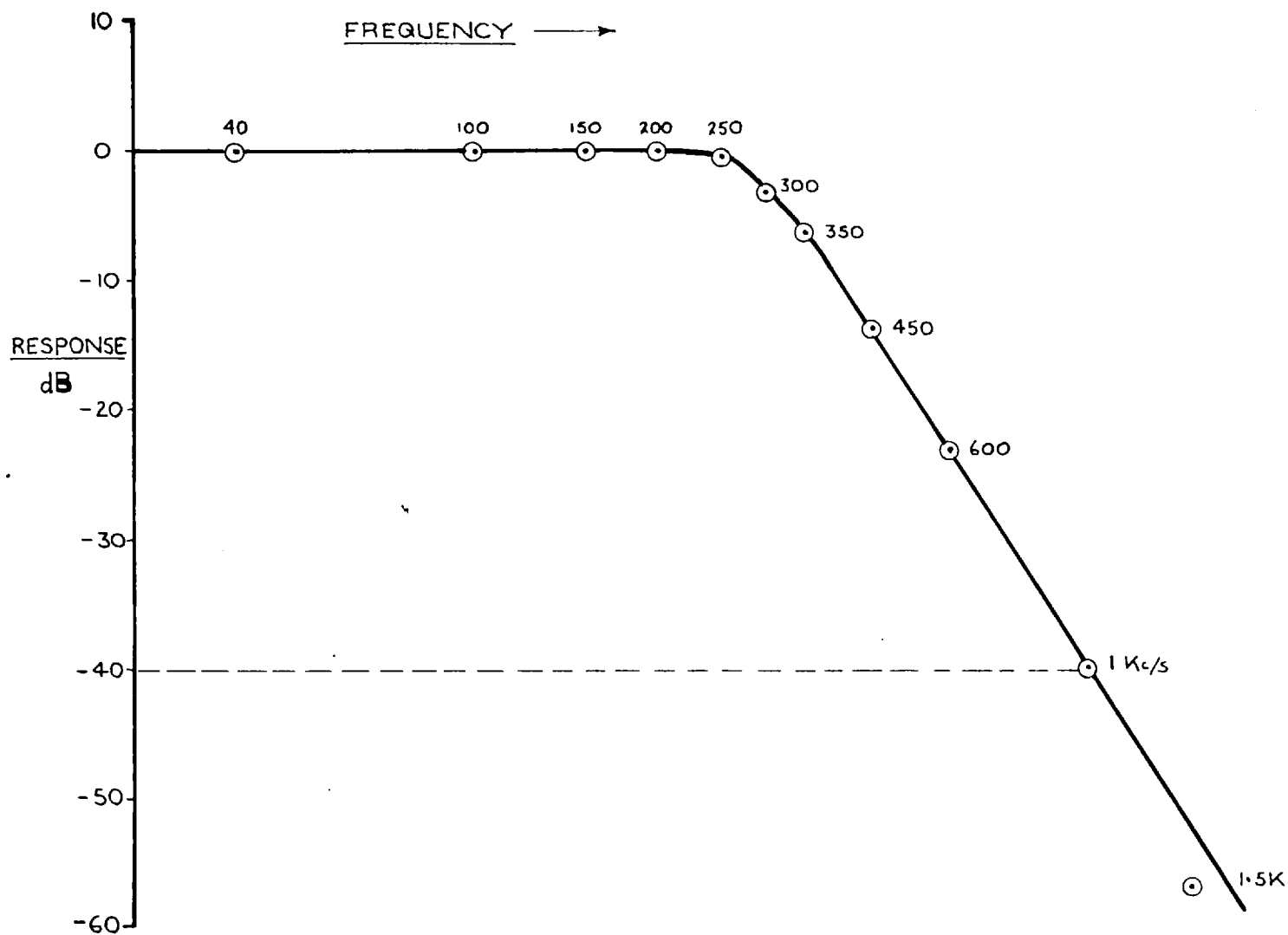


FIG. 19. PULSE RATE MODULATION SYSTEM - FREQUENCY RESPONSE

A summary of the main technical details is given below:-

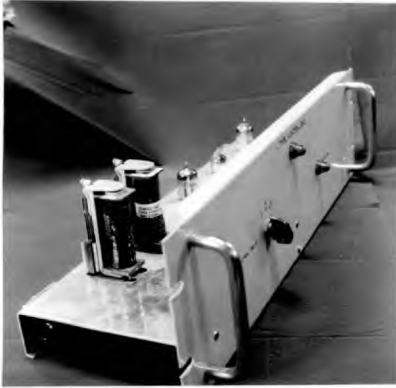
Recording current	=	± 50 mA
Recording frequency (square waves)	=	1 Kc/s - 7 Kc/s
Deviation ratio	=	75 %
Tape speed	=	$7\frac{1}{2}$ "/sec.
Total wow and flutter	=	0.3 %
Tape (instrument quality)	=	1200' x $1\frac{1}{4}$ '
Maximum playback signal from head	=	100 microvolts
Minimum allowable signal from head	=	5 microvolts
Ratio; maximum signal/minimum signal	=	20
Total harmonic distortion	=	- 40 dB
Hum level	=	- 50 dB
Signal amplitude	=	5 volts
Overall gain	=	1
Signal frequency band	=	5 - 250 c/s

The construction of the 20 pre-amplifiers and 20 demodulators was contracted outside the College, with the co-operation of the Admiralty. Full drawings and a working prototype of each were provided. The record/playback system has proved very satisfactory in use.

5.3. Tape Control Unit for Automatic Play-Rewind

It was anticipated that in a complicated problem it might be necessary to play the tape record up to a few thousand times. Under these circumstances it is very desirable that the play and rewind of

the record should be automatically controlled by the machine. A tape control unit, fig.20, was built for this purpose. Throughout the length



of the record the control track is recorded with pulses occurring at a constant 1 Kc/s rate. On playback there is then no output from the control track demodulator. On rewind the pulse rate builds up gradually to about 10 Kc/s due to the fast tape speed on rewind. Again there is no

fig.20. TAPE CONTROL UNIT output from the demodulator.

During automatic operation of the Learning Filter a signal on the control track causes the tape to reverse at the end of 'play' and 'rewind'. The control signal is a 5 volt peak 100 c/s sine wave. One signal burst is recorded at the end of the tape record. The tape is then allowed to rewind at high speed and another signal burst is recorded, in the correct position to stop the rewind, while the tape is running at high speed. After stopping, the tape plays backwards over the 'stop' signal giving the spurious signals shown in fig.21.

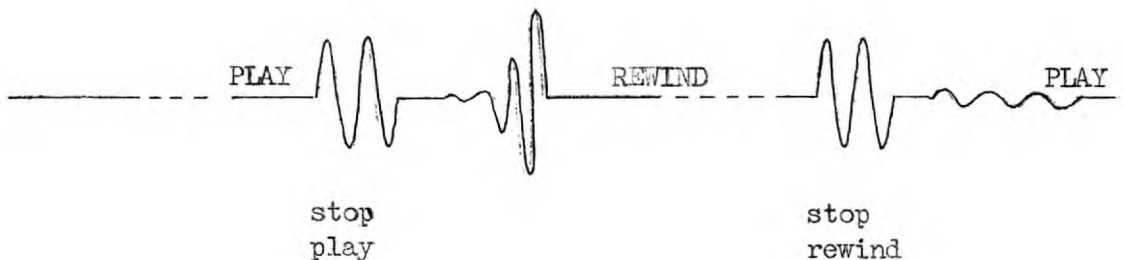


fig.21. SIGNALS FROM THE CONTROL TRACK

The tape control unit must discriminate against these. This is done in two ways. First of all the tape control unit has a selective gain stage with a peak in its frequency response at 100 c/s, obtained as shown in the circuit diagram fig.22. with negative feedback through a parallel T filter. Secondly, a diode clipper rejects all signals of less than 3 volts peak. A second diode clips all signals greater than 5 volts peak, to prevent the occasional large pulses from the demodulator (occurring when there is a complete 'dropout' on the tape) from actuating the control relay. The control signals reaching the integrating circuit (R_1C_1) correspond to the positive peaks of the original 100 c/s signal, and cause the build up of a negative charge on C_1 . The anode of V_{3a} becomes progressively more positive and this change is D.C. coupled to the grid of V_{3b} . After the control signal has been arriving for a short while V_{3b} is turned on and relay C operates. Contact c2 changes over and current flows through the relay into the $8\ \mu\text{F}$ condenser ensuring a snap-on action for the relay, and ensuring that it remains on for a fixed time. As soon as C operates the tape is stopped and may, after a slight delay, start to play backwards over the stop signal. If it starts to rewind, C will not operate again, since contact c1 has paralysed the amplifier for a few seconds by driving the grid of V_{1b} positive. If it starts to play, C is prevented from operating on the spurious signal (fig.21.) by the combination of frequency and amplitude discrimination discussed above. Even with sine wave inputs of full amplitude, C does not operate if the frequency lies outside the range

70 - 140 c/s, and the operate time delay increases greatly as the frequency deviates from 100 c/s.

The unit has proved very satisfactory and during the longest experiment so far performed (500 runs) it worked infallibly in controlling the automatic play-rewind. The position error, when stopping the tape from a rewind speed of about 100 "/sec. is about $\pm 1"$, after the unit has warmed up.

5.4. Coefficient Store

5.4.1. Servo-Controlled Potentiometers

The coefficients C_n , $C_{n_1 n_2}$, etc. are set up on precision 50 K.ohm potentiometers.(on loan from the Admiralty). There are 96 potentiometers, arranged in groups of 8 (fig.23.). The spindle of each potentiometer projects downwards into an oil



meter projects downwards into an oil bath, and each spindle has two magnetic clutches associated with it. The potentiometer wiper can be rotated clockwise or anticlockwise, depending on which clutch is energised, the motion being derived through gearing and the clutch, from a continuously rotating shaft.

fig.23. POTENTIOMETERS

5.4.2. Potentiometer Coupling Units

Coupling units have been designed for use with the potentiometers. Each potentiometer has one coupling unit which feeds it with

equal and opposite signals; enabling any coefficient between +1 and -1 to be obtained, as the potentiometer wiper is moved from one end of its track to the other. The coupling units are arranged in groups of 8 (fig.24.), the circuit diagram of one unit is given in fig.25. The

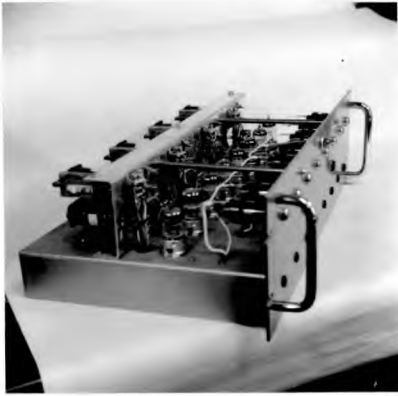


fig.24. COUPLING UNITS

input signal is applied to the grid of V_{1a} and this valve acts as a phase splitter, signals being taken from its anode and cathode. The signal from the anode of V_{1a} is fed to the cathode follower V_{1b} . Resistor R is then adjusted so that the signals at the two cathodes are equal and opposite. Relay K_n is normally not energised and the two signals go via relay contacts to the two ends of the 50 K.ohm potentiometer. Both cathodes are at +100 volts D.C. and no direct current flows in the potentiometer. The A.C. signal from the wiper passes through condenser C , to one input on the multi-input adder which performs the summation of all the terms in equation (2). The 150 K.ohm series resistor R' enables the signal from the wiper to be removed by applying earth to point E . This is done only when the influence curve of the potentiometer is being computed. The adder then forms the sum of all the terms with $C_s = 0$, and this is used to form y_0 for coefficient C_s .

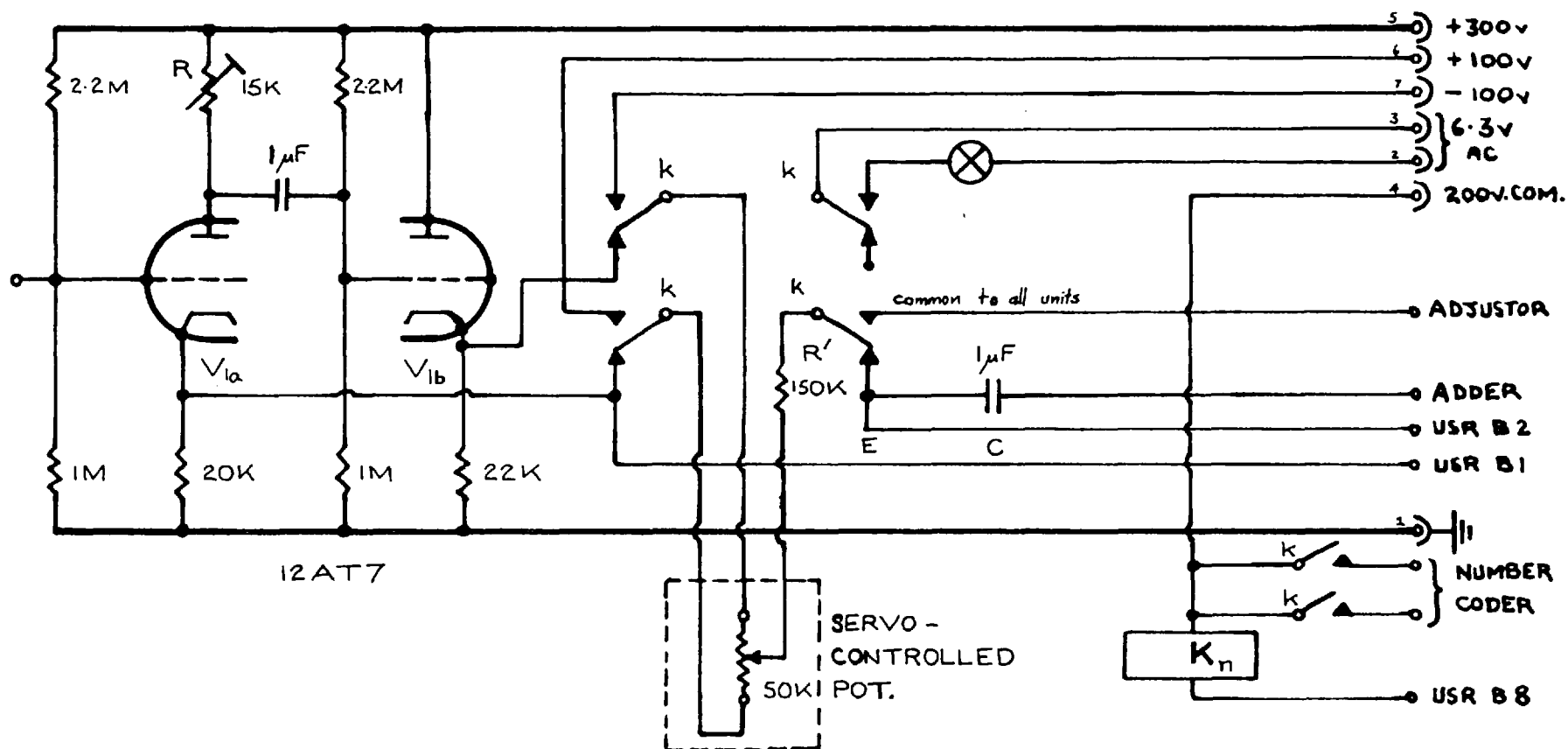


FIG. 25.

POT. COUPLING UNIT.

5.4.3. Potentiometer Adjustor

When it is desired to set the coefficient to a new value relay K_n is energised (fig.25.). The ends of the 50 K.ohm potentiometer are then disconnected from the phase splitter and connected to the ± 100 volts reference supply. The wiper is disconnected from the adder and connected to one input of the potentiometer adjustor (fig.26.)



fig.26. ADJUSTOR

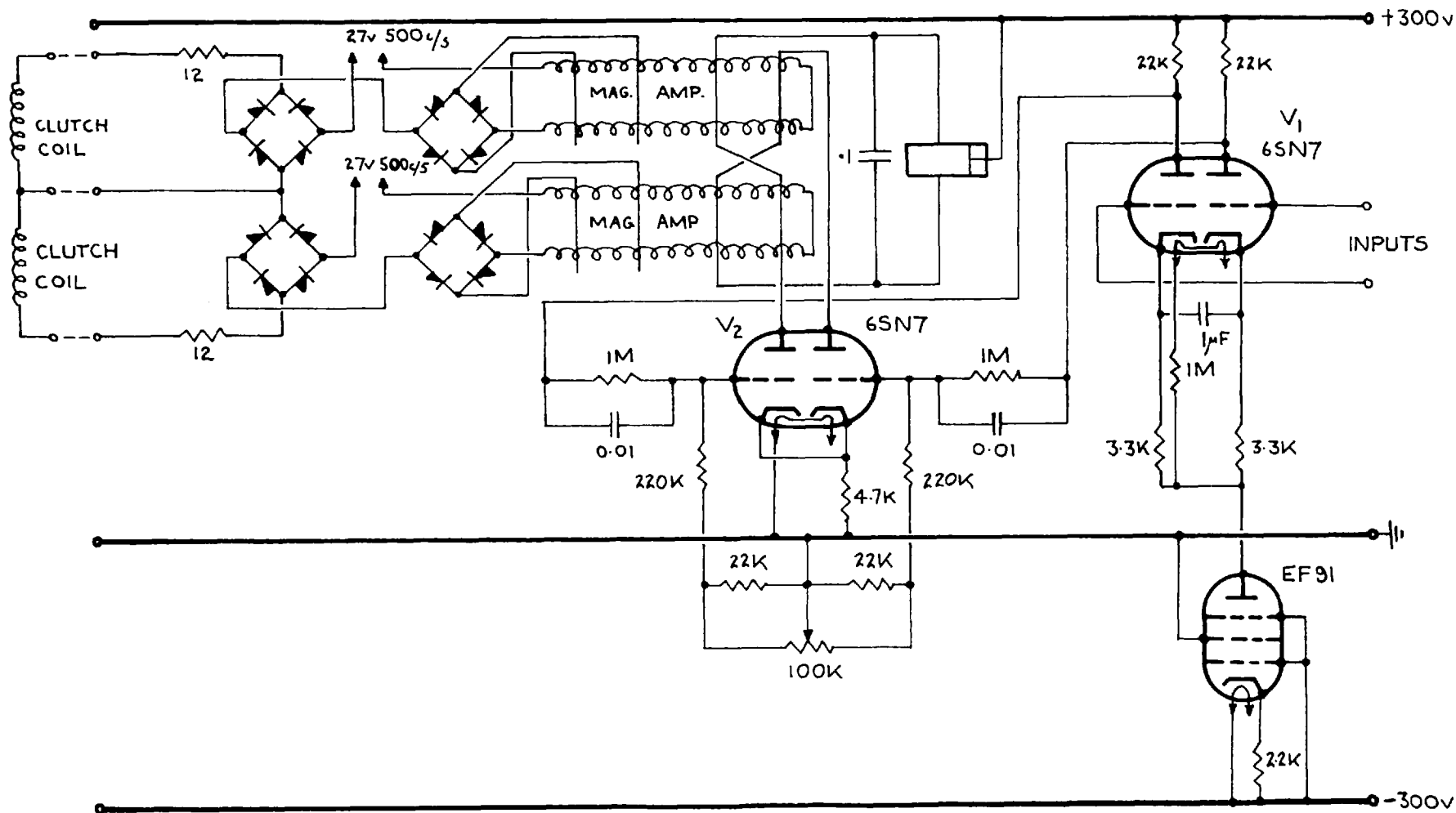
(on loan from the Admiralty). The other adjustor input is fed with a voltage which corresponds to the new coefficient. The difference amplifier V_1, V_2 , fig.27. obtains the difference between these two inputs, and the difference is amplified by a magnetic amplifier. The magnetic amplifier energises one or other of the

two magnetic clutches, and the potentiometer spindle is rotated in such a direction that the difference between the two adjustor inputs is minimized. Using this servo-system it is possible to set up coefficients on the potentiometers with an error less than 0.2 % of full scale.

5.5. Analogue Multiplication (General)

5.5.1. Specification

In the Learning Filter two types of analogue multiplier are required. The first, used for obtaining the products of the signal samples (e.g. $S_{n_1} S_{n_2}$) must multiply input signals lying in the frequency



-51-

FIG. 27. ADJUSTOR CIRCUIT

band 5 - 200 c/s, with an error less than 2 % of full scale, and since 76 are required these multipliers must be simple and not expensive.

The second type, required in the teaching apparatus, must be as accurate and stable as is practicable. Since they are used for squaring the filter error before integration it is essential that they should yield the D.C. component of the product. These multipliers can be fairly complex since only five are needed.

5.5.2. Review of Known Methods

At the start of the research work the above requirements appeared formidable. No suitable method of analogue multiplication was revealed by a thorough search of the literature, although many reports on a great variety of methods came to light. The most practical of these are tabulated in fig.28, with a very brief summary of their main features, together with performance figures for representative examples of each type. Analogue multipliers may be divided into two main groups,

- (a) those obtaining a product directly and,
- (b) those using a square law element and making use of the expression;-

$$xy = \frac{1}{4} \left[(x + y)^2 - (x - y)^2 \right]$$

Such a brief account as that given in fig.28. cannot do justice to the various authors referred to, but it makes clear the main point, that no existing multiplier came near to satisfying our requirements.

NO.	TYPE OF MULTIPLIER	REF.	NO. OF QUADRANTS	HIGH FREQ. LIMIT	TYPICAL ACCURACY %	NOTES ON APPARATUS REQUIRED 1 or 2 quadrant multipliers can always be arranged to give a 4 quadrant product, with increased complexity.
1	SERVO-DRIVEN GANGED POTENTIOMETERS	11	1	10 c/s	0.01	Precision electromechanical engineering.
2	CROSSED FIELDS C.R.T. MULTIPLIER	12	4	5 Kc/s	2	C.R.T., photocell, extensive auxiliary gear
3	HALL EFFECT MULTIPLIER	13	4	(a) 200 c/s (b) 1 Mc/s	1	Low wafer imp. (5 a) causes matching difficulties for D.C. working
4	DYNAMOMETER WATTMETER		4	1 c/s	0.1	Precision instrument
5	LOG. - ANTILOG. MULTIPLIER	14	1	200 Kc/s	5	About 12 valves, special nonlinear elements.
6	PULSE HEIGHT - PULSE WIDTH	15	1	20 c/s	0.01	About 20 valves
7	TIME SHARING MULTIPLIER	16	1	200 c/s	0.2	About 10 valves
8	PULSE COINCIDENCE	17	1	500 c/s	2	About 7 valves
9	A.M. - A.M. MODULATION	18	4	5 Kc/s	10	About 8 valves, standard I.F. components
10	F.M. - A.M. MODULATION	19	4	2 Kc/s	1	About 24 valves
11	BINARY STEP ATTENUATORS	20	4	2 c/s	0.1	About 100 valves, 20 high speed relays, 70 precision res.
12	STRAIN GAUGE MULTIPLIER	21	4	300 c/s	2	Special electromechanical device, matched strain gauges
13	THERMISTOR MULTIPLIER	22	2	(a) 1 c/s (b) 1 Mc/s	2	Simple amplifier, specially made thermistor unit
14	MAGNETIC AMPLIFIERS	23	4	100 c/s	5	Careful design needed, fairly complex
15	THYRISTE MULTIPLIER (SQ. LAW)	24	4	1 Kc/s	1	Specially chosen elements, fairly complex.
16	VALVE CHARACTERISTIC (" ")	25	1	200 Kc/s	5	Few valves, but preselection needed, frequent resetting.
17	TEMPERATURE LIMITED DIODE (" ")	26	1	0.1 c/s	3	Simple, $I_a = K I_h^2$
18	SPECIAL TUBES (" ")	27	4	80 Kc/s	1	7 valves, 2 special beam deflection tubes.
19	STRAIGHT LINE APPROX ⁿ TO A SQ. LAW	28	1	200 Kc/s	1	10 valves and 20 diodes, many preset pots.
20	TRIANGULAR WAVE	29	4	1 Kc/s	0.1	5 adding amplifiers, 4 diodes.
21	CIRCULAR BEAM TUBE	30	4	100 Kc/s	2	Circular beam deflected on a 4 quadrant target.

FIG. 28.

ANALOGUE MULTIPLIERS.

5.5.3. Triangular Wave Method

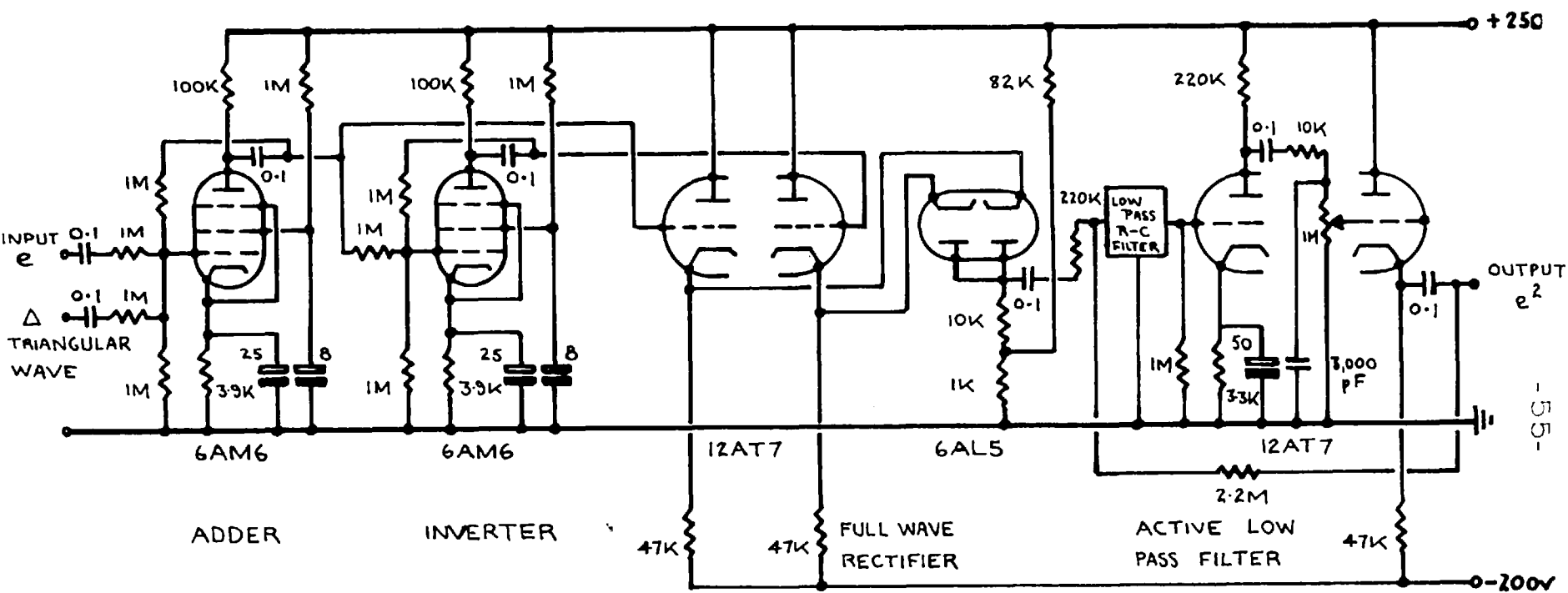
5.5.3.1. Experimental Squarer

The method of multiplication discussed by Meyer and Davies ²⁹ was considered to be the best method, for developing into the D.C. multipliers required in the teaching circuits. This method has the merit that only conventional circuit techniques and components are required, and that good accuracy is possible.

An experimental squarer was constructed to try out the method. The input signal is added to a triangular wave and the resulting waveform is rectified. The height of the triangles produced depends on the input signal and if the areas of the triangles are integrated, in a low-pass filter, an output signal proportional to the square of the input signal results.

The circuit of the squarer is given in fig.29, this was intended to handle A.C. signals only. The triangular wave frequency was 4 Kc/s, frequencies above 300 c/s were removed from the output using a low-pass R-C filter.

The performance of the circuit is shown in fig.30, and this was considered promising. A multiplier can be made by obtaining the sum and difference of the input signals, and by squaring the sum and difference signals by the triangular wave method. The squared signals are subtracted, and the carrier wave frequency components are filtered out. The development of a multiplier using the triangular wave method, and D.C. coupled circuits, was undertaken by Mr.R.Woodcock in August '57.



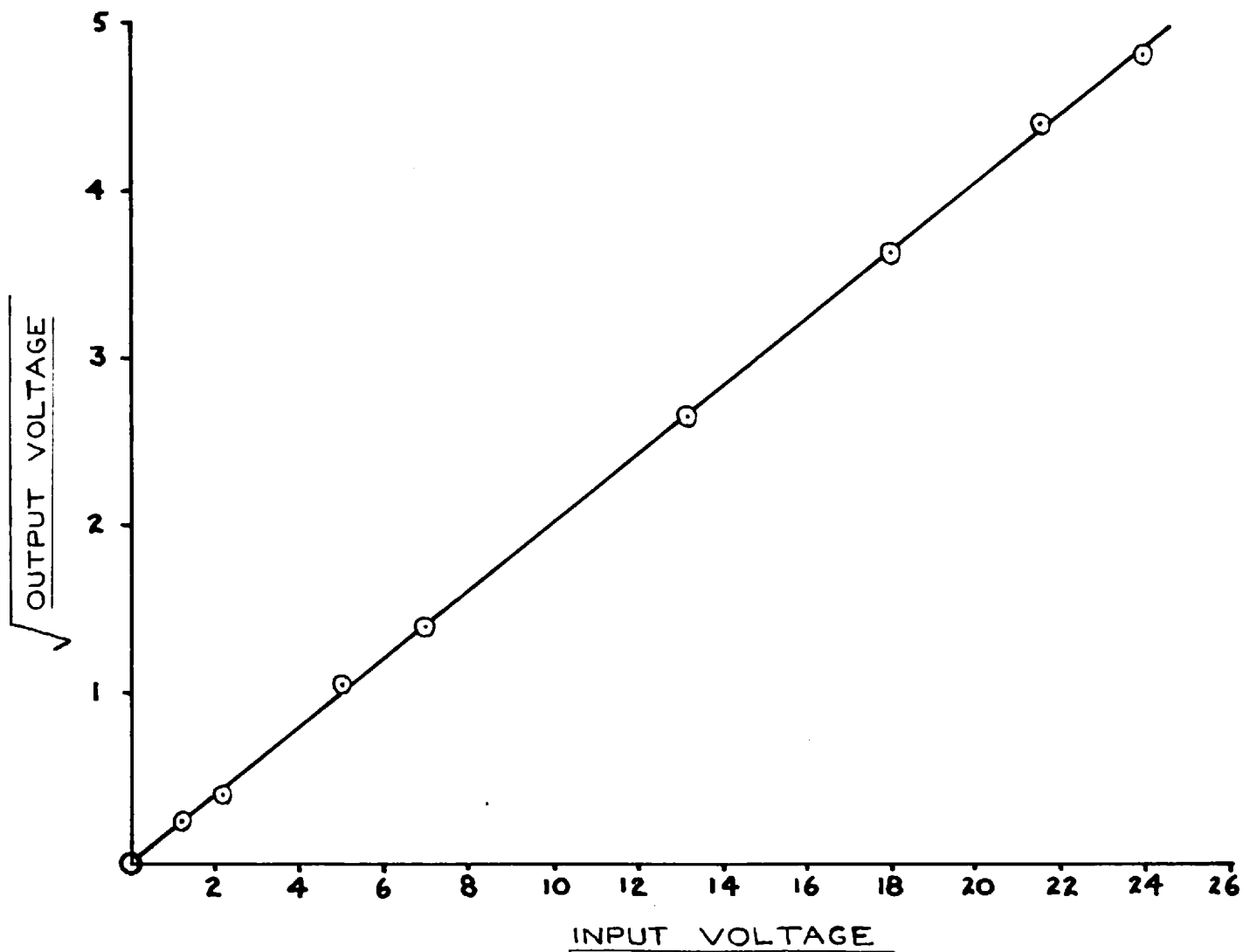


FIG. 30. CHARACTERISTIC OF ANALOGUE SQUARER.

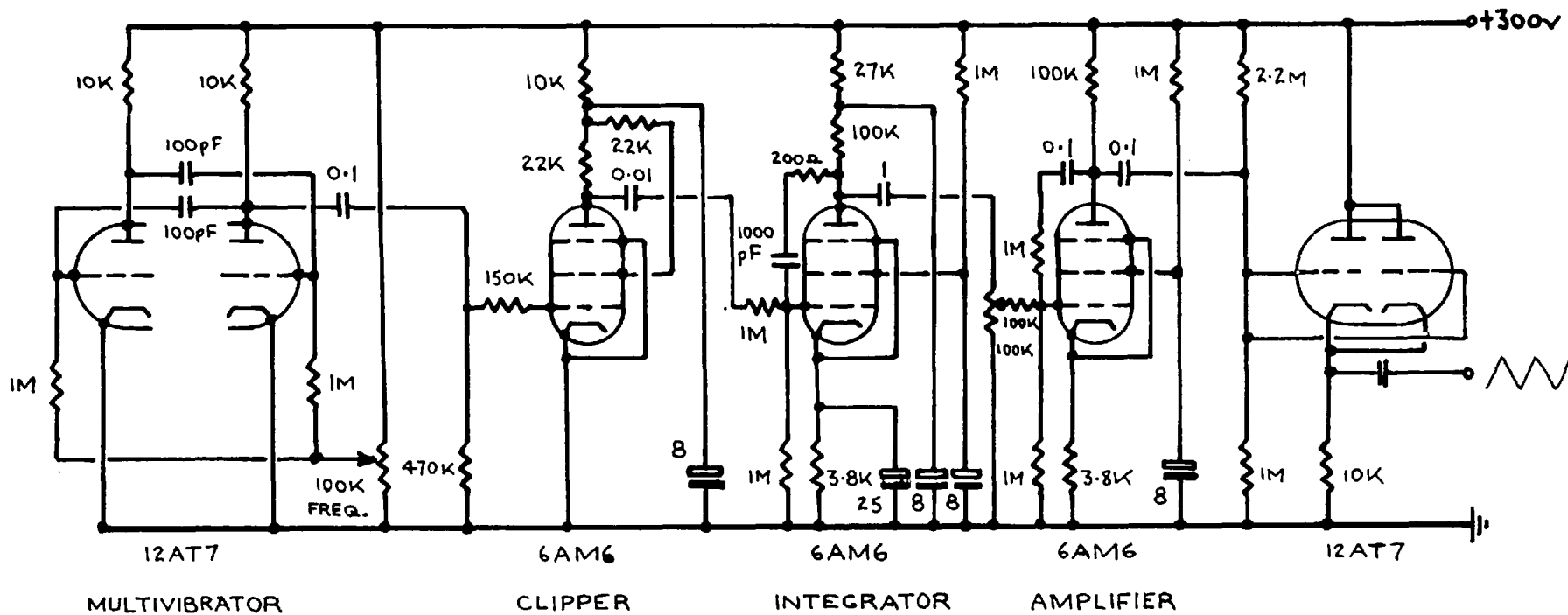


FIG. 31.

TRIANGULAR WAVE GENERATOR .

5.5.3.2. Triangular Wave Generator

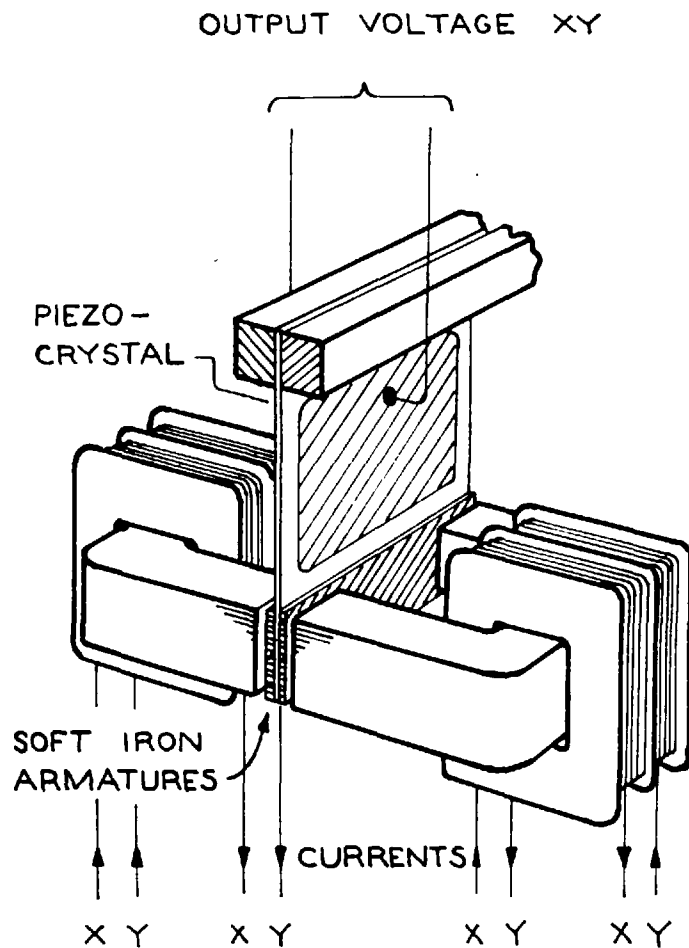
A triangular wave generator is required for use with the triangular wave multipliers. The circuit of the unit constructed is shown in fig.31. It consists of a symmetrical multivibrator followed by a pentode clipper which produces accurate square waves. These are integrated by a Miller integrator to give a triangular wave. The triangular wave is amplified by a pentode amplifier, with negative feedback to stabilize its gain, and the resulting signal is fed to the output through a cathode follower. The performance of this unit has proved to be satisfactory.

5.6. 'Piezo-Magnetic' Analogue Multiplier

5.6.1. Principle of Operation

The basic arrangement of this new type of analogue multiplier (Patent Application 3172/59), developed for use in the Learning Filter, is shown in fig.32. It consists of a bender type piezo-electric crystal, carrying two soft iron armatures at one end, one on each side of the crystal. Opposite each of the armatures, and separated from it by an air gap, is a laminated silicon iron C-core. The two C-cores carry four identical coils, one on each limb.

The input quantities to be multiplied are represented by currents X and Y , and the coils are connected as shown in fig.32. On one C-core, the coils are connected in the same sense, and the magnetic field produced is proportional to $(X + Y)$; the magnetic force exerted



FIELDS	$X + Y$	$X - Y$
FORCES	$(X + Y)^2$	$(X - Y)^2$
FORCE ON CRYSTAL	$(X + Y)^2 - (X - Y)^2 = 4XY$	

FIG. 32. PIEZO-MAGNETIC ANALOGUE MULTIPLIER

on the corresponding armature is proportional to $(X + Y)^2$.

On the other C-core the coils oppose, and the magnetic field is proportional to $(X - Y)$; the magnetic force exerted on the corresponding armature is proportional to $(X - Y)^2$. The two forces act in opposite directions and subtract, giving a net force on the piezo-crystal proportional to $(X+Y)^2 - (X-Y)^2$, which is equal to $4 XY$. The voltage out from the piezo-crystal is proportional to the net force, and is therefore proportional to XY .

To ensure satisfactory operation the piezo-crystal must be sufficiently stiff, so that the movement of the armatures is very small compared with the length of the air gap.

The maximum flux density, and the air gap length, must be chosen so that saturation in the iron part of the magnetic circuit does not spoil the linearity of the device. Both these requirements can be adequately met.

5.6.2. Development of the Prototype

In designing the prototype an attempt was made to obtain a multiplier, suitable for use in the Learning Filter, using only readily available materials and techniques.

The crystal chosen was a square bimorph of Rochelle salt, type 33D41, made in quantity by Cosmocord Ltd. Rochelle salt was used since it is very cheap, and also because it gives a greater voltage output than any other known material, when used as a force-to-voltage transducer. It also enables the transducer to be made in the form of

a "corner loaded square bimorph". The crystal is then fixed at three of its corners, which is a mechanically stable arrangement, and the force is applied to the fourth corner.

Grain orientated silicon iron C-cores were chosen for the magnetic circuits. These are readily available and are supplied with their pole faces ground off flat, which makes it easy to define the air gaps. The armatures used were cut from a silicon iron transformer lamination, 0.014 " thick.

For the prototype, bobbins were fabricated from paxolin sheet 1/32 " thick, and each bobbin was wound with 250 turns of 40 s.w.g. enamelled copper wire.

The prototype, under test, is shown in fig.33. The electromagnets are cast into "Araldite" epoxy resin blocks, and the crystal is



mounted on spacers on one of the blocks, and in the gap between the blocks. The air gaps, between the armatures and the pole faces, measure 0.004 " .

The performance of this first prototype was promising, the zero errors were less than 2 % of the maximum output and the departure from linearity was less than

fig.33. PROTOTYPE 1 % over the working range. The maximum usable output frequency was found to be 1000 c/s. With 45 mA peak input currents, the peak output was 500 mV.

It was estimated that the sensitivity of this prototype could be increased to a suitable value by using thinner wire for the coils, and a greater number of turns. Overload tests indicated that it would be undesirable to decrease the length of the air gaps, to increase the sensitivity, since saturation in the iron would then become more serious. Movement of the armatures was not a limiting factor, as can be judged from the fact that the crystal sensitivity is 5 volts/micron deflection at the tip. Decreasing the stiffness of the crystal would however lower the mechanical resonant frequency of the crystal-armature assembly, and this in turn the upper frequency limit.

It was decided to make the Araldite blocks semi-cylindrical, to save space, and to make the unit suitable for mounting on an octal plug.

5.6.3. Construction

5.6.3.1. Component Parts

The component parts of the "production" multiplier are shown



fig.34. COMPONENT PARTS

in fig.34, they are also shown drawn full size in fig.35. The C-cores are made by Telcon Ltd. They have the same centre size as the standard C-core (HWR 3-4) but the build-up is less, being 1/16" instead of 1/4". The armatures are cut from

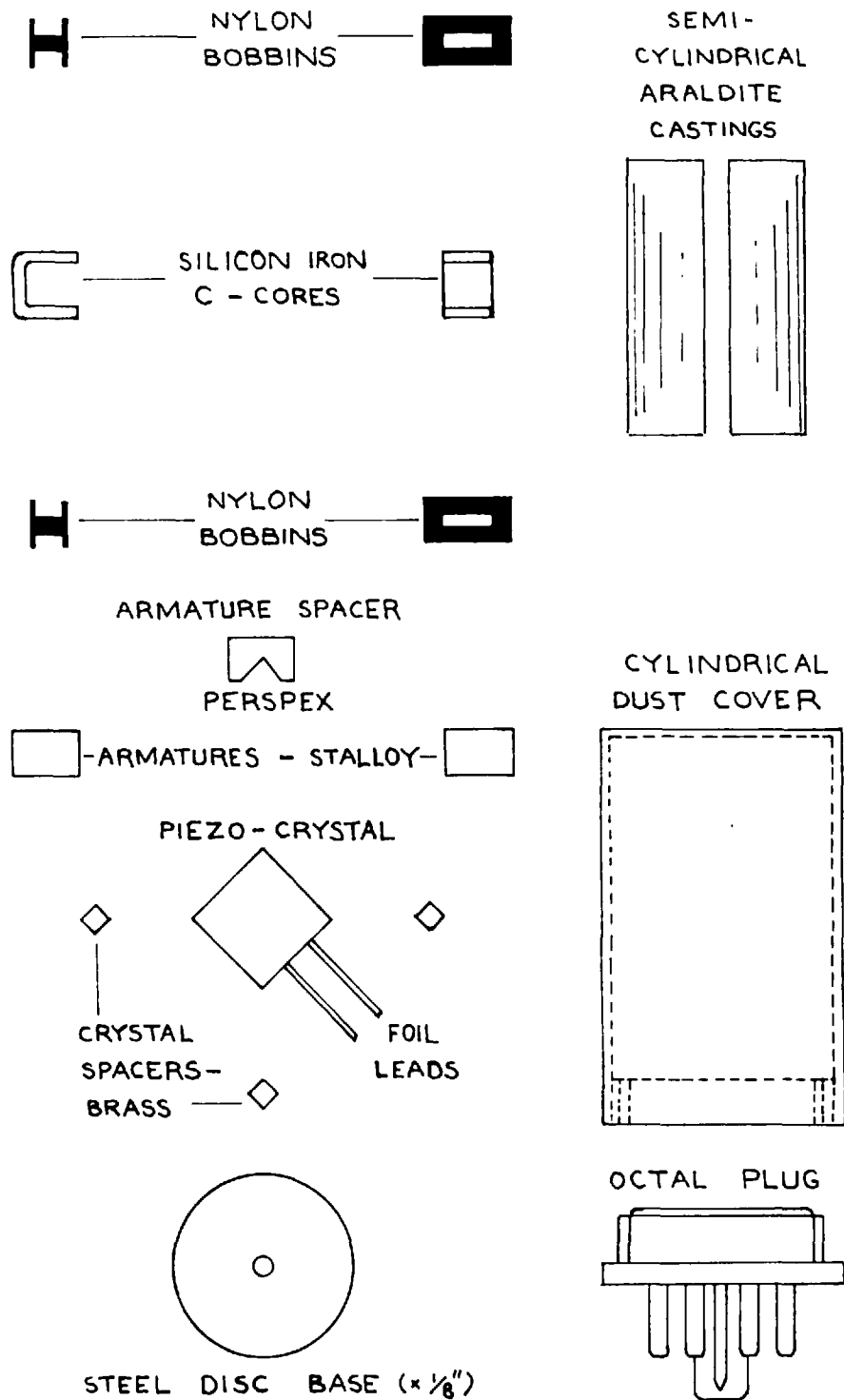


FIG. 35.

PIEZO-MAGNETIC MULTIPLIER - COMPONENT PARTS

SCALE - FULL SIZE

"Stalloy" 0.014" lamination material, they are afterwards pressed flat until 0.014" thick to remove turned up edges and burrs. The crystal spacers are cut from brass sheet 0.018" thick and are also pressed flat after cutting. These spacers define the air gaps on one side of the crystal.

The armature spacer, with its notch to take the corner of the piezo-crystal, is made by cutting a V-notch in a perspex bar, and by cutting slices off the bar to give spacers 0.048" thick.

The bobbins used are so small that it was decided that the only suitable method of manufacture would be to mould them from a plastic material. The bobbins were moulded in nylon, and an injection mould was made specially for this job by Prestwarë.Ltd.

5.6.3.2. Assembly

Each bobbin is machine wound with 250 turns of 46 s.w.g. enamelled copper wire. The winding is terminated at small holes suitably placed in the bobbin, and soldered to stout leads at these points. The soldering is done with a very small soldering iron, using indium, to avoid overheating the bobbin material. This is a suitable method for use with the very fine wire and small bobbins. Great mechanical strength is not necessary since the bobbin is subsequently potted in Araldite. The Araldite casting is made in the semi-cylindrical mould shown in fig.36. A C-core, with a wound and terminated bobbin on each limb, is stuck lightly in its correct position, on the inner surface of the rectangular plate. It is positioned using the template shown in fig.36.

The inner and lower surfaces of the mould are lightly greased with a mould release agent made by mixing vaseline and white spirit. The

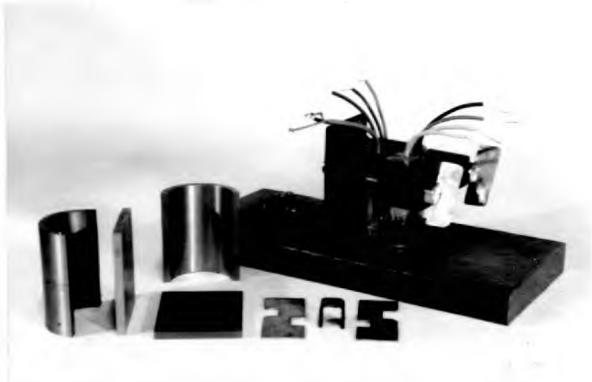


fig.36. MOULDS, TEMPLATE & ASSEMBLY JIG

lapped surfaces of the mould are held together during the casting operation by placing two moulds together, with their flat faces touching, and encircling them with a rubber band.

The moulds stand on a flat plate, and their weight is enough to

ensure a liquid proof contact at the base. The resin used is Araldite Type D, with hardener type 951. The curing time of this combination is 24 hours at 20°C, and 20 minutes at 100°C.

The crystal with its armature assembly and three spacers is now stuck onto one of the pair of Araldite blocks (fig.37.), so that



fig.37. MULTIPLIERS PARTLY ASSEMBLED

the armature is directly over the C-core pole faces. All the assembly operations are done using Araldite which gives a very strong bond, provided the surfaces to be bonded are first de-greased using carbon tetrachloride.

A screened lead is brought out from the crystal through a

hole drilled in the Araldite block. The earthed braid is connected to the metal foil on the two outer faces of the crystal, the centre conductor is connected to the foil which splits the bimorph into two. The block with its crystal is then stuck onto one half of the circular steel base plate.

The other half of the multiplier can now be put into position, and this operation is a critical one so far as the inherent accuracy of the multiplier is concerned. To ensure accurate operation of the multiplier the effective air gaps on the two sides of the crystal must be exactly balanced. This is achieved using a special jig, fig. 36, the electrical balance of the multiplier being observed on an oscilloscope. All the coils are connected in series and an alternating current of 4 mA peak is passed through them. With the coils connected in this way, when the air gaps are equal, the forces on the crystal are equal and opposite, and there is no output from the crystal. With the free half of the multiplier remote from the crystal it has little effect, and the other half acts as a squarer. As the free half of the multiplier is moved inwards the output voltage from the crystal is reduced. The last part of the inward travel is accomplished by forcing the free half against two rubber buffers, by turning the two wing nuts (fig. 36.). This affords a delicate final adjustment of the free half, to the position where there is minimum output. This minimum output is very near to zero (e.g. 0.3 % of full output).

After the free block has been stuck in its final position,

the jig, and the rubber buffers, are removed. This generally allows a very slight movement but the zero error remains less than 2 % of full output. The completed multiplier (fig.38.) measures 1" diameter and $1\frac{1}{2}$ "



long. It is wired to an octal plug and is fitted with a screw-on cylindrical metal dust cover.

The long curing time of the Araldite used for assembly is not a disadvantage since a rate of production of one multiplier per day is satisfactory for our pur-

poses. If a number of multipliers are constructed simultaneously, one operation

fig.38. COMPLETED MULTIPLIER
WITH DUST COVER

can be made on each multiplier each day, and one multiplier is completed each day. Production of the 76 multipliers needed is under way at present, and should be completed by May '60.

5.6.4. Performance

The main features of the multipliers produced are:-

Peak inputs	=	± 5 mA
Peak output	=	± 250 mV
Input resistance	=	200 ohms
Input inductance	=	70 mH
Resonant frequency	=	1.5 Kc/s
Errors	=	less than ± 1 % of full output
Power consumption with maximum sine wave inputs	=	5 mW

It is interesting to compare these figures with those for a

typical Hall Effect multiplier, using indium arsenide:-

Wafer input current	=	200 mA
Wafer input resistance	=	4 ohms
Field input current	=	40 mA
Field input inductance	=	1 H
Peak output	=	150 mV
Maximum frequency	=	250 c/s
Errors (after compensating field pick-up and p.d. between Hall contacts)	=	less than $\pm 1\%$ of full output
Power consumption, including field series resistor	=	2,600 mW

The multiplier is affected by strong magnetic fields unless magnetically screened, and is somewhat susceptible to vibration, but neither of these features limits its application to the Learning Filter. The Rochelle Salt crystal is sensitive to temperature changes, the output falls rapidly when temperature exceeds 100°F , and the crystal is ruined if the temperature exceeds 130°F . At normal room temperatures the output voltage does not change greatly with temperature, provided the crystal feeds a high resistance load. The change is approximately 2.5% between 70°F and 80°F and this is tolerable in the Learning Filter. The crystal capacitance changes markedly with temperature and this modifies the low frequency response somewhat, but the change is not important.

If a final adjustment to minimum error, after assembly, is considered desirable the two halves of the multiplier may be joined by a bolt, which enters a tapped hole in the fixed half, at the lead end.

This provides a fine screw adjustment of the air gaps.

5.6.5. Drive and Output Circuits

5.6.5.1. Valve Circuit

The drive and output circuits designed for use with the piezo-magnetic multiplier are shown in fig.39. The current to the coils is provided by a pair of cathode followers. Each can provide a peak current of ± 4 mA, in the 12 K resistor in series with the coils. The output from the crystal is amplified by the 12AX7 feedback amplifier. The negative feedback from the anode of V_{2b} to the cathode of V_{2a} , stabilizes the gain and enables the unit to feed up to four other multipliers. The anode load of V_{2b} is split giving two outputs, one is ± 50 volts peak and is used to feed other multipliers, the other is ± 5 volts peak and is used to feed a potentiometer coupling unit. The parallel T filter between the gain stages is put in to compensate for the peak in the output frequency response, due to crystal resonance. Three 750 pF trimming capacitors are provided for adjusting the stop frequency of the parallel T circuit. In actual fact the resonant frequency of different multipliers has turned out to be remarkably constant ($\pm 3\%$), and fixed capacitors could probably be used.

The overall response of the complete multiplier unit is given in fig. 40, and it will be seen that there is good compensation. Some compensation against the peak due to resonance could probably be built-in to the plug-in multiplier, (c.f. gramophone pick-ups, some crystal microphones) but this would have required a much longer

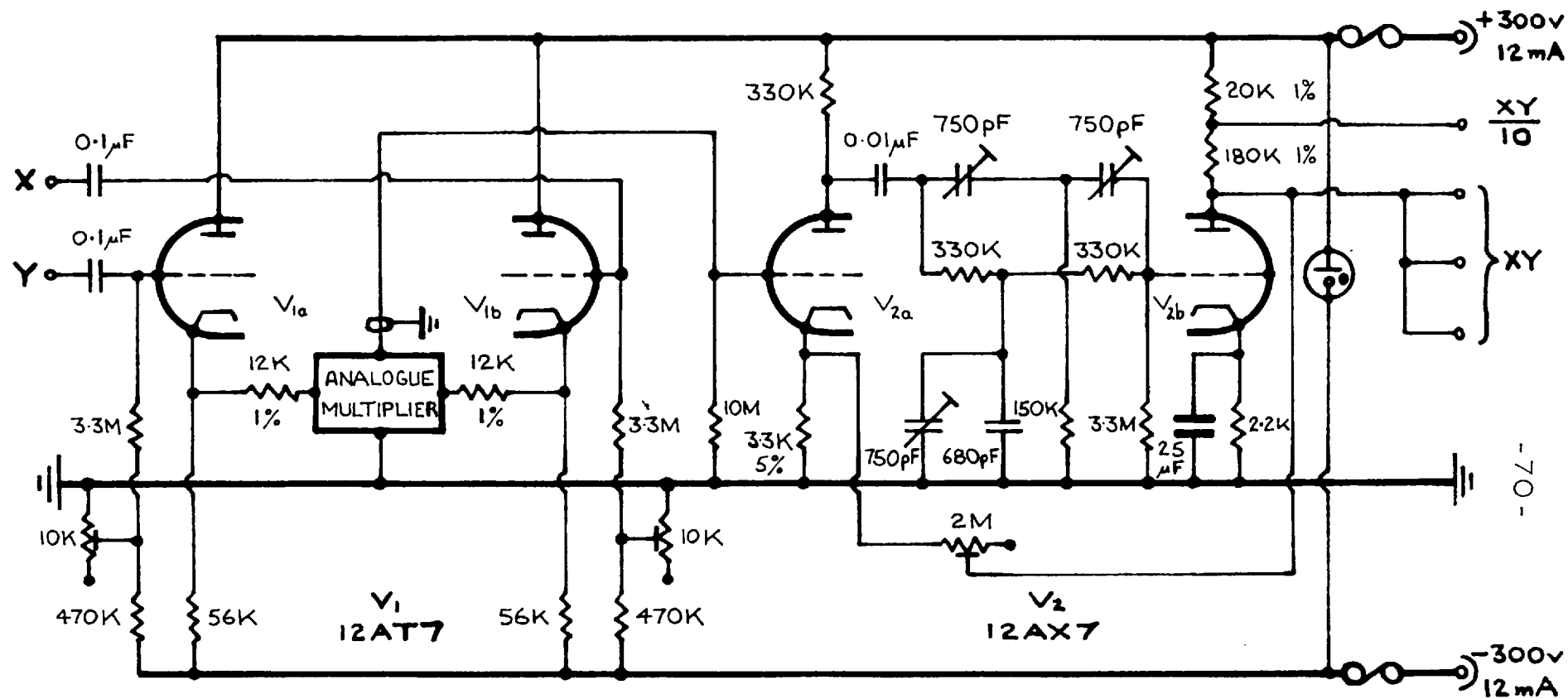
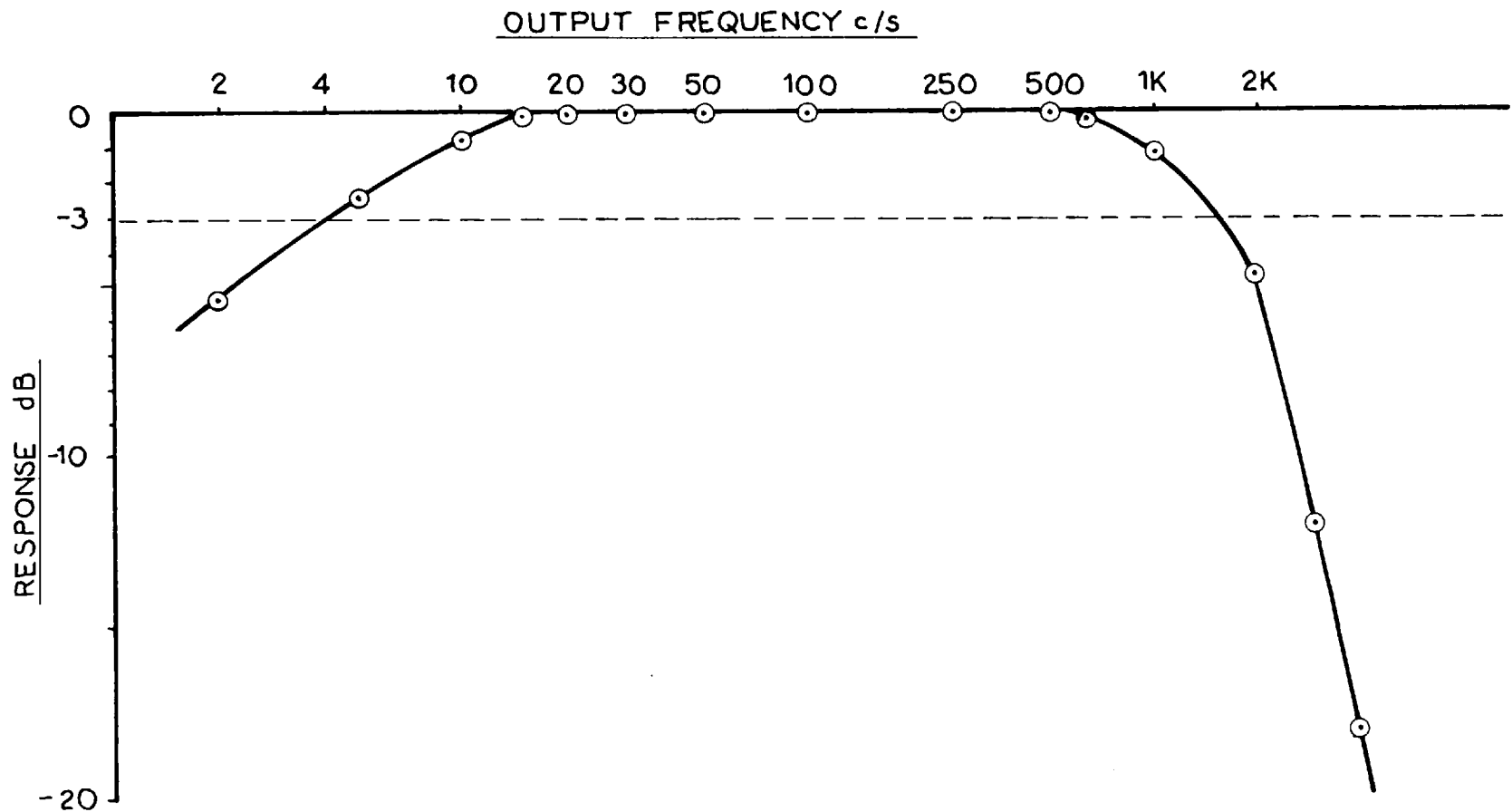


FIG. 40.

PIEZO - MAGNETIC MULTIPLIER

OVERALL FREQUENCY RESPONSE, WITH R-C FILTER



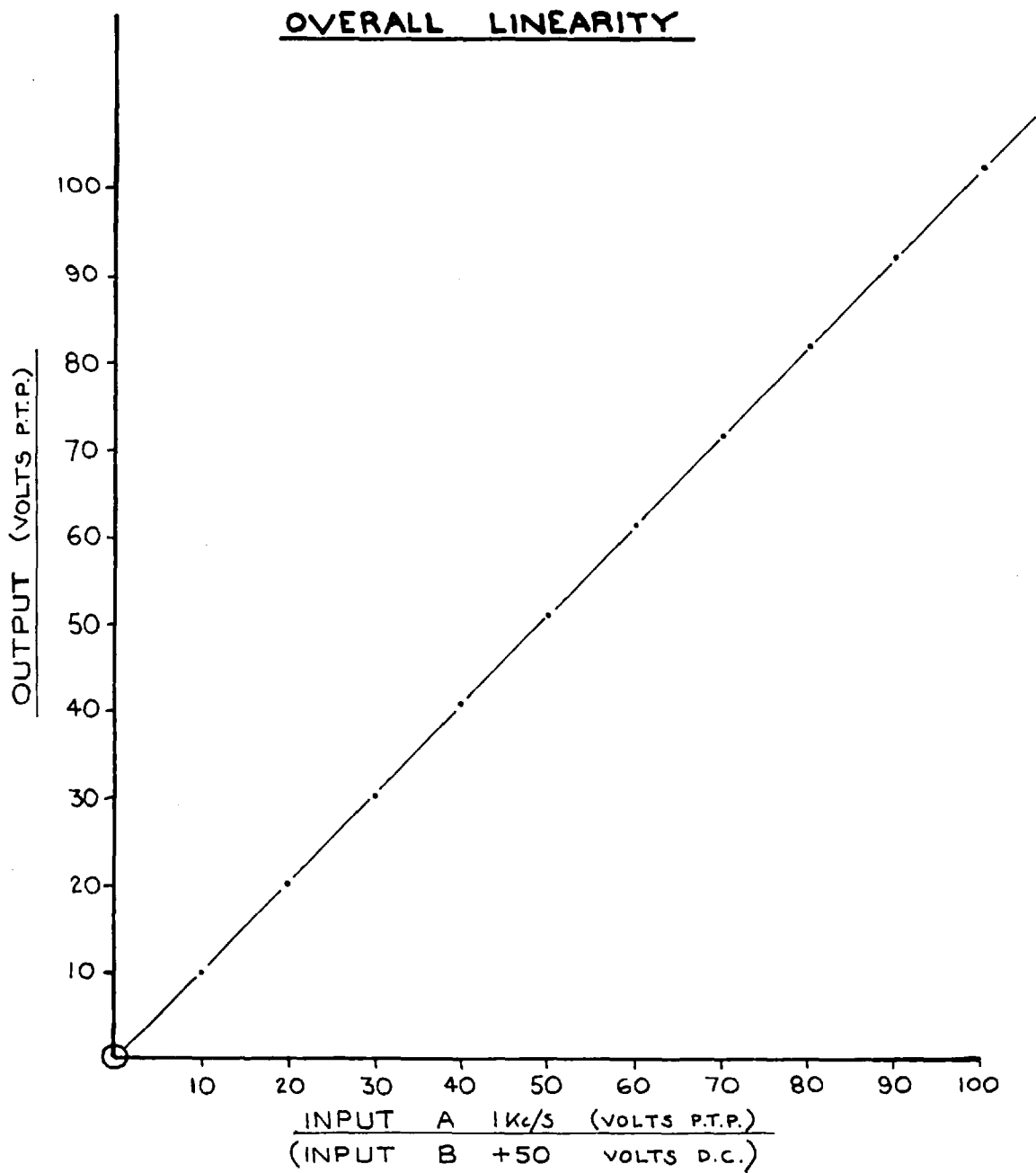


FIG. 41.

PIEZO - MAGNETIC MULTIPLIER

development programme. The overall linearity of the multiplier unit is shown in fig.41.

The drive and output circuits are contained in a chassis measuring $17'' \times 1\frac{3}{4}'' \times 3''$ (fig.42). This chassis contains two complete multiplier units. In the photograph one unit is shown with its valves



and plug-in multiplier in position.

A working prototype of the chassis, and full drawings were supplied to Gresham Developments Ltd., who undertook the job of producing 40 of the twin chassis for use in the Learning Filter.

fig.42. TWIN CHASSIS

The performance of a complete multiplier unit is summarized as follows:-

Peak inputs	= ± 50 volts
Peak output	= ± 50 volts or ± 5 volts
Errors	= less than $\pm 1\%$ of full output
Maximum output frequency	= 1 Kc/s
Long term amplitude stability (over 1 month under normal laboratory conditions)	= less than 2% change
Valves	= 1 12AX7, 1 12AT7
Power supplies	= ± 300 volts 12 mA

5.6.5.2 Transistor Circuit

An experimental transistor circuit was also made using six OC71 transistors. Input and output voltages were 1 volt peak and the supply voltages ± 3 volts. The total power consumption was only 0.1 watt. The currents for the multiplier inputs were obtained directly from two emitter followers. An emitter follower had to be used as a high input impedance stage after the crystal, and this was followed by two gain stages and an output stage. (fig.43)

This circuit was found to be less stable, and more expensive than the valve circuit, though otherwise satisfactory. The stability can of course be improved by more refined circuit design.

It was concluded that though transistors are very well suited to the piezo-magnetic multiplier in terms of impedance level, power consumption, and physical size, these features offered no particular advantage in the Learning Filter, and the development of the transistorized unit was discontinued.

5.7. Teaching Apparatus

5.7.1. Original Scheme

The teaching apparatus is that part of the Learning Filter which computes the influence curve of a given coefficient, and the value of the coefficient required to give minimum mean square error. The original plan for doing this is shown in fig.44. The contributions from all the potentiometers except the one whose influence curve is

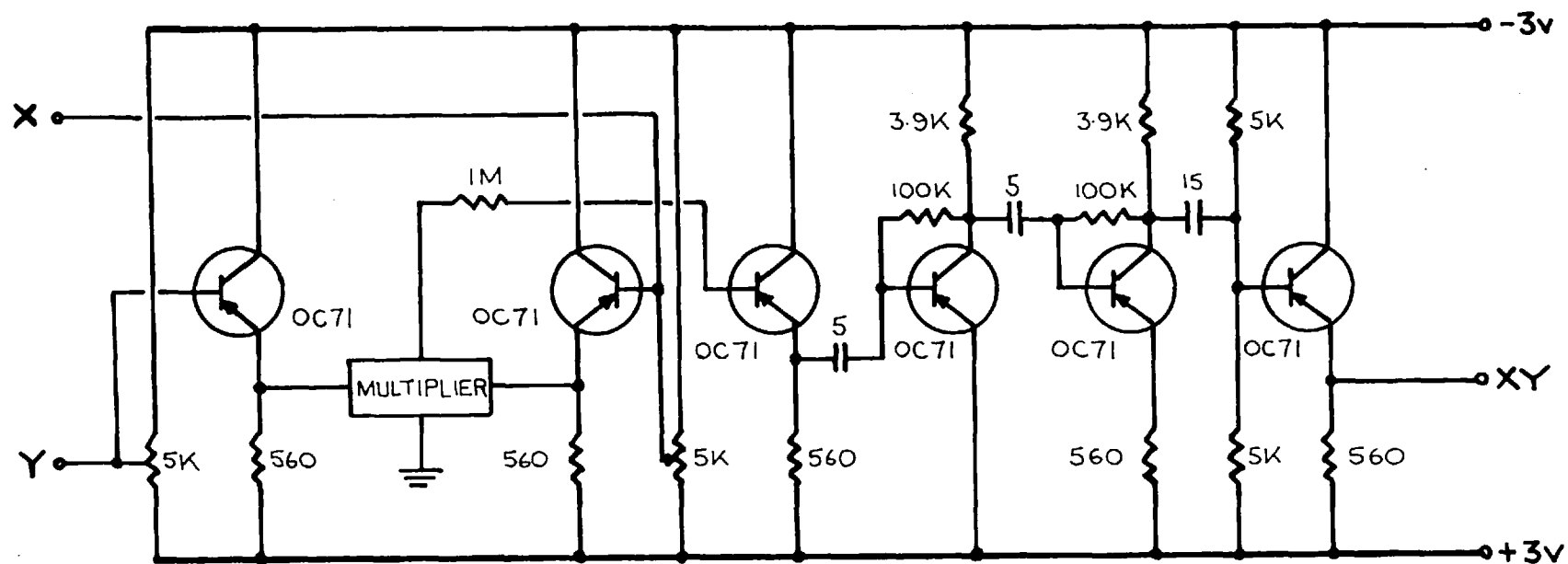


FIG. 43.

PIEZO-MAGNETIC MULTIPLIER - TRANSISTOR CIRCUIT.

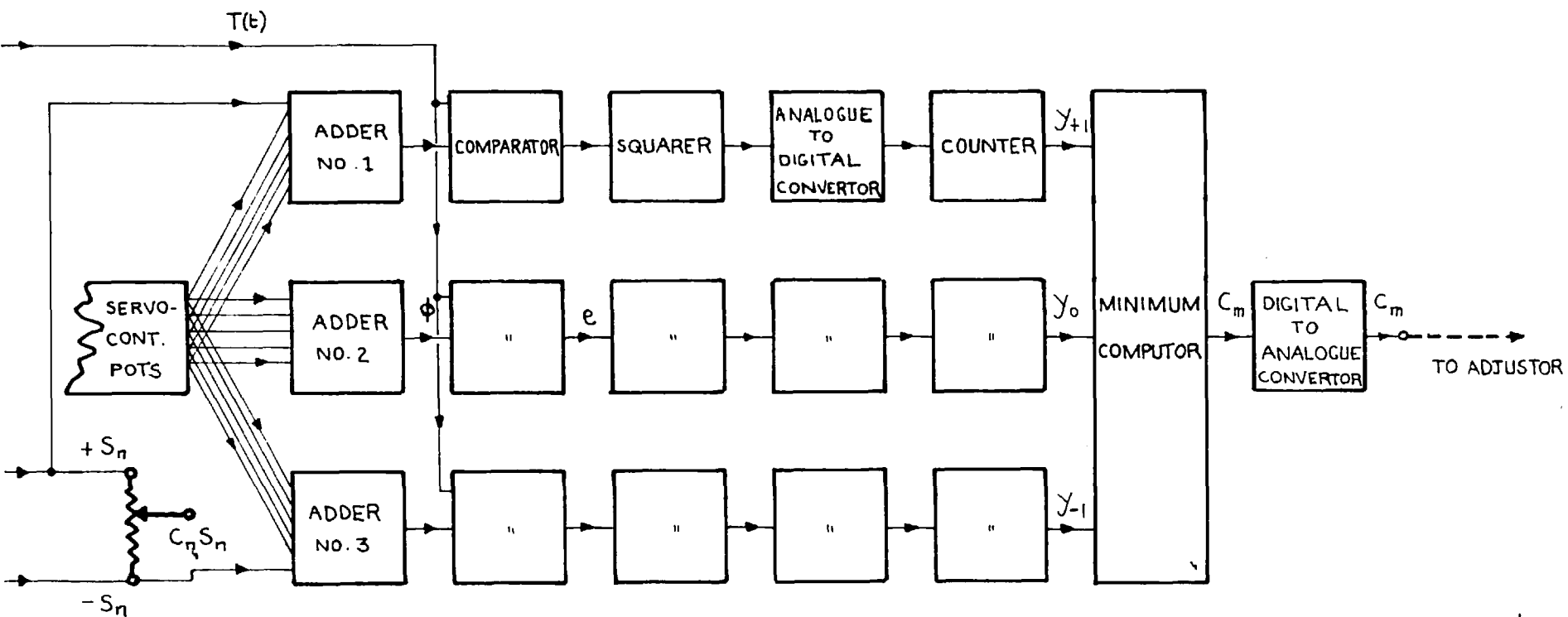


FIG. 44. TEACHING APPARATUS , FIRST BLOCK DIAGRAM

required, are added up in the three adders. The signal at the positive end of the selected potentiometer is added into adder no.1 and the signal at the negative end of the potentiometer is added into adder no.3. If the signals from the adders are subtracted from the teaching signal, and if the result is squared and integrated over the length of a run, this gives three values for the mean square error y_{+1} , y_0 , y_{-1} corresponding to the coefficient values +1, 0, -1. These points lie on a parabola and it is easily shown that the coefficient for a minimum value of the mean square error is given by:-

$$C_m = \frac{1}{2} \cdot \frac{y_{-1} - y_{+1}}{y_{-1} - 2y_0 + y_{+1}} \quad \text{.....Appendix 1}$$

and the value of the minimum by:-

$$Y_m = y_0 - \frac{1}{2} (y_{-1} - y_{+1}) \cdot C_m$$

Analogue to digital converters are shown in the diagram, it was intended that the integration and the computation of the minimum should be done digitally to reduce the magnitude of errors. The possibility was noted of using reversible counters as integrators, forming directly the quantities $(y_{-1} - y_{+1})$, y_0 , and $(y_{-1} - 2y_0 + y_{+1})$ needed for computation of C_m and Y_m .

Though sound enough algebraically the above scheme has several disadvantages when applied in practice. The scheme is relatively complicated, and to design and construct the necessary equipment would be very time consuming. With a large number of coefficients the

parabola is likely to be a very shallow one. This causes difficulties since we are then computing by subtracting large numbers which are nearly equal. The errors in the three channels shown in fig.44 may well be of the order of 1%, unless extreme pains are taken over the design of the adders, comparators and analogue-to-digital converters. It is therefore pointless to attempt to calculate C_m by the method of fig.44.

Errors occurring prior to the three channels shown do not have a serious effect on the calculation of C_m . Thus errors in the demodulators, multipliers, and potentiometer coupling units, add almost equally to all three ordinates. This shifts the whole parabola up or down in the Y direction but makes little difference to the calculated C_m . It is no use doing three consecutive runs, using one channel to determine y_{-1} , y_0 , and y_{+1} as the system error might change during the time of three runs.

The distinction drawn between errors affecting all ordinates equally, and uncorrelated errors in the three channels of fig.44, which affect the three ordinates separately, gives the clue to a way out of the difficulty. This is to compute a large common part of the three ordinates in one channel, for convenience this can be y_0 ; and to calculate the amounts by which y_{-1} and y_{+1} differ from y_0 , D_{-1} and D_{+1} , in two further channels.

The differences D_{-1} and D_{+1} should then be obtained with less than 1% error.

5.7.2. Modification of the teaching apparatus to work with differences.

The optimum can be found directly from the differences using the expressions:-

$$C_m = \frac{1}{2} \frac{D_{-1} - D_{+1}}{D_{-1} + D_{+1}} \quad \dots \text{Appendix 1}$$

and
$$Y_m = y_o - \frac{1}{4} C_m (D_{-1} - D_{+1})$$

The computed value of Y_m is still subject to the main y_o channel error and this is unavoidable. The important quantity is C_m and provided we can compute this accurately we can be confident that the Learning Filter will optimize itself. We only require Y_m as a check on the progress of the optimization and in this case a small error in Y_m does not matter, provided we can keep the error sufficiently constant from one run to the next. We have attempted to do this by the use throughout the machine, wherever necessary, of stabilized H.T. supplies, constant voltage heater supplies, high stability components, and by the liberal use of negative feedback.

We do however require at the n^{th} run $(Y_{m(n-1)} - Y_{m_n})$ if we intend to automatically miss out coefficients which do not significantly reduce Y_m . R. Woodcock has pointed out that although it is impossible to calculate Y_m from the differences alone, it is possible to calculate $(Y_{m(n-1)} - Y_{m_n})$ from the differences and from the coefficient C_m and its previous value C_m' , both of which are available at the end of the n^{th} run, before adjustment of the coefficient.

This method avoids taking the differences between almost equal quantities and makes it unnecessary to carry forward Y_m from one run to the next. Instead C'_m is used since this has to be stored anyway as the setting of a potentiometer.

It can be shown that:-

$$(Y_{m(n-1)} - Y_{m_n}) = \frac{1}{2} (D_{-1} + D_{+1}) (C'_m - C_m)^2 \dots \text{Appendix 1}$$

The revised block diagram for computation using D_{-1} and D_{+1} is as shown in fig.45. It is based on the fact that if the signal in the y_0 channel after the comparator is e then the signal in the Y_{+1} channel should be $(e+S)$, where S is the contribution from the positive end of the selected potentiometer. After squaring, the y_{+1} channel should contain $(e^2 + 2eS + S^2)$. The difference between y_{+1} and y_0 , D_{+1} , is therefore equal to $(2eS + S^2)$. Similarly for y_{-1} and D_{-1} . It will be seen that the analogue-to-digital conversion has been discarded, the integrations over the length of the run can now be done sufficiently accurately using analogue integrators. The modification to work with differences has therefore greatly simplified the design and construction of the teaching apparatus; the minimum computer has also been changed from a digital to an analogue device.

5.7.3. Simplification of the analogue teaching apparatus.

A further refinement, suggested by R.Woodcock, has been embodied in the design of the teaching apparatus. In this method instead of forming D_{-1} , y_0 , and D_{+1} on the three integrators we form

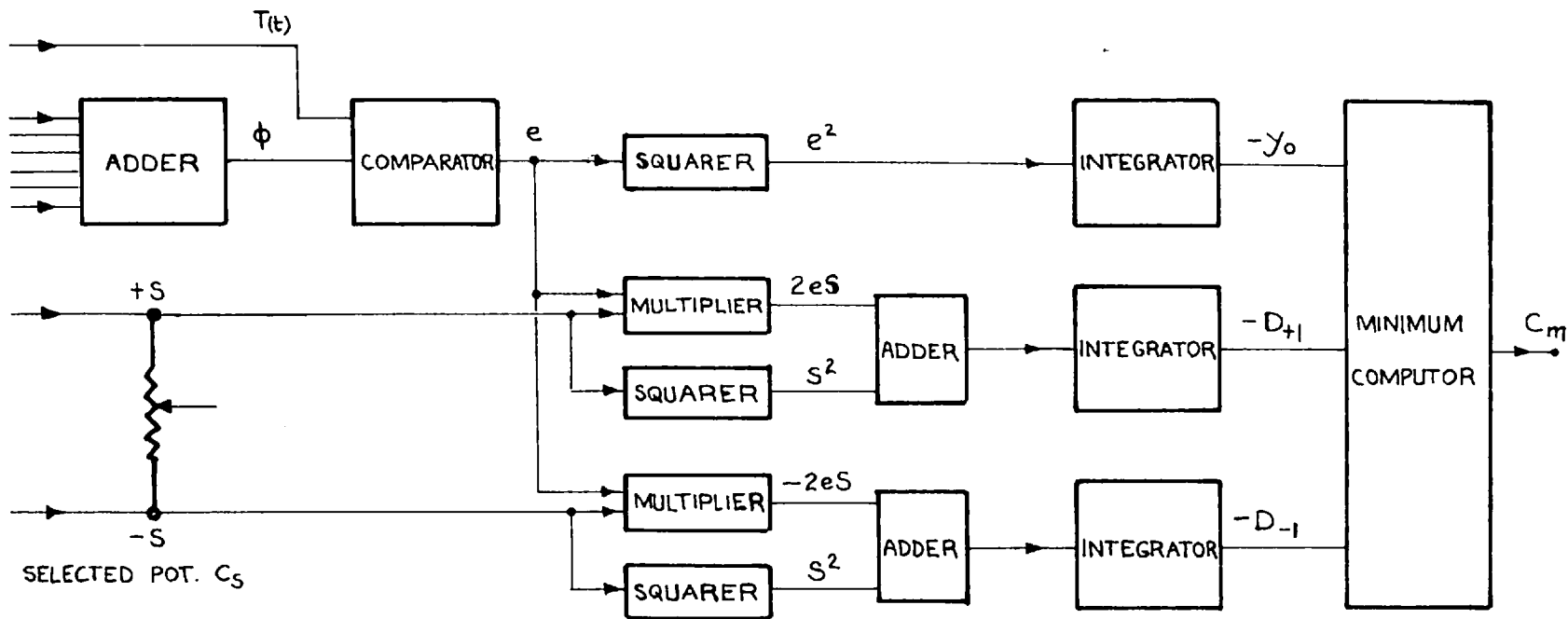


FIG. 45. TEACHING APPARATUS , MODIFICATION TO WORK WITH DIFFERENCES

$\frac{1}{2}(D_{-1} - D_{+1})$, y_0 , and $(D_{-1} + D_{+1})$, these are the quantities needed for computation of C_m and Y_m . From above:-

$$D_{+1} = \frac{(2eS + S^2)}{S^2}$$

$$\text{and } D_{-1} = \frac{(-2eS + S^2)}{S^2}$$

$$\text{therefore } \frac{1}{2}(D_{-1} - D_{+1}) = \frac{(-2eS)}{S^2}$$

$$\text{and } (D_{-1} + D_{+1}) = \frac{(2S^2)}{S^2}$$

These expressions lead to the final form of the teaching apparatus shown in fig.46. It will be seen that the complete teaching apparatus has now evolved into a fairly simple analogue computing network.

5.7.4. Realization of the teaching apparatus.

Five high gain analogue computing amplifiers are needed in the teaching apparatus. One is used as the 94-input adder which sums all the terms in equation 2; three as integrators forming y_0 , $\frac{1}{2}(D_{-1} - D_{+1})$ and $(D_{-1} + D_{+1})$, and one is used with a multiplier to perform the division yielding C_m . These five amplifiers are Solartron D.C. amplifiers, type AA 621, with the performance indicated below:-

Gain at D.C.	= approximately 4.5×10^6
Gain in D.C. channel	= not less than 10^4 at 1 Kc/s
Long term drift referred to input	= less than 100 mV
Effective grid current	= 1×10^{-10} amps
Noise at output	= less than 1 mV R.M.S.
Output voltage	= ± 100 volts
Output current	= ± 10 mA

A detailed analysis using these figures indicates that errors introduced by the amplifiers due to finite gain, drift and grid current, should be considerably less than 0.1 %. High stability, high quality, resistors and condensers have been used with the amplifiers. The multipliers shown are 'triangular wave' multipliers (section 5.5.) specially developed for these circuits by Mr. R. Woodcock. Each multiplier has a response within 1 % from D.C. to 600 c/s, the inputs may vary between ± 5 v. and the maximum output is ± 5 v. The long term errors are less than ± 0.5 % and the attainable error if the multiplier is adjusted after it has reached thermal equilibrium is less than ± 0.1 % of full output. The triangular wave generator used to drive the multipliers is as described in section 5.5.

The divider is a high gain amplifier with a multiplier in its feedback loop (fig. 46). From fig. 46. it will be seen that:-

$$V_o Y + X = V_o / G$$

Therefore if G, the gain of the D.C. amplifier, is very large:-

$$V_o = -X / Y$$

With this type of divider if $(D_{-1} + D_{+1})$ were to become negative the feedback around the high gain amplifier would be positive instead of negative and the circuit would be unstable. However since $(D_{-1} + D_{+1}) = \overline{2S^2}$ it must always be positive and no trouble occurs. If $(D_{-1} + D_{+1})$ becomes small trouble may occur, since, if the other input is finite the divider output will tend to infinity as $(D_{-1} + D_{+1})$

tends to zero. Again in practice there is no difficulty since $\overline{2S^2}$ does not reduce as the optimization proceeds. The divider will not saturate unless C_m lies outside the range -1 to $+1$. If it does saturate a warning light indicates that the teaching track may need to be rescaled.

The outputs from the teaching apparatus, y_o , $\frac{1}{2} (D_{-1} - D_{+1})$, $(D_{-1} + D_{+1})$, Y_m , and C_m , are connected to the control console, and their values may be read on the reference meter. The adjustor is supplied with C_m for setting the selected coefficient C_s equal to C_m before the next run is commenced.

At the time of writing the circuit for computing $(Y_{m_{n-1}} - Y_{m_n})$, fig.46, has not been completed though most of the parts required are ready. Full provision for automatically missing out coefficients has been made in the design of the control circuits (section 5.8.).

5.8. Sequence Control and Switching Circuits

5.8.1. Requirements

For long teaching runs lasting many hours it is very desirable that the operation of the Learning Filter should be automatic. The basic arrangement for fulfilling this requirement is shown in fig.47. The cycle is started by a push button on the control console, and this operates relays in the sequence control unit which control the 20 track tape deck. The tape starts to play at $7\frac{1}{2}$ "/sec. and computation of C_m and Y_m for a particular coefficient commences. When the end of the tape record is reached the signal from the control track passes

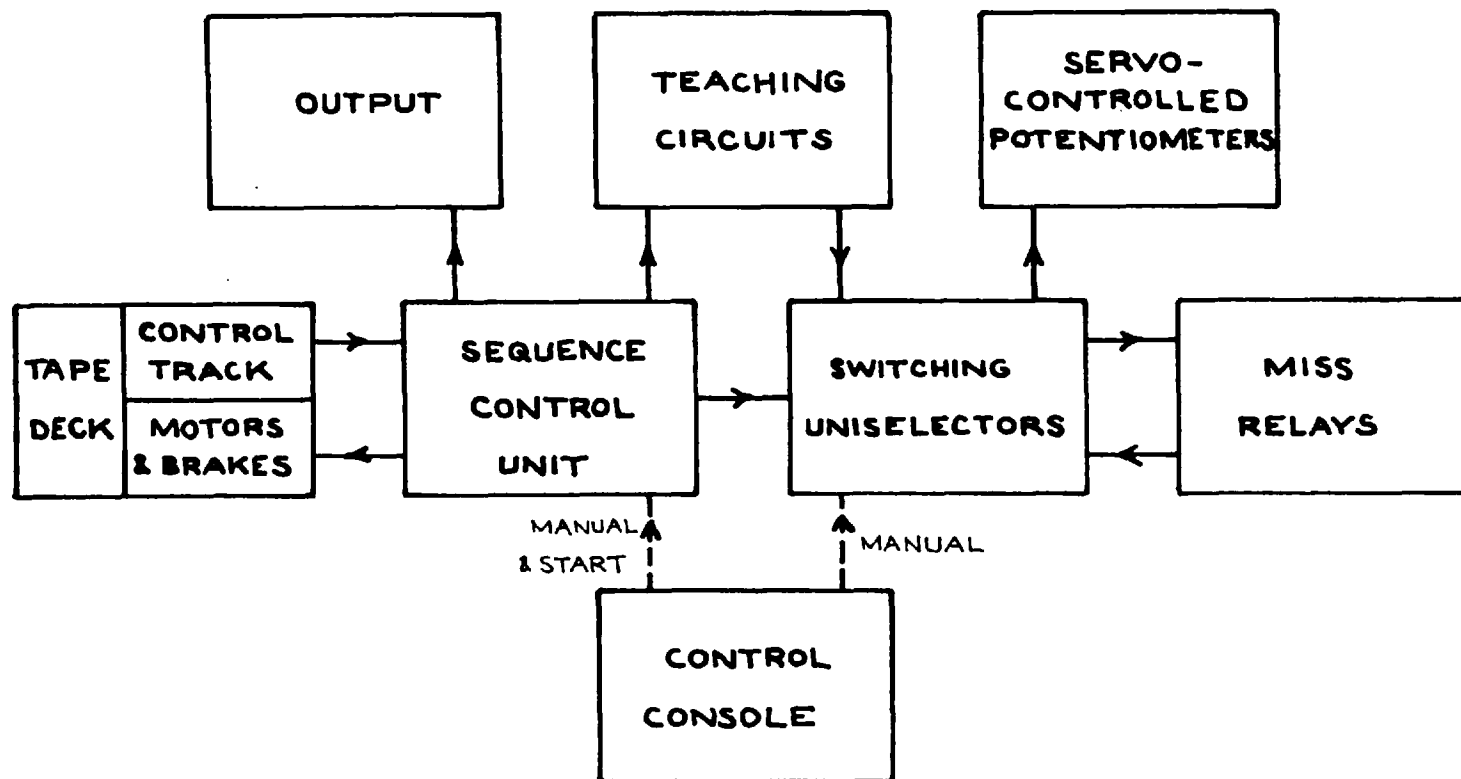


FIG. 47.

CONTROL BLOCK DIAGRAM.

to the sequence control unit which then stops the play of the tape and initiates the rewind, and the adjustment of the coefficient to C_m . The inputs to the integrators in the teaching circuits are disconnected and the integrators hold their charge. At the end of the rewind a further signal from the control track stops the tape, ready for another run of the record. When both rewind and adjustment are completed the uniselectors switch to the next significant coefficient and the integrators in the teaching circuits are discharged. After a short delay another run of the tape commences. There is one 'miss' relay associated with each coefficient and if this relay is energised the corresponding coefficient is omitted from the learning process.

In addition to the automatic procedure outlined briefly above, it is essential that manual control of the process should also be possible.

5.3.2. Sequence Control Unit

The operation of the sequence control unit may be understood with reference to the simplified circuit diagram fig.48. and the sequence diagram fig.49. There are four possible modes of operation, as detailed below.

1 Continuous - Automatic

In this mode all the switches and contacts are initially in the positions shown in fig.48, and no relays are energised. Nothing happens until the start button on the control console is pressed whereupon st2 closes and current flows through the 'play master relay' PM.

WHEN THE MACHINE IS SWITCHED TO CONTINUOUS-AUTOMATIC
THE SEQUENCE SHOWN BELOW OCCURS, IF THE
ADJUSTMENT TAKES LONGER THAN THE REWIND.

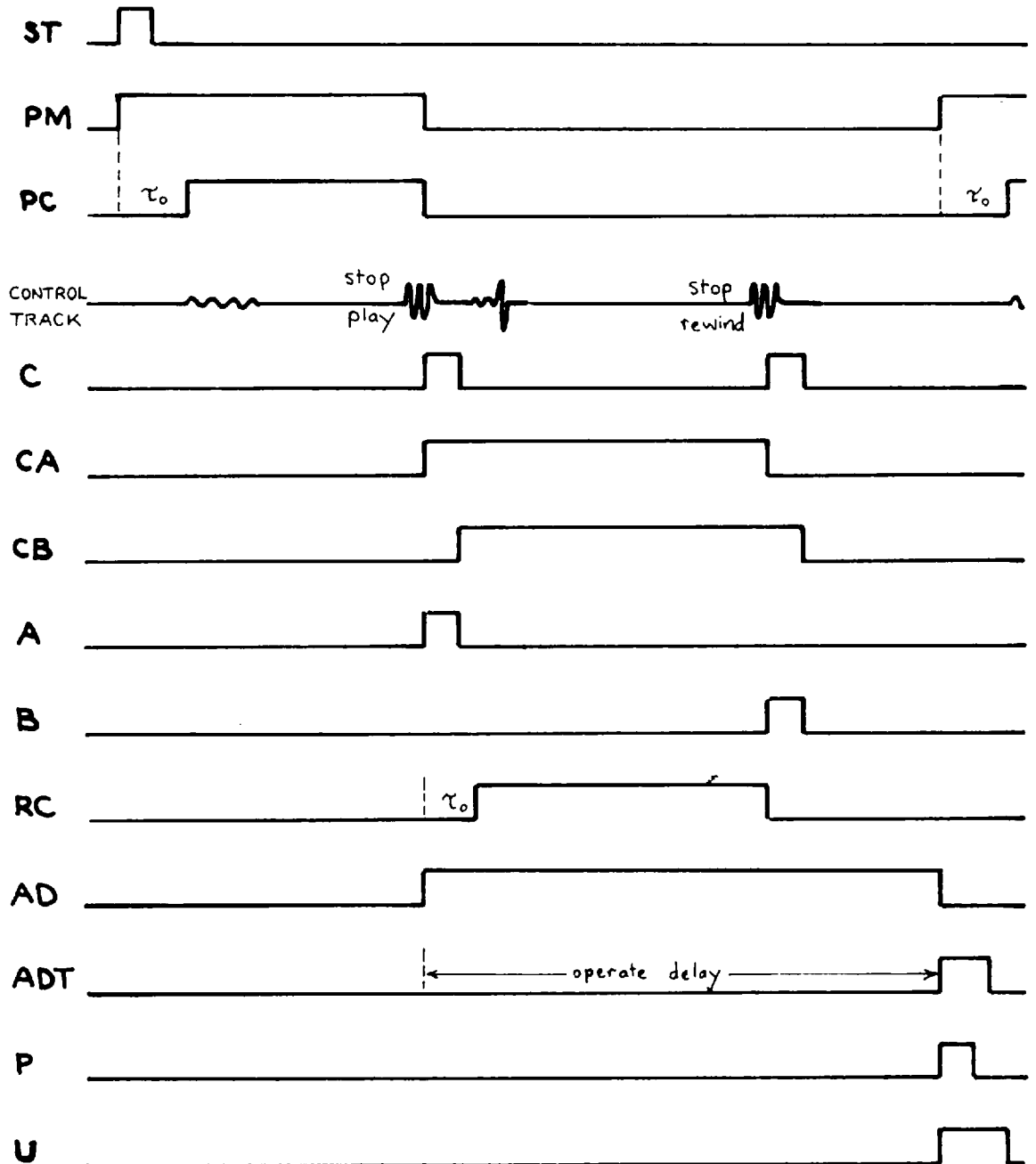


FIG. 49.

SEQUENCE CONTROL UNIT, SEQUENCE DIAGRAM

Contact pm1 closes and PM holds on. Contact pm2 changes over and current flows into the $32 \mu\text{F}$ condenser through the 47 K. resistor. No current flows through the 'play control relay' PC until the voltage at X reaches earth potential, because of the diode in series with PC. Once the voltage rises above earth potential the diode is biased in the forward direction and PC operates. The operate time delay for PC is about 2 seconds.

When PC operates the tape recorder motors are switched on and the tape starts to move at $7\frac{1}{2}$ "/sec. Nothing further happens until the stop signal on the control track is encountered. The stop signal causes relay C in the tape control unit to operate and contact cl closes for a short time. When cl closes, relay CA is energised via cb2, cl, and cbl. Contact cal closes but no current flows in relay CB since it is shorted by the path through cal, cb2, cl, and cbl. When cl opens again a short time later this short circuit is broken and CB is energised in series with CA. During the short time that CA is on and CB off relay A is energised via ca2 and cb5 and comes on in sympathy with C. Relay A therefore marks the end of the 'play'. Contact al opens, breaking the supply to PM and PC which fall out, and the magnetic tape comes to a halt as the motors are switched off and the brakes automatically applied. The PC operate delay timing condenser is discharged by pm2. Contact ca3 has changed over and the 'rewind control relay' RC is energised after a time delay of about 1 sec. The circuit is similar to that already described for PC. When RC operates, the rewind

motor is switched on, and the tape starts to rewind at high speed. While relay A comes on, a2 closes and the 'adjust relay' AD is energised holding on via ad3. Relay AD operates the appropriate coupling relay K_n (section 5.4.) which connects the reference supply of ± 100 v. across the potentiometer corresponding to the coefficient being dealt with, and switches on the motors driving the magnetic clutches. Relay AD connects the adjustor to the wiper of the potentiometer and to C_m , available as a voltage at the output of the minimum computer (section 5.7.). Relay AD remains on until the 'adjustment timer' operates, opening contact adt which breaks the supply to AD. Contact ad4 changes back and the $16 \mu F$ condenser discharges through relay P which comes on for a short time. Relay P therefore marks the end of the adjustment.

Rewind of the tape continues until the other stop signal is encountered on the control track, and this causes cl to close again for a short time. Prior to the rewind stop signal both CA and CB are energised. When cl closes a short circuit is provided around CA via cal, cbl, cl, and cb2, and CA falls out, opening cal. Contact ca3 changes back and RC falls out stopping the rewind and applying the brakes on the tape deck. Relay CB remains energised through cb2, cl, and cbl. When cl opens a short time later this path is broken and CB also falls out. During the short time that CA is off and CB on, relay B is energised through ca2 and cb6 and comes on in sympathy with C. Relay B therefore marks the end of the rewind.

The operation of relay U, which marks the end of one cycle

in the learning process, depends on which lasts longest, rewind or adjustment. If adjustment lasts the longest the sequence is as shown in the sequence diagram fig.49. Rewind has finished so cb7 is in the position shown in fig.48, and when pl closes at the end of the adjustment U is energised. Contact u1 changes over and U remains energised for a short time after pl opens because of the discharge current flowing through U from the $32\mu\text{F}$ condenser. This condenser ensures that there are no difficulties due to pulse overlapping, if rewind and adjustment finish at almost the same instant. If the adjustment lasts longest then b2 cannot operate U when the rewind ends, since ad7 is open.

If rewind takes longer than adjustment then ad7 is closed when b2 closes, and U is energised when the rewind ends. In this case pl cannot operate U at the end of the adjustment since cb7 is open. When U operates, the uniselector switching circuit jumps on one place ready to deal with the next coefficient. Contact u2 closes and PM is energised. Relay PC comes on after 2 seconds starting the next run. The integrators in the teaching circuits are discharged immediately before the run, while PM is on and PC off.

The start button is electrically interlocked so that pressing it at the wrong time cannot cause confusion. If it is pressed during the run it has no effect since PM is energised anyway. During the rewind cb4 is open and during the adjustment ad2 is open so the start button has no effect at these parts of the cycle either. The

adjust button can have no effect when the machine is switched to 'automatic' adjustment. The operate delays on PC and RC have been provided so that the machine does not attempt to reverse the magnetic tape before the tape has come to rest.

2 Single cycle-Automatic

If switch S_1 is switched over to single-cycle, when $u2$ closes at the end of the cycle it has no effect on PM. Nothing further happens until the start button is pressed again whereupon $st1$ initiates another complete cycle of play-rewind and adjustment. The start button is interlocked as before.

3 Continuous-Manual

With S_1 switched to continuous, and S_2 switched to manual, the sequence described in (1) above takes place, but the machine pauses after play and rewind. The adjustor is now connected to a potentiometer on the control console, on which any coefficient value can be set up, instead of to C_m from the minimum computer, to allow manual computation of C_m . The circuit is shown in fig.50. When the adjust button is pressed the cycle continues and the coefficient is adjusted to the value of C_m set up on the control panel. Switch contact s_2^2 is open and prevents $b2$ from operating U at the end of the rewind. Relay U is operated by $p1$ at the end of adjustment and initiates the next cycle.

4 Single cycle- Manual

In this condition human intervention is necessary to start play of the tape, and also to initiate adjustment. In this condition

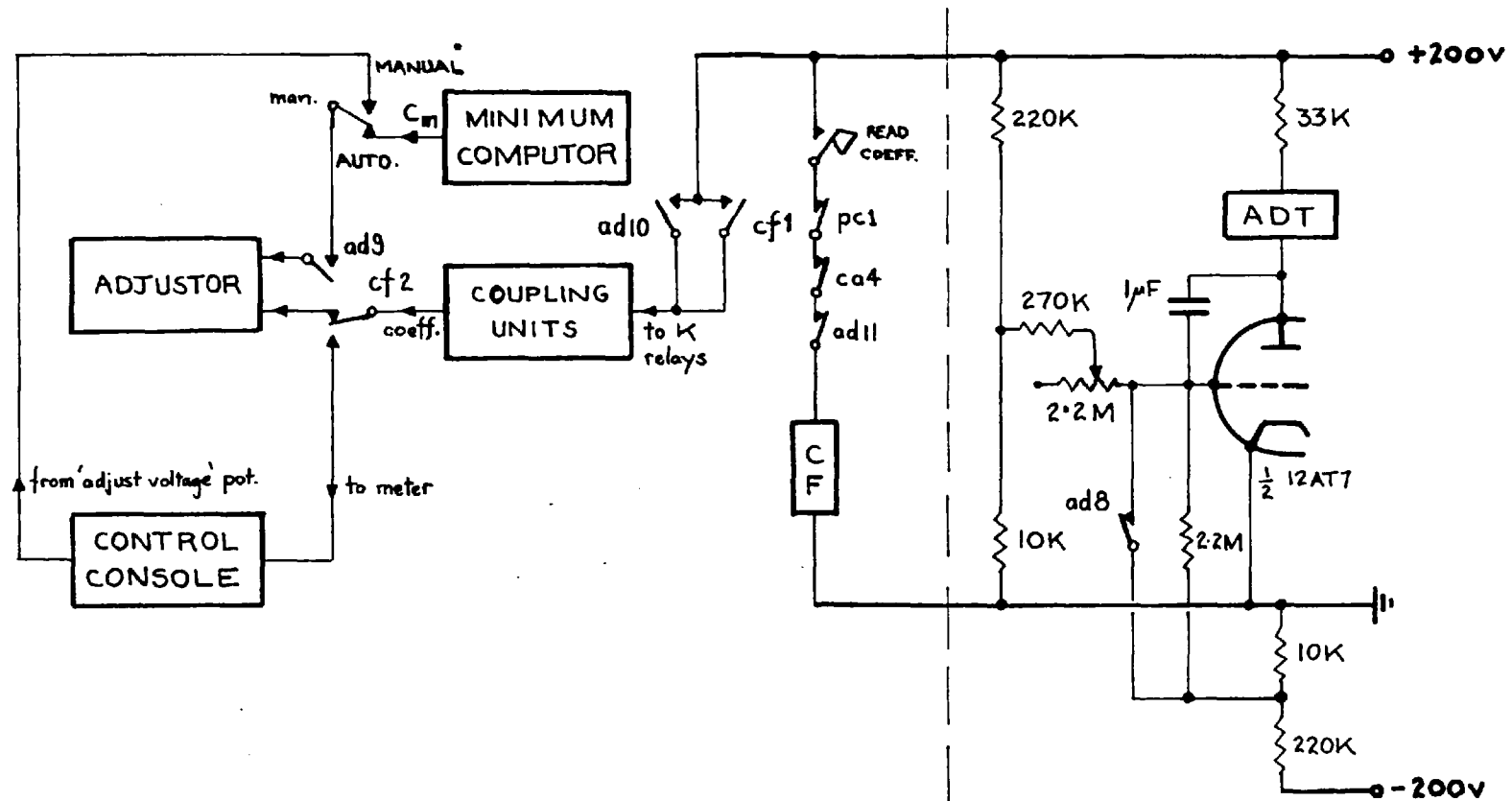


FIG. 50. 'READ COEFFICIENT', 'MANUAL ADJUST', AND 'ADJUST TIMER' CIRCUITS.

the 'adjust' cycle can be operated independently of the 'play-rewind' cycle, and this enables the coefficients of the Learning Filter to be set up to any required initial conditions.

'Read coefficient' circuit

It is essential to be able to read the value to which any particular coefficient is set, and the arrangement for doing this is shown in fig.50. When the 'read coefficient' button is pressed one of the K relays is energised, through cfl, (which one depends on the position of the coefficient switching uniselectors). The reference supply of +100 volts is put across the corresponding potentiometer and the wiper voltage can then be read on the meter on the control console. The contacts pcl, ca4, and ad11 are then wired in series with the pushbutton so that the coefficient can only be read if the play-rewind-adjust cycle is not taking place.

Adjustment timer

The circuit used as the adjustment timer is shown in fig.50. It is a Miller integrator circuit with a variable rate of charge. When relay AD is not energised, contact ad8 is closed, holding the grid of the triode negative so that the valve is cut off. When ad8 opens the condenser starts to charge, and the current in the valve slowly increases until ADT operates. This causes AD to fall out and ad8 closes resetting the circuit.

The time range obtainable by varying R_t is from 2 to 30 seconds. The timing accuracy and stability are rather poor but this

is not important. Two seconds is adequate for the automatic computation and adjustment, but the longer times are useful when checking the operation of the automatic computing and adjusting circuits.

Ready-adjust indicator

When using the machine in the single cycle-manual and continuous-manual modes an indication is needed as to whether the adjustment has been performed or not. Consequently the 'ready-adjust' relay RA has been installed. This operates an amber light on the control console which reminds the operator that play-rewind has taken place and that the adjustment has not been performed. The amber light goes out as soon as the 'adjust' button is pressed. A $1/\text{F}$ condenser (fig.48.) has been used to 'remember' whether the adjust relay has been in or not. On each cycle, contact a2 closes at the end of the 'play' and discharges the $1/\text{F}$ condenser, thus priming the 'ready adjust' circuit. If the machine is on 'manual-adjust' nothing happens until b1 closes at the end of the rewind. A pulse of current then flows into the condenser via b1 and relay RA is energised, holding on via ra1, ra2, and ad6. The amber light controlled by RA remains on until the adjust button is pressed, whereupon ad6 opens and RA releases.

If the machine is on 'automatic adjust' there are two possibilities. First suppose that adjustment is longer than rewind. After a2 has primed the circuit the next thing that happens is that ad5 changes over charging the condenser and ad6 opens. When b1 closes at the end of the rewind it now has no effect. When AD releases RA is not

energised. If rewind takes longer than adjustment ad5 and ad6 are in the positions shown, but the condenser is charged. When bl closes it has no effect, since equal voltages are applied to the two ends of RA and there is no current pulse.

Contact ra2 is provided to isolate the $1\frac{1}{2}$ F condenser from the positive supply.

5.8.3. Control Console.

The control console shown in fig. 51 has been provided so that the machine can be controlled from one central position. A diagram



fig. 51 CONTROL CONSOLE.

of the control panel is given in fig. 52 and the circuit diagram is given in fig. 53. There are 95 neon indicator lamps, one for each of the 94 coefficients, and one labelled FC (filter complete) representing a position where no coefficient is selected. The coefficient

whose influence curve is being computed is indicated by the corresponding neon flashing continuously. Coefficients which are being missed out of the learning process are indicated by their neons glowing steadily. The pair of decade switches in the centre, together with the 'operate' button and the function switch on the right enable the play-rewind-adjust cycle to be started at any chosen coefficient. They also

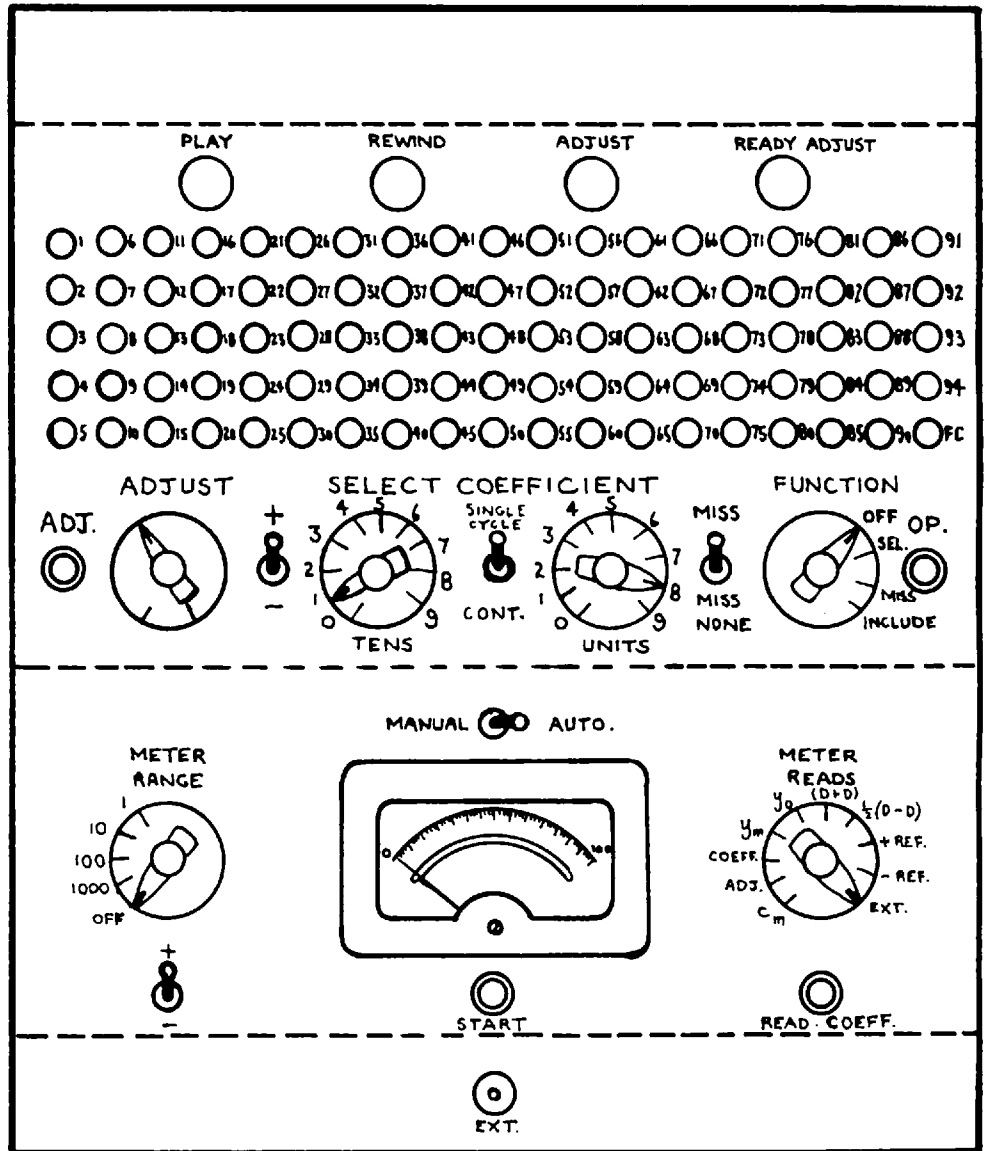


FIG. 52.

CONTROL PANEL

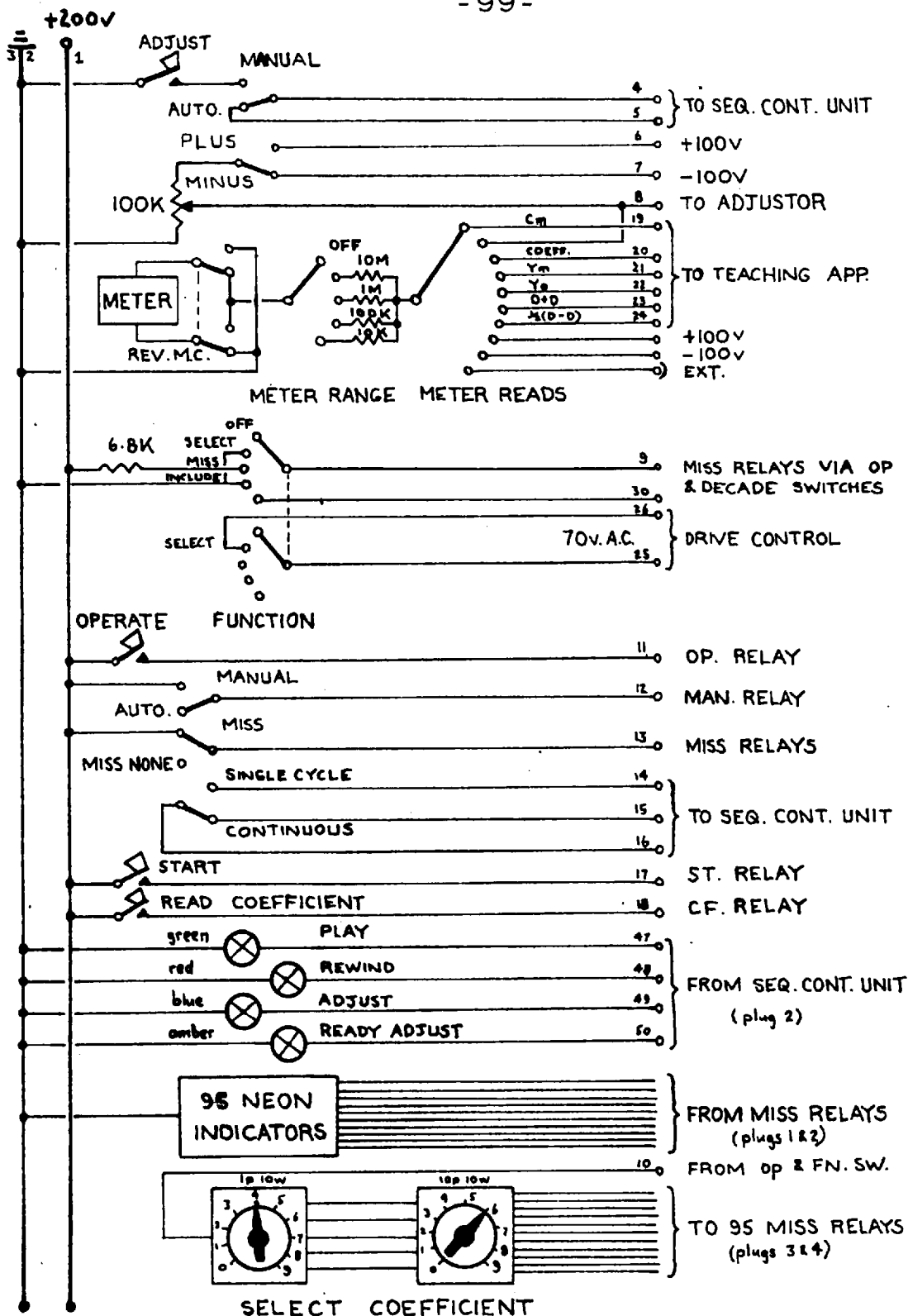


FIG. 53. CONTROL CONSOLE CIRCUIT DIAGRAM

enable any coefficient to be missed out of the learning process, or included again if it has been previously missed. The start button, read coefficient button and the single cycle/ continuous, and automatic/ manual switches referred to in explaining the sequence control unit are shown on the control panel.

For manual adjustment of a coefficient the 'adjust voltage' control is used to provide the input to the adjustor, the voltage being measured on the reference meter. When the voltage has been set up the adjust button is pressed to initiate the adjust cycle.

The 'meter range' switch on the left and the 'meter reads' switch on the right are self explanatory. The four indicator lights at the top of the panel show the state of the machine.

5.8.4. Coefficient Switching Uniselectors

Four uniselectors are used for switching the teaching apparatus to different coefficients. The drive control and bank switching circuit is shown in fig. 54. The starting position is with all the uniselectors on level 1 and in this condition none of the relays shown are energised. From fig. 54 it will be seen that the bank wiper lead is then connected to level 1 on uniselector CSA.

At the end of each cycle, when u closes CSA is stepped on one place by the discharge current from the $16\mu\text{F}$ condenser, and this continues until level 25 is reached. The high speed stop relay SA is then energised and sa changes over, transferring the drive to the second uniselector CSB. The bank switch relay BA also operates and ba

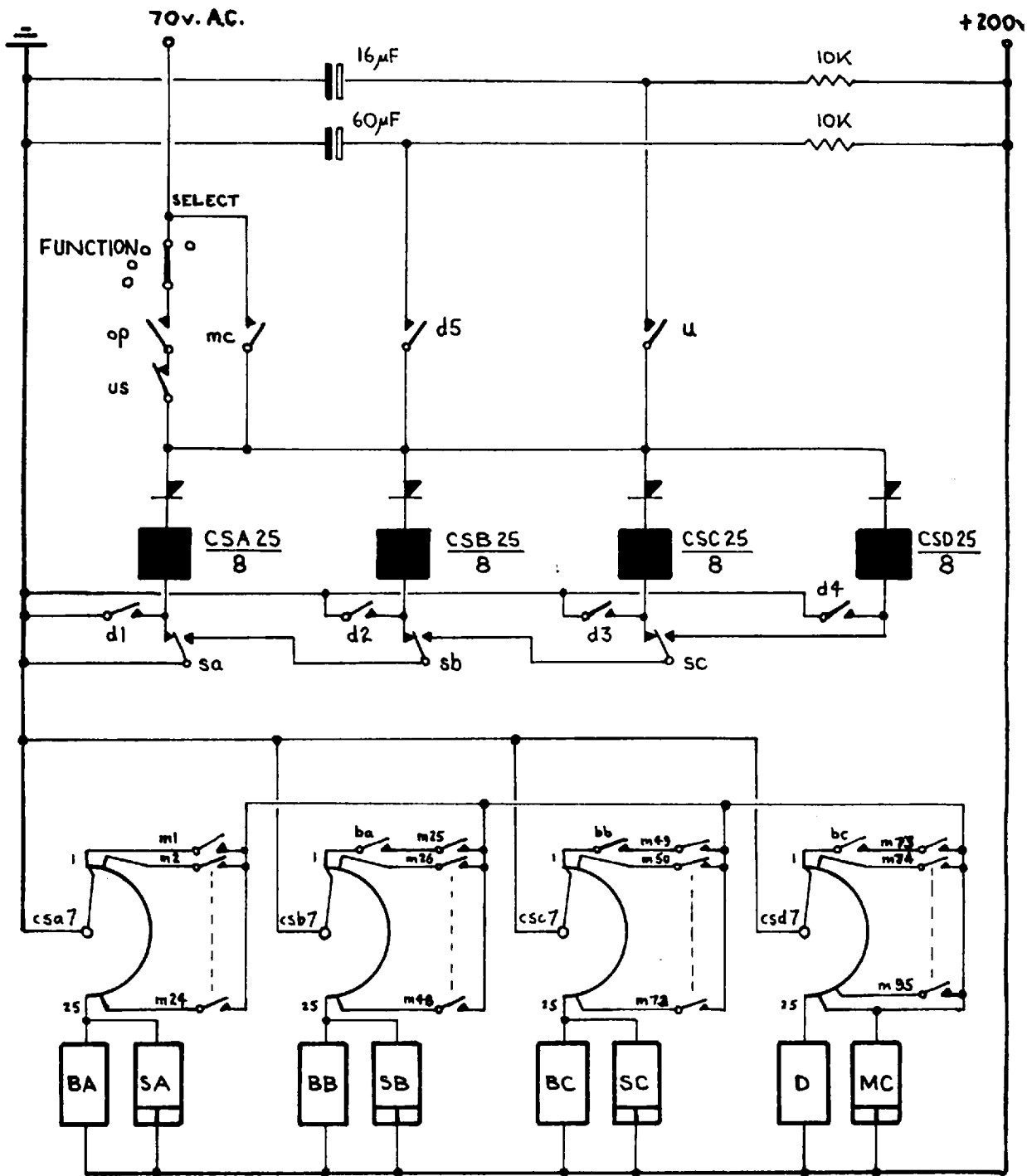


FIG. 54. DRIVE CONTROL AND BANK SWITCHING CIRCUIT

changes over transferring the wiper lead to level 1 (coefficient 25) on uniselector CSB.

When CSB reaches level 25 the drive is transferred to CSC, and when CSC reaches level 25 it is transferred to CSD. When level 25 on CSD is reached relay D is energised. Contacts d1 to d4 close and also the 'late make' contact d5. The $60\mu\text{F}$ condenser is then discharged through the four driving magnets in parallel and all the uniselectors step one place, enabling the cycle to be started again.

The procedure is modified if any of the miss relays M_n are energised. Suppose m3 is closed. When csa7 reaches level 3 the high speed relay MC is energised and mc closes. The uniselectors are then driven round by the A.C. supply until a position is found where the corresponding miss relay is not energised.

The bank circuits of the four uniselectors are wired as shown in fig.55. Only the bank connections of uniselector CSA are shown, the connections to the other three uniselectors are similar.

Bank 1 connects the signal corresponding to a coefficient of +1 to the teaching apparatus.

Bank 2 applies earth to the wiper of the selected potentiometer, so that its contribution is not added in the multi-input adder, this corresponds to a coefficient of zero.

Banks 3 and 4 connect the magnetic clutch coils of the selected potentiometer to the adjustor.

Bank 5 controls the 95 neon lamps on the control panel.

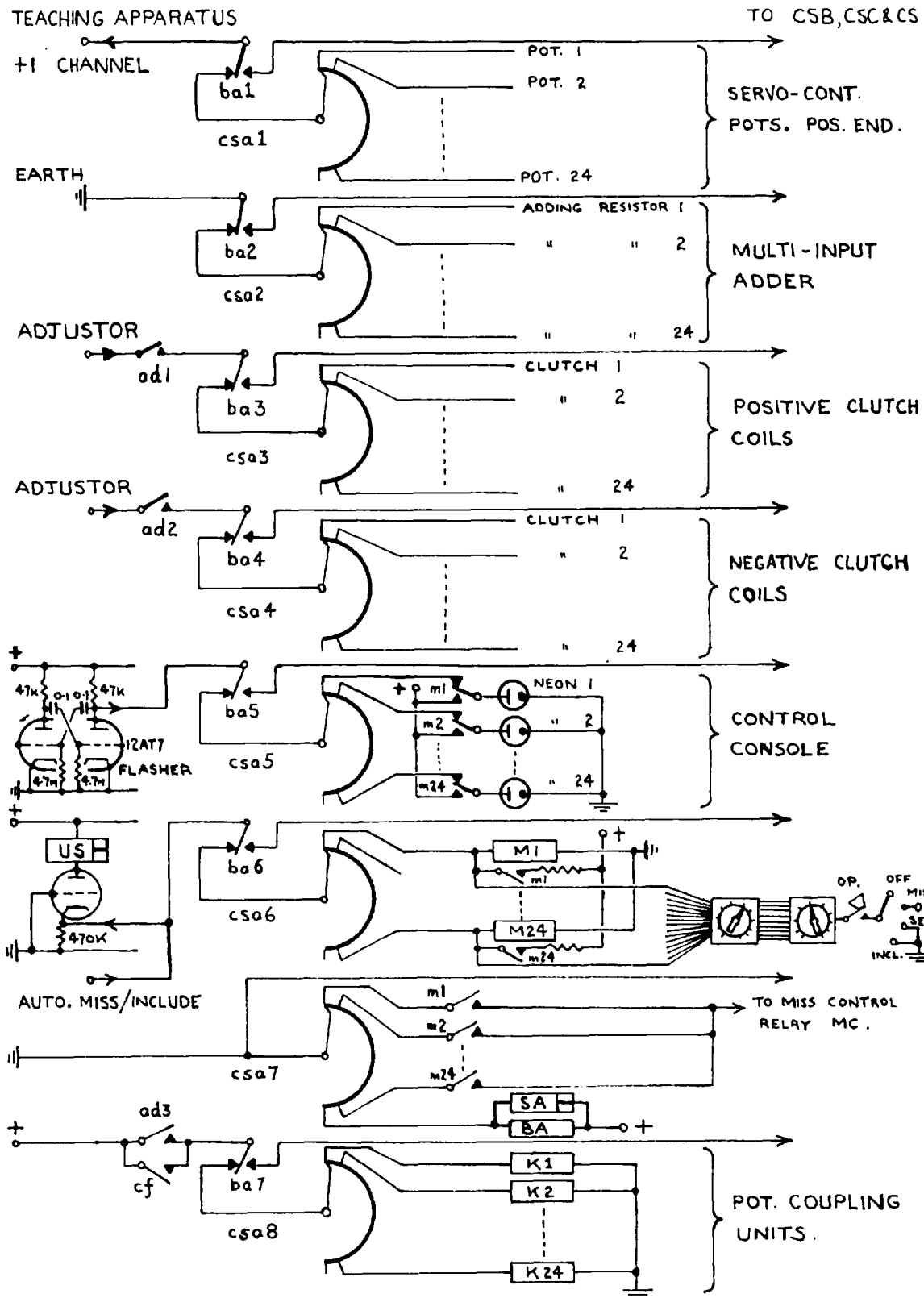


FIG. 55.

COEFFICIENT SWITCHING UNISELECTORS

BANK CIRCUITS

The neon corresponding to the selected coefficient is connected via a bank wiper to one of the anodes of a slow multivibrator, and the neon flashes on and off in sympathy with the anode voltage. The usual neon flashing circuit, where the neon is used to discharge a condenser which charges up slowly through a resistor, could not be used because the neon indicators employed were fitted with internal series resistors. An oscillating relay flasher was used successfully, but was discarded in favour of the multivibrator because it was noisy.

If a particular miss relay is energised then the corresponding neon is connected to H.T.+ and glows steadily.

Bank 6 is used for the miss relays, and also the coefficient selection circuit. Suppose a particular coefficient is set up on the pair of 'select coefficient' decade switches, and that the 'function' switch is set to 'miss'. When the 'operate' button OP is pressed H.T. is applied to the miss relay and the miss relay comes on, holding on through one of its own contacts and a resistor, and thereafter the corresponding coefficient will get left out of the learning routine. If at some later time it is desired to include the coefficient again, then the function switch is set to 'include'. Button OP now applies earth to the live end of the relay coil, and the miss relay falls out.

After a particular run the corresponding miss relay can be controlled from the bank 6 wiper in a similar manner, according to the instructions received from the automatic miss circuit, which decides whether the reduction in Y_m that has been produced is significant or

not. If a number of miss relays are holding on, they can be released simultaneously by operating the 'miss-none' switch, which breaks the H.T. supply to the relays.

If the function switch is switched to 'select' it applies earth to the tag chosen by the decade switches. (and if the miss relay was on, it releases). The switch also connects the A.C. supply to the drive circuit (fig. 54) and the uniselectors run round until the bank 6 wiper picks up the earthed bank tag. The high speed 'uniselector stop' relay US then operates stopping the scan, and the coefficient chosen on the decade switches is then selected. This enables the learning routine to be started at any chosen coefficient.

The bank 8 wiper energises one of the K_n relays in the potentiometer coupling units, when it is desired to adjust the potentiometer, or to read the value of the coefficient set up on it.

5.9. Teleprinter Output.

5.9.1. General

When the Learning Filter is working automatically it is useful to have a check on the decrease in mean square error, and a record of the coefficient settings made by the machine.

A Creed type 7B teleprinter (fig.56) has been installed to print out the required information at the end of each run. Circuits have been designed so that the quantities: run number (1-1000), potentiometer number (1-94), coefficient value (C_m) and magnitude of the

mean square error (Y_m) can be printed out. Each of these is first represented in teleprinter code (fig.57), the five code digits being



available simultaneously. The code is then scanned by a parallel to series converter which reads out the code digits sequentially adding the start and stop pulses required by the teleprinter. The quantities C_m and Y_m appear as voltages at the output of the

fig.56. TELEPRINTER

teaching apparatus and this makes nec-

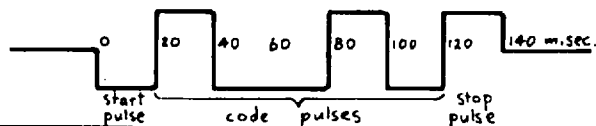
essary an analogue to teleprinter code converter, (hereafter called the voltage coder). The results are printed out in four columns (fig.57), provision has been made for automatically printing out the column headings every 25 runs.

Each set of numbers is printed out in $12\frac{1}{2}$ seconds, which is of the same order as the rewind time, and therefore the time for one run is not significantly increased.

5.9.2. Print out Control Circuits

Two uniselectors are provided on which the codes to be fed to the teleprinter are set up. From fig.58 it will be seen that unselector TPA is wired with the codes corresponding to the column headings. Unselector TPB is supplied with the codes corresponding to the quantities to be printed out. Both have codes for spaces, line feed, carriage return etc., as appropriate. The scan of these uniselectors is controlled

TYPICAL TELEPRINTER SIGNAL -



COMBINATION	LETTERS	FIGURES	CODE
1	A	-	1 1 0 0 0
2	B	?	1 0 0 1 1
3	C	:	0 1 1 1 0
4	D		1 0 0 1 0
5	E	3	1 0 0 0 0
6	F	%	1 0 1 1 0
7	G	£	0 1 0 1 1
8	H	J	0 0 1 0 1
9	I	8	0 1 1 0 0
10	J		1 1 0 1 0
11	K	(	1 1 1 1 0
12	L	)	0 1 0 0 1
13	M	.	0 0 1 1 1
14	N	,	0 0 1 1 0
15	O	9	0 0 0 1 1
16	P	0	0 1 1 0 1
17	Q	1	1 1 1 0 1
18	R	4	0 1 0 1 0
19	S	!	1 0 1 0 0
20	T	5	0 0 0 0 1
21	U	7	1 1 1 0 0
22	V	=	0 1 1 1 1
23	W	2	1 1 0 0 1
24	X	/	1 0 1 1 1
25	Y	6	1 0 1 0 1
26	Z	+	1 0 0 0 1
27	CARRIAGE RETURN		0 0 0 1 0
28	LINE FEED		0 1 0 0 0
29	LETTER SHIFT		1 1 1 1 1
30	FIGURE SHIFT		1 1 0 1 1
31	SPACE		0 0 1 0 0
32	ALL SPACING		0 0 0 0 0

TYPICAL	OUTPUT	RECORD
25	10	+632 787
RUN	POT	C Y
26	15	-210 775
27	17	+497 762
28	20	+015 754

FIG. 57. TELEPRINTER CODE, AND OUTPUT RECORD.

TPA BANKS						TPB BANKS							
A B C D E						A B C D E							
Level						Level							
1	0	0	1	0	0	Space	1	0	0	1	0	0	Space
2	1	1	1	1	1	Letter Shift	2	1	1	0	1	1	Figure Shift
3	0	1	0	1	0	R	3						RUN COUNTER AND CODER
4	1	1	1	0	0	U	4						
5	0	0	1	1	0	N	5						
6	0	0	1	0	0		6	0	0	1	0	0	
7	0	0	1	0	0		7	0	0	1	0	0	POT. NUMBER CODER
8	0	1	1	0	1	P	8	0	0	1	0	0	
9	0	0	0	1	1	O	9						
10	0	0	0	0	1	T	10						
11	0	0	1	0	0		11	0	0	1	0	0	± DETECTOR VOLTAGE CODER C _m
12	0	0	1	0	0		12	0	0	1	0	0	
13	0	0	1	0	0		13	0	0	1	0	0	
14	0	0	1	0	0		14						
15	0	0	1	0	0		15						VOLTAGE CODER Y _m
16	0	1	1	1	0	C	16						
17	0	0	1	0	0		17						
18	0	0	1	0	0		18	0	0	1	0	0	
19	0	0	1	0	0		19	0	0	1	0	0	VOLTAGE CODER Y _m
20	0	0	1	0	0		20	0	0	1	0	0	
21	0	0	1	0	0		21						
22	1	0	1	0	1	Y	22						
23	0	0	1	0	0		23						Line Feed Carriage Ret.
24	0	1	0	0	0	Line Feed	24	0	1	0	0	0	
25	0	0	0	1	0	Carriage Ret.	25	0	0	0	1	0	

FIG. 58. CODE WIRING OF UNISELECTORS TPA AND TPB

by the circuit of fig. 59. Uniselectors PS is the parallel to series code converter. It makes use of the fact that when the unselector is energised from the 50 c/s supply through a diode, the unselector steps round at 50 steps per second, remaining on each bank tag for 20 m.sec. This is exactly the length required for each code pulse in the teleprinter code. Bank 1 on PS is therefore wired as shown, tag 1 is the rest position. Tag 2 is marked with -80 v, each of the tags 3 - 7 are marked with plus or minus 80 v. depending on the code to be read out. Tag 8 is marked with +80v. When wiper psl scans round, a complete teleprinter code with start and stop pulses is fed to the teleprinter. Operation of the changeover switch CH allows the teleprinter to be used as a typewriter from its own keyboard.

The relay contacts heal-7, only change over when it is desired to print out the column headings, so they are normally in the positions shown. Consider now the operation of the circuit of fig.59 during a typical run. Initially all the uniselectors are on level 1 and none of the relays are energised. When the play of the tape finishes, the teaching circuits have finished computing, the results can be printed out and the coefficient adjusted to its new value. When the adjust relay operates, ad changes over and C_4 discharges through the 'start print' relay SP, energising it for a short time. Contact spl closes and the 'print out' relay PR is energised, holding itself on through ftpl and prl. Unselector PS can now be driven by the half wave A.C. supply, through contact pr2, and PS scans round, psl reading out the code on the

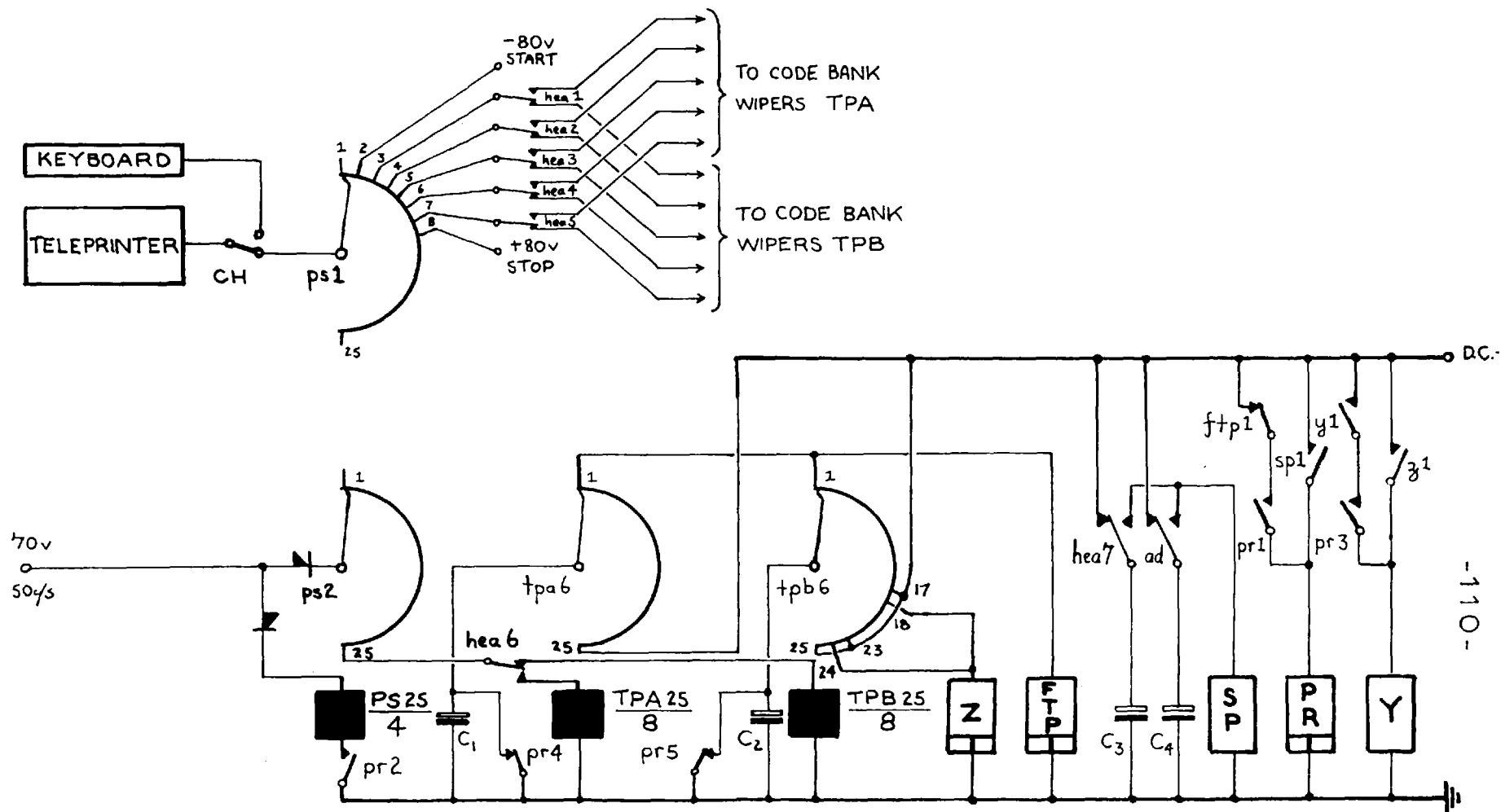


FIG. 59. PRINT-OUT CONTROL CIRCUIT

first level of TPB (00100). When ps2 reaches level 25 it makes a path for the next half cycle to energise TPB, and when PS steps onto level 1, TPB steps onto level 2. The code on level 2 of TPB (11011) is then read out by ps1 as it scans round. This sequence proceeds, the codes on successive levels of TPB being read out in turn. When level 17 on TPB is reached wiper tpb6 charges condenser C_2 . When level 18 is reached C_2 discharges through relay Z, holding it on for 500 m.sec. until TPB steps on. Relay Z resets the voltage coder, since C_m has already been coded and printed out, and initiates the coding of Y_m . When contact z1 closes relay Y is energised, holding itself on through y1 and pr3. Relay Y transfers the voltage coder from C_m to Y_m . When TPB reaches level 23 C_2 is recharged and when TPB steps onto level 24, relay Z is again energised, resetting the voltage coder ready for the next run. Condenser C_2 is recharged on level 25 and when TPB steps onto level 1 the 'finished with teleprinter' relay FTP is energised by C_2 . Contact ftpl opens and PR falls out, opening pr2 and stopping uniselectors PS and TPB on level 1 ready for the next run. Contact pr3 also opens releasing Y. Condenser C_2 is discharged by pr5 and FTP releases ready for a new cycle of operation.

If relay HEA is energised (which can only happen during 'play' not 'adjust') the action is similar to that described above, but the sequence is initiated by hea7 when play of the tape commences. Unisector TPB remains stationary and TPA steps round. Unisector PS now reads out the codes on the banks of TPA and the column headings are

printed out. When tpa6 reaches level 1 again, FTP is energised as before, stopping the scan. In this cycle, relays Z and Y do not operate.

A new run must not start until print-out is complete. This is ensured by setting the adjustment timer to give a delay of more than $12\frac{1}{2}$ seconds. A further refinement would be to interlock the PR and AD circuits.

5.9.3. Run Counter and Coder.

The circuit of the run counter is shown in fig.60, it includes a circuit which initiates print out of the column headings every 25 runs. The circuits are initially as shown in fig.60. Relay U closes at the end of each complete play-rewind-adjust cycle, and contact ul discharges C_1 through the driving magnet of uniselector RU, stepping RU onto level 2. Uniselector RU steps once each time ul closes until a count of 9 is reached. When wiper ru6 steps onto level 10 the half wave A.C. supply drives RU round until it reaches level zero. While wiper ru7 is on level 10 it makes a path to the A.C. supply for uniselector RT. Consequently when RU steps from 10 to 11, RT steps from 0 to 1. The carry from the tens to the hundreds is arranged in a similar way. The number of runs counted, in decimal form is therefore given by the positions of the wipers on RH, RT, and RU. There are five code banks on each of these uniselectors. The wipers of the code banks read out the teleprinter code corresponding to the particular decimal digits indicated. The wipers are connected to the code banks of uniselector TPB. The counter is reset to the 001 position by pressing the pushbuttons R_{1-3} .

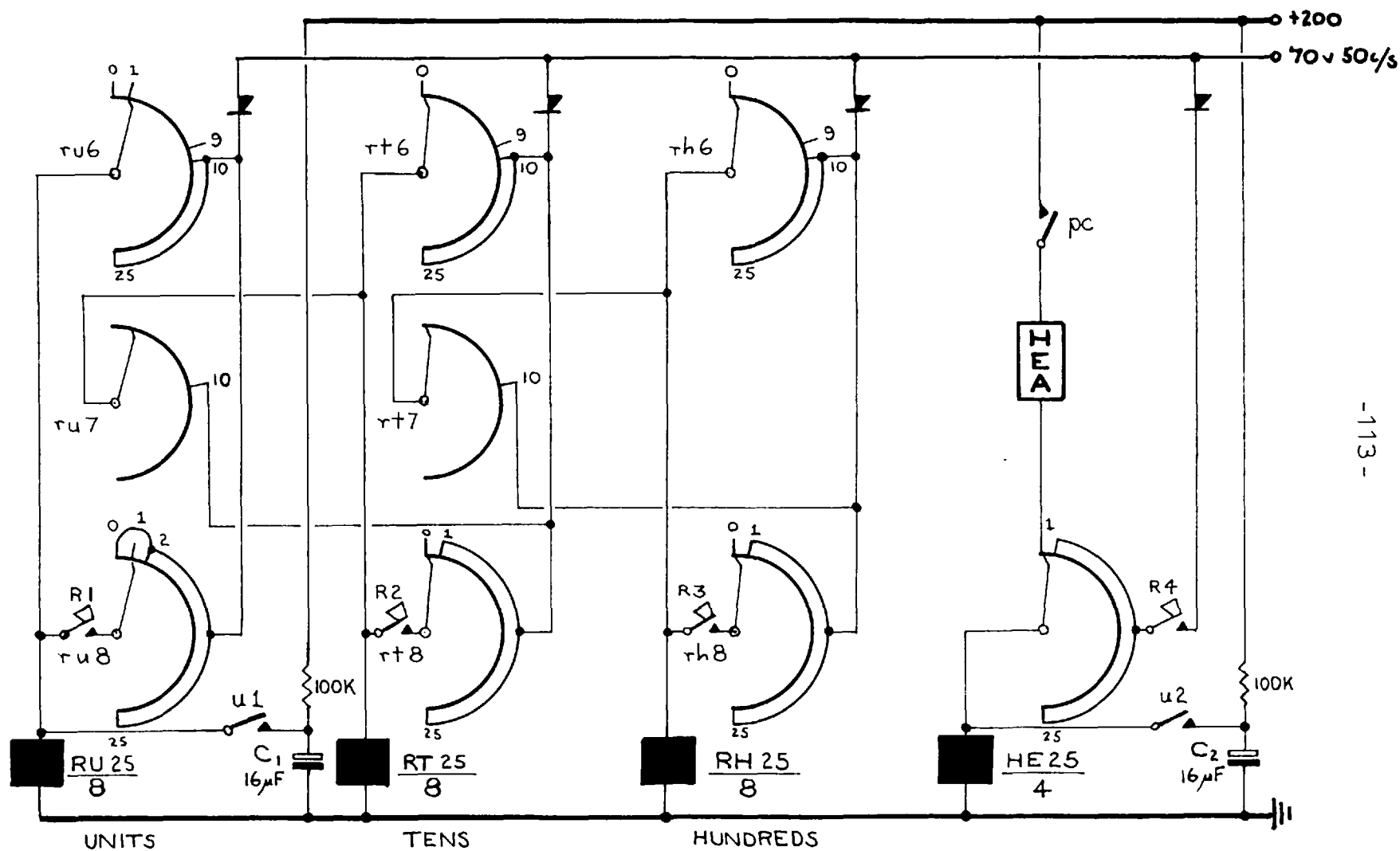


FIG. 60.

RUN COUNTER, CONTROL BANKS.

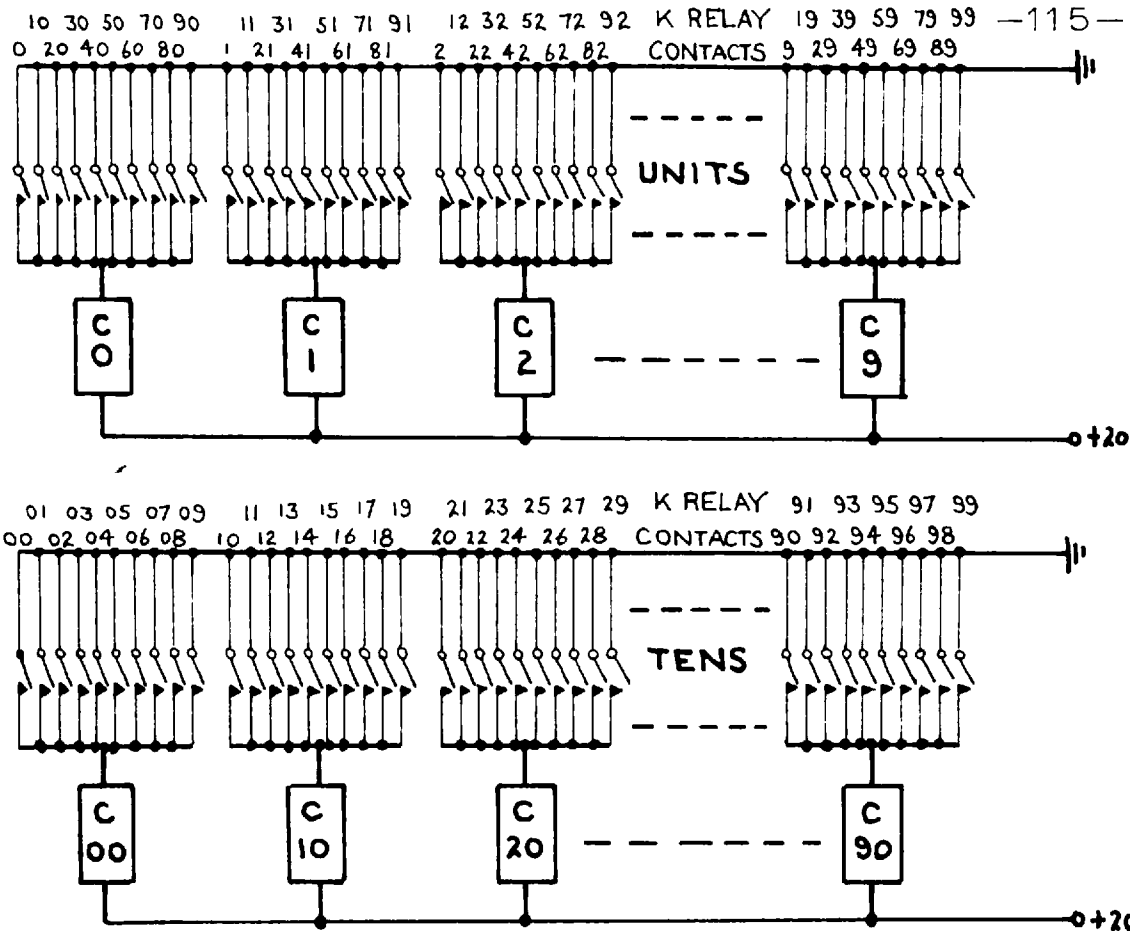
A separate uniselector HE is used to control print-out of the column headings every 25 runs. The uniselector is stepped one position at the end of each cycle, when u2 closes. Each time HE passes onto level 1, relay HEA is energised at the start of the 'play' when pc closes. Relay HEA then initiates print-out of the column headings from uniselector TPA. Uniselector HE can be reset by pressing pushbutton R_4 .

5.9.4. Potentiometer Number Coder

There is one relay (K_n) associated with each servo-controlled potentiometer. This relay is energised when the corresponding potentiometer is being adjusted to the new coefficient value C_m . Two make contacts on each K_n relay are used to operate 20 coding relays, which convert the potentiometer number into teleprinter code. The circuit is given in fig.61. Only one K relay can be energised at any one time and this energises one of the units relays (C-0 to C-9) and one of the tens relays (C-00 to C-90). For example if K_{21} is energised then relays C-1 and C-20 are also energised. Each of the coding relays has 5 make contacts, as shown in the lower part of fig.61, and when one of the coding relays is closed the appropriate teleprinter code appears on the five output lines, and is fed to the code banks of uniselector TPB.

5.9.5. Plus-Minus Detector

The voltage representing the calculated coefficient C_m may lie anywhere between ± 100 volts. By providing a circuit which recognises whether this voltage is positive or negative it is possible to arrange that only positive voltages need be translated into teleprinter



CODING CONTACTS. SAME FOR TENS AND UNITS.

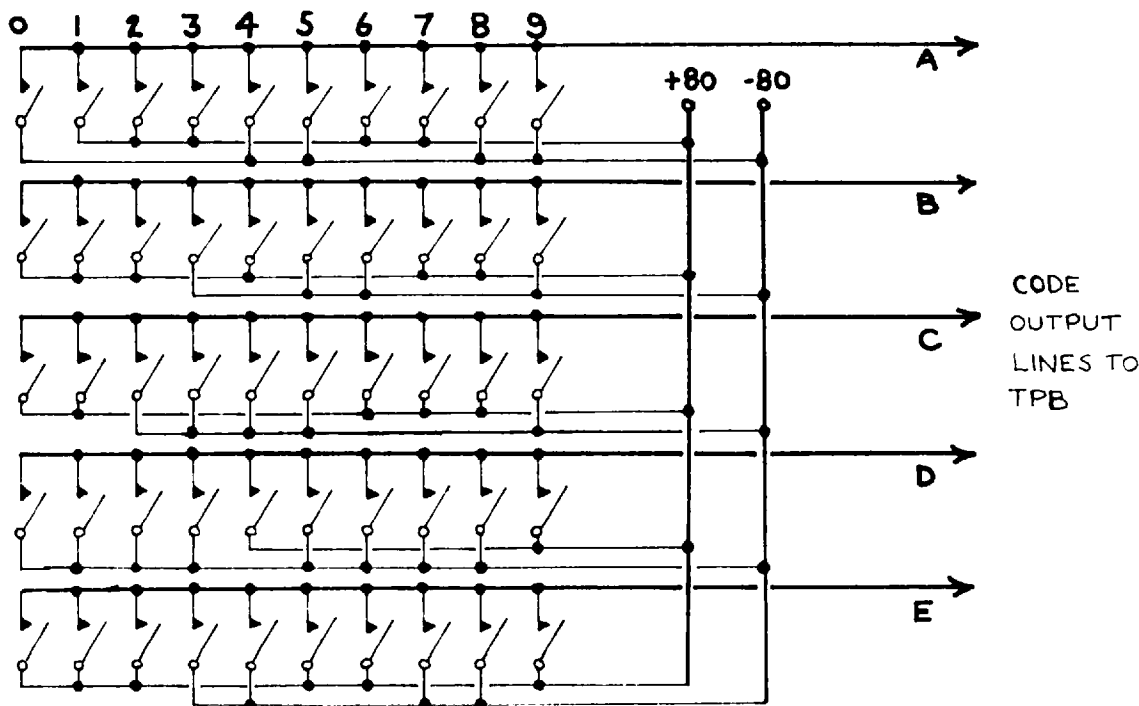


FIG. 61. COEFFICIENT NUMBER CODER.

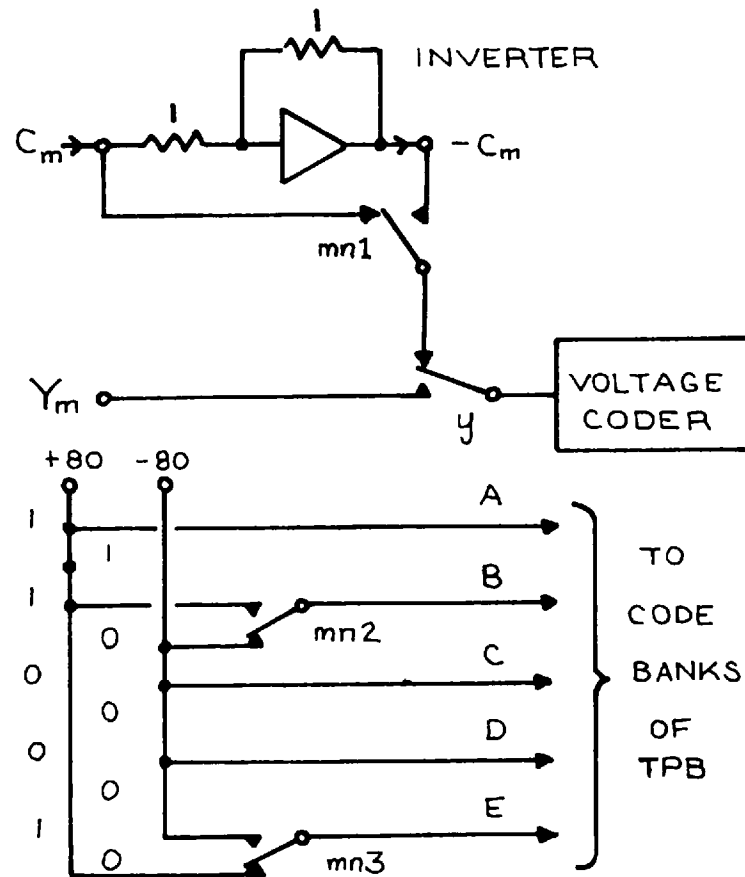
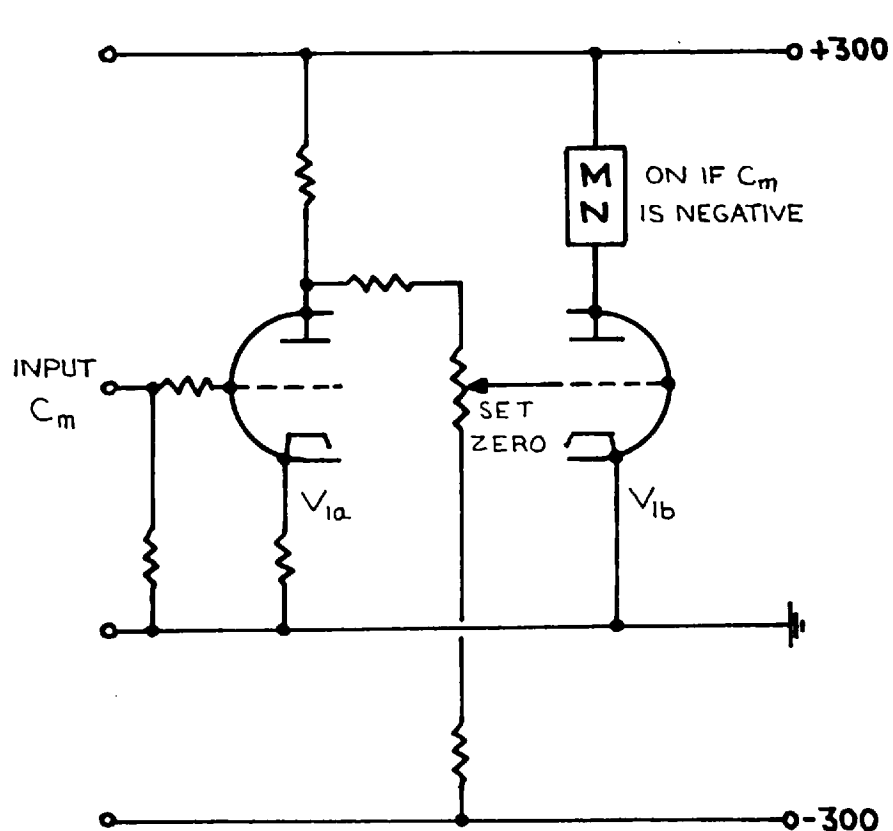


FIG. 62. PLUS - MINUS DETECTOR.

code. The circuit is shown in fig.62. Relay MN is operated by a 2-stage D.C. amplifier. The set zero control is adjusted so that V_{1b} is almost cut off. A very small negative voltage at the input terminal is then sufficient to operate MN. Contact mn1 then changes over and the voltage coder is fed with $-C_m$ (which will be positive) from the inverter output. Hence the voltage coder has a positive input whether C_m is positive or negative. The contacts mn2 and mn3 ensure that the correct sign precedes C_m when it is printed out. After $\left| C_m \right|$ has been coded contact y changes over and, after resetting, the voltage coder codes Y_m (which is always positive).

5.9.6. Voltage Coder

This unit accepts any voltage lying between 0 and +100 volts, and supplies the teleprinter codes corresponding to the three most significant decimal digits of the voltage, to the code banks of TPB. The circuit, using three uniselectors, is given in fig.63 and fig.64. The comparator shown, operates a high speed relay Q, when the voltage V_w picked up by the wipers of the coder is less than the input voltage, V_c . Initially the wipers of the three uniselectors CH, CT, CU, are all on level 1. Consider an input of 72.55 volts as a typical case. From fig.64, it will be seen that when contact pr closes, there is a path from the A.C. supply through CH, hal, hbl, q, and z, to ground. Unselector CH therefore steps round at 50 steps/second. When wiper ch7 (fig.63) picks up 70 volts, Q operates, and q changes over stopping CH. Relay HA is then energised through hb2, q, and z; and hal opens,

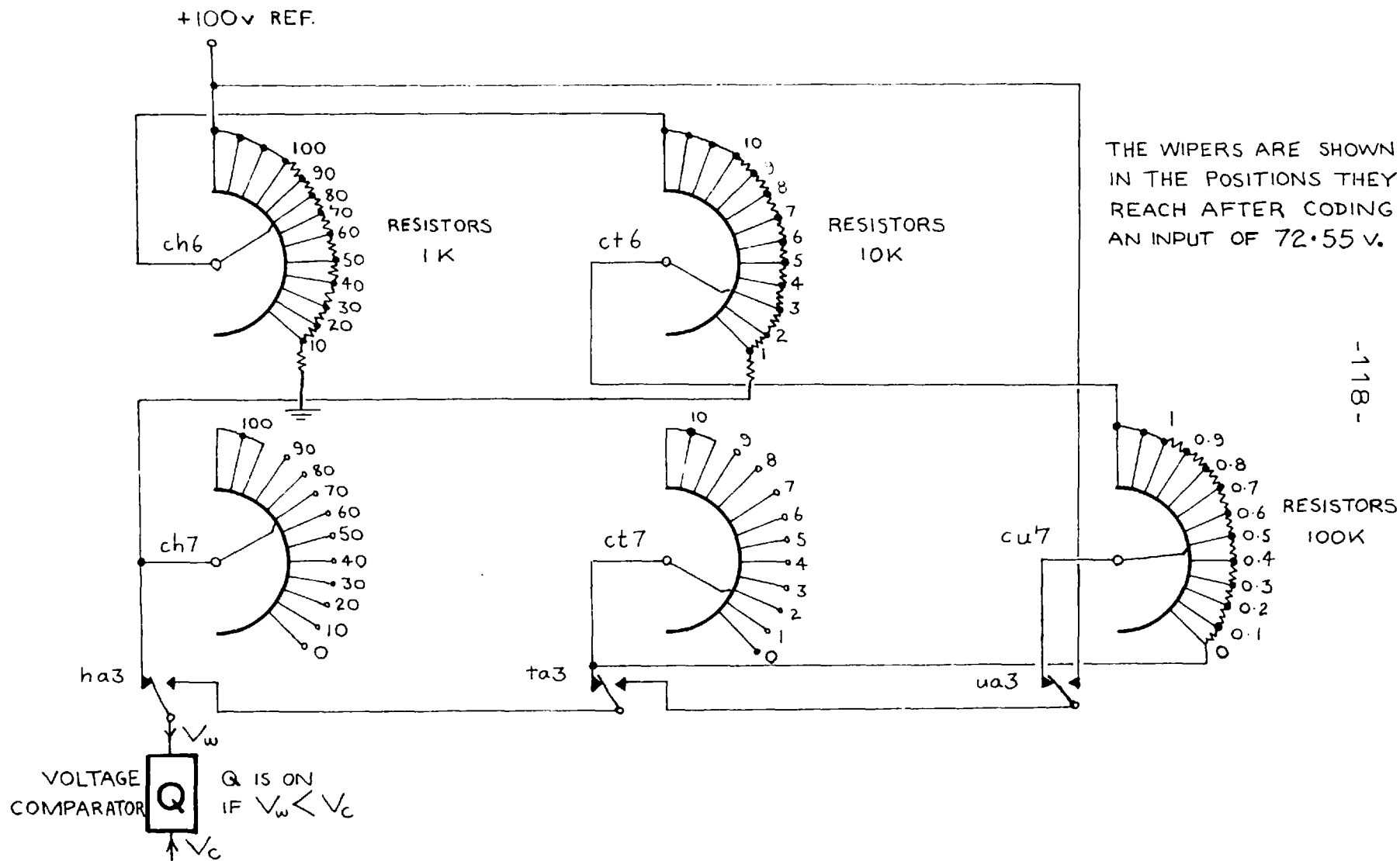


FIG. 63.

VOLTAGE CODER.

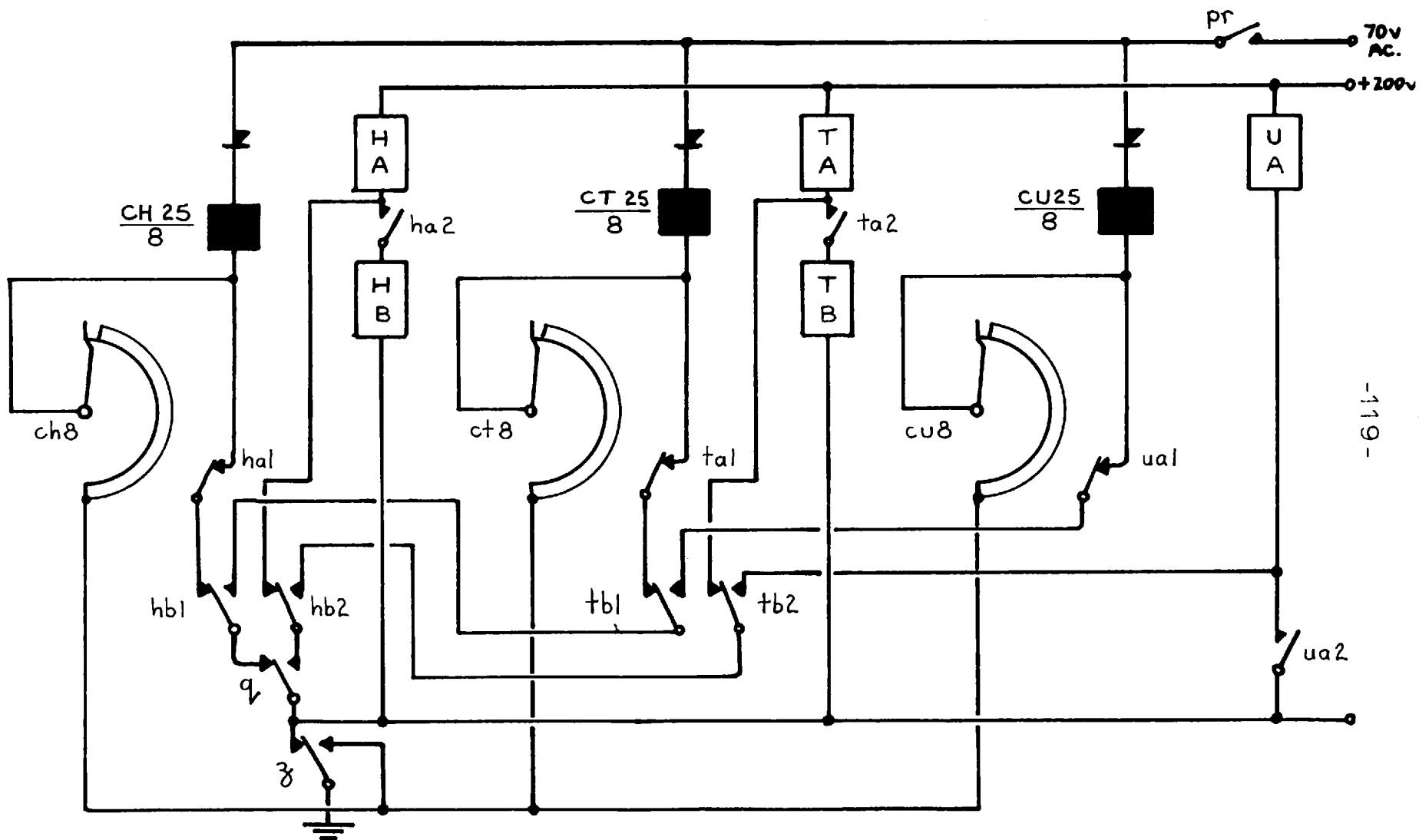


FIG. 64.

VOLTAGE CODER , DRIVE AND RESET CIRCUITS.

ensuring CH does not step again. Contact ha2 closes but relay HB is not energised because of the short circuit around it through hb2 and q. The voltage comparator is arranged so that q falls back after about 40 m.secs., and when it does so the short circuit is broken and HB comes on in series with HA. Contacts hbl and hb2 then change over transferring q to the next stage. Contact ha3 has already connected Q to wiper ct7. Wiper ch6 always picks up a voltage which is 10 v. greater than that picked up by ch7 and this 10 v. increment is divided into 10 equal parts on the bank tags of CT.

When contact hbl closes, CT runs round until wiper ct7 picks up the 72 volt point, whereupon Q operates and q changes over stopping CT. The second stage operates in the same way as the first and relays TA and TB transfer Q to the third stage.

Wiper ct6 always picks up a voltage which is 1 volt greater than that picked up by ct7. This 1 volt increment is divided into 10 equal parts on the bank tags of uniselector CU. When Q releases, uniselector CU runs round until wiper cu7 picks up the 72.5 volt point. Contact q then changes over stopping CU and relay UA operates holding on through contact ua2. Contact ual opens ensuring that CU does not step again when Q releases, and this prevents any 'jitter' on the last digit. The wipers of CH, CT, and CU are shown in fig.63 in their final positions for an input of 72.55 volts, which is coded as 725. If the input is C_m this is read as 0.725, and if the input is Y_m it is read as 72.5.

There are five code banks on each of the uniselectors. The wipers of the code banks read out the teleprinter codes, corresponding to the decimal digits indicated by the final positions of the wipers, and the codes appear on the code banks of TPB. In the above example, coding requires less than one half second. The coding time varies directly with the size of the input voltage.

The voltage coder is reset when contact z changes over. When Z operates there is no D.C. path for relays HA, HB, TA, TB, and UA, which have been holding on, and they all release. Contact z also provides a circuit via the homing wipers ch8, ct8, and cu8; and CH, CT, and CU run round to their start positions on level 1. When Z releases after 500 m.secs. the circuit is ready to code another voltage.

5.9.7. Voltage Comparator

For use with the voltage coder, it is essential to have a comparator capable of detecting a difference of not more than 0.1 volts, between two input voltages which may lie in the range 0 to 100 volts. The comparator must be unidirectional, that is, it operates if B is less than A, but not if A is less than B. The circuit used was developed from the basic 'long tail pair' difference amplifier. In this circuit using two triodes, the input signals are fed to the two grids, and the difference signal is obtained between the two anodes. A usual refinement is to replace the common cathode resistor by a pentode. The pentode serves to keep the total current in the two triodes constant if both input grids go up or down together (common mode).

This circuit alone does not have sufficient gain to give the 0.1 volt resolution mentioned above, if a relay is put between the two anodes. It is therefore necessary to amplify the difference signal further and this is simplified if the circuit can be modified to give an output relative to earth. This has been achieved as shown in fig.65, by adding the resistors marked R_d and $R_L / 2$. If the input grids are at the same voltage then since twice as much current flows in $R_L / 2$ as flows in R_L , the positive voltage at X is equal in magnitude to the negative voltage at Y. This results in very good common mode rejection. The amplified difference signal is available at point Z, relative to ground.

A complete analysis shows that the gain to the difference signal is given by:-

$$G_d = \frac{\mu}{2} \cdot \frac{R_L}{R_p + R_L}$$

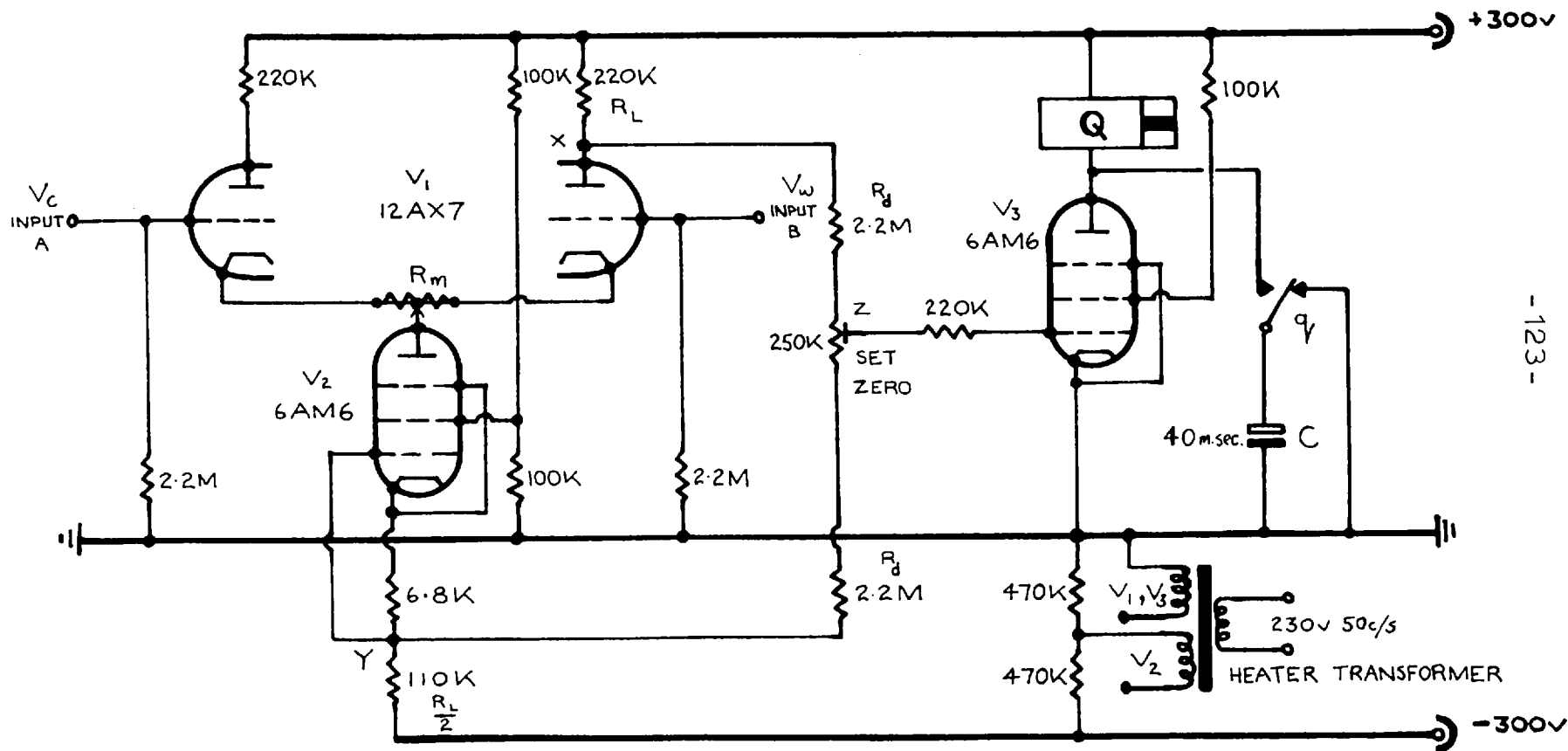
and that the gain is independent of the pentode impedance.

The common mode gain is given by:-

$$G_s = \frac{1}{2} \cdot \frac{R_L}{R_p + R_L} \cdot \frac{\mu_2 - \mu_1}{\mu_2 + \mu_1}$$

The measured common mode rejection ratio G_d / G_s (with $R_m = 0$) was 3300/1 (this may be improved even further by adjusting R_m).

The difference signal at Z is further amplified by the pentode V_3 which operates the high speed relay Q. When contact q changes over it connects condenser C to the anode of V_3 causing a rush of current



**FIG. 65. VOLTAGE COMPARATOR FOR ANALOGUE → TELEPRINTER
CODE CONVERTER**

through Q which ensures a positive snap-on action. As soon as Q operates relay contacts ha₃, ta₃ or ua₃ (fig.65.) ensure that V_z is cut off again by the new input conditions. Relay Q is held on by the charging current of condenser C for 40 m.secs. and then releases, C being discharged by q ready for another operation.

Performance

Using stabilized H.T. supplies and high stability resistors the resolution obtainable with the circuit of fig.65 is rather better than 0.1 volts over the entire range from 0 to +100 volts.

The circuit was also tested under typical operating conditions, by using it to control a voltage coder, fed with a succession of input voltages chosen according to a random number table. In each case the voltage V_w picked up in the final position of the coder was compared with the input voltage, V_o . The result of this test indicated that the average resolution of the coder was 0.1 volts.

All the teleprinter output arrangements described above have been tested in experimental form, but the installation of the circuits is not yet complete.

5.10. Auxiliary Output

In addition to the teleprinter output, it will be an advantage in some problems to have the results in another form. The best form will depend on the nature of the problem.

For example, in prediction problems it will be an advantage

to be able to look at fairly long lengths of record, so that the teaching track and the filter output, can be compared. This can be done by superimposing the signals on a dual trace oscilloscope, one trace being broken by blanking pulses for identification. Alternatively the oscilloscope time base may be switched off, and the results recorded on film using an oscilloscope camera.

In other problems, such as the filtering of signals from noise, it is an advantage to be able to hear the input, and the output after filtering. In this case because of the low frequency band used, the signals must be speeded up, until they lie in the audio band. They can then be played back through an ordinary audio amplifier and loudspeaker. This process is facilitated by the use of a second, portable, twin-track tape recorder with a wide range of speeds and special record-playback circuits. Such a tape recorder would also be useful on the input side for translating problems into the frequency band accessible to the machine. The pulse rate modulation system already described is suitable, with a few modifications, for use with the portable tape recorder. Thus if a very slow input signal is to be speeded up, the timing condenser can be increased to lower the modulator centre frequency to a suitable value, and during record the tape speed is made correspondingly low. It can be arranged that when the tape is played back at $7\frac{1}{2}$ "/sec. the centre frequency becomes the standard value of 4 Kc/s.

It is a useful feature that changing the tape speed on playback does not require complicated compensation in the demodulator.

Since all frequencies are changed in direct proportion to the tape speed, the signal frequency and also the signal amplitude are changed in the same proportion. This necessitates a change in the scale of the signal on playback but this is readily achieved.

The portable tape recorder would also be useful in collecting problems for the Learning Filter from various sources with a minimum of trouble, and for recording permanently on $\frac{1}{4}$ " tape the results achieved by the Learning Filter. The second tape recorder has not yet been installed.

5.11. Power Supplies

The H.T. supplies needed by the Learning Filter are considerable. The requirement for the valve circuits, excluding the 40 twin multiplier chassis, is +300v, 2 amps, and -300v, 0.3 amps, both well stabilized, and this has been provided. A further ± 1 amp. will be required for the multiplier chassis when their installation is complete.

The stabilizers used are of the type described by Attree,³¹ these use a 12AX7 cascode D.C. amplifier with a gain of about 1000, which is used to control the grids of a number of E.L.34 series valves; the reference voltage being provided by a neon type C.V.449.

The relay circuits are fed from the 200v. D.C. mains supply, and can consume a maximum current of 3 amps. The total valve heater supply is about 150 amps at 6.3v. A.C.; the current for D.C. coupled circuits is obtained from constant voltage transformers.

6. INTERCONNECTION OF LEARNING FILTER PARTS.

The disposition of the various parts of the Learning Filter may be seen from the assembly plan of the machine fig. 66, and the floor plan fig. 67. The stage reached by October 1959 is shown in the photograph of the complete machine, fig. 68. Wherever possible units have been interconnected using multi-way plugs and sockets, so that they can be removed for servicing. High grade connectors, with gold plated phosphor bronze pins and sockets, have been used, and no trouble has been experienced due to faulty connections. Screened lead has been used where necessary and the multiple cable forms have been colour coded using the decimal system, so that circuits can be easily traced. Low impedance output stages have been freely used to avoid trouble due to hum pick up and crosstalk on long cable runs.

Power cables have been kept as far as possible from signal cables. Power circuits which only need to be live when signals are not being processed (e.g. the 500 c/s supply to the clutches, and the power to the potentiometer driving motors) are connected by relays when required, under the control of the Sequence Control Unit, and this reduces the possibility of pick up from these circuits.

The relay and uniselector circuits generate strong interference, but since these only switch when the signals from the tape are not being processed, they cause no trouble.

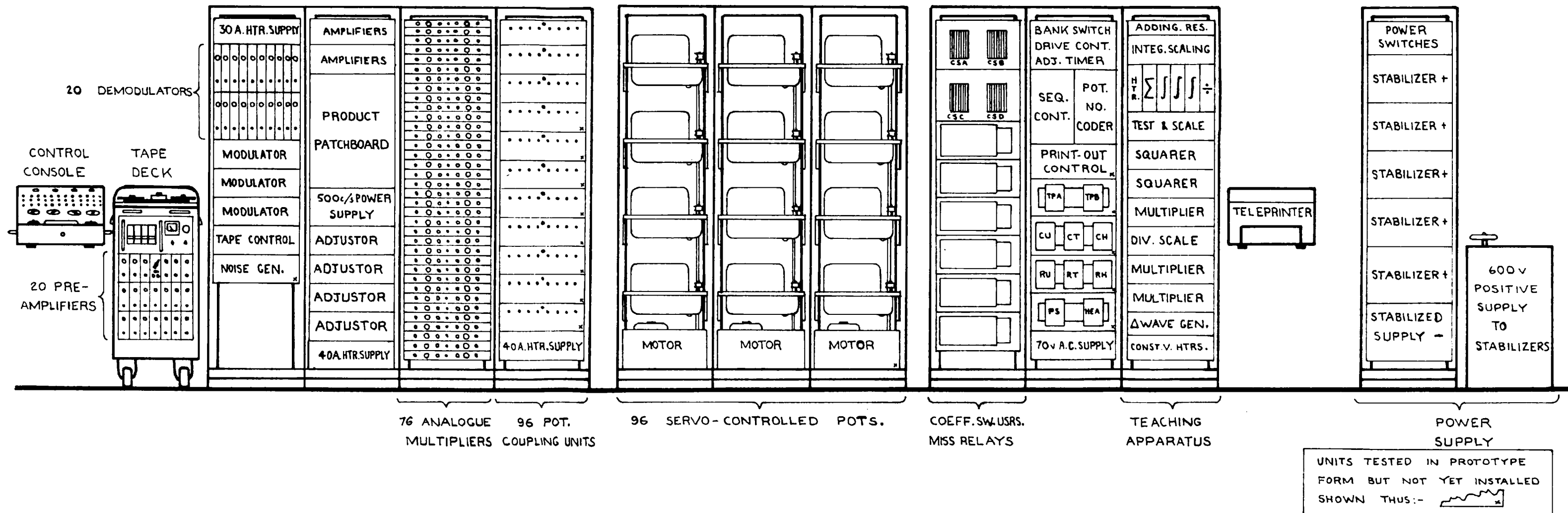
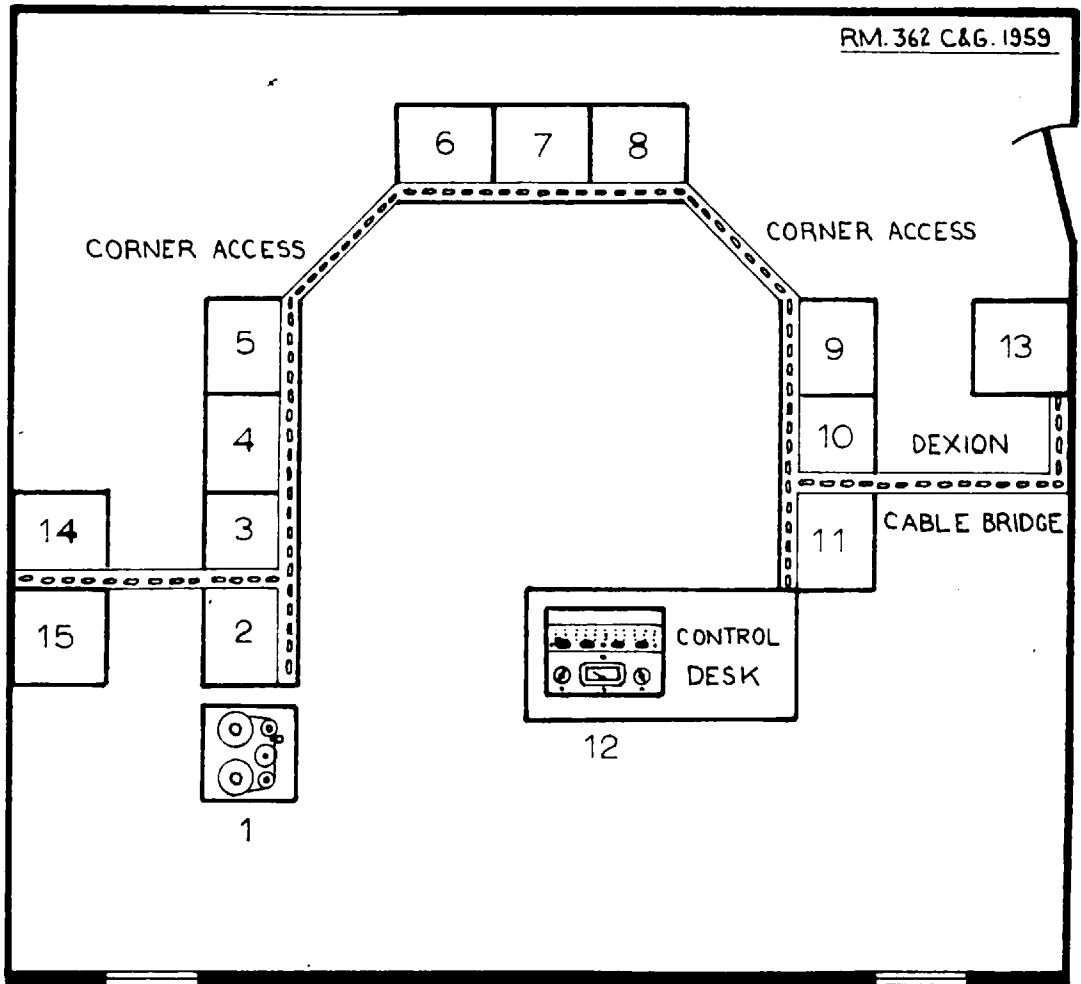


FIG. 66. ASSEMBLY PLAN, LEARNING FILTER.



KEY

- 1 TAPE DECK CONSOLE
- 2 MODULATORS , DEMODULATORS , TAPE CONTROL UNIT , INPUT.
- 3 MULTIPLIER INTERCONNECTION PATCHBOARD , ADJUSTOR
- 4 76 PIEZO-MAGNETIC MULTIPLIERS
- 5 POTENTIOMETER COUPLING UNITS
- 6 } SERVO- CONTROLLED POTENTIOMETERS
- 7 }
- 8 }
- 9 COEFFICIENT SWITCHING UNISELECTORS , MISS RELAYS
- 10 SEQUENCE CONTROL , DRIVE CONTROL , OUTPUT CODING & CONTROL
- 11 TEACHING APPARATUS
- 12 CONTROL CONSOLE
- 13 OUTPUT TELEPRINTER
- 14 H.T. SUPPLY STABILIZERS
- 15 H.T. POWER PACK

FIG. 67. FLOOR PLAN , LEARNING FILTER



FIG. 68. LEARNING FILTER

7. RESULTS OBTAINED WITH THE LEARNING FILTER

7.1. Scaling of a Sine Wave

This experiment was designed to check that the complete machine was working in the manner intended, and in particular to check the performance of the teaching apparatus. Eight linear terms were used and the input signal was a 100 c/s sine wave. The teaching signal was also a 100 c/s sine wave, in phase with the input signal but of half the amplitude. The 8 coefficients were initially set to +1, and the pick up heads were not staggered for playback. The 8 coefficients should therefore be regarded as one coefficient of 8 times the range.

The result achieved is shown in fig. 69. The sum of the 8 signals was initially much larger than the teaching signal and the first coefficient was set by the machine to -1, the maximum possible value. The calculated value was such that if it could have been applied the mean square error would have immediately been reduced to zero. The second and third coefficients were similarly set to -1. This meant that 6 of the sine waves were cancelling one another out, and this enabled the fourth coefficient to be set to a value such that the output of the filter was exactly equal to the teaching signal. It will be seen from fig. 69 that the filter computed the correct value for the fourth coefficient and reduced the mean square error to zero.

The result of this very simple experiment was reassuring in that it demonstrated that the machine was at least capable of computing accurately and capable of organising itself in the intended manner.

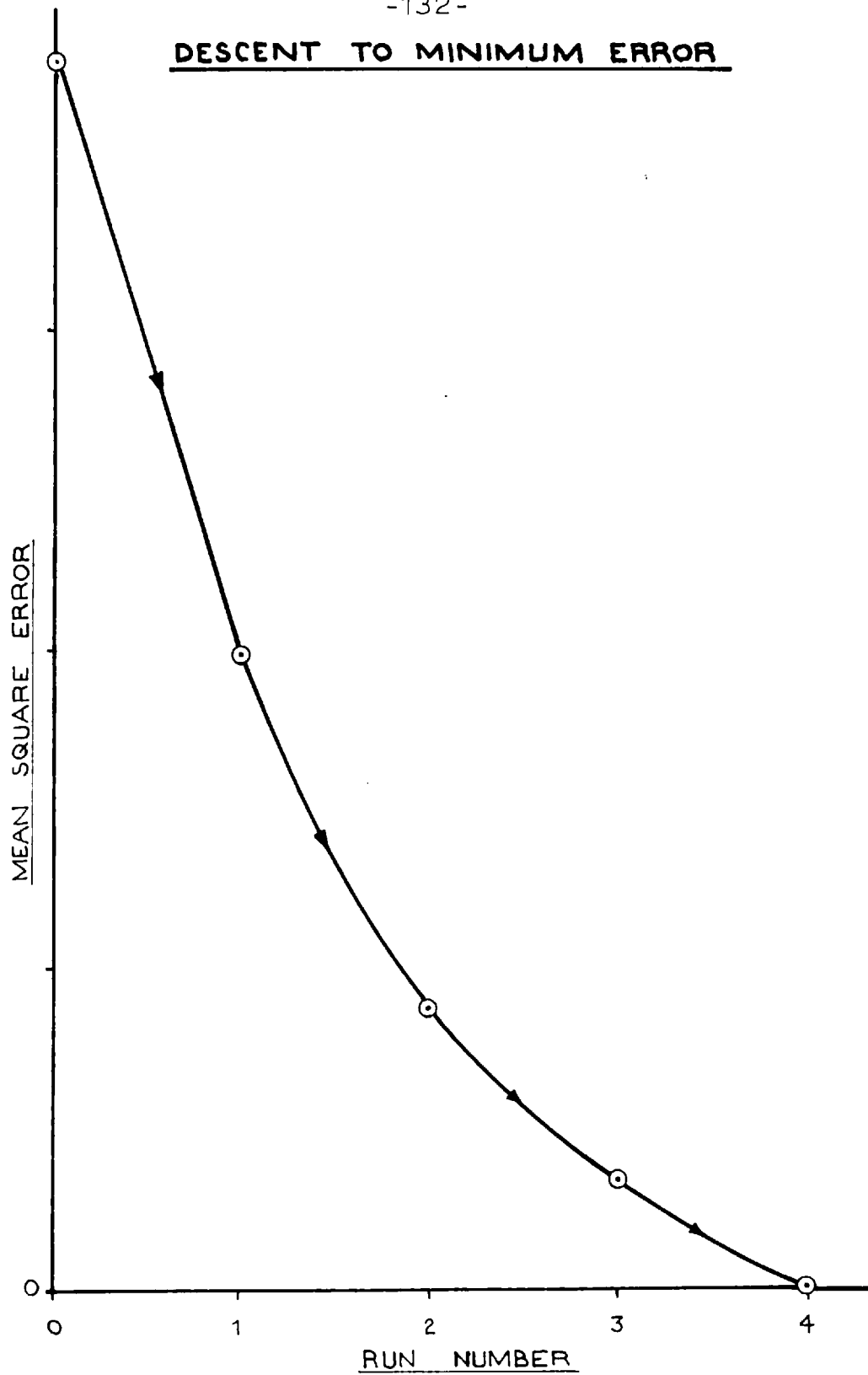


FIG. 69. SCALING OF A SINE WAVE

7.2. Phase Shifting and Scaling

This experiment was designed to utilize two coefficients and to demonstrate in the simplest possible way the action of the machine in optimizing itself by successive approximations.

The input was again a 100 c/s sine wave but the pick up heads were staggered to give a phase lag of 120° in the second channel. The teaching signal was a smaller 100 c/s sine wave, arranged to have a phase lead of approximately 120° on the first channel. The Learning Filter therefore had to produce a sine wave of a particular phase and amplitude, by adding two other sine waves in correct proportion.

It was known that the correct result would require two equal negative coefficients. The coefficients were initially set to +1 (as wrong as possible) and the machine was allowed to optimize itself, adjusting the coefficients alternately. The result achieved is shown in fig. 70. It will be seen that the reduction in the mean square error is very rapid initially and that an error as near to zero as required can be obtained.

The progress of one of the coefficients to its final value is shown in fig. 71. It is notable that starting the coefficients from -1 instead of +1 does not appreciably speed up the optimization, since the initial reduction in the mean square error is so rapid.

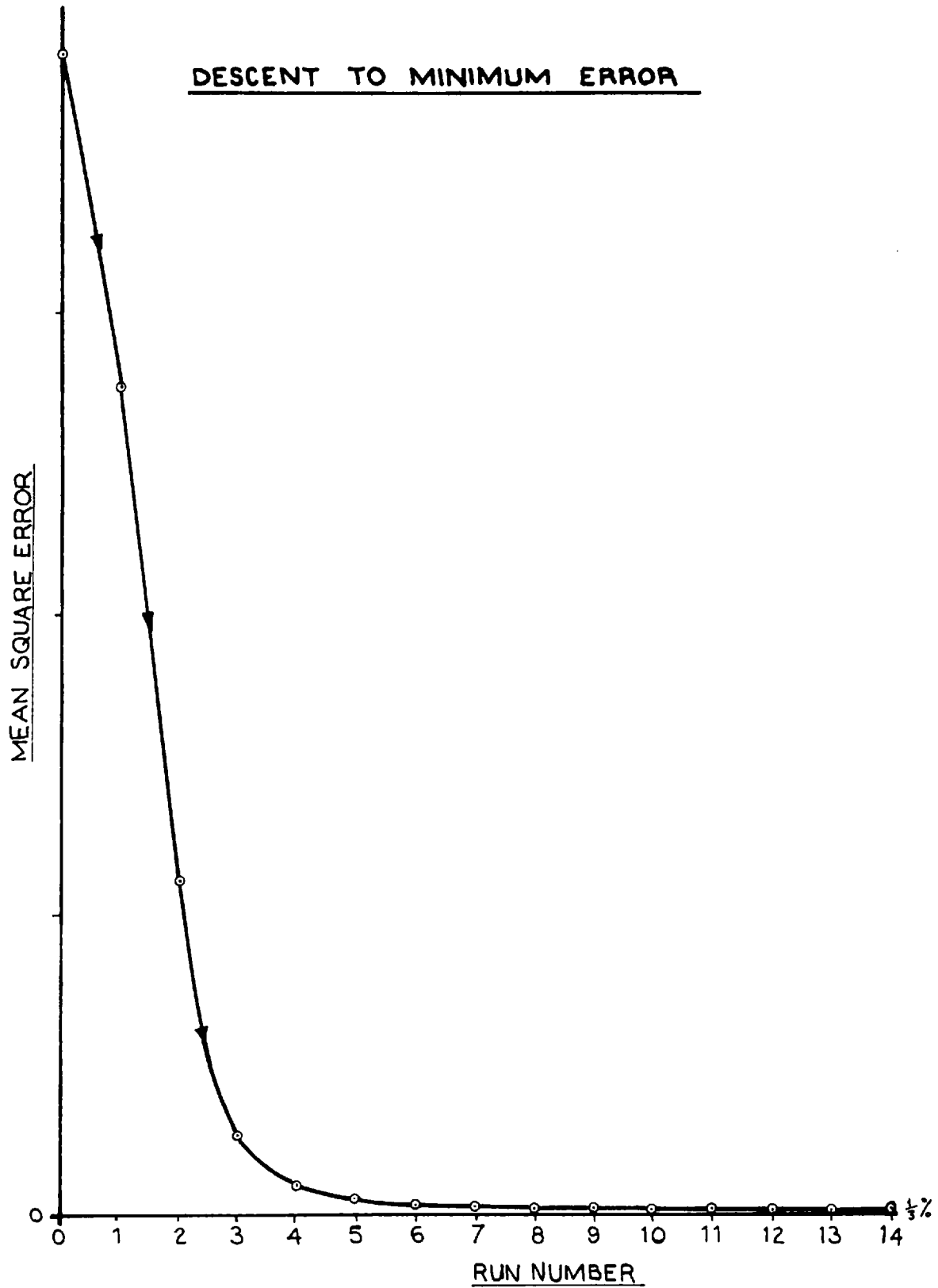


FIG. 70. PHASE SHIFTING AND SCALING

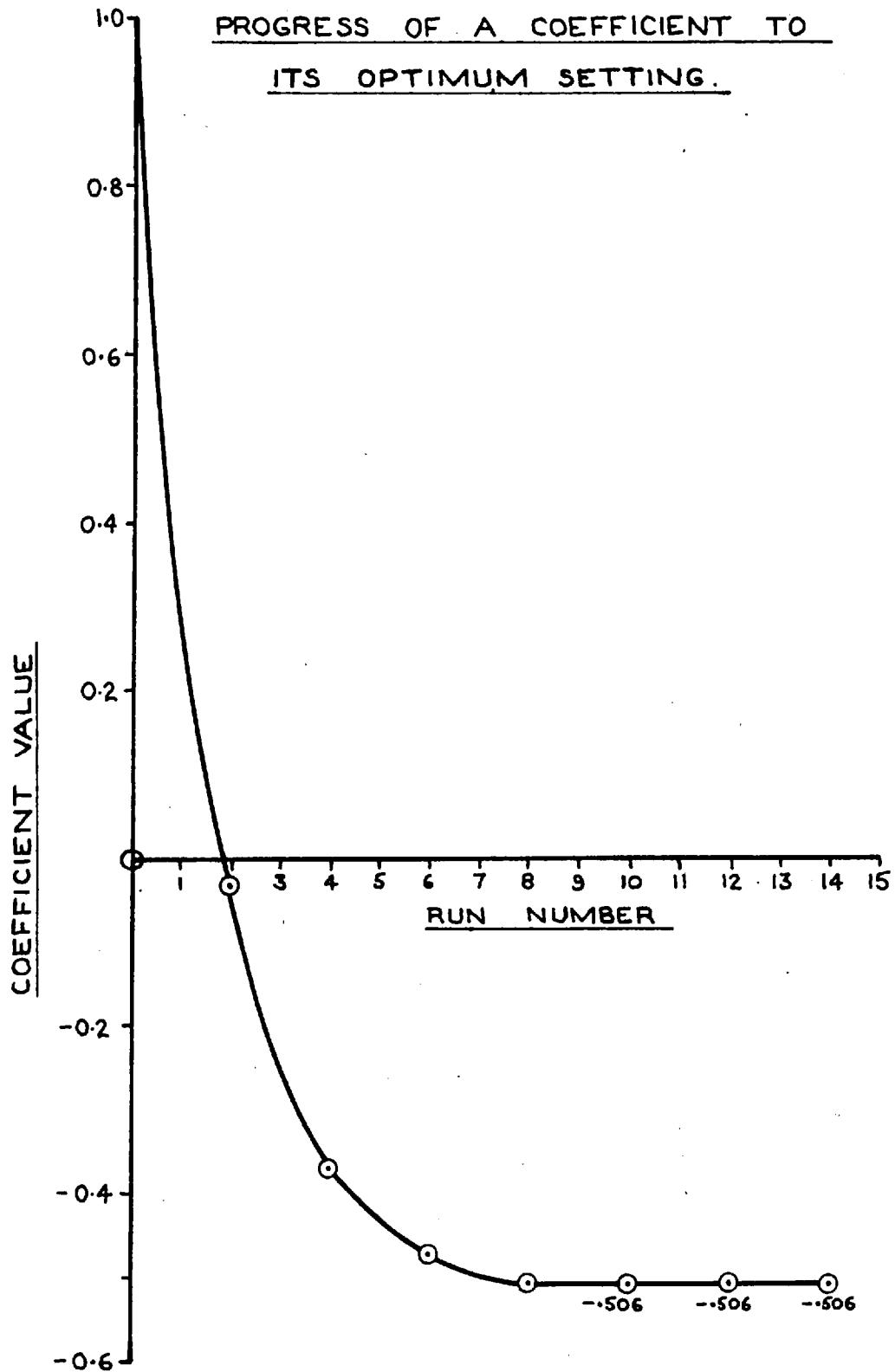


FIG. 71.

PHASE SHIFTING AND SCALING

7.3. Filtering of a Sine Wave, from a Sine Wave plus Noise

For this rather more interesting experiment a 200 c/s sine wave was added to a random noise signal of equal power (band limited to 250 c/s). The noise was generated by passing a small current through a neon tube, and amplifying the fluctuations in the voltage across the tube. Eight coefficients were used and the heads were staggered to give equal delays of 2.5 m.sec. between the channels. The teaching signal was a 200 c/s signal in phase with the signal in the first channel.

In this case the theoretically correct result is known. Suppose we have eight equal sine waves, all in phase, and each with random noise of a given power added to it. If we now add the eight signals together the sine wave amplitude will be increased eight times and the sine wave power will be increased sixty four times. Addition of the noise signals on the other hand, since they are random, increases the noise power by only eight times. The signal to noise ratio is therefore increased by a factor of eight, if we make the eight filter coefficients equal and alternating in sign. We must note here however that in the filter the noise waveforms present are not from independent sources. They are in fact delayed versions of the same noise waveform and if the delays are small there will be correlation between them. Since the rate at which the waveform can change is limited by the bandwidth, if the delays are greater than $1/F$ there will be little correlation. (In fact for noise which has passed through a low pass filter, the autocorrelation function is of the form $(\sin x)/x$ with zeros at

intervals of $1/F$, $2/F$, $3/F$ etc.)³² Since our frequency response is not the ideal rectangular one necessary for analysis it is rather difficult to compute the exact effect of this correlation, but we must expect a gain in signal to noise ratio of rather less than eight times in practice.

The filter in attempting to separate the 200 c/s signal from the noise forms itself into a narrow band filter centred at 200 c/s. The most interesting method of observing the results is therefore to plot the frequency response of the Learning Filter at a number of stages in the learning process, and this has been done in fig. 72. The progress is rapid and beyond the 24th run there is little reduction in the mean square error. However for the sake of interest, and because we expected good resolution in the calculation of the coefficients, due to the arrangement of the teaching apparatus, the experiment was continued up to 328 runs. The frequency response at this point is shown in fig. 73. It will be seen that two subsidiary peaks are prominent and a third is beginning to show itself. We can calculate the ideal frequency response obtained with the coefficients equal and alternating in sign (Appendix 2), and this is plotted in fig. 74, together with the actual curve obtained by setting the coefficients to the theoretical values. The small discrepancy is probably due to errors in the settings of the signal delays. The value of Y_m obtained after 24 runs was 39, corresponding to an improvement in signal to noise ratio of 4.6 times. The value of Y_m obtained with the coefficients set to their theoretical

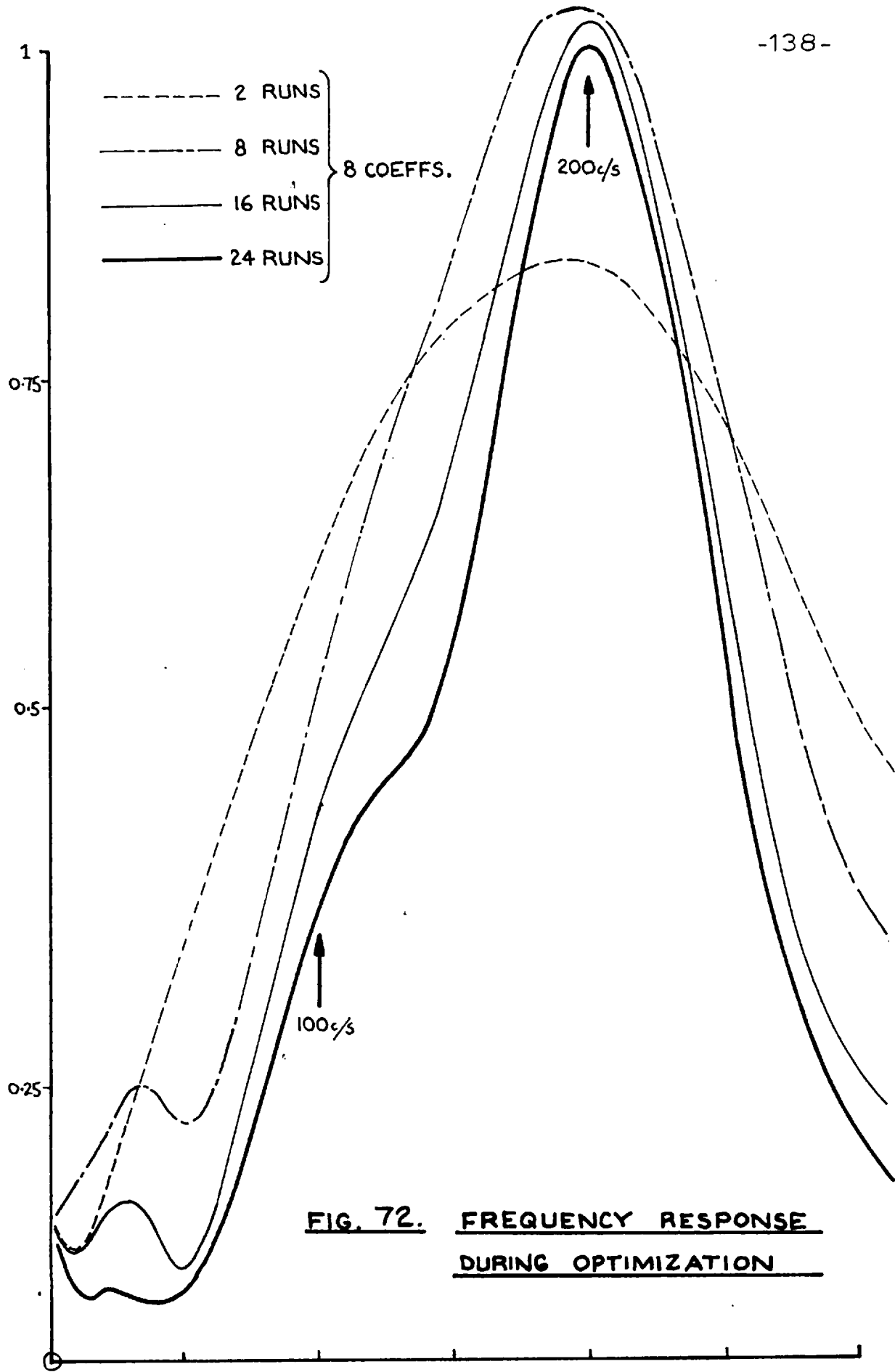
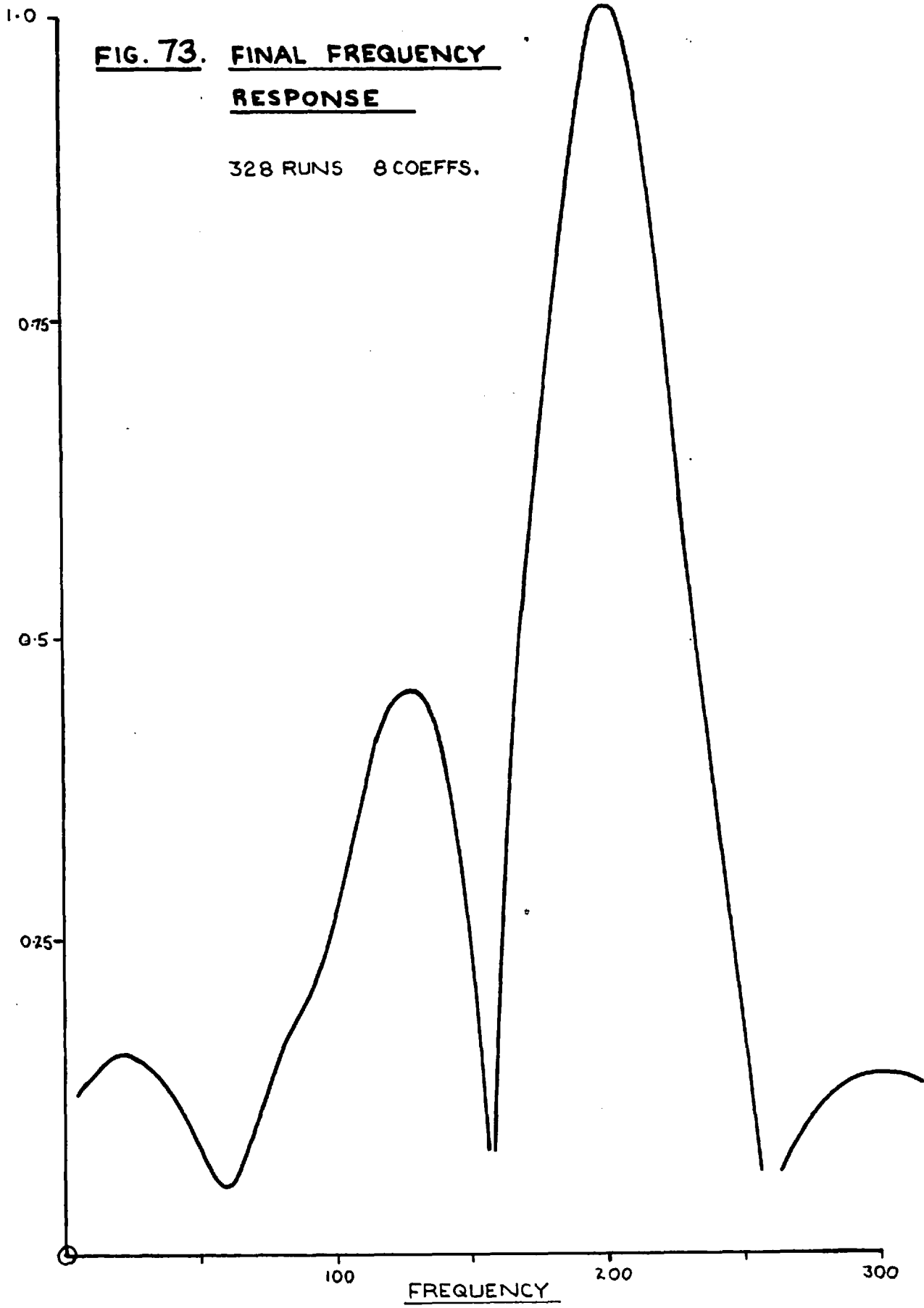


FIG. 72. FREQUENCY RESPONSE DURING OPTIMIZATION

FIG. 73. FINAL FREQUENCY
RESPONSE

328 RUNS 8 COEFFS.



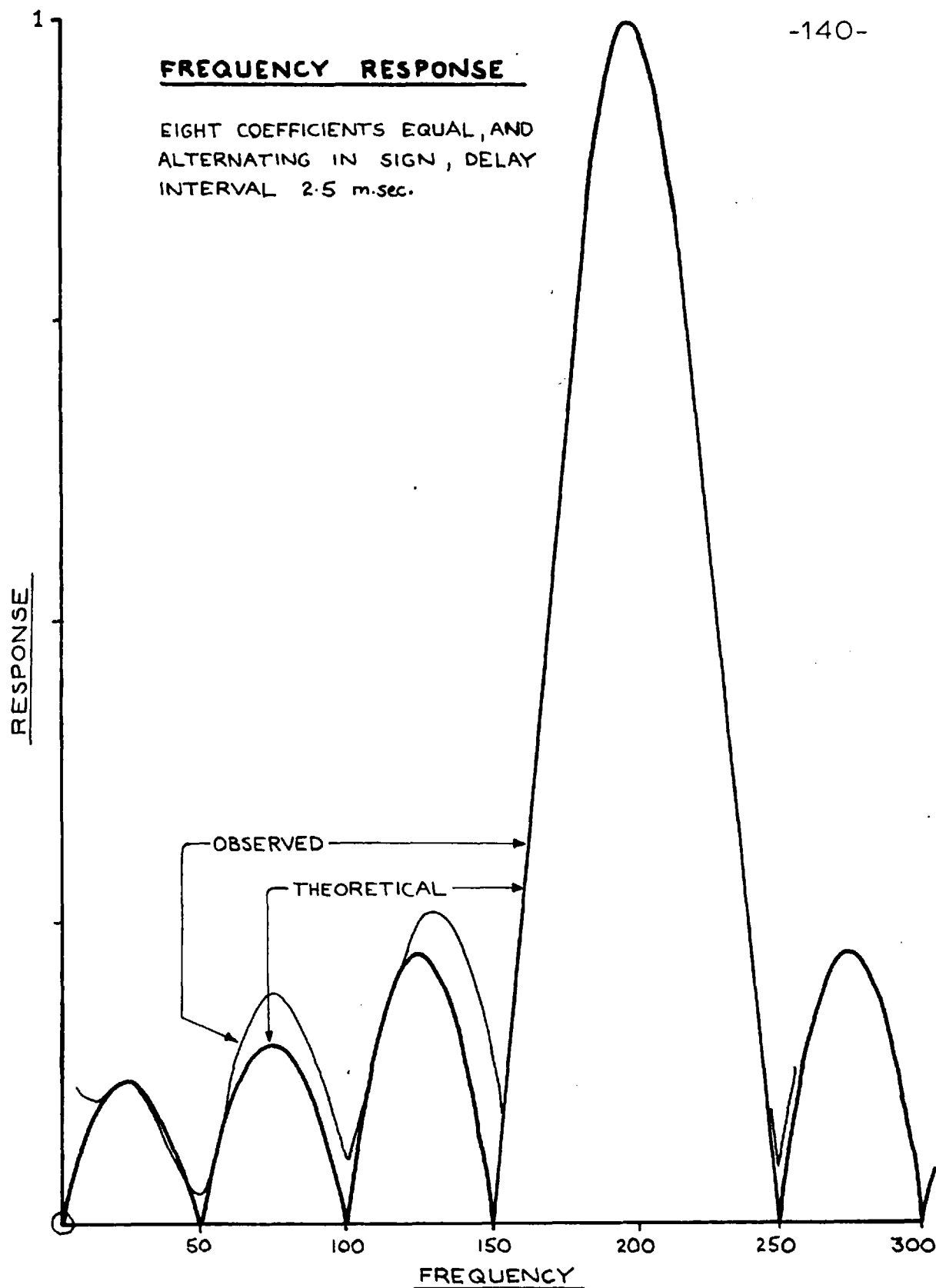


FIG. 74. IDEAL NARROW BAND FILTER

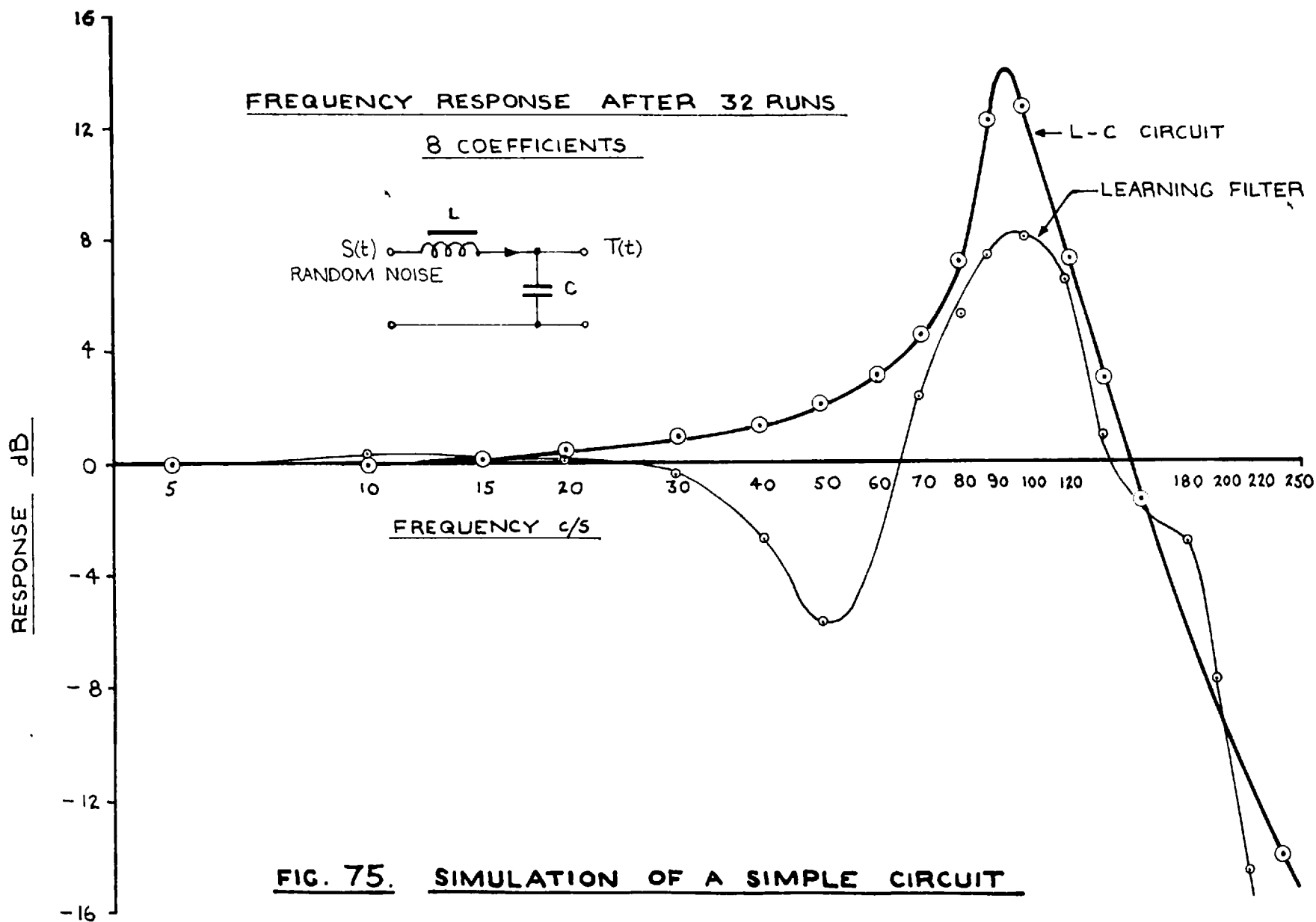
values was 38.2. This emphasises that looking at the frequency response is a very critical way of estimating the filter performance. To get a "filter's eye view" of the problem we should only consider mean square errors. Thus the filter considers the final curve of fig. 72, "just as good" as the curves of figs. 73 & 74. The filter works in the real world and we must not expect it to take too much notice of our idealised theories.

7.4. Simulation.

One of the simplest simulator problems, is to arrange for the machine to imitate a simple circuit of some kind. The L-C circuit of fig. 75 was chosen for simulation, the resonant frequency of the circuit was 100 c/s. The training records for the filter were obtained by feeding random noise into the input of the circuit, and using the noise output from the circuit (modified by the circuit frequency response) as the teaching signal.

The filter then attempts to duplicate the effect produced on the noise by the circuit, using eight coefficients with signal delays of 2.5 m.sec.

The frequency responses of the circuit, and of the Learning Filter after 32 runs, are shown in fig. 75. The experiment was unfortunately marred by the occurrence of severe dropouts caused by creases in the magnetic tape, and it was not taken further than 32 runs. The result is promising as far as it goes and has been included here because of the importance of experiments of this type, and because it



demonstrates that the filter can learn to adjust its coefficients to form an optimum simulator, using information which would be useless to any other kind of machine. It is hoped to try more complicated experiments of this type in the near future.

7.5. Survey of Experimental Work and Present State of the Machine

The results given above are for simple, almost trivial, problems. They were experiments done when first bringing the Learning Filter into operation as a complete system. They serve to demonstrate that the Learning Filter works reliably in the intended manner, with acceptable accuracy.

At present the calculated values of Y_m show sporadic errors of 1 or 2% some ten times greater than the anticipated errors, the cause can no doubt be found and eliminated. Calculated values of C_m are consistent to 0.1 or 0.2% which is acceptable.

The automatic control and adjusting apparatus has worked accurately and very reliably.

An immense amount of experimental work remains to be done, using the Learning Filter as a research tool, and two postgraduate students will be continuing the work for at least another three years.

8. DISCUSSION

8.1. Design

The design of the machine was greatly influenced by what could be borrowed or adapted. The co-operation of the Admiralty enabled us to borrow 128 servo-controlled potentiometers and though these had to be rewired they added considerably to the feasibility of the project. A large number of relays and uniselectors on standard racks were also loaned by the Admiralty, and some of these were stripped down and rewired into the Learning Filter.

8.2. System Performance and Limitations.

The Learning Filter has proved to be fairly reliable so far and it is interesting to try and estimate the long term reliability. The machine contains some 500 valves, and on the basis of data relating to modern valves, we can expect not more than about 5 failures per year. Wherever possible a large amount of negative feedback has been used to permit considerable change in valve characteristics. We can confidently expect 10^6 operations from a typical relay and so these should give little trouble. All the relays are protected by dust covers. The failures so far have amounted to 5 valves, 1 potentiometer, 2 resistors, 3 dry joints, and this suggests that there will be no difficulty in performing long teaching runs without any failures at all.

The accuracy of the machine has been chiefly limited by cost. Thus a computing amplifier stage for errors rather less than 1% can be made, with an open loop gain of 200 and an output impedance of 200 ohms,

using 3 triodes. This also allows the use of 1% high stability resistors. At this sort of accuracy the cost goes up very rapidly as the allowable error goes down, and it would have been unrealistic to aim at errors less than 0.1%.

It is difficult to form a clear estimate of how well the filter will perform, in most practical cases. Some idea of its ability can be gained by considering two simple problems. First consider the prediction of a function of order n . We might expect that given enough terms the filter can predict such a function exactly. It turns out that using n coefficients the filter can exactly predict a function of order $(n-1)$. (Appendix 3). Thus it can exactly predict a parabola using only three of its 18 coefficients. Secondly consider the action of the filter in forming itself into a narrow band filter. With 18 coefficients and a centre frequency of 200 c/s, the response will be zero at 178 c/s and 222 c/s, and the amplitude of the first "side lobe" is 20% of the main peak.

The delay unit has proved successful in providing variable signal delays, but the staggering of the heads is not done in quite the way originally planned. The method adopted is to record a reference sine wave on the magnetic tape, and to display the sine wave from the head which is being staggered, with the sine wave from track 1, on a dual trace oscilloscope. The signal delay can then very readily be made equal to an integral number of half wavelengths of the reference sine wave. This is a closed loop method of adjustment, as opposed to the

vernier scale method of adjustment which is open loop and makes no allowance for zero errors in the positions of the heads. The vernier scales could therefore be dispensed with in a future machine. It would be an advantage to replace them by an adjustment mechanism having a mechanical advantage, since there is considerable friction between the discs of the delay unit which makes fine adjustment difficult.

8.3. Some Alternative Methods

8.3.1. Delay Unit

The present form of the delay unit necessitates the use of rather wide magnetic tape, and since 18 of the 20 tracks are recorded with identical signals this is largely wasted. We can avoid this if we separate the delay unit from the tape recorder. The function of the delay unit is then to accept a single input signal and to produce 18 delayed versions of it. One way of doing this is to use a magnetic drum. We record simultaneously and continuously on 18 tracks on the drum and use staggered pick up heads to playback the signal with different delays. The signal is continuously erased after it has passed by the pick up heads. If necessary the delays can be changed by altering the speed of the drum. Such a unit would also be suitable for "on line" operation where we do not wish to pre-record the signals on magnetic tape.

Alternatively a magnetic disc could be used with the recording heads in a radial line. The smallest delay could then be arranged on the outermost track, thus allowing greater physical separation

between the first pick-up head and the recording heads. ~~Acoustic~~ delay lines of the type developed for digital computers might be considered, thus a mercury delay line in a 5 foot tube gives a delay of about 1 m.sec., but this would not be an economic solution.

8.3.2. Generation of Non-linear Terms

The piezo-magnetic multipliers described in section 5.6. have proved very suitable for use in the Learning Filter, and may well be useful in other applications. It is an advantage that all the parts are readily available, and that the construction can be varied widely to suit particular requirements. Both inputs and the output can be isolated from ground if required, and the multiplier gives a true four-quadrant product. A limitation is that a D.C. output is not directly obtainable, but a D.C. product can be obtained using two such multipliers with an A.C. carrier, which is afterwards filtered out. The simplicity of the unit makes it quite feasible to consider this approach. It should be fairly easy to make a multiplier of this type for operation up to frequencies of 10 Kc/s, and such a unit would be very useful for processing speech and music signals. The frequency response could be further extended by making the multiplier into a closed loop force-balance device. The output of the crystal is then amplified and fed back to the crystal via a linear transducer so as to oppose the original movement. The feedback could be applied using a second crystal cemented to the first, or a polarised magnetic circuit.

The piezo-magnetic realization of the new multiplier is not

the only one possible. Thus magnetic forces can be replaced by electrostatic forces, and other transducers of force to voltage might be considered. Quadratic forces might be obtained using magnetostriction in nickel or electrostriction in unpolarised barium titanate. This would allow a very high frequency solid multiplier to be made, though considerable development would be necessary.

The figure of merit = output voltage x bandwidth / input power for the piezo-magnetic multiplier used in the Learning Filter, is 3000 times the corresponding figure for its most suitable alternative, the Hall multiplier.

In spite of this the cost of the multipliers and their circuits is a significant fraction of the total cost of the machine. It is therefore interesting to consider whether it is possible to generate the non-linear terms in some other way. In specifying the filter operator as a polynomial, with multipliers forming all possible products of the signal samples, we are playing safe and providing a completely general filter operator. In practice however we are more concerned with deciding whether 'non-linearity' effects some improvement. Suppose instead of a multiplier we provide some cheap non-linear element such as a non-linear resistor, or a diode circuit. The signal in the "sample product" channel will now be some other non-linear function of the signal samples. From the theory however it is clear that this does not affect the optimization process. We still have $C_m = \overline{eS} / \overline{S^2}$ and if there is correlation between the filter error

and the signal resulting from our non-linear operation, the filter will use this signal, weighted appropriately, to effect some improvement in its performance. This is an approach which may be worth investigation in the future.

8.3.3. Digital Realization of the Learning Filter

The success of high-speed digital computers has been so phenomenal that it has become almost common practice to write programmes for these machines in preference to building special purpose equipment.

It has also become established practice to transistorize most new equipment. We have chosen to ignore both these methods, but with good reason.

It is somewhat surprising to find that the Learning Filter is considerably faster than a digital computer, but a few figures quickly prove this. Taking figures for a modern machine using a word length of 40 binary digits, we find that addition requires 1 m.sec. and multiplication 2 m.sec. This is for a serial machine with a special fast multiplier unit. The Learning Filter continuously performs some 100 additions and 200 multiplications and to carry out these operations on the digital computer would require 0.5 seconds. This would enable the input signal to be sampled twice per second instead of 400 times per second, so the digital computer is 200 times slower. Even if we sacrifice the greater accuracy and speculate about a digital machine with only 8 binary digits, the analogue machine is still much faster. When it is considered that in some problems the optimization time for the

analogue machine might be twenty hours then it is clear that the cost in digital computer time would be prohibitive.

The use of transistors in a precision analogue machine, is not so advantageous as is their use in a digital machine, where there are no rigorous requirements with regard to stability and distortion. Transistors are definitely not economic for use in the Learning Filter at present and they were even less economic three years ago when the project was commenced. In the future their use will be attractive, if the price of transistors continues to fall as rapidly as at present, and they would permit the construction of a much smaller machine with a low power consumption.

8.4. Future Development of the Existing Machine

The Learning Filter has now been developed to a state where it can successfully and conveniently be used on a great variety of problems. A great deal can be learned by synthesizing problems for the machine to solve, and this can be done using little extra equipment.

Once these initial exploratory experiments have been done it will be desirable to try the machine on a few practical problems, and at this stage the problem of special input equipment will have to be faced. For example we may have tables of figures, punched cards, punched tape, magnetic tape, pen records on paper, or film. When we also consider that there is a wide range of standards within each group, the problem of converting from all these forms into a form suitable for the Learning Filter, is a hopeless one. Thus to take

magnetic tape alone, this may be recorded at any speed from $1\frac{7}{8}$ "/sec. to 200 "/sec. The tape width can vary from $\frac{1}{4}$ " to several inches. The recording may be amplitude modulated, frequency modulated, pulse width, pulse rate, or pulse position modulated. It is clearly out of the question to try and provide general purpose conversion equipment at the input to the Learning Filter. Such an undertaking would require an effort considerably greater than that expended in constructing the Filter itself. Consequently any input equipment provided must be well chosen and limited in scope. Those forms of recording suitable for replay, will no doubt be found in conjunction with replay equipment belonging to other organizations, and it should be possible to borrow or hire time on this equipment to play back the records and re-record them on a portable tape recorder. If it is found that considerable useful data is in the form of pen records then it may be worthwhile to install special replay equipment. This will have to be considered carefully however because at present there is every indication that there will be no lack of problems for the machine. It may be more economic to repeat some experiments yielding useful data, than to consider translating previously recorded data from an unsuitable form.

Considerable time has already been spent on design and construction and what is required most now that the machine is working, is a reasoned programme for investigating the capabilities of the machine, on a sound experimental basis. It is probable that until this has been done no sound decisions can be made regarding the form

of further extensions to the Learning Filter.

8.5. The Use of Learning Filters

In the seventeen years which have elapsed since the original work of Kolmogoroff and Wiener on the optimum linear filter and predictor, their work has inspired a tremendous effort into the theoretical task of obtaining the corresponding non-linear device.

In spite of this mathematical effort and the resulting flood of literature on the subject, it is astonishingly hard to find significant practical results, where non-linearity is deliberately introduced into a system (except simple diode circuits, on-off servo-mechanisms and the like). This is of course a *raison d'être* of the Learning Filter, and it is to be hoped that the machine will do something to redress the balance.

The Filter should be capable of tackling two problems, of great current interest, the simulator problem (suggested by Professor A. Tustin) and the optimum control system problem. The simulation of an unknown system, which the Filter can perform using only the data contained in the normal input and output records of the system, may be very useful. It is an advantage that it can do this without a human operator having to try and puzzle out what the mathematical equations are which adequately represent the system (which in some cases is impossible anyway). Also when a system analogue is constructed on a conventional analogue computer, the results obtained do not refer to the real system, they refer to what has been wired up on the analogue

computer. There is always a danger that such results may be misleading, due to incorrect formulation of the equations, or non-valid approximations. There is no such danger in the case of the Learning Filter since it is in direct touch with the real world, and its internal structure is directly modified by its environment.

As a simple example of the application of the Learning Filter to control systems, consider its use in optimising a straight follower with prediction. The Learning Filter processes the input signal, and the processed signal is fed into the control system. During the teaching period the output of the system is compared with an advanced value of the input signal and the Filter tries to minimize the mean square error between the two. In doing so it improves the control system as much as it can, if necessary it can compensate for errors in the control system. Again we have the great advantage that we need never know what the actual transfer function is, learning has taken over from design in this respect.

In some control system applications it would be an advantage if the machine could learn continuously, this would enable it to adapt and compensate for slow changes in the system (e.g. the lightening of a ship as its fuel is burnt). This can be arranged if instead of training the machine on one long sample of the input-output records, which is used over and over again, we train it on successive sufficiently long samples of the real input and output. The Learning Filter is then used "on line" instead of "off line" during the training period (for

prediction training we can compare the live input with a delayed version of the output). A "sufficiently long sample" is used in the statistical sense and it may be wondered how we can be sure that the sample is long enough. It is known that if the Filter is operating on a stationary time series, then the coefficients will be constants. All we need do therefore is to observe the calculated value of C_m as a function of time. As time goes on C_m should converge to a steady value and when it does so we can conclude that the sample is long enough. This was well demonstrated during the sine wave plus noise experiment (section 7.3.) for which the run length was chosen to be 100 seconds. A particular calculated C_m is plotted as a function of run time in fig.76, and it can be seen that 100 seconds is adequate.

On-line operation makes it undesirable to switch the potentiometer on which the coefficient is set up, out of circuit, when it has to be adjusted, as is done in the present machine. This could be overcome by using double ganged potentiometers. Alternatively a special stand-by potentiometer might be switched in to replace the one being adjusted.

One application of the machine as an optimum filter, is to the problem of the reception of radar echoes, which are generally heavily overlain with noise. The machine can achieve "correlation reception" using its linear part only, and it will be interesting to see what improvement is effected by the non-linear terms (in a dummy experiment at low speed).

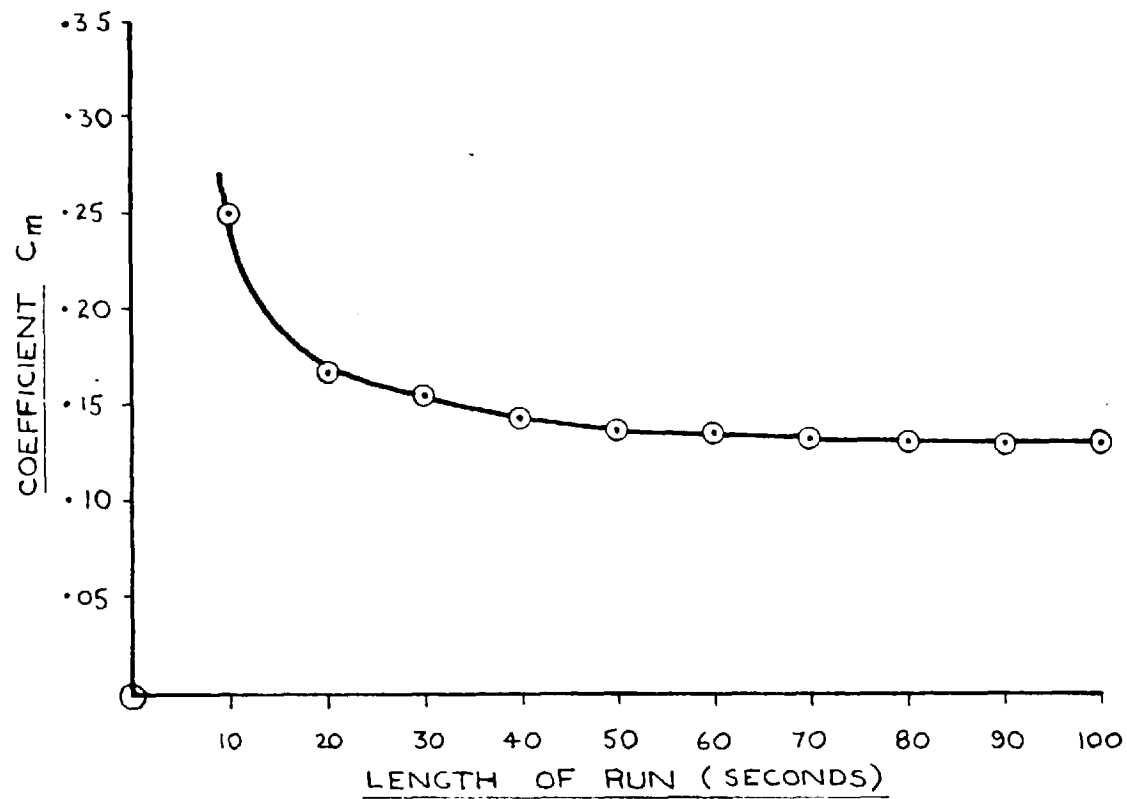


FIG. 76.

CONVERGENCE OF C_m

The Learning Filter is most satisfactory as a practical machine in that it designs itself automatically and accurately using only raw data. We do not even have to set up initial conditions on the coefficients and if we do so the machine will rapidly correct our (probably misguided) attempts. An occasional error is not fatal since its effect will be removed by the learning process. Complete loss of part of the operator polynomial would probably not be catastrophic, though it would impair the performance to some extent, and in these respects the machine duplicates some of the many desirable features of the human brain.

If the machine is to be widely applicable it should not be too costly or too complex. This should not be a severe limitation if we restrict ourselves to functional copies which do not themselves learn. In this case the whole of the teaching apparatus is dispensed with and the servo-controlled potentiometers can be replaced by fixed resistors. If simple non-linear elements can be used instead of multipliers, very little remains apart from one adder and the delay unit. In some linear cases it will be more economic to dispense with this realization and to synthesize L-C-R networks which will give an equivalent effect. This can be done given a knowledge of the coefficients and the delays.

So far we have only discussed a machine having a single input and a single output, but in many practical cases there will be numerous input signals, all contributing to the resulting output. The

machine has been designed to allow problems having two inputs to be tried out, and there is no difficulty in principle in the multi-input case. Another possibility might be a multi-input, N-output machine, though this can always be considered to be N multi-input, single output machines in parallel. We need not go so far as this however, since the machines could have common delay units and sample product channels, with a complete set of coefficients and an adder for each output channel.

Such a machine if it could be made sufficiently complex might be able to help with the prediction of such notoriously unpredictable quantities as the weather or the stock exchange, though it may be unrealistic to consider these at present.

8.6. Status of the Learning Filter as an " Intelligent Machine."

There has been considerable interest in the general problem of making machines "more intelligent "³³ and many machines have been built exhibiting desirable properties of one sort or another.³⁴ From the work that has been done so far a reasonably clear pattern is beginning to emerge, and although it is out of the question to discuss all this work in detail, some general conclusions can be drawn. From studies in such fields as pattern recognition and speech recognition, it is clear that it is relatively easy to evaluate some principles of operation, and to build simple machines which can recognize primitive patterns and simple sounds. However when one tries to scale up these machines to deal with situations which the human brain finds commonplace, the

complexity and size of the machine tends to increase beyond all reasonable bounds. The same will probably be true of the Learning Filter if we try to emulate some of the remarkable filtering achievements of the human brain. Again it is easy to build equipment which can simulate the conditioned reflex behaviour of living organisms, in a realistic way, but if we try to use this approach to synthesize complex behaviour patterns the amount of equipment required gets unreasonably large. Another approach has been to write "learning programmes" for digital computers, in an attempt to make up with speed what we lose in complexity. Surprisingly little can be gained like this, though if it proves possible to make computers work at microwave frequencies they will take us a long way.

We can conclude that in terms of required complexity and physical size we are at present falling short by many orders of magnitude of what is required to emulate the human brain. We can also conclude that this being so the major part of the "intelligence" in our machines will have to be supplied directly or indirectly by human brains, for many years to come. In any case as a practical measure we do not wish to replace the human brain, merely to assist it. This means that we should concentrate on special purpose machines. One of the main attributes of man, which has ensured his dominance, is this use of special purpose machines. If we wish to chop down a tree we use an axe and do not adopt the solution of the beaver, assiduous gnawing for a few million years, though no doubt we should be equally well equipped

at the end of it. In other words we have remained general purpose, and we have collected a vast range of special purpose attachments instead of adapting our own physical form. All successful machines compensate for some deficiency in our own physical make up. The same is true of our modern data processing machines, and their success is more due to this than to any intelligence we have managed to endow them with. When the brain is used for processing data it suffers from a number of defects. Common complaints are "not too fast, one thing at a time, can't remember," and on such points machines can score heavily. Clear appreciation of this point is handicapped by misleading nomenclature (and sometimes by extravagant claims of "intelligence"). Thus if we talk of a "universal" digital computer we may just as well talk of a "universal" telescope, "which can look at anything." The computer is a striking example of a special purpose machine, for doing arithmetic, and it is a most successful one. The main attributes of the Learning Filter are patience, an infallible memory, and an ability to work with data in rather unpromising forms. We might also say that it has a mania for improving itself, which is something not found in all human brains.

A successful machine then, should be fitted to the task it has to perform; as much a priori knowledge should be built in as possible, and it should not be more general purpose or adaptive than it has to be, otherwise it is merely inefficient.

On the broader question of whether a "super intelligent" general purpose machine is a possibility, we can only say that the

machine would be faced with so many conflicting requirements that it would probably end up showing the same mixture of imprecision and indecision as the human brain.

9 REFERENCES

- 1 WIENER, N. The Extrapolation, Interpolation, and Smoothing of Stationary Time Series.(circulated in report form 1943), John Wiley, New York, 1949.
- 2 KOLMOGOROFF, 'A. Interpolation and Extrapolation of Stationary Series, Bull. Acad. Sci. U.S.S.R. Ser. Math. 2, 3 1942.
- 3 WIENER, N. Response of a Nonlinear Device to Noise, Report No. 129, Radiation Laboratory, M.I.T. 1942
- WIENER, N. Nonlinear Problems in Random Theory, John Wiley, New York, 1958
- BOSE, A.G. A Theory of Nonlinear Systems, Report No. 309, Res. Lab. of Electronics, M.I.T., May, 1956.
- ZADEH, L.A. Optimum Nonlinear Filters, J. Appl. Phys., 24, 396-404, 1953.
- ZADEH, L.A. & An Extension of Wiener's Theory of Prediction, J. App. Phys. 21, 7, 645-55, July 1950.
- RAGAZZINI, J.R. Optimum Filters for the Detection of Signals in Noise, Proc. I.R.E. 40, 1223-31, Oct., 1952.
- SINGLETON, H.E. Nonlinear Systems with Sampled Input, Report No. 160, Res. Lab. of Electronics, M.I.T., 1951.
- WHITE, W.D. I.R.E. Convention 1951, summary Proc. 2, 505, Feb., 1954.
- BRILLIANT, M.B. Theory of the Analysis of Nonlinear Systems, Tech. Rep. 345 Res. Lab. of Electronics, M.I.T., March, 1958.

- 4 GABOR, D. Communication Theory and Cybernetics, I.R.E. Trans. on Circuit Theory, CT-1 4 19-31 Dec.1954.
- 5 SHANNON, C.E. Communication in the Presence of Noise. Proc. I.R.E. 37, 10, 1949.
- 6 KUHRT, F. Properties of Hall Generators, Siemens Z. 28, 370-6, 1954.

HARTL, W. Application of Hall Generators, Siemens Z. 23, 376, 1954.
- 7 SOUTHWELL, R.V. Relaxation Methods, Oxford Univ. Press, 1946.

TEMPLE, G. General Theory of Relaxation Methods applied to Linear Systems. Proc. Roy. Soc. Vol.CXLIX, 476, 1939.
- 8 ERATH, L.W. & SMITH, F.C. A Slope Modulator for F.M. Recording, I.R.E. Trans. PGTRC-2, 20-24, Nov. 1954.
- 9 EARNSHAW, J.B. Diode Pump Integrator, Electronic Engng. 28, 26-30, Jan. 1956.
- 10 THIELE, The Design of Filters using only R-C sections and Gain Stages, Electronic Engng. 28, 31-6 Jan, 30-2 Feb, 1956.

LINVILL, R-C Active Filters, Proc. I.R.E. 42, 555-64 March, 1954

SHUMARD, Design of H.P., J.P., & B.P., Filters using R-C Networks and D.C. Amplifiers with Feedback, R.C.A. Review, 11, 534, 1950.
- 11 WASS, C.A. Introduction to Electronic Analogue Computers, Pergamon Science Series, 1955.
- 12 DEELEY, E.M. & MacKAY, D.M. Multiplication and Division by Electronic Analogue Methods, Nature 163, 4147 p650 Apr.1949.

DEELEY, E.M. The Design of an Electrodynamic Multiplier, Proc. I.E.E., IV, 101, 187-91, 1954.
- 13 KUHRT, F. Properties and Applications of Hall Generators, V.D.E. Fachber, 19(1) 1-8, 1956. (see also Ref.6)

LONGREN, L. An Analogue Multiplier based on the Hall Effect, Jl. of Appl. Phys. 29, 2, 158-66 Feb. 1958.

- HILSUM, C. Multiplication by Semiconductors, Electronic Engng. 30, 369, p664, Nov. 1958.
- CHASMAR, R.P. & An Electric Multiplier Utilizing the Hall Effect
COHEN, E. in Indium Arsenide. Electronic Engng. 30, 369, p661, Nov. 1958.
- GRUBBS, W.J. Hall Effect Devices, Bell System Tech. Jl. 38, 3, 853-76, May, 1959.
- 14 SAVANT, C.J. Multiplier for Analogue Computers, Electronics, 27, 9, 144-7, Sept. 1954.
- 15 STERNBERG, An Accurate Electronic Multiplier, R.C.A. Rev. 16, 4, 618-34, Dec. 1955.
- 16 FREEMAN, & High Speed Time Sharing Analogue Multiplier,
PARSONS, Trans. I.R.E. EC-3, 11-17, March 1954.
- 17 CZAJKOWSKI, Z. Electronic Methods of Analogue Multiplication, Electronic Engng. 28, 342, July & August, 1956.
- 18 MEHRON, & Instantaneous Multipliers for Computers,
OTTO, Electronics, 27, 2, 144-8, Feb. 1954.
- MEYER & Two New Electronic Analogue Multipliers, Rev.
FULLER Sci. Instr. 25, 12, 1166-73, Dec. 1954.
- 19 Mc.COOL, An A.M.-F.M. Electronic Analogue Multiplier, Proc. I.R.E. 41, 1470-7, Oct. 1953.
- 20 GOLDBERG, E.A. Step Multiplier in Guided Missile Computer. Electronics, 24, 120, 1951.
- 21 HALL, A.C. A Generalized Analogue Computer for Flight Simulation, Trans. Am. I.E.E. 69, 303, 1950.
- 22 DAVIDSON, G., Subminiature Thermal Computer Element,
DJINIS, W., & Electronic Equip. 3, 1955.
SAVET, P.
- 23 FINZI & An Application of Magnetic Amplifier Circuits to
MATHIAS, Perform Multiplication & other Analytical Operations, Comm. & Electronics, 8, 455-61, Sept. 1953.
- 24 KOVACH, L.D., & An Analog Multiplier using Thyrite, Trans. I.R.E. on Elect. Comp. 2, 42-5, June, 1954.
COMLEY, W.

- 25 ROSS, H., & SHUFFREY An Electronic Square Law Circuit, J.Sci.Instr., 25, 200, 1948.
- EL SAID Novel Multiplying Circuits with Application to Electronic Wattmeters, Proc. I.R.E., 37, 1003, 1949.
- SYDNOR et.al. Analog Multipliers and Squarers using a Multi-grid Modulator, I.R.E. Trans., EC-5, 2, 82-5, June, 1956.
- 26 DRAPER, J.H.P., & TUCKER, D.G. A Square Law Circuit, J. Sci. Instr., 24, 257, 1947.
- 27 MILLER, SOLTES, & SCOTT Wide Band Analogue Function Multiplier, Electronics, 28, 2, 160-3, Feb., 1955.
- 28 MARSHALL An Analogue Multiplier, Nature, 167, 29, 1951.
- CHANCE et.al. Quarter Squares Multiplier using Segmented Parabolic Characteristics, Rev. Sci. Instr., 22, 683-8, Sept., 1951.
- FISHER, M.E. A Wide Band Analogue Multiplier using Crystal Diodes, Electronic Engng., Dec., 1957.
- 29 NORSWORTHY, K.H. A Simple Electronic Multiplier, Electronic Engng, 26, 312, 72-5, Feb., 1954.
- MEYER, & DAVIES Triangular Wave Analog Multiplier, Electronics, 29, 8, 182-5, August, 1956.
- 30 ANGELO, E.J. An Electron Beam Tube for Analog Multiplication, Rev. Sci. Instr., 25, 280-4, March, 1954.
- 31 ATTREE, V.H. A Cascode Amplifier Degenerative Stabilizer, Electronic Engng., 27, 175, April, 1955.
- 32 SCHWARTZ, M. Information Transmission, Modulation and Noise, p.434, Mc.Graw Hill, 1959.
- 33 WIENER, N. Cybernetics, J.Wiley, New York, 1948.
- TURING, A.M. Computing Machinery and Intelligence, "Mind", 59, 433-60, 1950.
- NEUMANI, J.v. The General and Logical Theory of Automata, J.Wiley, New York, 1951.

- MacKAY, D.M. Mindlike Behaviour in Artefacts, Brit. J. Phil. Sci., 2, 105-21, 315-17, 1951.
- BOX, G.E.P., & WILSON, K.B. On the Experimental Attainment of Optimum Conditions, J. Roy. Statistical Soc., 13 B, 1, 1951.
- YOUNG, J.Z. Doubt and Certainty in Science, Oxford Univ. Press, 1951, (Reith Lectures 1950).
- SHANNON, C.E. Computers and Automa, Proc. I.R.E., 41, 1234-41, 1953.
- UTTLEY, A.M. The Classification of Signals in the Nervous System, E.E.G.Clin. Neurophysiol., 6, 479, 1954. Conditional Probability as a Principle of Learning, Int. Cong. on Cybernetics, Namur, Pub. Gauthier-Villars, Paris, 1956.
- SHANNON, C.E. & McCARTHY, J. Automata Studies, Princetown Univ. Press, Princetown, New Jersey, 1956.
- BOX, G.E.P. Evolutionary Operation: a Method for Increasing Industrial Productivity, Applied Statistics, 4, 81, 1957.
- NEUMANN, J.v. The Computer and the Brain, Yale Univ. Press, New Haven, 1958.
- 34 ASHBY, W.ROSS. Design for a Brain, Chapman & Hall, London, 1952.
- SHANNON, C.E. Programming a Computer for Playing Chess, Phil. Mag., 41, 256 - 75, 1950. Presentation of a Maze-Solving Machine, Trans. VIII, Cybernetics Conf., p. 173-80, New York, 1952.
- OETTINGER, A. Programming a Digital Computer to Learn, Phil. Mag., 43, 1243-63, 1952.
- WALTER, W.GREY. The Living Brain, Duckworth, London, 1953.
- TAYLOR, W.K. The Electrical Simulation of some Nervous Function Activities, IIIrd. Information Theory Conf., London, Sept., 1955.
- KALMAN, R.E. Design of a Self Optimising Control System, Trans. A.S.M.E., p. 468, Feb., 1958.

- UTTLEY, A.M. The Design of Conditional Probability Computers;
Information and Control, U.S.A., 1958.
- FRIEDBERG, R.M. A Learning Machine, I.B.M. Journal, 2, 2, 1958,
(computer programme).
- GRIMSDALE, R.L., A System for the Automatic Recognition of Pat-
SUMNER, F.H., terns, Proc. I.E.E. B, 106, p. 210, March, 1959,
TUNIS, C.J., & (computer programme).
KILBURN, T.

ACKNOWLEDGEMENTS

I wish to thank Professor D.Gabor F.R.S. for his constant enthusiasm, encouragement, and supervision; and the Admiralty for generously supporting this project. I must also thank Mr.J.Anderson, C.B.E., (Chief Scientist, Admiralty Signal and Radar Establishment, Portsmouth) for his interest in the work.

I have also greatly profited from the advice and continued interest of Mr.R.Benjamin (A.S.R.E. Portsmouth), and Professor A. Tustin (Imperial College).

My thanks are also due to the workshop staff of Imperial College for their aid on matters of detail, in particular to Mr.J.C. Ganley and Mr.C.Mills. I must also thank Mr.J.Vaudrey for his reliable constructional work on some of the new apparatus.

APPENDIX 1

It is known (Sect. 2) that the mean square error y , as a function of any one of the filter coefficients C , is a parabola. We have available three values of Y ; y_{-1} , y_0 , y_{+1} , corresponding to coefficient values of -1 , 0 , and $+1$. We may write the general equation of the parabola as:-

$$Y = aC^2 + bC + d$$

where a , b , and d are constants. We require the value of C which gives a minimum value of Y .

By substitution:-

$$y_0 = d$$

$$y_{-1} = a - b + y_0$$

$$y_{+1} = a + b + y_0$$

$$\text{Therefore } 2a = y_{+1} + y_{-1} - 2y_0$$

$$\text{and } -2b = y_{-1} - y_{+1}$$

For a minimum:-

$$dY/dC = 2aC + b = 0$$

$$\text{i.e. } C_m = -b/2a = \frac{1}{2} \frac{y_{-1} - y_{+1}}{y_{+1} + y_{-1} - 2y_0}$$

The corresponding value of Y is given by:-

$$\begin{aligned} Y_m &= aC_m^2 + bC_m + y_0 \\ &= y_0 + C_m(aC_m + b) \\ &= y_0 + \frac{1}{4}C_m(y_{-1} - y_{+1}) - \frac{1}{8}C_m(y_{-1} - y_{+1}) \end{aligned}$$

$$\text{i.e.} \quad \underline{Y_m = y_o - \frac{1}{4}C_m(y_{-1} - y_{+1})}$$

If instead of knowing y_{-1} , y_o , and y_{+1} , we know y_o , and D_{-1} and D_{+1} , the amounts by which y_{-1} and y_{+1} differ from y_o :-

Then substituting

$$y_{-1} = y_o + D_{-1} \text{ we have } 2a = (D_{-1} + D_{+1})$$

$$\text{and } y_{+1} = y_o + D_{+1} \text{ we have } 2b = -(D_{-1} - D_{+1})$$

Which gives:-

$$C_m = \frac{1}{2} \frac{D_{-1} - D_{+1}}{D_{-1} + D_{+1}}$$

$$\text{and } \underline{Y_m = y_o - \frac{1}{4}C_m(D_{-1} - D_{+1})}$$

We also wish to know how much the mean square error is reduced when the coefficient is altered from its previous value C'_m to its computed value C_m . That is we require $(Y_{m(n-1)} - Y_{m_n})$ at the n^{th} run.

$$\text{Now } Y_{m(n-1)} = \frac{1}{3}(D_{-1} + D_{+1})C_m'^2 - \frac{1}{3}(D_{-1} - D_{+1})C_m' + y_o$$

$$\text{and } Y_{m_n} = y_o - \frac{1}{4}C_m(D_{-1} - D_{+1})$$

Therefore:-

$$\begin{aligned} (Y_{m(n-1)} - Y_{m_n}) &= \frac{1}{3}(D_{-1} + D_{+1})C_m'^2 - \frac{1}{3}(D_{-1} - D_{+1})C_m' + \frac{1}{4}C_m(D_{-1} - D_{+1}) \\ &= \frac{1}{3}(D_{-1} + D_{+1})C_m'^2 - (D_{-1} + D_{+1})C_m C_m' + \frac{1}{3}(D_{-1} + D_{+1})C_m^2 \\ &= \underline{\underline{\frac{1}{3}(D_{-1} + D_{+1})(C_m' - C_m)^2}}} \end{aligned}$$

APPENDIX 2 Narrowband filter obtained by adding delayed samples of
a signal, with equal coefficients and delays

Let the input to the filter be $u(t)$ and let the sampling interval be T . Then the output $v(t)$ is given by:-

$$v(t) = u(t) + u(t-T) + u(t-2T) + \dots + u(t-nT)$$

For 8 equal coefficients, alternating in sign, and taking the Laplace Transform:-

$$V(p) = U(p)(1 - e^{-pT} + e^{-2pT} - \dots - e^{-7pT})$$

Therefore the filter operator:-

$$F(p) = (1 + j^2 e^{-pT} + j^4 e^{-2pT} - \dots + j^{14} e^{-7pT})$$

And, summing the series:-

$$\begin{aligned} F(p) &= \frac{1 - j^{16} e^{-8pT}}{1 - j^2 e^{-pT}} \\ &= \frac{1 - e^{-8pT}}{1 + e^{-pT}} \end{aligned}$$

The frequency response $F(j\omega)$ is obtained by putting $p = j\omega$ in $F(p)$

$$\text{i.e.} \quad F(j\omega) = \frac{1 - e^{-8j\omega T}}{1 + e^{-j\omega T}}$$

APPENDIX 3 Prediction of a Function of Order n

We make use of the fact that for a function of order n, the n^{th} difference D_n of the signal samples is a constant.

e.g. for a quadratic function we have:-

$$\begin{array}{ccccccc} S_2 & & S_1 & & S_0 & \dots & 3S_0 - 3S_1 + S_2 \\ D_1 & S_1 - S_2 & & S_0 - S_1 & & \dots & 2S_0 - 3S_1 + S_2 \\ D_2 & & S_0 - 2S_1 + S_2 & & & \dots & \end{array}$$

The terms on the right of the dotted line have been constructed by working back from D_2 . We therefore have, predicting one sampling interval into the future:-

$$S_{-1} = 3S_0 - 3S_1 + S_2$$

We can obtain similar expressions for greater distances into the future.

$$\begin{array}{lcl} \text{e.g.} & S_{-2} & = 6S_0 - 8S_1 + 3S_2 \\ & S_{-3} & = 10S_0 - 15S_1 + 6S_2 \end{array}$$

Following the above procedure for cubic and quartic functions we obtain:-

$$\begin{array}{lcl} S_{-1} & = & 4S_0 - 6S_1 + 4S_2 - S_3 \quad \text{cubic} \\ \text{and} & S_{-1} & = 5S_0 - 10S_1 + 10S_2 - 5S_3 + S_4 \quad \text{quartic} \end{array}$$

The pattern of the coefficients is now clear, and we can infer that for a function of order n :-

$$\begin{aligned} S_{-1} = & (n+1)S_0 - \frac{(n+1)n}{2!}S_1 + \frac{(n+1)n(n-1)}{3!}S_2 \\ & - \frac{(n+1)n(n-1)(n-2)}{4!}S_3 + \dots + S_n \end{aligned}$$

Thus $n+1$ linear terms are required to predict a function of order n.