

THE DESIGN OF LINEAR SERVO-MECHANISMS HAVING PRESCRIBED TRANSIENT
 RESPONSES WITH AN EXPERIMENTAL INVESTIGATION INTO THE PREDICTION
 OF TRANSIENT RESPONSE FROM FREQUENCY RESPONSE.

PART I.

REVIEW OF BASIC THEORY.

	Page No.
Chapter 1. Introduction.	1
2. Transient Analysis of Linear Systems.	7
3. Frequency Analysis of Linear Systems.	25
4. Stability Criteria.	38
5. Criteria of Relative Stability. Design Procedures.	57

PART II.

THE DESIGN OF LINEAR SERVO-MECHANISMS HAVING PRESCRIBED
 TRANSIENT RESPONSES.

Introduction.	79
Chapter 6. The Design of $\frac{\theta_o(\rho)}{\theta_i}$ Transfer Functions.	80
7. Numerical Work Illustrating the Use of the $\frac{\theta_o(\rho)}{\theta_i}$ Design Charts.	107
8. The Design of $\frac{\epsilon(\rho)}{\theta_i}$ Transfer Functions.	122
9. Use of the $\frac{\epsilon(\rho)}{\theta_i}$ Design Charts. Comparison of $\frac{\theta_o(\rho)}{\theta_i}$ and $\frac{\epsilon(\rho)}{\theta_i}$ Methods.	133

ProQuest Number: 13838566

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



ProQuest 13838566

Published by ProQuest LLC (2019). Copyright of the Dissertation is held by the Author.

All rights reserved.

This work is protected against unauthorized copying under Title 17, United States Code
Microform Edition © ProQuest LLC.

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 – 1346

PART III.

EXPERIMENTAL INVESTIGATION OF TRANSIENT AND FREQUENCY RESPONSES
OF METADYNE SERVO-MECHANISM.

	Page No.
Introduction.	145
Chapter 10. Description of Apparatus and Methods of Testing.	148
11. Experimental Step and Frequency Responses. Step Responses calculated from Frequency Responses.	167
12. Discussion of Results.	179
13. Conclusion. Further Work.	202

BIBLIOGRAPHY.

- Appendix I. The Exponential Fourier Series for a Periodic Function.
- Appendix II. Tables to Assist in the Calculation of Nyquist Diagrams at Real and Complex Frequencies.
- Appendix III. Tables of $\frac{1}{n} \sin n\theta$ and $\frac{1}{n} \cos n\theta$.
- Appendix IV. Harmonic Analysis Method of obtaining Step Response from Frequency Response.

The Design Charts pertaining to Part II of the Thesis are
separately bound.

THE DESIGN OF LINEAR SERVO-MECHANISMS HAVING PRESCRIBED TRANSIENT RESPONSES,
WITH AN EXPERIMENTAL INVESTIGATION INTO THE PREDICTION OF TRANSIENT RESPONSE
FROM FREQUENCY RESPONSE.

Introductory Note.

Towards the end of 1946, considerable variety existed in the technical literature on the subject of servo-mechanisms. With the intention of reviewing this literature and thus obtaining guidance for future work, the author, in October 1946, began a programme of study and research in the Electrical Engineering Department. Shortly after this, the volume of published work on servo-mechanisms released at Conferences in the post-war years, and appearing also in text-book form, increased to an extent which made it quite clear that the review originally planned could not be carried out in a work of this nature. The author, therefore, turned to the more restricted problems of the design of servo-mechanisms with specified transient responses and the correlation of frequency and transient response in an experimental system. At the same time, it was thought that these might be prefaced by a review of the basic theory of the subject, and more particularly of the relationship of relative stability to the transient response.

The Thesis, therefore, is divided into three Parts. Part I contains a review of the basic theory. Chapters 4 and 5 of this Part give a full discussion of stability and relative stability not to be found elsewhere in the literature of the subject, and some points, which it is believed, are

not so well-known. Part II deals with the specific problem of designing for a prescribed overshoot in the step-function response. The whole of this Part is original and follows up the idea, due to Campbell, of the principal mode representation of the error quantity. Part III gives the description and results of an experimental investigation into the transient and frequency responses of a metadyne-controlled servo-mechanism.

The work was carried out during the period October 1946 to October 1951 in the Electrical Engineering Laboratory. The author wishes to thank Professor Bernard Hague, D.Sc., for his permission to use the equipment contained therein and for his encouragement to the author while in pursuance of the work. The author also wishes to acknowledge the benefit derived from a short visit made in 1948 to the Electrical Engineering Department, University of Birmingham and to thank Professor A. Tustin for kindly permitting this.

PART I.

REVIEW OF BASIC THEORY.

CHAPTER I.

INTRODUCTION.

Part I of this thesis provides a review of the basic theory governing the operation and design of servo-mechanisms. Following a general description of the closed-sequence type of control, the development of servo-mechanism theory is outlined. Transient and frequency analyses are dealt with in the succeeding chapters and a discussion of stability is then given. The Part concludes with a consideration of relative stability and design procedure.

1.1. Closed-Sequence Control Systems.

A broad division of control systems can be made according to whether the control is of the open-sequence or closed-sequence type, the fundamental difference between the two being the effective cause which operates the system. In the first case, the quantity operating the system in no way depends on the value of the controlled quantity, but is fixed in size and variation merely by previous design. In the second case, the effective cause operating the system is the difference between a prescribed quantity and another quantity which is proportional to the actual value of the controlled quantity of the system, with the direction of operation so arranged that this difference always tends to zero. The schematic diagram of a closed-sequence control system is therefore as given by Fig. 1.

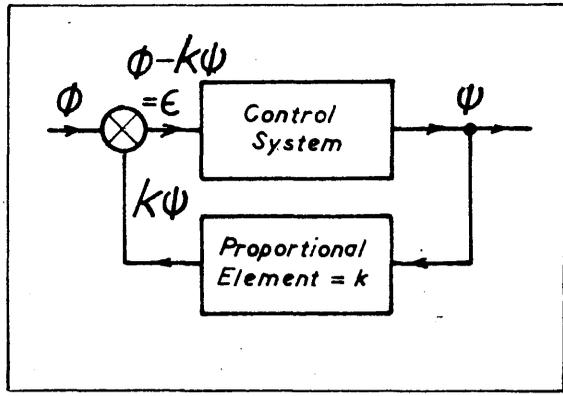


Fig.1. Closed-sequence control system.

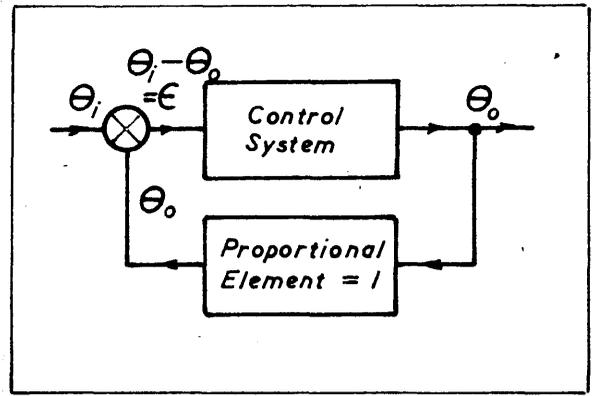


Fig.2. Angular position control servo-mechanism.

In the diagram, ψ is the controlled quantity, e.g. voltage, temperature, speed, etc. It is measured by some means which produces a proportional quantity $k\psi$ for comparison with ϕ , the quantity corresponding to the desired value of the controlled variable. The value of ϕ will in general not be pre-determined.* The difference ϵ is now arranged to move ψ in such a way that ϵ tends to zero, that is, that ψ tends to ϕ/k , which is the desired value of the output quantity.

The sub-division of closed-sequence control systems usually designated by the term servo-mechanisms refers to such systems in which the quantities ϕ and ψ are of a mechanical nature and a one-to-one relationship is desired between them. Agreement on this definition is, however, by no means universal. If this statement is adopted, however, the diagram of Fig. 2 results. The prescribed value of the controlled quantity or output quantity θ_o , is given by θ_i and is designated the input quantity. The result of the

* If ϕ is pre-determined, some type of regulator results. In the case of ϕ being constant, it is usually referred to as the datum quantity of the regulator.

control action is that θ_o tends to the value θ_i under both steady-state and changing conditions. Such a system occurs, for example, in the following by a gun-mounting of its prescribed aiming angle as provided by the angle of rotation of a light computing shaft. This particular problem of fire-control is one of remote positioning of a large mass. Such angular position controls form the great bulk of servo-mechanism applications, and it has become customary to represent the input and ^{output} quantities, therefore, as the shaft angles θ_i and θ_o .

The adoption of closed-sequence control immediately confers two main advantages compared with the more simple open-sequence type. These are, firstly, the accuracy of correspondence of output and input which is obtained, and secondly, the speed of response of the output when following a changing input. Both these advantages are secured merely by inserting sufficient amplification into the main control sequence from the error quantity to the output quantity. It is, of course, fundamental that power amplification be present in order to move an output having inertia, and possibly resisted by external torques. While, in the open-sequence type of control, accuracy can only be achieved by a power amplification which is both high and constant, a closed-sequence control system only requires that it should be high. In practice, it is impossible for an open-sequence control to retain its initial calibration in the face of normal temperature and load changes. It is, therefore, essential that fast and accurate control systems be of the closed-sequence type, and it is with the subdivision of servo-mechanisms that this thesis is concerned, although little extension is required to cover the general case.

The advantages mentioned in the preceding paragraph are not, however, obtained without sacrifice in the simplicity and operation of the system. In particular, self-excited oscillations requiring no external input signal, will usually appear unless steps have otherwise been taken to counteract such behaviour. This possibility is, of course, common to all systems having feedback over an energy source and constitutes the main undesirable feature of a closed-sequence control. The design problem is therefore the achievement of accuracy and speed of response, which require high amplification, without such amplification causing undue loss of stability.

1.2. Development of Servomechanism Theory.

Servomechanism theory has taken two main lines in its development. These are firstly, the response of the system to a transient input and secondly, the response of the system to a steady-state sinusoidal input.

In the first case, the input quantity is given a discontinuity, such as a step-function of displacement, and the speed and accuracy with which the output reproduces this step, are measures of the performance of the system. Such a procedure is known as transient analysis and it has developed on account of the approximation to such an input which occurs in practice, as for example, in fire-control, when a fresh target is engaged. A second reason is the relative ease with which a laboratory test may be carried out to check a particular design. As an alternative to the step-function, the system may be given a pulse of short duration and the time taken for the system to revert to zero used as an indication

of the stability. As it is not possible to achieve in practice the measurement of the true impulse response, the relation of pulse-width, i.e. duration, to the time constants of the system is highly important. Since this will certainly not be a simple matter to take into account when attempting to assess the speed of response, it is generally easier, both analytically and in practice, to obtain the response to a step-function rather than use the pulse-response. The true impulse-response may of course be obtained by differentiation of the step-response.

There are normally other practical requirements to be met by a servo-mechanism. The maximum permissible errors under specified maximum input velocity and acceleration are usually stated, and a further requirement upon minimum smooth output speed may have to be satisfied. There is also the problem of designing the apparatus in a manner which will reduce mechanical resiliences in the drive or platform of the mounting as far as possible, or will arrange these in the best manner to prevent anti-stabilising influences. It is only in addition to these factors, that the step-response may be considered as specifying the performance of the system.

The steady-state response to a sinusoidal input is, on the other hand, not so intimately connected with conditions encountered in practice, nor is it particularly easy to arrange a test in the laboratory. The method is justified, however, in the matter of analysis and design, affording as it does, an analytical means of breaking down the system into a number of simpler components, the individual effects of each of which appear explicitly in the final overall output-input relationship.

This fact and the existence of previous work on feedback amplifiers applicable with little modification to servo-mechanisms, has brought the frequency response method into prominence. Here it is a question of obtaining the amplitude response of the system, that is, the ratio of output to input magnitudes, for frequencies of the input variation from zero to as large a value as possible. The presence of inertia at the output will cause this response to fall off as the frequency is raised, and the bandwidth of a very small and fast servo will very often not exceed 30 c/s. Large systems have, of course, correspondingly smaller bandwidths. Whatever the bandwidth may be, however, reproduction of input signals will only occur if this band extends to frequencies representing the highest important frequency component in any of these input signals, and if constancy of amplitude response is maintained within the band. Peaks within this band cause certain frequencies to be magnified out of proportion, and the result on applying a step-function is to obtain an output containing an excessive oscillation at approximately this frequency. A technique has, therefore, arisen of restricting the maximum value of the amplitude-response to a figure varying from 1.2 to 1.6 and largely determined by past operating experience, in an attempt to keep this resonant oscillation, which contributes to the overshoot in a step-response, to acceptable values. The explicit effect on a step-response of any deficiency in the frequency response is not in general predictable. The general trend of these effects, however, is indicated at the conclusion of Sec. 3.2.

CHAPTER 2.

TRANSIENT ANALYSIS OF LINEAR SYSTEMS.

In the present Chapter, formulation of servo-mechanism performance is considered first of all from the classical point of view. The Laplace Transform technique is then used in deriving the transfer function, and in establishing an important theorem relating the minimum overshoots and undershoots in the step-response to the steady-state performance of the system. The Chapter concludes with a classification of basic servo-mechanism types.

2.1. Output-Input Relationship of Linear Systems.

Throughout this thesis we shall be dealing with linear systems, that is, systems whose performance is mathematically described by a linear differential equation with constant coefficients. In practice no system is truly linear. Small motions may be greatly dependent on stiction, variable friction, and backlash in gearing, and large errors may place the operating field currents of generators and motors in the saturation range. While the assumption of linearity may only hold, therefore, over a very limited range, it is, analytically, the simplest one to make; and it at least provides some results which can be later modified to account for certain non-linear features occurring in practical systems. With the assumption of linearity, the superposition principle and the related frameworks of operational calculus and Fourier analysis become available.

The formulation of the behaviour of any system consists in breaking the system down into simpler units, each of which may be described by a simple linear differential equation. For servo-mechanisms, most of these units or components will be uni-directional, that is, without mutual effects. Certain bi-directional elements do occur, however, the most important of these being the motor-load combination. Consider, for example, the simple position control system of Fig. 3.

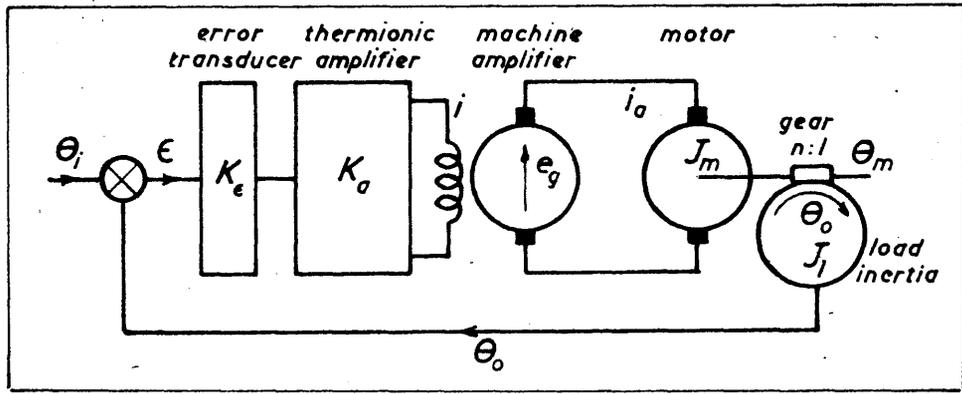


Fig. 3. Simple position control system.

- Let K_e = error transducer constant, V/radn
 K_a = amplifier gain up to output stage grids
 μ = amplification factor of output-stage
 R = total output stage resistance in each half circuit (including valve resistance)
 $T = \frac{L}{R}$, the output stage time constant
 K_E = generator e.m.f. per incremental field ampere
 R_a = total armature circuit resistance
 K_B = motor back e.m.f. per radn/sec.
 K_T = motor torque per armature ampere
 J = total inertia referred to motor shaft, slug-ft².
 B = viscous friction coefficient of load, lb-ft per radn/sec.
 n = gear ratio, motor to load.

Then $L \frac{di}{dt} + Ri = \mu K_a K_e \epsilon$ - output stage e.m.f. equation (1)

$e_g = K_E i$ - generator e.m.f.

$K_B \frac{d\theta_m}{dt} + i_a R_a = e_g$ - armature circuit e.m.f. equation (2)

$J \frac{d^2\theta_m}{dt^2} + B \frac{d\theta_m}{dt} = K_T i_a$ - torque equation at motor shaft (3)

$\theta_m = n\theta_o$
 $\epsilon = \theta_i - \theta_o$ - fundamental error equation (4)

Eliminating in order to find the overall differential equation relating the output to the input, we obtain

$$\frac{d^3\theta_o}{dt^3} + \left[\frac{1}{T_o} + \frac{1}{T} \right] \frac{d^2\theta_o}{dt^2} + \frac{1}{T_o T} \frac{d\theta_o}{dt} + \frac{K}{T_o T} \theta_o = \frac{K}{T_o T} \theta_i \quad (5)$$

where
$$T_o = \frac{J}{\left[\beta + \frac{K_B K_I}{R_a} \right]}, \quad T = \frac{L}{R}, \quad K = \frac{\mu K_a K_e K_T K_E}{n R_a R \left[\beta + \frac{K_T K_B}{R_a} \right]}$$

According to the various forms which the input signal θ_i takes, so the appropriate solution of (5) will yield the performance of the system. Thus for the response to a step-function,

we set $\theta_i = H(t)$, where
$$H(t) = 1, \quad t > 0$$

$$= 0, \quad t < 0. \quad (6)$$

In terms of the classical solution, we obtain the particular integral of $\theta_o = 1$, giving the "steady-state" output, and then the complementary function which determines the transient part of the output. To do this, the roots of the system characteristic equation,

$$m^3 + \left[\frac{1}{T_o} + \frac{1}{T} \right] m^2 + \frac{1}{T_o T} m + \frac{K}{T_o T} = 0 \quad (7)$$

are required, and we shall suppose these are m_1, m_2 and m_3 , for the present taken as distinct. The response of the system is now

$$\theta_o = 1 + A_1 \varepsilon^{m_1 t} + A_2 \varepsilon^{m_2 t} + A_3 \varepsilon^{m_3 t}$$

subject to the values of A_1, A_2, A_3 being determined by the initial conditions. This concludes the formulation and principle of classical solution of the above relatively simple system taken for the purposes of illustration.

2.2. Stability and Root Location.

The simple system above yielded a third-order differential equation relating the output and its various time derivatives to the input quantity.

A more general system will also include certain time derivatives of the input quantity, in addition to being of higher order. Thus a relation of the form

$$(a_0 D^r + a_1 D^{r-1} + \dots + a_{r-1} D + a_r) \theta_0 = (b_0 D^s + b_1 D^{s-1} + \dots + b_{s-1} D + b_s) \theta_i \quad (8)$$

will include all particular cases. For the majority of systems, r and s do not exceed 7 and 2 respectively. As above, the roots of the characteristic equation

$$a_0 m^r + a_1 m^{r-1} + \dots + a_{r-1} m + a_r = 0 \quad (9)$$

determine the form of the transient part of the response. The terms which may arise are given below, with the particular root-location to which they correspond. The constants A, R, ϕ are for the moment arbitrary and are of no consequence, as we are interested only in the manner in which the term depends on t .

Root-location.	Term in transient response.
$\pm \alpha$	$A e^{\pm \alpha t}$
$+j\beta$	$R \sin(\beta t + \phi)$
$+\alpha \pm j\beta$	$R e^{\alpha t} \sin(\beta t + \phi)$
$-\alpha \pm j\beta$	$R e^{-\alpha t} \sin(\beta t + \phi)$
0^*	$A e^{0t} = A$

Any of the above roots occurring twice or more will cause their transient term to contain some positive power of t as a factor. The conclusion can therefore be stated that stability will only be assured if roots of the type $\alpha, \pm j\beta, \alpha \pm j\beta$ and a second or higher order root at the origin are excluded.

* This will not occur in practice. From the viewpoint of stability, it is permitted however, and is included for the sake of completeness.

Alternatively, for stability, only those roots having negative real parts and a simple root at the origin, can be admitted. This is shown in

Fig. 4.

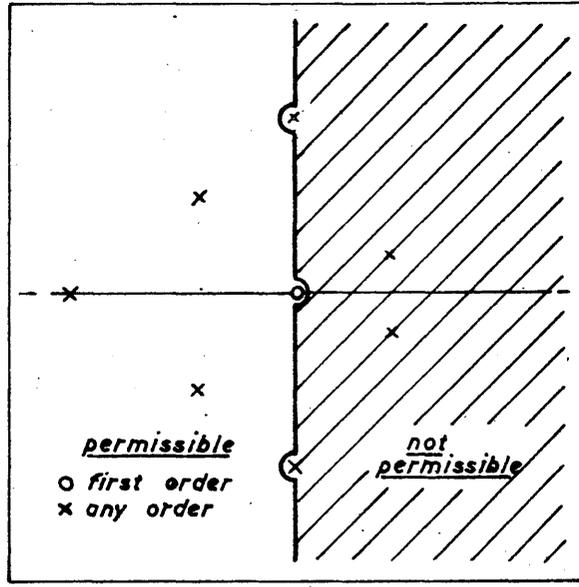


Fig. 4. Location of roots of characteristic equation.

This figure makes clear a fundamental distinction between the right and the left halves of the complex plane, namely that the former confers unstable, and the latter, stable properties. In contrast to this, no fundamental difference exists between the top and bottom halves of the plane. This follows from the fact that the coefficients of the characteristic equation are wholly real, and that complex roots can occur only in conjugate pairs, each of which defines one real frequency together with its damping coefficient.

The actual location of the roots of the characteristic equation is therefore a matter of great importance, but from the point of view of stability only, it is sufficient that the real parts of all the roots be negative. Criteria exist which answer this question, but in practice the existence of merely limiting stability is insufficient. For a suitable response, we

require in fact that the real part of any complex root be relatively large compared with the imaginary part, in order to reduce the number of oscillations occurring while a certain decay in the term takes place. For large absolute damping, i.e. damping irrespective of the oscillatory frequency, the real parts of all the roots should be large.

If an actual time solution of the output is required, the evaluation of these roots and the corresponding multiplying constants, requires an excessive amount of arithmetical work. Further, the alteration of any design parameter means that a complete new calculation is required. In an attempt to surmount these difficulties, recourse has been made to charts^{19,20} for a third order system, but the performance of a system represented by a high order differential equation cannot be so represented. Design and synthesis by this method has in the past been limited to very simple systems. The method of "Standard Forms" given by Whiteley²¹ partially relieves the situation and this is outlined in Sec. 5.2. The theory given in Part II of this thesis takes the synthesis problem a stage further.

2.3. The Laplace Transform and the Transfer Function.

The definition of the Laplace Transform adopted here is that if $x(t)$ is a time function, known for $t > 0$, the Laplace Transform $\bar{x}(p)^*$ of $x(t)$ is

$$\bar{x}(p) = \int_0^{\infty} e^{-pt} x(t) dt, \quad (10)$$

where p is a complex number with $Re(p)$ large enough to make the integral

* The notation and definition are those of Jaeger^{12,13}. Use will also be made of the symbols \mathcal{L} and \mathcal{L}^{-1} to denote Laplace Transformation and inverse Laplace Transformation respectively, as in the book by Gardner and Barnes¹¹. Thus $\mathcal{L} x(t) = \bar{x}(p)$ and $\mathcal{L}^{-1} \bar{x}(p) = x(t)$.

converge. In addition to transforming time functions, certain mathematical operations may also be transformed, as in the following two theorems .

Theorem I.

$$\mathcal{L}D^n x = \rho \mathcal{L}D^{n-1} x - D^{n-1} x_{t=0} \quad , \text{ and in particular}$$

$$\mathcal{L}Dx = \rho \bar{x}(\rho) - x_{t=0} \quad , \text{ for } n=1$$

$$\mathcal{L}D^2 x = \rho^2 \bar{x}(\rho) - Dx_{t=0} - \rho x_{t=0} \quad , \text{ for } n=2$$

Theorem II

$$\mathcal{L} \int_0^t x(\tau) d\tau = \frac{\bar{x}(\rho)}{\rho} \quad , \text{ given that } \varepsilon^{-\rho t} \int_0^t x(\tau) d\tau \rightarrow 0 \quad \text{as } t \rightarrow \infty .$$

With the aid of the above theorems, the differential equation (8) relating the input and output of a linear system, is first transformed into an algebraic equation in ρ , with the appropriate initial conditions inserted into the problem. Thus, taking the system as initially at rest and subject to an input $\theta_i(t)$ at $t=0$, we obtain

$$(a_0 \rho^r + a_1 \rho^{r-1} + \dots + a_{r-1} \rho + a_r) \bar{\theta}_0(\rho) = (b_0 \rho^s + b_1 \rho^{s-1} + \dots + b_{s-1} \rho + b_s) \bar{\theta}_i(\rho) \quad , \quad s < r \quad ,$$

and

$$\bar{\theta}_0(\rho) = \frac{b_0 \rho^s + b_1 \rho^{s-1} + \dots + b_{s-1} \rho + b_s}{a_0 \rho^r + a_1 \rho^{r-1} + \dots + a_{r-1} \rho + a_r} \cdot \bar{\theta}_i(\rho) .$$

Let

$$Q(\rho) = \frac{b_0 \rho^s + b_1 \rho^{s-1} + \dots + b_{s-1} \rho + b_s}{a_0 \rho^r + a_1 \rho^{r-1} + \dots + a_{r-1} \rho + a_r} \quad , \quad (11)$$

then $Q(\rho)$ is a rational fractional function relating the transform of the output to the transform of the input and therefore expresses the performance of the system in every respect. $Q(\rho)$ is known as the system transfer function, and it is deducible in the above manner or more directly, from the combined effect of the separate transfer functions of the sequence, as follows.

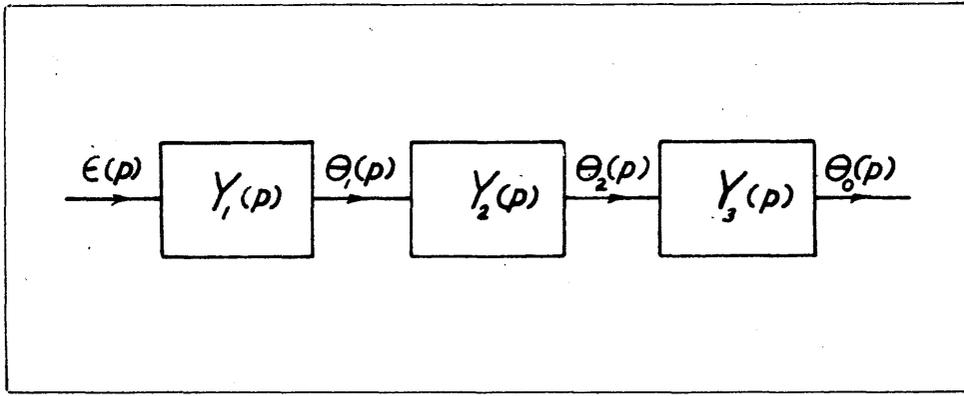


Fig. 5. Transfer functions in cascade.

If the connection of any one component does not load* the previous member of the chain, then

$$\frac{\bar{\theta}_0(p)}{\bar{\epsilon}(p)} = \frac{\bar{\theta}_0(p)}{\bar{\theta}_2(p)} \cdot \frac{\bar{\theta}_2(p)}{\bar{\theta}_1(p)} \cdot \frac{\bar{\theta}_1(p)}{\bar{\epsilon}(p)} = Y_1(p) Y_2(p) Y_3(p) \text{ etc.};$$

thus, for the example of Fig. 3, the component equations yield the transfer functions

$$\frac{\bar{i}}{\bar{\epsilon}}(p) = \frac{\mu K_a K_e}{R} \cdot \frac{1}{(1+pT)}$$

$$\frac{\bar{e}_g}{\bar{i}}(p) = K_E$$

$$\frac{\bar{\theta}_m}{\bar{e}_g}(p) = \frac{K_T}{R_a} \cdot \frac{1}{[B+K_T K_B] p(1+pT_0)}$$

$$\frac{\bar{\theta}_0}{\bar{\theta}_m}(p) = \frac{1}{n}$$

Hence

$$\frac{\bar{\theta}_0}{\bar{\epsilon}}(p) = \frac{\mu K_a K_e K_E K_T}{n R_a R} \cdot \frac{1}{p(1+pT)(1+pT_0)} = Y(p), \text{ say.} \quad (12)$$

$Y(p) = \frac{\bar{\theta}_0}{\bar{\epsilon}}(p)$ is known as the open-loop transfer function or simply loop transfer function. The basic diagram of a servo-mechanism can now be depicted as in Fig. 6.

* If mutual interaction takes place, the two components cannot be analytically separated. In this case a general transfer function for the two in combination is obtained.

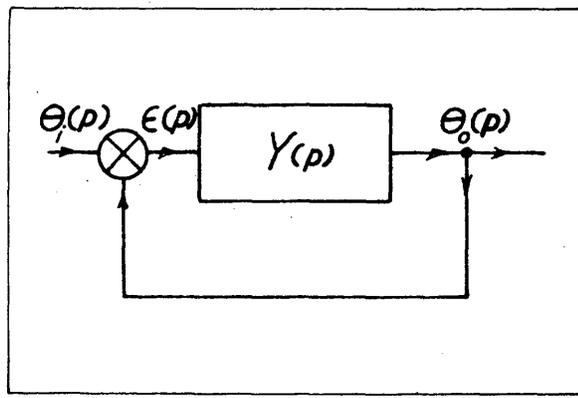


Fig. 6. Basic diagram of a servo-mechanism.

Equation (12), with the error equation $\bar{e}(p) = \bar{\theta}_i(p) - \bar{\theta}_o(p)$, yields the basic servo equations, namely

$$\frac{\bar{\theta}_o(p)}{\bar{\theta}_i} = \frac{Y(p)}{1 + Y(p)} \quad (13)$$

$$\frac{\bar{e}}{\bar{\theta}_i}(p) = \frac{1}{1 + Y(p)} \quad (14)$$

When $Y(p)$ is written out, (13) gives the equation

$$(T_o T p^3 + [T_o + T] p^2 + \rho + K) \bar{\theta}_o(p) = K \bar{\theta}_i(p)$$

in agreement with (5).

2.4. Transient Analysis. Test Signals.

Having formulated in the previous section the system transfer function,

$$Q(p) = \frac{\bar{\theta}_o(p)}{\bar{\theta}_i} = \frac{Y(p)}{1 + Y(p)},$$

the response to a particular input $\theta_i(t)$ is

$$\theta_o(t) = \mathcal{L}^{-1} \frac{Y(p)}{1 + Y(p)} \bar{\theta}_i(p). \quad (15)$$

Once $\theta_i(t)$ is assigned, and $\bar{\theta}_i(p)$ substituted, the time solution is obtained

as indicated in standard text-books¹¹⁻¹³ on the subject. It should be

emphasised that (15) is the response of the system from a state of rest,

prior to the application of the input signal. It is usually designated

the normal response of the system. The term transient response hitherto

used in a general sense, is strictly the complementary function part only, of the classical solution.

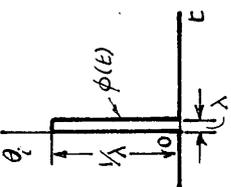
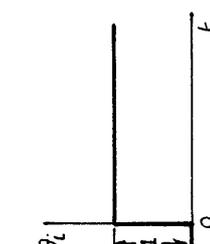
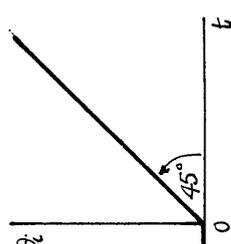
The normal response to a unit step-function is very often required. As it applies an infinite rate-of-change of input signal, it is a convenient way of assessing the speed of response of the system. The unit impulse-function might also be used for this purpose but as the response shows more overshoots and undershoots than the step-response, it is easier to work with the latter. In addition to the unit impulse-function, $\delta(t)$, and the unit step-function $H(t)$, a further test signal is the unit ramp-function $tH(t)$. The normal responses to the above will be denoted by $r_s(t)$, $r_o(t)$ and $r_l(t)$, and the errors by $e_s(t)$, $e_o(t)$ and $e_l(t)$ all respectively. The definitions of these signals and their responses are summarised in Table 1.

It is to be noted that the transform of the response to a unit impulse-function is

$$\bar{r}_s(p) = \frac{\gamma(p)}{1 + \gamma(p)} \quad (16a)$$

the R.H.S. of which has already been defined as the system transfer function. It follows that the response $r_s(t)$ to a unit impulse-function can be used to characterise a system in every respect just as the system transfer function does or the governing differential equation. This response has been called the weighting function of the system by certain writers².

TABLE I
INPUT TEST SIGNALS

Input	Definition	Diagram	Transform of Input	Transform of Output	Transform of Error.
Unit impulse-function.	$\delta(t) = \lim_{\lambda \rightarrow 0} \phi(t)$ $\phi(t) = 0, t \leq 0$ $= 1/\lambda, 0 < t < \lambda$ $= 0, t \geq \lambda$		$\mathcal{L} \delta(t) = 1$	$\bar{r}_3(\rho) = \frac{Y(\rho)}{1 + \gamma(\rho)}$ <p style="text-align: right;">(16a)</p>	$\bar{\epsilon}_3(\rho) = \frac{1}{1 + \gamma(\rho)}$ <p style="text-align: right;">(16b)</p>
Unit step-function.	$H(t) = 0, t < 0$ $= 1, t \geq 0$		$\mathcal{L} H(t) = \frac{1}{p}$	$\bar{r}_0(\rho) = \frac{Y(\rho)}{[1 + \gamma(\rho)]p}$ <p style="text-align: right;">(17a)</p>	$\bar{\epsilon}_0(\rho) = \frac{1}{[1 + \gamma(\rho)]p}$ <p style="text-align: right;">(17b)</p>
Unit ramp-function.	$tH(t)$		$\mathcal{L} tH(t) = \frac{1}{p^2}$	$\bar{r}_1(\rho) = \frac{Y(\rho)}{[1 + \gamma(\rho)]p^2}$ <p style="text-align: right;">(18a)</p>	$\bar{\epsilon}_1(\rho) = \frac{1}{[1 + \gamma(\rho)]p^2}$ <p style="text-align: right;">(18b)</p>

2.5. Two Theorems relating Test Signal Responses.

A generalisation of the test signals of the previous Section will now be considered. Let us take the test signal whose Laplace Transform is

$$\bar{\theta}_i(p) = \frac{A}{p^{n+1}} \quad (19)$$

This has the time form, $\theta_i(t) = A \frac{t^n}{n!} H(t)$, of which the n^{th} derivative is A . It will be termed an n^{th} -order input. For $n=0$ and $n=1$ respectively this gives the zero-order input or step-function of height A , and the first-order input having the constant slope A . The n^{th} -order response and error will be given by the inverse transforms of

$$\bar{r}_n(p) = \frac{Y(p)}{1+Y(p)} \cdot \frac{A}{p^{n+1}} \quad (20a)$$

$$\bar{e}_n(p) = \frac{1}{1+Y(p)} \cdot \frac{A}{p^{n+1}} \quad (20b)$$

For $n=0$ and 1 respectively we obtain the zero-order error or displacement error and the first-order error or velocity error. The corresponding responses are the zero-order response or response to a step-function of height A , and the first-order response or response to a ramp-function of slope A .

The first theorem concerns a relationship between the error quantities namely,

Theorem 1 - the area under the time plot of the $(n-1)^{\text{th}}$ -order error from $t=0$ to any particular instant, is equal to the n^{th} -order error at that instant. Conversely the derivative of the n^{th} -order error at any instant is equal to the $(n-1)^{\text{th}}$ -order error.

The proof is simple and practically self-evident from a consideration of Theorems I and II of Section 2.3. For the area under the $(n-1)^{\text{th}}$ -order error

is $\int_0^t \epsilon_{n-1}(t) dt$, and the transform of this by Theorem II, is

$$\begin{aligned} \frac{\bar{\epsilon}_{n-1}(p)}{p} &= \frac{1}{1+Y(p)} \cdot \frac{A}{p^n} \cdot \frac{1}{p} \\ &= \frac{A}{1+Y(p)} \cdot \frac{1}{p^{n+1}} \end{aligned}$$

that is, the transform of the n^{th} - order error. The theorem is thus

proved. It applies also, of course, to the responses. The importance of this theorem was first pointed out by Ludbrook¹⁸ as follows.

Theorem 2. - A servo-mechanism having zero n^{th} - order steady-state error, necessarily has a minimum total of n overshoots and undershoots in the zero-order response or response to unit step-function.

This can be argued from the fact that if the n^{th} order steady-state error is zero there must be at least one overshoot in the $(n-1)^{\text{th}}$ order response, for initially the error (given by $\theta_i - \theta_o$) must be positive and if there were no overshoot, the integrated $(n-1)^{\text{th}}$ - order error up to infinity, would be positive. This contradicts Theorem 1. We are therefore obliged to have at least one overshoot in order that the two areas may be equal and opposite, and the total integrated error up to infinity be zero. By differentiation of the $(n-1)^{\text{th}}$ - order error the $(n-2)^{\text{th}}$ - order error is obtained and by the nature of the differentiating process, a further undershoot in the $(n-2)^{\text{th}}$ - order response is obtained. Proceeding in this manner, we arrive at a minimum total of n overshoots and undershoots in the $(n-n)^{\text{th}}$ - order or zero-order response. Fig. 7 below illustrates the theorem for a servo-mechanism having zero 2^{nd} -order steady-state error. This represents

the optimum condition for this type of servo-mechanism and implies no

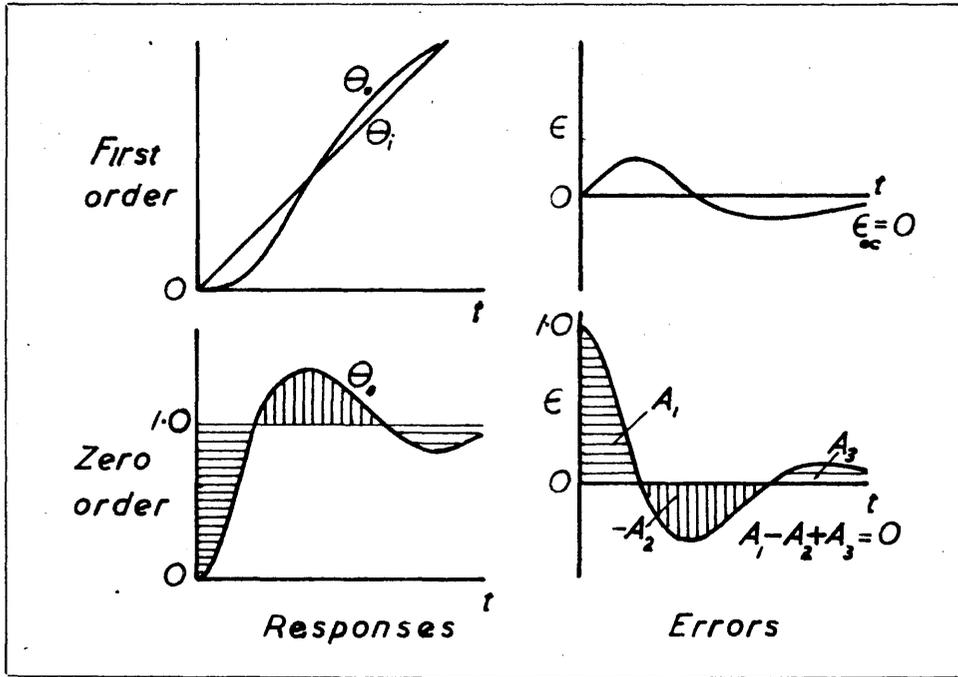


Fig. 7. A minimum of two over- and undershoots in the unit step-response of a servo-mechanism having zero order steady-state error.

overshoot in the 2nd-order response itself. Having thus related zero steady-state following properties to the number of over-and undershoots in the response to a unit step-function, it is now logical to enquire what conditions are necessary for these zero steady-state errors. This forms the subject of the next Section.

2.6. Steady-state Properties of Servo-mechanisms.

The steady-state properties of servo-mechanisms when subjected to the input signals $AH(t)$, $AtH(t)$, $A\frac{t^2}{2}H(t)$... $A\frac{t^n}{n!}H(t)$ are easily established by the Final Value Theorem¹¹ of Laplace Transform theory, namely

Theorem III.

If $\bar{x}(p) = \mathcal{L}x(t)$

then $\lim_{p \rightarrow 0} p\bar{x}(p) = \lim_{t \rightarrow \infty} x(t)$

Applying this to the n^{th} order error transform, in order to establish zero steady-state error, we have $\bar{e}_n(p) = \frac{1}{1+Y(p)} \cdot \frac{A}{p^{n+1}}$

and
$$\lim_{t \rightarrow \infty} e(t) = \lim_{p \rightarrow 0} \frac{A}{[1+Y(p)]p^n} \tag{21}$$

The R.H.S. of (21) will tend to zero if $Y(p)$ tends to the form K/p^r as p tends to zero, where the integer r is greater than n . The result may be expressed by the statement,

Theorem 3 - A servo-mechanism has zero, constant or infinite n^{th} order steady-state error, provided that

$$\lim_{p \rightarrow 0} Y(p) = \frac{K}{p^r}$$

where the integer r is greater than n , equal to n , or less than n respectively.

Corollary - When r is equal to n , the steady-state n^{th} order error is A/K , for $n \geq 1$ and A/tK , for $n=0$, where A is the magnitude of the n^{th} derivative of the input.

In the majority of servo-mechanisms $r=1$, as in the example of Fig. 3. These types therefore will sustain a steady-state first-order or velocity-error, limited only by the size of K . In the linear case they will have no zero-order or displacement error.* The factor K is termed either the

* In practice, stiction causes such types to have a small displacement error.

gain-factor, figure of merit or in the case of the systems being considered, i.e. $r=1$, the velocity-error constant. Here it is dimensionally seconds⁻¹, since for an error of α degrees,

$$\alpha = \frac{A}{K}$$

i.e. $[K] = \text{degrees per second per degree, i.e. seconds}^{-1}$.

It is therefore important from the point of view of steady-state errors, to have K as large as possible. This is done by increasing the thermionic amplification, which can be done without seriously affecting the frequency dependent part of the sequence $Y(p)$. K in fact comprises all such constants, gains, gear ratios etc. which are frequency independent, so arranged that the frequency dependent part, exclusive of the factor $1/p^r$ tends to unity as $p \rightarrow 0$. To take the example of Fig.3, whose open-loop transfer function was deduced in Sec.2.3, namely

$$\frac{\bar{\theta}_o(p)}{\bar{e}} = Y(p) = \frac{\mu K_a K_e K_E K_T}{n R_a R} \frac{1}{p(1+pT)(1+pT_o)}, \quad (12)$$

the gain-factor here is $K = \frac{\mu K_a K_e K_E K_T}{n R_a R}$. The factor $1/p^r$ is a simple pole at the origin, i.e. $r=1$, and the remaining part

$$\frac{1}{(1+pT)(1+pT_o)}$$

expresses the effect of two time-lags. This part, which in future will be denoted by $g(p)$, tends to unity as $p \rightarrow 0$. In general $g(p)$ will be of the

form
$$\frac{(1+pS_1)(1+pS_2) \dots (1+pS_l)}{(1+pT_1)(1+pT_2) \dots (1+pT_m)}$$

where S_1, S_2, \dots and T_1, T_2, \dots may be either real or complex, and $r+m > l$.

This last condition ensures that at infinite frequency $Y(p) \rightarrow 0$

The results of Theorem 3 will now be used to classify servo-mechanism types according to the type of input which they will follow without any steady-state error. This is done in Table II.

TABLE II

$\lim_{p \rightarrow 0} Y(p)$	Input sustained without steady-state error.	Type	Name of Type
$\frac{K}{p^0} = K$	None	0	displacement-error
$\frac{K}{p}$	steady displacement or step-function	1	zero displacement-error
$\frac{K}{p^2}$	steady-velocity or ramp-function	2	zero velocity-error
$\frac{K}{p^3}$	steady acceleration or 2nd order input.	3	zero acceleration-error.

Types 1 and 2 are also said to be displacement, velocity- and acceleration-controlled, respectively. For instance in a Type 1 servo-mechanism, a steady-input velocity is followed only as a result of an error being present to maintain the output at a constant velocity equal to the input-velocity and analogously with the other Types.

In concluding this Section, it should be stated that a Type 3 servo-mechanism has, as far as the author is aware, not yet been constructed.

The theoretical form is useful, however, as Type 2 servo-mechanisms having high gain do, in fact, take on Type 3 characteristics. In connection with stability, it is possible to use Type 3 transfer function forms (see Secs. 5.2 and 9.2).

CHAPTER 3.

FREQUENCY ANALYSIS OF LINEAR SYSTEMS.

The present Chapter is intended to explain the theoretical basis of frequency analysis and its use in stating design requirements. The deficiencies of frequency response characteristics are discussed with special reference to the input signal frequency spectrum, and in conclusion brief consideration is given to the effect of external disturbances.

3.1. Basis of Frequency Analysis.

In Sec. 1.2 it was stated that the technique of frequency analysis design methods was the control of the amplitude-response of the servo-mechanism in order to avoid peaks greater than 1.2 - 1.6, these figures being quite empirical. The fundamental principles on which such frequency response investigations should be based, in order to secure accurate reproduction of the input signal by the output, form the subject of this and the following three Sections.

From elementary considerations, the precept upon which perfect, and consequently non-physical, reproduction would be achieved, is that all frequency components in any input signal should be instantaneously and faithfully represented in the output signal. This immediately gives the obvious requirement of unity amplitude-response over an infinite bandwidth and zero phase-shift at all frequencies, or in terms of the overall transfer function, $Q(p) = \bar{\theta}_o / \bar{\theta}_i(p)$,

$$|Q(j\omega)| = \left| \frac{\theta_o(j\omega)}{\theta_i} \right| = \frac{\hat{\theta}_o}{\bar{\theta}_i} = 1 ,$$

and the angle of $Q(j\omega) = 0$, for all values of ω . Here $|Q(j\omega)|$ denotes the magnitude of $Q(j\omega)$, that is the ratio of the magnitudes (r.m.s. or peak) of the output and input sine-waves at any frequency.

A less unreasonable, but nevertheless still non-physical, demand is that the output shall be identical with the input after the elapse of a certain time in the transmission, say t_d seconds. This requirement can be expressed in terms of the frequency response by means of a theorem of Fourier Transform Theory¹⁴, as follows.

Theorem - If $\mathcal{F}x(t) = X(\omega)$,*
 then $\mathcal{F}x(t \pm t_d) = \epsilon^{\pm j\omega t_d} X(\omega)$

Applying the Theorem, since $\theta_o(t)$ is to be given by $\theta_i(t - t_d)$, then the Fourier Transforms $\Theta_o(\omega)$ and $\Theta_i(\omega)$ are related by

$$\Theta_o(\omega) = \epsilon^{-j\omega t_d} \Theta_i(\omega).$$

Hence $Q(j\omega) = \frac{\Theta_o(j\omega)}{\Theta_i(j\omega)} = \frac{\Theta_o(\omega)}{\Theta_i(\omega)} = \epsilon^{-j\omega t_d}$, that is, $Q(j\omega)$ should have amplitude unity and the phase-shift $-\omega t_d$, for all values of ω .

These requirements are shown in Fig. 8. The quantity t_d is known as the delay time.

* $\mathcal{F}, \mathcal{F}^{-1}$ denote the process of Fourier Transformation and inverse Fourier Transformation respectively, that is

$$\mathcal{F}x(t) = \int_{-\infty}^{+\infty} x(t) \epsilon^{-j\omega t} dt = X(\omega), \text{ say}$$

and
$$\mathcal{F}^{-1}X(\omega) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} X(\omega) \epsilon^{j\omega t} d\omega = x(t).$$

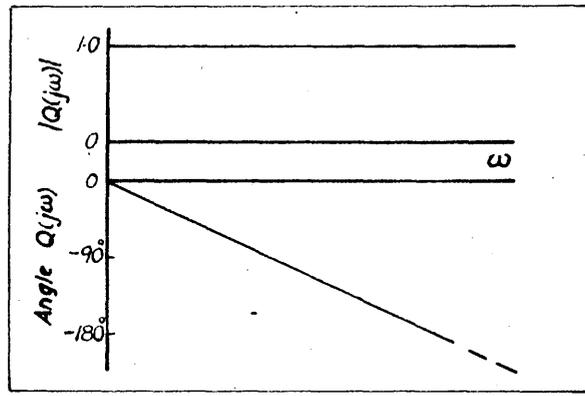


Fig. 8. Frequency responses for perfect reproduction after a finite delay of t_d seconds.

The frequency characteristics of Fig. 8, although in fact non-physical are sometimes taken as the aim of communication circuit design, but for servo-mechanisms, the presence of the finite delay t_d is undesirable. As we know, however, that the frequency response of a servo-mechanism only covers a finite bandwidth,* it is reasonable to enquire what will be the resulting deformation of the input signal, supposing that these so-called ideal characteristics were satisfied within a finite, instead of an infinite bandwidth. Suppose, for instance, that the frequency characteristics of a servo-mechanism, shown in Fig. 9a, were approximated by those shown in Fig. 9b, how is the reproduction of an input signal now affected, compared with simply the finite delay in transmission when an infinite bandwidth exists? The correct answer to this question must take into account the input signal itself, for the loss of accuracy of following will differ with different

* The bandwidth is usually defined only for servo-mechanisms showing a resonant peak. It is the band up to that frequency at which the amplitude response again becomes unity.

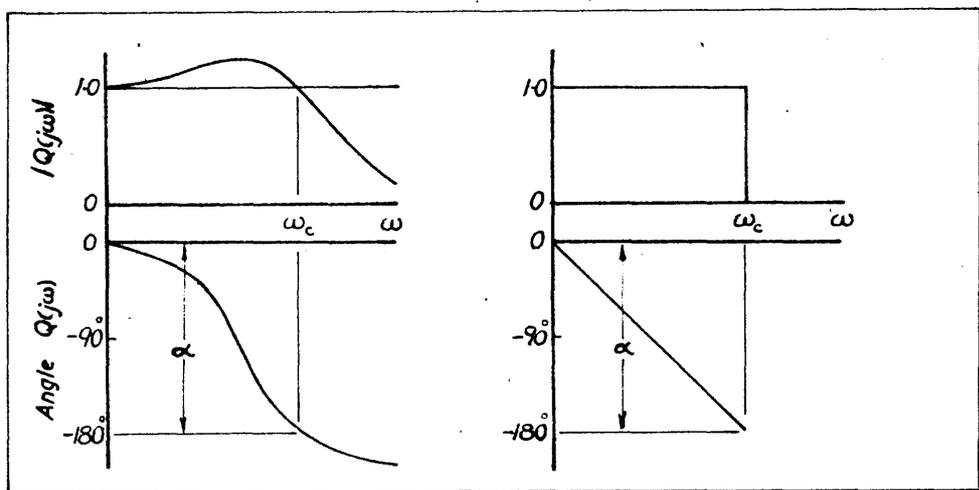


Fig. 9. Actual and approximate frequency characteristics of servo-mechanism or low-pass filter.

types of input signal, and it is precisely the form of the Fourier Transform of the input, which determines this point. From an elementary viewpoint it can be at once seen that if the input signal can be represented or even approximated by a periodic wave and if this is analysed into its frequency components, those types of input signal in which the magnitudes of the frequency components decrease rapidly as the order of the component increases, will suffer the least distortion in passing through a system having a finite bandwidth. For in those cases, only the small magnitude frequency components will be cut out and they will make the least contribution to the wave shape as a whole:

For the unit step-function the output from the finite-band system* of Fig. 9b is

$$\theta_o(t) = \frac{1}{2} + \frac{1}{\pi} Si \left[\omega_c \left(t - \frac{\alpha}{\omega_c} \right) \right] \quad (22)$$

where $Si(x)$ is the sine-integral of x given by

$$Si(x) = \int_0^x \frac{\sin y}{y} dy$$

the function (22) is illustrated in Fig. 10b and a typical response for an actual physical servomechanism is shown in Fig. 10a.

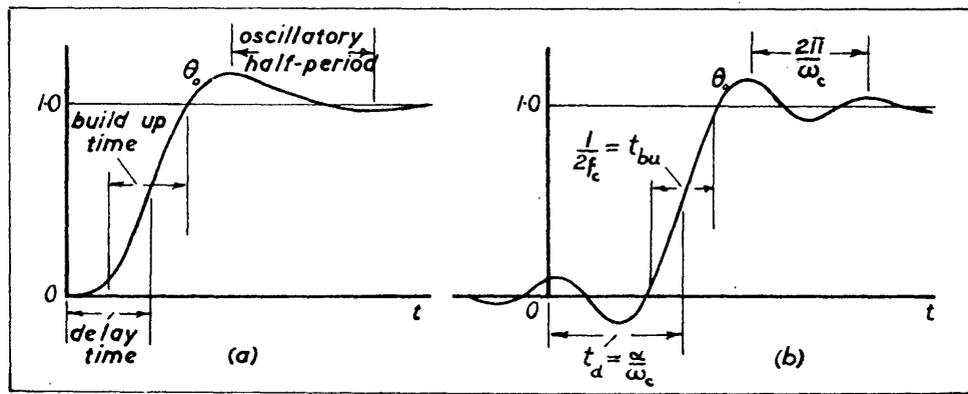


Fig. 10. Responses to unit-step function of actual servomechanism and approximate finite-frequency band system.

There are a number of relations concerning the delay-time t_d , the build-up time t_{bu} and the period of the decaying oscillation $\frac{2\pi}{\omega_c}$, but as they all pertain to a non-physical system they are of limited use. All lead to the

* It will be noticed that the output is not zero before $t = -0$. This is a further indication that the postulated system is non-physical and it is due to incompatible amplitude and phase characteristics.

conclusion that ω_c should be large, which is self-evident. It is useful to bear in mind the relation

$$t_{bu} = \frac{1}{2f_c}$$

however, as a preliminary design aid. It at least will give an approximate indication of how fast an actual servo-mechanism possessing a certain bandwidth can be expected to respond.

3.2. The Input Signal Frequency Spectrum.

The importance of the frequency spectrum of the input signal in relation to the finite frequency band of system was noted in the previous Section. In this Section the frequency spectra of various input signals are compared. This is done by replacing the single transient input signal by its periodic wave equivalent and merely performing a Fourier Analysis for each.

In each case $p(t)$ represents the equivalent periodic wave and $P(n\omega)$ the relative amplitude and phase of the n^{th} harmonic. We have therefore

$$p(t) = \frac{1}{T} \sum_{n=-\infty}^{n=+\infty} P(n\omega) \epsilon^{jn\omega t}$$
$$P(n\omega) = \int_{-\frac{T}{2}}^{+\frac{T}{2}} p(t) \epsilon^{-jn\omega t} dt$$

where

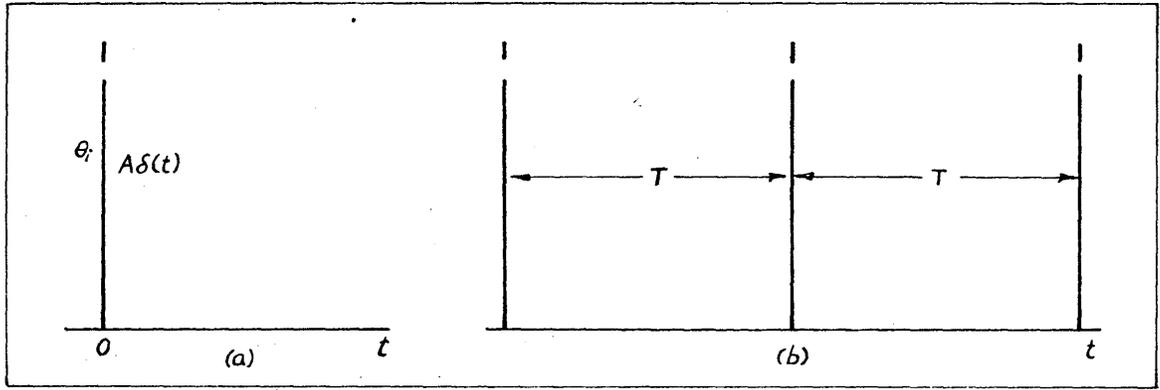
This is commonly known as the exponential Fourier Series¹⁴ and is derived in Appendix I. It is also shown in the Appendix that the n^{th} harmonic is

instantaneously given by

$$\frac{2}{T} |P(n\omega)| \cos(n\omega t - \phi_n)$$

and therefore has the amplitude $\frac{2}{T} |P(n\omega)|$ and the phase $\phi_n =$ angle of $P(n\omega)$. The following results then hold for the test functions discussed in Chapter 2.

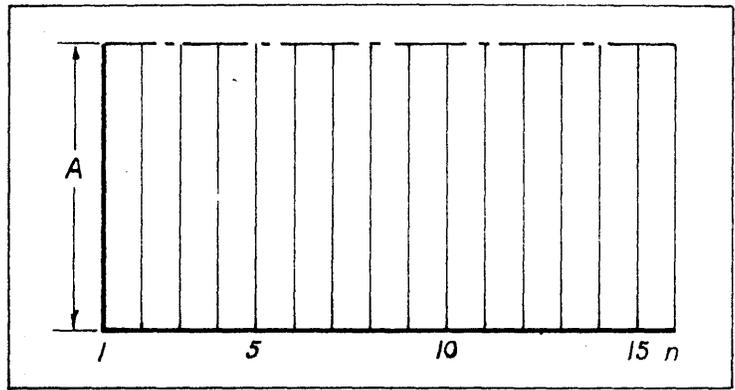
1. Impulse Function Input



(a) Impulse Input.

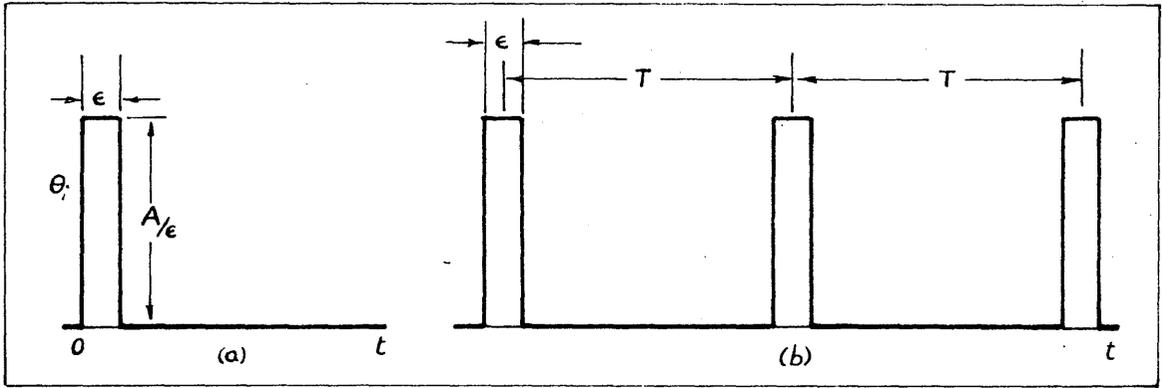
(b) Equivalent periodic input.

$|P(n\omega)| = A$
i.e. all components
present to the same
extent.



(c) Amplitude line spectrum
of periodic input
 $n > 0$ only is shown,
as $|P(-n\omega)| = |P(n\omega)|$.

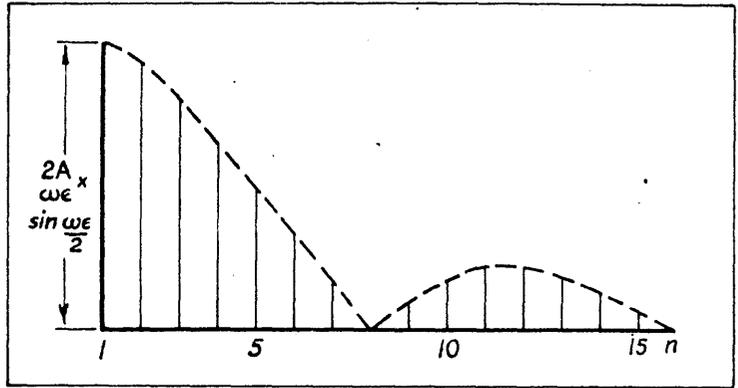
2. Pulse Input (practical test form of impulse)



(a) Pulse input.

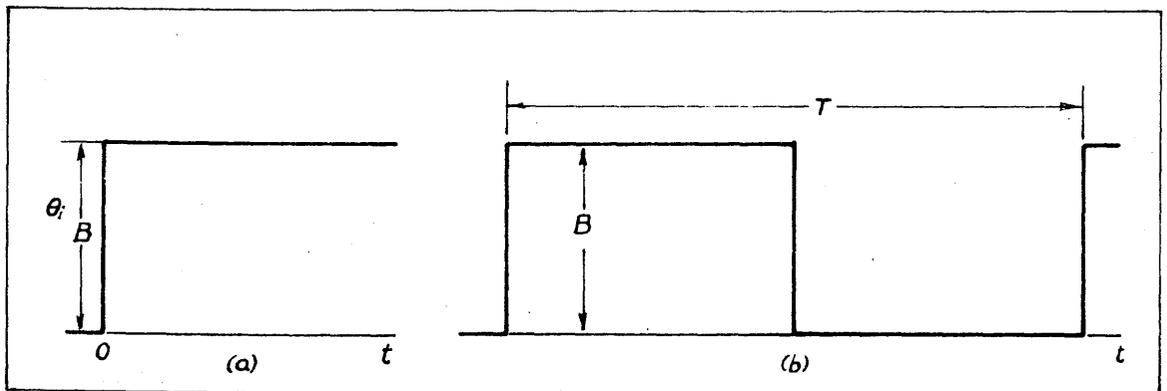
(b) Equivalent periodic input.

$$|P(n\omega)| = A \left| \frac{\sin \frac{n\omega\epsilon}{2}}{\frac{n\omega\epsilon}{2}} \right|$$



(c) Amplitude line spectrum.

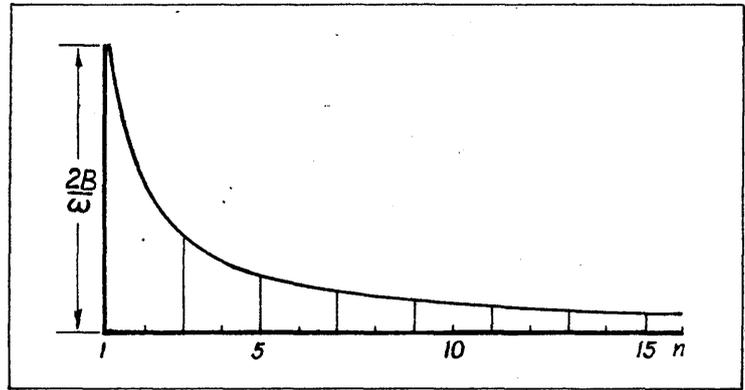
3. Step or Zero-Order Input.



(a) Step input.

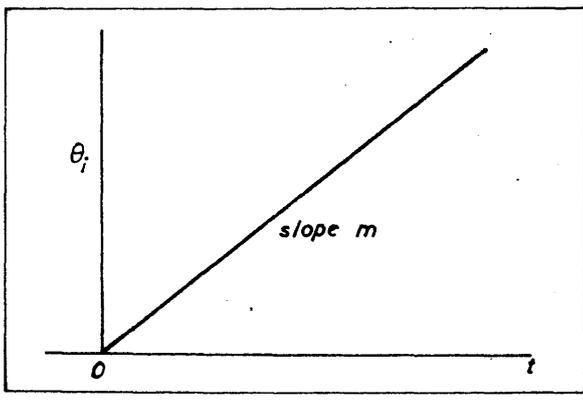
(b) Equivalent periodic input.

$$\begin{aligned}
 |P(n\omega)| &= \frac{BT}{2}, n=0 \\
 &= \frac{BT}{\pi n}, n \text{ odd.} \\
 &= 0, n \text{ even.}
 \end{aligned}$$



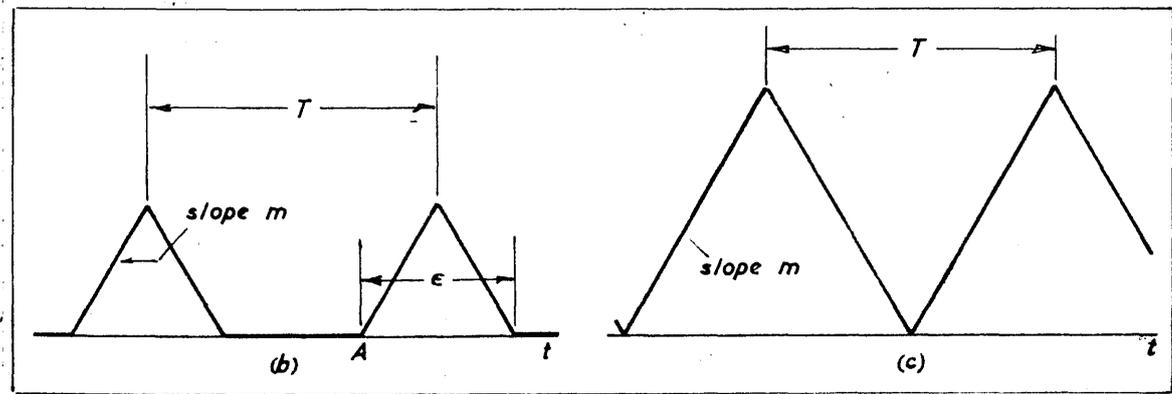
(c) Amplitude line spectrum.

4. First-Order Input.



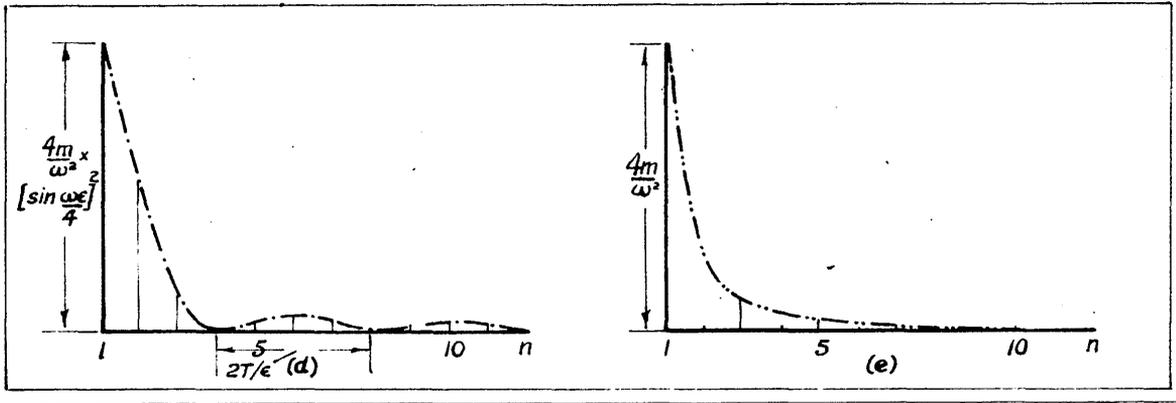
(a) First-order input.

No exact periodic equivalent is possible. Approximations are as follows. The correct conditions hold only at A, in figure (b).



(b)

(c)



(d) Amplitude line spectrum of (b).

(e) Amplitude line spectrum of (c).

$$|P(n\omega)| = \frac{m\epsilon^2}{4} \left[\frac{\sin \frac{n\omega\epsilon}{4}}{\frac{n\omega\epsilon}{4}} \right]^2$$

$$|P(n\omega)| = \frac{mT^2}{4}, \quad n = 0$$

$$= \frac{mT^2}{\pi^2 n^2}, \quad n \text{ odd}$$

$$= 0, \quad n \text{ even}$$

The results 1 to 4 may be compared by expressing the n^{th} harmonic relative to the fundamental. Thus we have

n^{th} harmonic amplitude fundamental.	Impulse	pulse	step	first-order approximation.
	1	$\frac{1}{n} \left[\frac{\sin \frac{n\omega\epsilon}{2}}{\frac{n\omega\epsilon}{2}} \right]$	$\frac{1}{n}$	$\frac{1}{n^2} \left[\frac{\sin \frac{n\omega\epsilon}{4} \right]^2 / \left[\frac{\sin \frac{\omega\epsilon}{4} \right]^2$ (b)
				$\frac{1}{n^2}$ (c)

The relative amplitudes are shown plotted against n , the order of the harmonic, in Fig. 11. In calculating the curves T/ϵ has been taken equal to 8 for the pulse wave and equal to 2 for the input 4(b). In the second case the ratio T/ϵ actually requires to be taken large enough to obtain the conditions illustrated at the point A, in the diagram of 4(b). As this ratio is increased, keeping the slope m constant, the envelope of the spectrum will move towards the right.

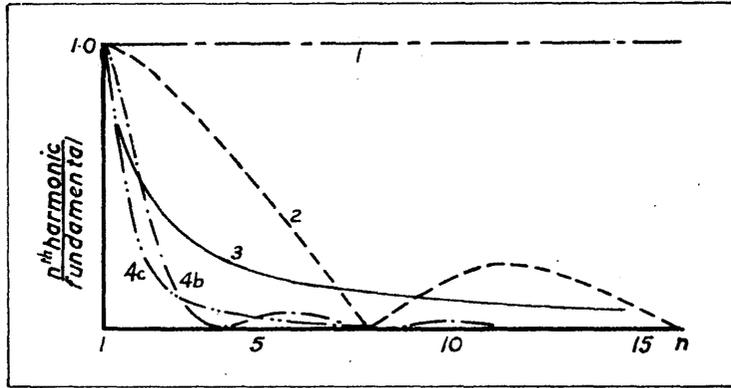


Fig. 11. Relative harmonic amplitudes of periodic waves equivalent to test-signals 1 - 4.

A study of Fig. 11 indicates in a general manner that the importance of the higher harmonics increases as the discontinuities (or sharpness) of the periodic waves become greater. For the impulse input all are required in equal amount; for the step input the harmonics drop off as $1/n$; for the triangular wave they decrease as $1/n^2$. Conversely, it can be concluded that, should these higher harmonics be removed from a wave, it will lose its sharpness. For instance, a square wave passing through a system having the amplitude response I, in Fig. 12a will be reproduced in a rounded form as in I, Fig. 12b. The characteristic II, however, containing more of the higher harmonics will result in the sharper reproduction shown. The curve II, Fig. 12a has been shown with a resonance peak of about 1.2 and will be expected to give a well damped and sharp square-wave response. Curve III is too resonant and produces a damped train of oscillations.

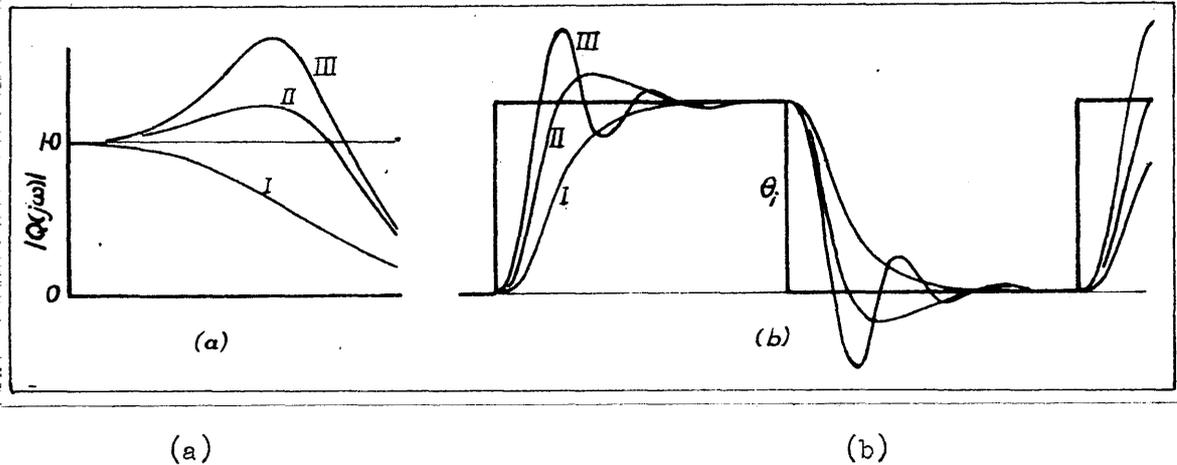


Fig. 12. Square-wave responses, corresponding to three contrasting types of frequency response.

3.3 The System Frequency Response in relation to External Disturbances.

Thus far it has been argued that the aim of the design should be to extend the frequency band of the servo-mechanism as far as possible. This is quite correct provided no external disturbances are considered, but in any practical system this will not be the case. There may be spurious noise originating in the apparatus or coming in with the input information. There may be effects of a mechanical nature, such as the existence of unbalanced torques at the output due to wind, or the roll of a ship, and mechanical resonances also may be present in the structure of the load itself. The elimination of unwanted signals in the input is at present the object of much research but, in the past, the tendency has been to avoid having the bandwidth greater than is necessary for the estimated types of input signal. By this means, the high frequencies associated with such inputs are smoothed out. Whenever possible, the sources of any

mechanical resilience should be investigated and if resonances are likely to be in the operating range of the system, it may be possible to utilise the effect in a stabilising manner. The influence of a sinusoidal disturbing torque on the output shaft at some low frequency well within the band of the servo-mechanism, which is approximately the effect produced on an unbalanced mounting by the roll of a ship, is most serious. If the open-loop amplification is kept at as large a value as possible at that frequency, that is, if near-unity overall amplitude response can be carried up to that frequency, a partial solution is obtained, but the difficulty of maintaining adequate stability increases. Such practical consideration as the above take away some of the value of the conclusions drawn from fundamental theory. By extending this theory to the disturbing influences themselves, however, it may become possible to modify the system design in the direction to minimise their effects.

This Section concludes the explanation of the theoretical background of frequency analysis. The use of the sinusoidal technique in design is explained in detail in text-books^{2,4,5} on the subject. A brief review of the design procedures in this method is given in Sec. 5.2 following considerations of relative stability.

CHAPTER 4.

STABILITY CRITERIA.

In this and the succeeding Chapter, a review is given of the methods relating to the investigation of stability. Various authors* have taken up this question and numerous criteria have been stated. These however merely result from different methods of expressing the problem. The fundamental relationships are explained below and the extension to finite time-lags considered.

4.1. The Fundamental Theorem on the Zeros of an Analytic Function; Application to Stability Investigation.

The result of complex variable theory upon which all the criteria of stability are based is sometimes known as the "principle of the argument"[†] and is as follows:-

1. Let the function $F(p)$ of the complex variable p be regular within and on a closed contour C except for a finite number of poles.
2. Let $F(p) \neq 0$ for values of p on C .

Then the excess of the number of zeros, Z , over the number of poles, P , of $F(p)$ within C is the net number of times $F(p)$ encircles the origin in a positive (anti-clockwise) direction as p traces the contour C once in the positive direction. This includes multiple poles and zeros counted according to their order.

* References. 22-28, 47.

† See for instance, Copson "Functions of a Complex Variable", p.119.
Bode, "Network Analysis", p. 149.

The application to stability is straightforward. Since we are to investigate the possible occurrence of roots of the characteristic equation anywhere in the right half-plane, the contour required is one enclosing this area, such as in Fig. 13. It is understood that the radius R will be allowed to increase indefinitely.

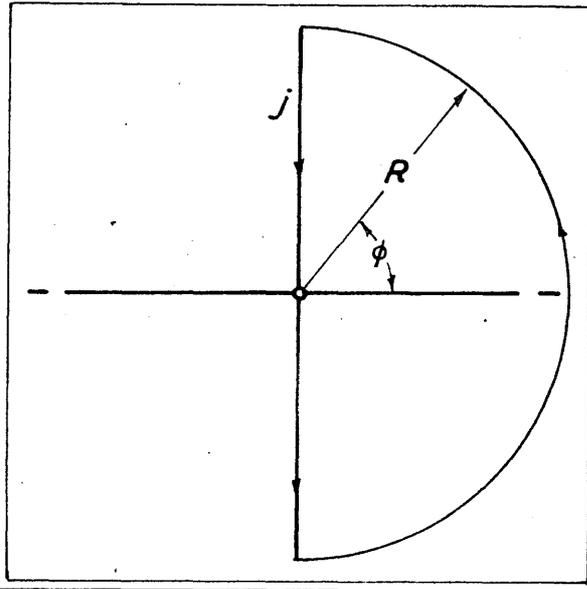


Fig. 13. Contour bounding the right half-plane as $R \rightarrow \infty$.

Having stated the theorem and its hypotheses, it must now be ascertained if the functions, whose zeros are to be investigated, satisfy these conditions. Accordingly the general system having any number of feedback loops is examined, Fig. 14.

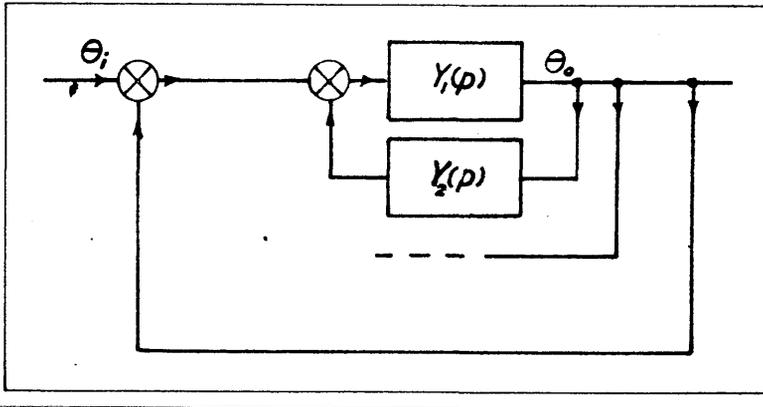


Fig. 14. m-loop coupled system.

The system of Fig. 14 has been given m coupled* loops and will give the resultant transfer function $\theta_o/\theta_i(p) = Y(p)$, given by

$$Y(p) = \frac{Y_1(p)}{1 + Y_1(p)Y_2(p) + Y_1(p)Y_3(p) + \dots} \quad (23)$$

and the overall differential equation relating θ_i and θ_o

$$\left(1 + \frac{1}{Y(p)}\right)\theta_o(p) = \theta_i(p) \quad ; \quad (24)$$

the characteristic equation is therefore

$$1 + \frac{1}{Y(p)} = 0 \quad (25)$$

Before examining the form of $Y(p)$, the individual transfer functions $Y_1(p)$, $Y_2(p)$, etc. require to be described. If we exclude finite time lags for the moment, $Y_1(p)$, $Y_2(p)$ etc. will be quotients of rational integral functions of p of low degree and possibly containing some power of p as a factor. The transfer functions will be stable by themselves and comprise no "non-minimum

* If a transfer function is common to two or more loops, the loops are said to be coupled.

phase^{*} elements. This being so, $Y_1(\rho), Y_2(\rho)$ etc. will have neither poles nor zeros in the right half-plane.

The various functions we can use to discuss stability are, therefore, as follows.

$$1. \quad 1 + \frac{1}{Y(\rho)} = 1 + \frac{1 + Y_1(\rho)Y_2(\rho) + Y_1(\rho)Y_3(\rho) + \dots}{Y_1(\rho)}$$

This may have zero anywhere in the ρ -plane, but has no poles in the finite right-half plane. As was shown in Sec. 2.6, however, $Y(\rho)$ will behave like K/ρ^r as ρ tends to zero, and therefore $1/Y(\rho)$ will have an r^{th} order zero at $\rho=0$. As ρ becomes infinite, $Y(\rho)$ will likewise tend to K'/ρ^s and hence $1/Y(\rho)$ will have an s^{th} order pole at $\rho=\infty$; r and s are both low positive integers - 1, 2, 3 etc. These points are illustrated by the example in Sec. 4.2 under the heading Generalised Nyquist Criterion.

2. Since $1 + 1/Y(\rho) = (1 + Y(\rho))/Y(\rho)$, the zeros of $1 + Y(\rho)$ may be investigated, that is,

$$\text{the function} \quad 1 + \frac{Y_1(\rho)}{1 + Y_1(\rho)Y_2(\rho) + Y_1(\rho)Y_3(\rho) + \dots}$$

This may have zeros and poles anywhere in the ρ -plane, and if, using the theorem, we are to look for possible zeros in the right half-plane, the number of poles in this area must previously be found out.

3. The zeros of $1 + 1/Y(\rho)$ are the zeros of $1 + Y_1(\rho) + Y_1(\rho)Y_2(\rho) + Y_1(\rho)Y_3(\rho) + \dots$.

If the individual transfer functions are $Y_1(\rho) = \frac{f_1(\rho)}{g_1(\rho)}$, $Y_2(\rho) = \frac{f_2(\rho)}{g_2(\rho)}$, ... etc., the zeros of $1 + 1/Y(\rho)$ will simply be the zeros of $g_1(\rho)g_2(\rho)\dots + f_1(\rho)g_2(\rho)\dots = 0$,

* Components whose phase-shift exceeds that deducible from their amplitude-frequency characteristic, which is the minimum possible.

that is, a rational integral function of ρ . This is, in fact, the left-hand side of the overall equation relating θ_o to θ_i , say

$$a_0 \rho^n + a_1 \rho^{n-1} + a_2 \rho^{n-2} + \dots + a_{n-1} \rho + a_n \quad (26)$$

The function (26) has a n^{th} order pole at infinity, but no poles in the finite ρ -plane.

To summarise, therefore, for multi-loop systems, forms 1 and 3 have no poles in the right half-plane. For single-loop systems, $Y(\rho)$ becomes $Y_1(\rho)$, and therefore none of the three forms have poles in the right half-plane. The application of these results to stability criteria now follows.

4.2. Stability Criteria.

1. Simple Nyquist Stability Criterion.

Direct form.

Let $Y_1(\rho) = \theta_o/\theta_i(\rho)$ be the loop transfer-function of a single-loop servo-mechanism; $Y_1(\rho)$ has no poles in the right half-plane and $Y_1(\rho) \rightarrow 0$ as $\rho \rightarrow \infty$. Then as ρ goes once round the contour of C of Fig. 13 in a positive direction, only values of $Y_1(\rho)$ for $\rho = j\omega$ need be considered. If N is the number of positive revolutions made by $1 + Y_1(j\omega)$ about the origin, then $Z = N$ is the number of unstable roots of $1 + Y_1(\rho) = 0$. For stability therefore, $1 + Y_1(j\omega)$ should make no net rotation about the origin; alternatively $Y_1(j\omega)$ should make no rotation about the point $[-1, 0]$. Further, the real coefficients of ρ in $Y_1(\rho)$ mean that $Y_1(-j\omega) = \text{conj. } Y_1(j\omega)$, that is, the part of the locus for negative frequencies is the mirror image in the real axis of the part of the locus for positive frequencies. The stability criterion is therefore as

follows:-

If the locus $Y_i(j\omega)$ for all values of ω from 0 to ∞ encloses the point $[-1, 0]$, the system is unstable; if the locus does not enclose the point $[-1, 0]$, the system is stable.

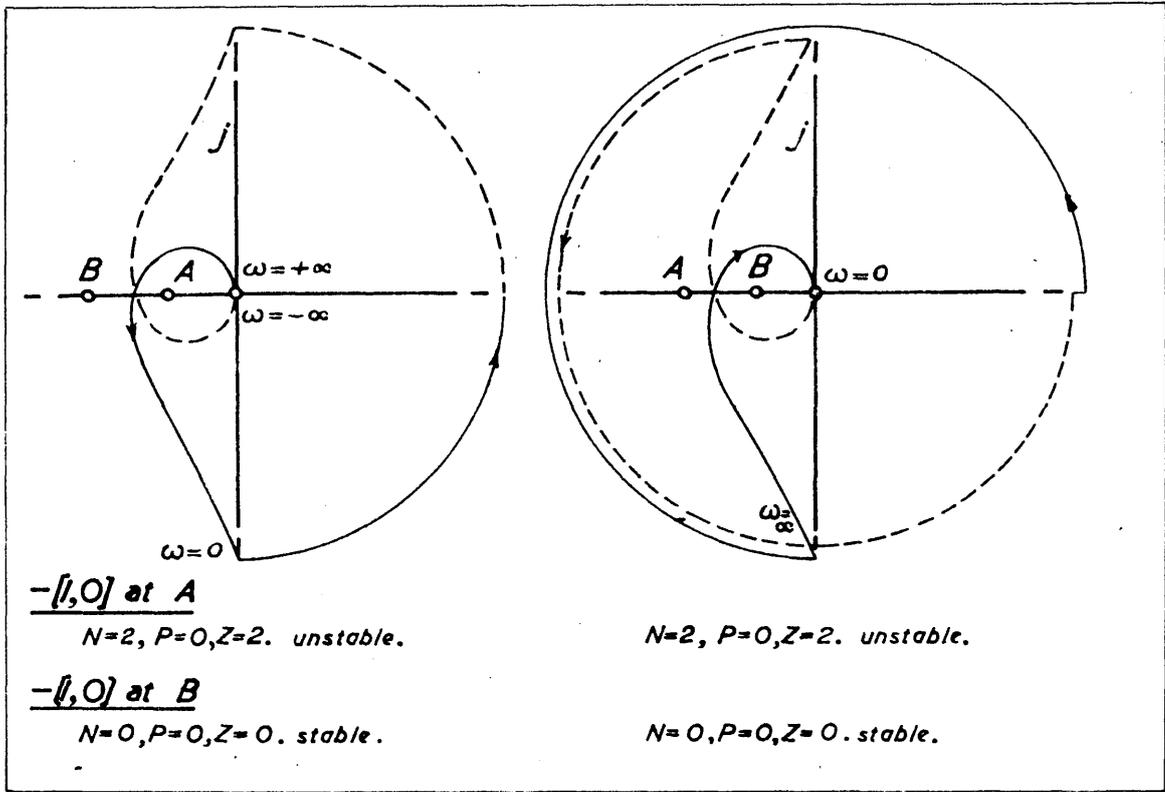
This is the simple criterion given by Nyquist for systems having no poles of the function $\theta_o/\epsilon(\rho)$ in the right half-plane. In the criteria which follow, the general statement on the number of unstable roots will be retained, and the function will be considered for all values of ρ on the contour. It is understood that for a stability criterion one need only set this number of unstable roots equal to zero.

Inverse form.

Let $1/Y_i(\rho)$ be the transfer function $\epsilon/\theta_o(\rho)$ of a single-loop servo-mechanism. $1/Y_i(\rho)$ has no poles in the right half-plane and as $\rho \rightarrow \infty$, $1/Y_i(\rho)$ is of the order of ρ^s , say, where s is a low positive integer. The inverse form of the rule is as follows.

If the function $1/Y_i(\rho)$ makes N anti-clockwise revolutions about the point $[-1, 0]$, as ρ moves once round the contour C in a positive direction, then there are $Z = N$ roots of $1 + 1/Y_i(\rho) = 0$ in the right half-plane.

Both forms are illustrated below for the function $Y_i(\rho) = K/p(1+\rho T_1)(1+\rho T_2)$. The point $[-1, 0]$ is placed first at A and then at B, to show an unstable and a stable locus respectively. In considering $Y_i(\rho)$ for values of ρ near zero, ρ is replaced by $\rho \epsilon^{j\theta}$ where θ changes from $+\pi/2$ to $-\pi/2$. This circumscribes the pole of $Y_i(\rho)$ at the origin. Similarly in considering $1/Y_i(\rho)$ for values of ρ along the infinite arc, ρ is replaced by $R \epsilon^{j\phi}$ where ϕ changes from $-\pi/2$ to $+\pi/2$. For an s^{th} order pole of $1/Y_i(\rho)$ at infinity, $1/Y_i(\rho)$ will rotate through $s\pi$ anti-clockwise.



(a) direct Nyquist criterion.

(b) inverse Nyquist criterion.

Fig. 15. Direct and inverse Nyquist criteria for single-loop servo-mechanism.

2. Generalised Nyquist Stability Criterion.

Direct form

Let $Y(p)$ be the resultant transfer function $\theta_o/c(p)$ of a multi-loop servo-mechanism. $Y(p)$ will be the general form

$$Y_1(p) / [1 + Y_1(p)Y_2(p) + Y_1(p)Y_3(p) \dots]$$

and will have poles in the right-half plane if $1 + Y_1(p)Y_2(p) + Y_1(p)Y_3(p) \dots$

has zeros in that region. Since $1 + Y_1(p)Y_2(p) \dots$ will have no

poles in that region, as the transfer functions $Y_1(p), Y_2(p), \dots$ are stable by themselves, an application of the simple Nyquist criterion above can be used to determine the zeros of $1 + Y_1(p)Y_2(p) + Y_1(p)Y_3(p) \dots$, i.e. the poles of $Y(p)$ in the right-half plane. Let this number be P. We have also the condition that $Y(p) \rightarrow 0$ as $p \rightarrow \infty$. The rule can therefore be stated as follows.

If the locus $Y(j\omega) = \theta_{\omega} / \epsilon(j\omega)$ makes N anti-clockwise revolutions about the point $[-1, 0]$, as ω varies from ∞ to $-\infty$, then there are $Z = N + P$ roots of $1 + Y(p) = 0$, in the right-half plane. For stability, therefore, $N = -P$.

Inverse form

Let $1/Y(p) = [1 + Y_1(p)Y_2(p) + Y_1(p)Y_3(p) \dots] / Y_1(p)$ be the transfer-function $\epsilon / \theta_e(p)$ of a multi-loop servo-mechanism. Then it has already been shown that $1/Y(p)$ has no poles in the right-half plane. $1/Y(p)$, however, will be of the order of p^5 as before, when $p \rightarrow \infty$. The criterion is thus identical with the inverse form for a single-loop system, namely:-

If the function $1/Y(p)$ makes N anti-clockwise revolutions about the point $[-1, 0]$ as p goes once round the contour C in a positive direction, then there are $Z = N$ roots of $1 + 1/Y(p) = 0$ in the right-half plane. For stability, therefore, $N = 0$.

The superiority of the inverse form in this case is evident as only the one diagram is required. This is demonstrated by consideration of the two-loop servo-mechanism having $Y_1(p) = K_1 / p(1 + pT_m)$ and $Y_2(p) = K_2 p^{4-3} / (1 + pT_0)^3$, with $T_0 > T_m^*$. In this case the resultant transfer function $\theta_e / \epsilon(p) = Y(p)$ is

* This example is taken from reference 2. In this and other works there is no mention of the advantage conferred by the inverse diagram in eliminating the poles of a multi-loop system.

$$Y(p) = \frac{K_1}{p(1+pT_m)} \cdot \frac{1}{1 + K_1K_2T_o^3p^3 / (1+pT_m)\lambda(1+pT_o)^3}$$

Poles of $Y(p)$ in the right-half plane can be ascertained by a simple Nyquist diagram to investigate the zeros of $1 + K_1K_2T_o^3p^3 / (1+pT_m)\lambda(1+pT_o)^3$, and it will be found that there are none if $K_1K_2 < 8$ and two if $K_1K_2 > 8$. Let us arrange for the second case, so that there are $P = 2$ poles of in the right-half plane. The direct and inverse diagrams showing $Y(p)$ and $1/Y(p)$ respectively are now shown in Fig. 16. If we place the point $[-1, 0]$ first at A, then at B, we can use the same locus to illustrate unstable and stable conditions in turn.

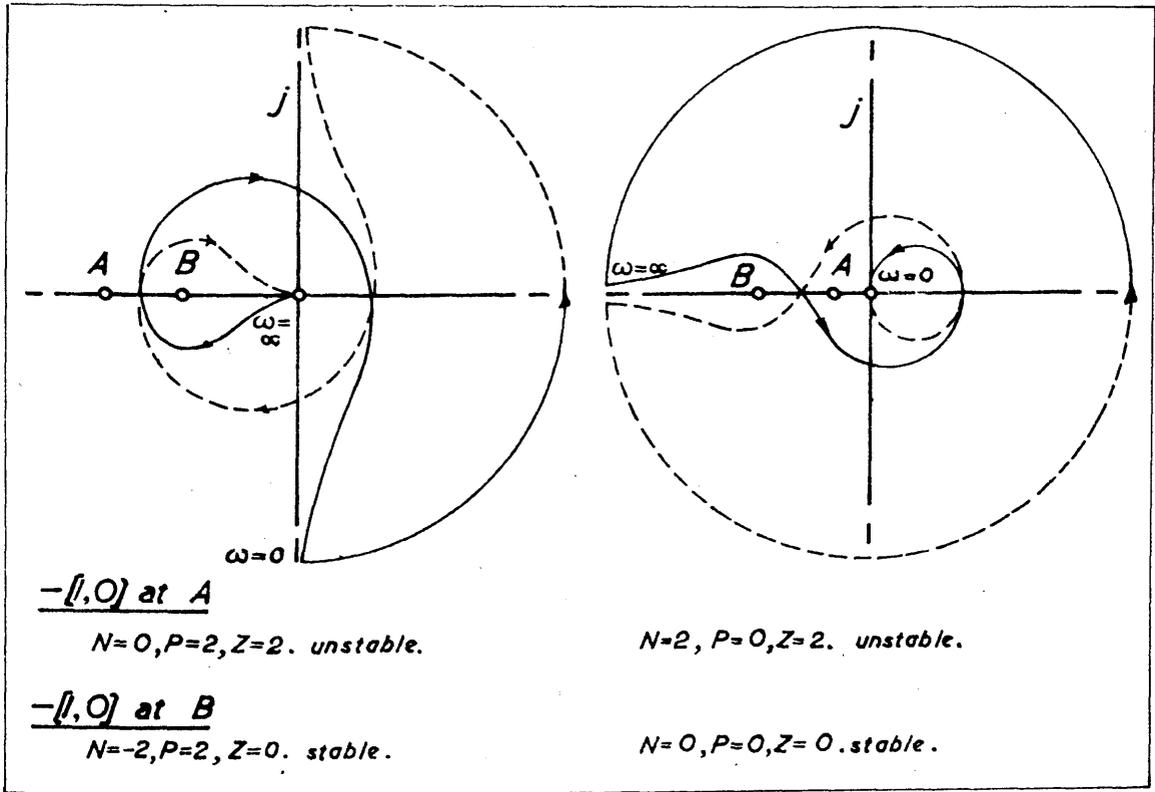


Fig. 16. Generalised Nyquist criteria for two-loop servo-mechanism having unstable subsidiary loop.

The simple direct Nyquist criterion cannot be used for examples similar to the above.

3. Leonhard²⁶ Criterion.

The Leonhard criterion uses the polynomial

$$a_0 p^n + a_1 p^{n-1} + a_2 p^{n-2} + \dots + a_{n-1} p + a_n = H(p) \quad (26)$$

as the function representing the left-hand side of the characteristic equation. This eliminates poles from the finite p -plane. According to the theorem, therefore, there are $Z = N$ roots of $H(p) = 0$ within the right-half plane, where N is the number of anti-clockwise revolutions made by $H(p)$ as p traces the contour C in a positive direction. Stability will be ensured if $N = 0$, i.e. no net rotation of $H(p)$ about the origin. This may be simplified further. For since $H(p) \rightarrow a_0 p^n$ as $p \rightarrow \infty$, the rotation made by $H(p)$ as p moves along the infinite arc from $R_\infty^{-j\pi/2}$ to $R_\infty^{j\pi/2}$ through $+\pi$ radians, is simply $+n\pi$. The total rotation of being zero for stability, this leaves $-n\pi$, i.e. $n\pi$ anti-clockwise radians, for the phase change of $H(p)$ as p goes from $-j\infty$ to $+j\infty$. Since $H(j\omega)$ is symmetrical about the real axis, this is equivalent to the following statement.

If $H(p)$ represents the left-hand side of the equation

$$a_0 p^n + a_1 p^{n-1} + a_2 p^{n-2} + \dots + a_{n-1} p + a_n = 0$$

in which $a_0 > 0$, there will be no unstable roots of the equation if $H(j\omega)$ rotates through n anti-clockwise quadrants as ω varies from 0 to ∞ .

The Leonhard criteria for a fourth order equation in the unstable and stable conditions are shown in Fig. 17.

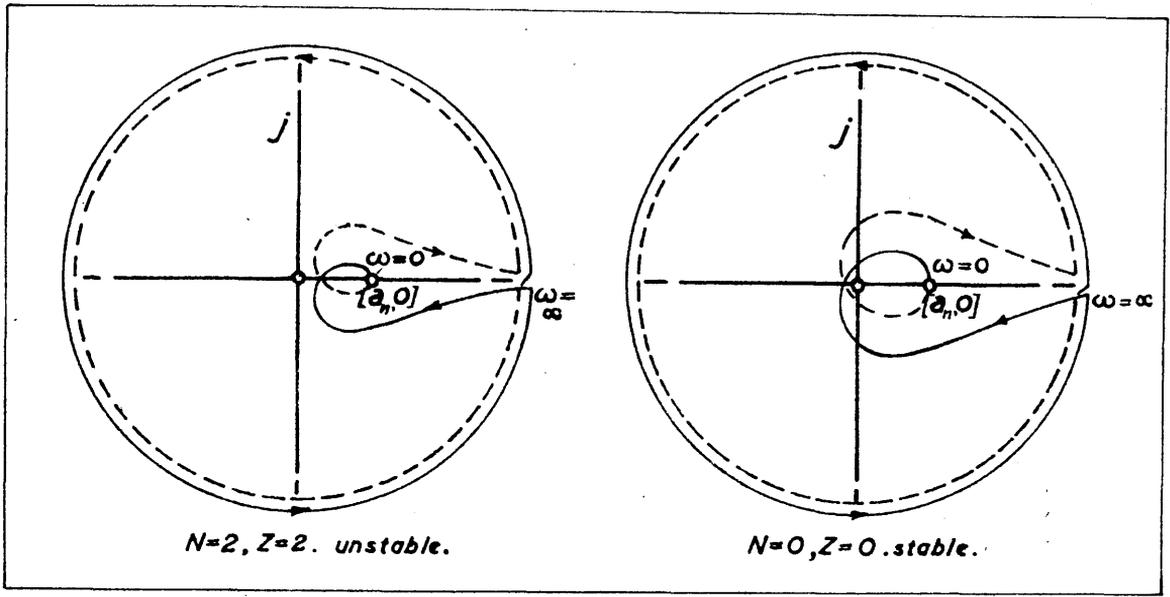


Fig. 17. The Leonhard criterion for a fourth-order differential equation.

The Leonhard diagram does not possess the advantages of the Nyquist diagram for design purposes, in that the connection of additional elements in the sequence is not explicitly shown in the new set of coefficients in the equation.

4. Routh²² Criterion: Hurwitz²³ Criterion.

The starting point of both these criteria, which apply only to polynomial equations of the form (26) is again the fundamental theorem of Sec. 4.1. Here the rotations of $H(j\omega)$ about the origin are counted by the net number of times the quotient $\frac{\text{real part } [a_0(j\omega)^n + \dots + a_n]}{\text{imag. part } [a_0(j\omega)^n + \dots + a_n]}$ changes its sign from + to - through zero. It is possible to put this into an analytical form using Sturm's Theorem. As the rule for forming the Routh test-functions is cumbersome, the same result due to Hurwitz is

given, namely:-

The necessary and sufficient conditions that the equation

$$a_0 p^n + a_1 p^{n-1} + a_2 p^{n-2} + \dots + a_{n-1} p + a_n = 0$$

(where a_0, a_1, \dots are real, a_0 being positive), should have no roots with positive real parts is that $\Delta_1, \Delta_2, \Delta_3 \dots \Delta_n$ should all be positive, where

$$\Delta_1 = a_1, \quad \Delta_2 = \begin{vmatrix} a_1 & a_3 \\ a_0 & a_2 \end{vmatrix}, \quad \Delta_3 = \begin{vmatrix} a_1 & a_3 & a_5 \\ a_0 & a_2 & a_4 \\ 0 & a_1 & a_3 \end{vmatrix}, \dots$$

$$\Delta_n = \begin{vmatrix} a_1 & a_3 & a_5 & \dots & a_{2n-1} \\ a_0 & a_2 & a_4 & \dots & a_{2n-2} \\ 0 & a_1 & a_3 & a_5 & a_{2n-3} \\ 0 & a_0 & a_2 & a_4 & a_{2n-4} \\ \dots & \dots & \dots & \dots & a_n \end{vmatrix}; \quad a_r = 0, r > n.$$

It is to be noted, however, that Routh stated a means of finding the number of unstable roots while Hurwitz merely obtained the condition that there should be none of them. A further point, which is time-saving, is that a necessary condition for none of the roots of the equation to lie in the right-half plane, is that all the coefficients a_0, a_1, \dots, a_n should be positive.

For a cubic, Hurwitz's criterion is

$$a_1 > 0, \quad \begin{vmatrix} a_1 & a_3 \\ a_0 & a_2 \end{vmatrix} > 0, \quad \text{i.e. } a_1 a_2 > a_0 a_3$$

Similarly the Leonhard criterion, see Fig. 18, states that at

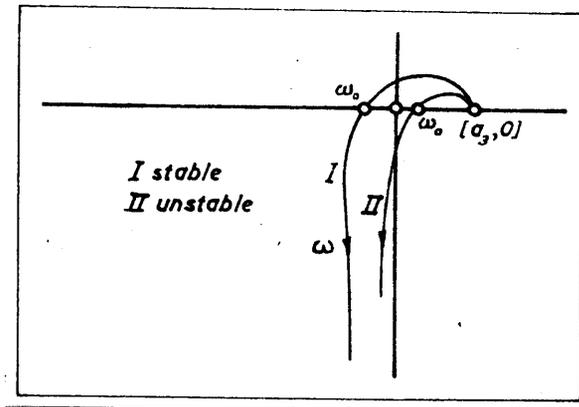


Fig. 18.

the frequency ω_0 at which the imaginary part of $H(j\omega)$ is zero, the real part should be negative.

Since
$$H(j\omega) = (-a_1\omega^2 + a_3) + j(-a_0\omega^3 + a_2\omega)$$

Then
$$-a_1\left[\frac{a_2}{a_0}\right] + a_3 < 0$$

ie.
$$a_1 a_2 > a_0 a_3$$

It is however impossible to extend this method of correlating the Leonhard and Hurwitz criteria to the general equation of the n^{th} order.

The use of either the Routh or Hurwitz criterion for design purposes is limited. They may be extended however to lay down conditions on the coefficients, if all the roots are to have either a prescribed ratio of real part to imaginary part (relative damping), or a prescribed value of the real part irrespective of the imaginary part (absolute damping).

This is taken up in Sec. 5.1.

In concluding this Section, mention should be made of one further

adaptation of the generalised Nyquist criterion which requires very little computation in its application, namely that of Demontvignier and Lefèvre.²⁸ This is based wholly on the phase-frequency characteristic, from which, with the aid of a Table, the number of half-turns which the function makes round the origin, can be found.

4.3. The Case of Finite Time Lag.

Thus far the transfer-functions which have resulted have been quotients of rational integral functions. If the possibility of finite time-lag is admitted, transfer functions of the type $\varepsilon^{-Tp} \cdot Y_i(p)$ are formed. Further, as the finite time-lag can occur in either the main sequence or in the subsidiary loops or both, a considerable variety of problems is presented. This can be organised either on the basis of control type, number of exponential lags and location of finite time-lags, or on the basis of the resulting overall differential equation. As the first arrangement, besides being more complex, leads to the same overall differential equation for a number of different connections of transfer functions, the second one is the more logical. The types of equation may therefore be listed as follows.

I. Having one exponential term

$$1. a_0 p^2 + a_1 p + a_2 + \alpha p^m \varepsilon^{-Tp}, \quad \text{for } m = 0, 1, 2^*$$

$$2. a_0 p^3 + a_1 p^2 + a_2 p + a_3 + \alpha p^m \varepsilon^{-Tp}, \quad m = 0, 1, 2, 3.$$

and so on for higher degree of the polynomial part of the equation.

* For $m > 2$ the equation always gives an unstable solution. Similarly in the cubic case, type 2, with $m > 3$.

As a subclass of I, may be listed those in which $m=0$ and the constant term of the equation is absent. This will be the case in the overall equation of a single-loop system having the transfer function

$$\frac{\theta_o(p)}{\epsilon} = \frac{\epsilon^{-Tp} \cdot K}{p(1+pT_1)(1+pT_2)}$$

II. Having two exponential terms, with different finite time-lags in general.

The possible combinations are very large, but it is likely that the majority will always be unstable. Practical cases could arise however, as in a two loop system having a main sequence transfer function

$Y_1(p) = K\epsilon^{-Tp}/p(1+pT_1)$, and a feedback transfer function $Y_2(p)$. For this

we have

$$\frac{\theta_o(p)}{\epsilon} = \frac{K\epsilon^{-Tp}}{p(1+pT_1)} \cdot \frac{1}{1 + K\epsilon^{-Tp}Y_2(p)/p(1+pT_1)}$$

and the characteristic equation

$$p^2T_1 + p + Y_2(p)K\epsilon^{-Tp} + K\epsilon^{-Tp} = 0$$

The types occurring in the classification II must necessarily be treated as they arise and some difficulty may be expected in the solution of the problem.

With regard to the slightly simpler class I, Gork²⁹ has treated the stability of the equations $10p + 1 + 9\epsilon^{-2p}$ and $10p + 1 + 6\epsilon^{-2p}$ showing that they are unstable and stable respectively. Sherman³² has given conditions for the stability of the equation

$$a_0p^2 + a_1p + a_2 + \alpha p\epsilon^{-p}$$

which agree with those of Minorsky.³³ Ansoff and Krumhansl³⁴ have extended

this case to cover the equation

$$a_0p^2 + a_1p + a_2 + \alpha p^m \epsilon^{-Tp}, \quad \text{for } m \geq 0$$

4.3. Numerical example.

In this Section, the equation

$$\rho^3 + 5\rho^2 + 5\rho + 1 + 10^{-2}\rho \quad (27)$$

is examined as an example illustrating the use of the fundamental theorem for transcendental functions. A few notes on the application of the theorem are first given.

1. Equations of the type

$$a_0\rho^n + a_1\rho^{n-1} + \dots + a_{n-1}\rho + a_n + \alpha\rho^m \varepsilon^{-T\rho} = 0, \quad m \leq n$$

may be treated directly in the form of the L.H.S. of the equation, i.e. a rational integral function in ρ plus the term $\alpha\rho^m \varepsilon^{-T\rho}$. Such a form tends to the value $a_0\rho^n$ for values of ρ on the infinite arc of Fig.13 but has no poles elsewhere in the finite part of the right half ρ -plane.

2. It is convenient to express the equation however as

$$a_0\rho^n + a_1\rho^{n-1} + \dots + a_{n-1}\rho + a_n \left[1 + \frac{\alpha\rho^m \varepsilon^{-T\rho}}{a_0\rho^n + a_1\rho^{n-1} + \dots + a_{n-1}\rho + a_n} \right] = 0,$$

and investigate the zeros of the expression in brackets. This has the advantage of separating out the effect of the time lag. It also in the author's opinion makes plotting much easier.

If
$$F(\rho) = \left[1 + \frac{\alpha\rho^m \varepsilon^{-T\rho}}{a_0\rho^n + a_1\rho^{n-1} + \dots + a_{n-1}\rho + a_n} \right] \quad (28)$$

then for values of ρ along the infinite arc $\rho = R\varepsilon^{j\phi}$, $-\pi/2 < \phi < \pi/2$, $F(\rho)$ tends to the value unity as R increases indefinitely. For values of ρ on an infinitely small semi-circle passing to the right of the origin, $F(\rho)$ behaves like $1 + \frac{\alpha\rho^m}{a_n}$. For indefinitely large values of ρ on the imaginary axis, $F(j\omega)$ tends to unity, for $m < n$, and to $1 + \frac{\alpha\varepsilon^{-jT\omega}}{a_0}$ for $m = n$.

If however $m > n$, say $m = n + q$, then $F(j\omega)$ tends to $1 + \frac{\alpha}{a_0} (j\omega)^q \xi^{-jT\omega}$ as $\omega \rightarrow \infty$ and hence becomes indefinitely large. Under these conditions the locus is bound to include the origin and the system will be unstable. This has been referred to already.

The important point that $F(p)$ may now have poles anywhere in the finite p -plane must also not be overlooked. This can easily be checked by applying the Routh criterion or the Leonhard criterion to the rational integral function in the denominator.

It is interesting to note that these poles cannot be taken out by inverting the expression $\alpha p^m \xi^{-Tp} / [a_0 p^n + a_1 p^{n-1} + \dots + a_n]$, i.e. by using the function

$$1 + \frac{a_0 p^n + a_1 p^{n-1} + \dots + a_{n-1} p + a_n}{\alpha p^m \xi^{-Tp}}$$

as an essential singularity occurs at infinity, and the theorem does not apply. This is in contrast to the case $1 + 1/\sqrt{G(p)}$ for a multi-loop system pointed out previously.

Consider then the equation,

$$p^3 + 5p^2 + 5p + 1 + 10\xi^{-2p} = 0$$

Application of the Routh criterion to $p^3 + 5p^2 + 5p + 1$ indicates that there are no zeros of this in the right half-plane, and that therefore

$$F(p) = 1 + \frac{10\xi^{-2p}}{p^3 + 5p^2 + 5p + 1} \quad (29)$$

has no poles in that area. We note also that $F(p)$ vanishes for infinitely large imaginary values of p and also along the infinite arc. The locus along the imaginary axis will therefore constitute the entire plot. In

Fig. 19, the locus

$$F(j\omega) - 1 = \frac{10\varepsilon^{-2j\omega}}{1 - 5\omega^2 + j(5\omega - \omega^3)}$$

is shown. For the entire range $\infty > \omega > -\infty$, the point $[-1, 0]$

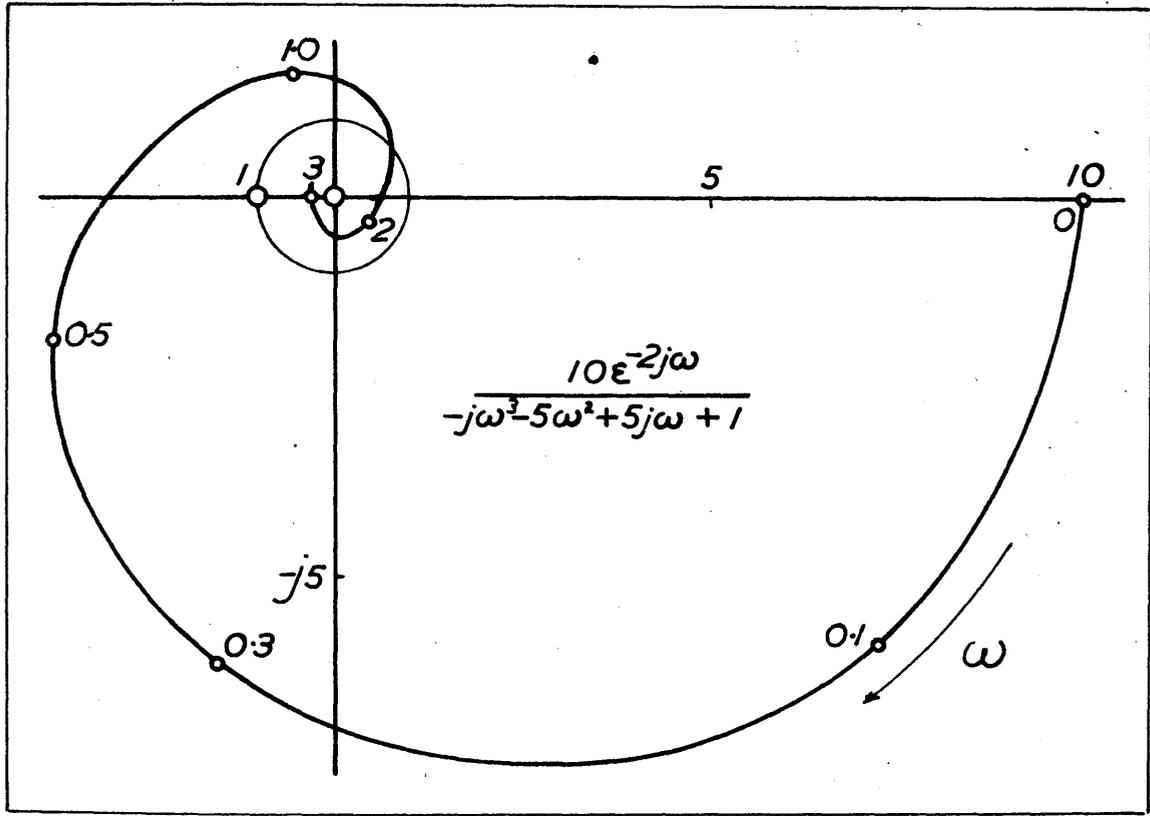


Fig. 19.

is enclosed twice in an anti-clockwise direction, and hence the equation has two unstable roots. More details can be taken from the diagram if required. As the locus merely moves radially as the coefficient of the exponential term alters, it can readily be estimated approximately

- (i) that $2.565\varepsilon^{-2p}$ renders the system just stable and that
- (ii) $45.6\varepsilon^{-2p}$ will introduce two more unstable roots.

By calculating the frequency value for which the unit circle intersects the curve and by estimating the rotation necessary to make this point coincide with $[-1,0]$, it can also be found

(iii) that $10\epsilon^{-0.256p}$ will produce limiting stability and that

(iv) $10\epsilon^{-4.44p}$ will introduce two more unstable roots.

The diagram illustrates in an excellent way the manner in which an infinitely large number of unstable roots will be obtained as the coefficient of the exponential term increases without limit, for one-by-one, the infinite number of turns which $F(j\omega)$ makes about the origin will expand to enclose the point $[-1,0]$. In view of these design advantages, this type of diagram is much to be preferred over the method of Görk.

This Section concludes the review of those criteria designed to test limiting stability.

CHAPTER 5.

CRITERIA OF RELATIVE STABILITY. DESIGN PROCEDURES.

The previous Chapter examined methods of ascertaining whether stable or unstable operation would take place in any particular system. Those methods were described as criteria of limiting stability, in that they only provided information concerning the existence or non-existence of stable operation, and if the former was the case, no indication was obtained of the degree or margin of stability. By a simple extension, however, this extra detail may be predicted. Such extensions of the previous methods may be termed criteria of relative stability. Theoretically the development is trivial but practically it is of some consequence. In this Chapter, a review of these criteria is first given and this is followed by a summary of the available design procedures.

5.1. Extensions of Previous Criteria of Limiting Stability.

The simple criteria already considered are, in effect, means of determining whether there are any roots of the characteristic equation of the system, which have negative damping, that is, roots having positive real parts. The criteria, which we now seek, are means of determining whether there are any roots having damping less than some prescribed positive value. In comparison with the contour of Fig. 13 which bounded the right half of the complex plane, the contour now considered extends over the imaginary axis as shown in Fig. 20.

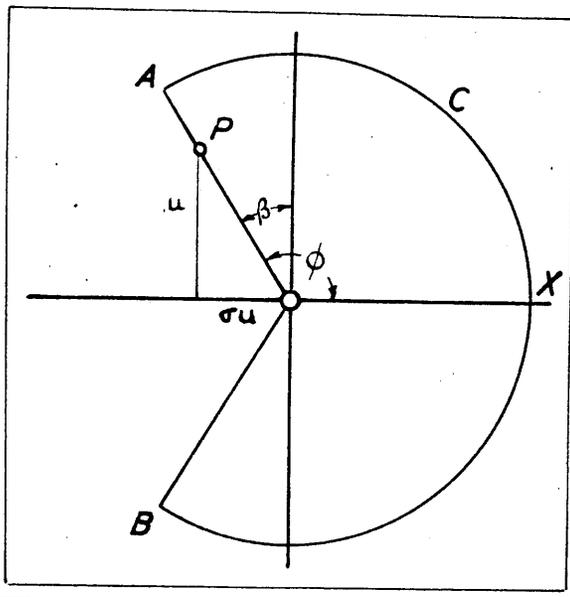


Fig. 20. Contour from which roots of characteristic equation should be excluded, for relative damping to exceed a prescribed value.

In the above contour, values of ρ along the ray making the angle β with the imaginary axis in the upper half-plane are given by $\rho = u(-\sigma + j)$, where $\sigma = \tan \beta$. As the parameter u is increased from zero, so P moves out along the ray. As β increases from zero to $\pi/2$ the ray pivots about the origin from a position of co-incidence with the positive imaginary axis, to co-incidence with the negative real axis, the lower ray remaining at all times the mirror image of the upper one in the negative real axis. In this way the whole of the left half-plane may be investigated. The values of ρ in this more general case may be said to represent the use of complex frequencies by analogy with the values of ρ on the imaginary axis only, which represent real frequencies. Further, the adoption of the above contour, on the finite portion of which the ratio of damping to frequency is constant, lays down a fixed minimum value of relative, as distinct from absolute, damping. A contour to represent the absolute

damping δ say, merely requires the infinite arc to be closed by a line parallel to the imaginary axis and displaced δ to its left. Technically, relative damping is more important as it assigns a definite value to the number of cycles or half-cycles of oscillation which are possible in a transient before it is damped to some fraction of its initial value.

For instance, if this fraction was given by $\epsilon^{-\pi}$ (about 4.3%) we have,

$$(\text{initial value}) \epsilon^{-\sigma u t} = (\text{final value}),$$

where
$$\frac{\text{final value}}{\text{initial value}} = \epsilon^{-\pi}$$

The number of half-cycles of oscillation occurring in the time $t = \pi/\sigma u$ is given by $\pi/\sigma u / \pi/u = 1/\sigma = \cot \beta$, which is constant for a fixed value of β . There are therefore 2, 1, and 0.5 half-cycles of oscillation while the amplitude reduces to 4.3% of its initial value, for $\beta = 26.5^\circ$, 45° and 63.43° respectively.

The extensions of the previous criteria, which are now briefly given, merely consist in the application of the fundamental theorem to the contour of Fig. 20. As analogous results may be derived for the case of absolute damping, details of this are omitted.

1. Extension in Nyquist Form.

Let $Y(\rho)$ be the resultant transfer function $\theta_o/\epsilon(\rho)$ of a multi-loop servo-mechanism. $Y(\rho)$ is of the general form

$$\frac{Y_1(\rho)}{1 + Y_1(\rho)Y_2(\rho) + Y_1(\rho)Y_3(\rho) \dots}$$

where $Y_1(\rho)$, $Y_2(\rho)$, . . . etc. are the transfer functions indicated in Fig.14 and have neither poles nor zero in the right half-plane. Let P' be the number of poles of $1 + Y_1(\rho)Y_2(\rho) + Y_1(\rho)Y_3(\rho) \dots$ within the contour C of Fig.20. P' can usually be determined by inspection. Let Z' be the number of zeros

of $1 + Y_1(p)Y_2(p) + \dots$ within C and N' the number of anti-clockwise revolutions of $1 + Y_1(p)Y_2(p) + \dots$ about the origin as p goes once round C in the positive* direction. Then by the theorem

$$\begin{aligned} Z' - P' &= N' \\ \text{and hence } Z' &= N' + P' \end{aligned}$$

Since the zeros of $1 + Y_1(p)Y_2(p) + \dots$ are the poles of $Y(p)$, the theorem may now be applied to $Y(p)$. Let N be the number of anti-clockwise revolutions of $Y(p)$ about the point $[-1, 0]$, as p goes once round the contour. Then the number of zeros Z , of $Y(p)$ within C is given by

$$\begin{aligned} Z - Z' &= N \\ \text{and hence } Z &= N + Z' \end{aligned}$$

For stability therefore $N = -Z'$.

The case of the single loop servo-mechanism is simpler, in that a preliminary application of the theorem to determine the poles of $Y(p)$ within C is unnecessary as these may be found by inspection. It is to be noted that poles of $Y_1(p)$ cannot occur in the right half-plane but they may occur within the contour C of Fig. 20. This is a slight complication compared with the Nyquist criterion of limiting stability for a single-loop system.

For the same reason the use of the inverse function $e/\theta_0(p)$ is here no simpler than the direct form using $\theta_0/e(p)$.

2. Extension in Leonhard Form.

The function representing the left-hand side of the characteristic equation is

$$H(p) = a_0 p^n + a_1 p^{n-1} + a_2 p^{n-2} + \dots + a_{n-1} p + a_n$$

* The interior of the contour being on the left hand side of the tracing point.

$H(p)$ has no poles in the finite p -plane. Let N be the number of anti-clockwise revolutions of $H(p)$ as p traces out the contour C in the positive direction and Z be the number of zeros of $H(p)$ within C . Then according to the theorem $Z = N$.

But since $H(p)$ behaves like $a_0 p^n$ as $p \rightarrow \infty$, $H(p)$ will rotate anti-clockwise by $n(\pi + 2\beta)$ radians as p moves round the infinite arc BCA . Let ψ be the anti-clockwise rotation (in radians) of $H(p)$ as p moves along the rays from A to O and from O to B . Then

$$Z = N = \frac{n(\pi + 2\beta) + \psi}{2\pi}$$

Let γ be the anti-clockwise radians rotation of $H(p)$ as p moves from O to A . Then $\gamma = -\psi/2$ and $\psi = -2\gamma$. The number of zeros of $H(p)$ within C is therefore given by

$$Z = -\gamma/\pi + n/2 + n\beta/\pi \quad (30)$$

and the number of zeros of $H(p)$ within the V-shaped area AOB is

$$n - Z = \gamma/\pi + n/2 - n\beta/\pi \quad (31)$$

The relations (30) and (31) are those given by Leonhard²⁷.

The use of complex frequencies in the manner of the above two extensions appears to have occurred to a number of different authors, which is not surprising in view of the trivial mathematical difference from the use of real frequencies. Vazsoni³⁰ in 1949 stated the Nyquist extension for the case of a single-loop servo-mechanism, without mention of the very similar work of Leonhard in 1948. Prior to this Campbell³⁵ and Profos³⁶ in 1945, had separately been using complex frequencies. In their work, the left hand of the p -plane is mapped using rectangular co-ordinates, that is, different fixed values of absolute damping.

By this method Campbell showed how an approximate unit-step function response could be derived from the Nyquist diagram. The Tables in Appendix II were prepared by the writer in 1948 to help in the calculation of a Nyquist diagram for complex frequencies. This in fact is one of the practical difficulties which beset the investigation of complex frequency loci.* Neither Vazsoni or Leonhard[†] made reference to an early paper of considerable theoretical and technical value by Lüthi²⁵ in 1942. In this apparently little known paper, Luthi has anticipated both Leonhard's contributions of 1944²⁶ and 1948²⁷, and furthermore has given an extension of the Hurwitz Criterion to the case of relative stability. For the sake of completeness this is given below, together with an improved result expressed in terms of Routhian test functions.[‡]

* Kusters and Moore⁴⁹ give charts for the decibel-modulus and phase-shift at complex frequencies in a very recent paper presented at the D.S.I.R. Conference in July 1951.

† Leonhard⁴⁸ refers to the work of Luthi in his recent paper presented at the above D.S.I.R. conference.

‡ Due to Mr. Babister of the Aeronautics Department, Glasgow University (as yet unpublished).

3. Extension in Hurwitz form.

Let the equation whose roots all lie outside the contour C of Fig.

20 be

$$H(p) = a_0 p^n + a_1 p^{n-1} + a_2 p^{n-2} + \dots + a_{n-1} p + a_n = 0 \quad (26)$$

If the roots of (26) are z_1, z_2, \dots, z_n , then

$$H(p) = a_0 \prod_{r=1}^n (p - z_r)$$

Let these roots first be rotated clockwise by the angle β , giving the polynomial

$$H_1(p) = a_0 \prod_{r=1}^n (p - z_r \varepsilon^{-j\beta})$$

Secondly, let the roots be rotated anti-clockwise giving the polynomial

$$H_2(p) = a_0 \prod_{r=1}^n (p - z_r \varepsilon^{j\beta})$$

Both the above operations will result in the roots being in the negative half-plane, although not occurring in conjugate complex pairs. $H_1(p)$ and $H_2(p)$ are therefore polynomials with complex coefficients. If however, the equation

$$H_1(p) \cdot H_2(p)$$

of degree $2n$ is formed, then its roots still lie in the negative half-plane and occur in conjugate complex pairs. Thus the equation

$$a_0^2 \prod_{r=1}^n (p - z_r \varepsilon^{-j\beta}) (p - z_r \varepsilon^{j\beta}) \quad (32)$$

represents a stable system and has real coefficients. It will satisfy the ordinary Hurwitz conditions.

The extended criterion applied to any equation*

$$a_0 p^n + a_1 p^{n-1} + a_2 p^{n-2} + \dots + a_{n-1} p + a_n = 0$$

of the n^{th} degree, is therefore as follows.

* All the coefficients a_0, a_1, \dots, a_n are positive, otherwise the system is unstable. The requirement $a_0, a_1, \dots, a_n > 0$ is a necessary but insufficient condition for stability.

- (i) Form the equation of degree $2n$, having n roots equal to the roots of the original equation rotated through $+\beta$ radians, and n roots equal to the roots of the original equation rotated through $-\beta$ radians. This is the equation

$$\left(\sum_{r=0}^n a_r \varepsilon^{jr\beta} p^{n-r} \right) \left(\sum_{r=0}^n a_r \varepsilon^{-jr\beta} p^{n-r} \right) = 0$$

Let this equation when written out be given by

$$b_0 p^{2n} + b_1 p^{2n-1} + b_2 p^{2n-2} + \dots + b_{2n-1} p + b_{2n} = 0$$

- (ii) Apply the normal Hurwitz Criterion to the set of coefficients b_0, b_1, \dots, b_{2n} of the above equation.

The most unfortunate feature of this extended criterion is the doubling of the degree of the original equation and consequent difficulty in evaluating the Hurwitz determinants. For the simple cubic equation, a sixth order determinant and its successive corner minors of order five, four, three etc., require to be evaluated. The calculation of the coefficients themselves is not inordinately great. For a cubic we have

$$\begin{aligned} b_0 &= a_0^2 & b_1 &= 2a_0 a_1 \cos \beta \\ b_2 &= a_1^2 + 2a_0 a_2 \cos 2\beta & b_3 &= 2a_0 a_3 \cos 3\beta + 2a_1 a_2 \cos \beta \\ b_4 &= a_2^2 + 2a_1 a_3 \cos 2\beta & b_5 &= 2a_2 a_3 \cos \beta \\ b_6 &= a_3^2 \end{aligned}$$

It is not likely, in view of this doubling of the order of the determinants, that this criterion will be of much practical value.

4. Extension in Routh form.

In this form the determinants which require to be greater than zero are all of the third order, and there are $\frac{n(n-1)}{2}$ of them to be evaluated for an equation of the n^{th} degree. The extended criterion is as follows:-

The necessary and sufficient conditions that the polynomial

$$H(\rho) = a_0 \rho^n + a_1 \rho^{n-1} + a_2 \rho^{n-2} + \dots + a_{n-1} \rho + a_n = 0$$

(where a_0, a_1, \dots are real, a_0 being positive) should have no roots* in the sector AOBCA of the complex plane (see Fig. 20), are that B_0, C_0, D_0, \dots should all be positive, where

$$(i) \quad A_0 = a_0 \sin^2 n\phi, \quad A_1 = a_1 \sin n\phi \sin(n-1)\phi, \quad A_2 = a_2 \sin n\phi \sin(n-2)\phi, \quad \dots$$

$$A_{n-1} = a_{n-1} \sin n\phi \sin\phi, \quad A_m = 0, \text{ for } m > n-1$$

$$(ii) \quad B_0 = a_1 \sin\phi, \quad B_1 = a_2 \sin 2\phi, \quad B_2 = a_3 \sin 3\phi, \quad \dots$$

$$B_{n-1} = a_1 \sin n\phi, \quad B_m = 0, \text{ for } m > n-1$$

$$(iii) \quad C_0 = - \begin{vmatrix} A_0 & A_1 & A_2 \\ B_0 & B_1 & B_2 \\ 0 & B_0 & B_1 \end{vmatrix}, \quad C_1 = - \begin{vmatrix} A_0 & A_1 & A_3 \\ B_0 & B_1 & B_3 \\ 0 & B_0 & B_2 \end{vmatrix}, \quad \dots \quad C_m = - \begin{vmatrix} A_0 & A_1 & A_{m+2} \\ B_0 & B_1 & B_{m+2} \\ 0 & B_0 & B_{m+1} \end{vmatrix}$$

$$C_m = 0, \text{ for } m > n-2$$

$$(iv) \quad D_0 = - \begin{vmatrix} B_0 & B_1 & B_2 \\ C_0 & C_1 & C_2 \\ 0 & C_0 & C_1 \end{vmatrix}, \quad D_1 = - \begin{vmatrix} B_0 & B_1 & B_3 \\ C_0 & C_1 & C_3 \\ 0 & C_0 & C_2 \end{vmatrix}, \quad \dots \quad D_m = - \begin{vmatrix} B_0 & B_1 & B_{m+2} \\ C_0 & C_1 & C_{m+2} \\ 0 & C_0 & C_{m+1} \end{vmatrix}$$

$$D_m = 0, \text{ for } m > n-3$$

The determinant for E_m can be deduced from that for D_m by replacing B by C and C by D . Similarly for F_m etc.

The array of test-functions is in the general case formidable. If we consider the cubic equation, and compare this method with the previous one, we obtain

* The number of roots within the sector may also be found as shown by Mr. Babister.

$$A_0 = a_0 \sin^2 n\phi \quad , \quad A_1 = a_1 \sin n\phi \sin (n-1)\phi \quad , \quad A_2 = a_2 \sin n\phi \sin (n-2)\phi$$

$$B_0 = a_1 \sin \phi \quad , \quad B_1 = a_2 \sin 2\phi \quad , \quad B_2 = a_3 \sin 3\phi$$

$$C_0 = - \begin{vmatrix} A_0 & A_1 & A_2 \\ B_0 & B_1 & B_2 \\ 0 & B_0 & B_1 \end{vmatrix} \quad , \quad C_1 = - \begin{vmatrix} A_0 & A_1 & 0 \\ B_0 & B_1 & 0 \\ 0 & B_0 & B_2 \end{vmatrix} \quad , \quad C_2 = 0$$

$$D_0 = - \begin{vmatrix} B_0 & B_1 & B_2 \\ C_0 & C_1 & 0 \\ 0 & C_0 & C_1 \end{vmatrix} \quad , \quad D_1 = 0$$

If B_0, C_0 and D_0 are all positive, there are no roots within the sector. Comparison with the previous criterion clearly shows the Routh extension to be superior, as it requires the calculation of six coefficients and the evaluation of three third order determinants. The Hurwitz extension requires about the same number of coefficients and the evaluation of a sixth order, fifth order, fourth order etc. determinant.

In a quartic equation the Routh extension results in six third order determinants, which is much preferable to the set of determinants up to the eighth order given by the Hurwitz extension.

5.2. Design Procedures.

The previous criteria are by themselves no guarantee that the performance of the system will be acceptable. In the response to a step-function input, a certain degree of damping will be known to exist, but the actual height of the overshoot will be as yet unknown. In the criteria obtained so far, no indication is present as to whether the roots are real or complex. It would be possible for a relatively small real root to be present and cause the response to persist longer than permitted.

This could, of course, be separately checked by applying an extended criterion of absolute damping (for instance by the substitution of $-k+p$ in place of p and application of the Hurwitz criterion). There are, however, several design procedures which deal with the problem more directly. It is proposed only to review the two main methods briefly. The first of these is based upon the frequency response of the system and may be carried out in a number of ways differing only in slight details of calculation and graphical work. The second method is based upon the response to a unit-step function input.

1. Frequency Response Basis. Q - contours.

A Q-contour is the locus of points on a Nyquist, inverse Nyquist, or other design plane, at which the value of $|\theta_o/\theta_i(j\omega)|$ is constant, the particular magnitude of $|\theta_o/\theta_i(j\omega)|$ being known as the Q-value. If a set of Q-contours for the values $Q = 1, 1.1, 1.2$ etc. are placed upon a Nyquist diagram, for example, the manner in which the Q-value is affected by changing values of $\theta_o/\theta_i(j\omega)$ is made evident. The maximum Q-value may, therefore, be controlled at the same time as the open-loop transfer function is designed and sufficient relative stability can be obtained. The Q-contours for the various design planes are now given. As a formidable literature^{2,4,5} exists on this subject, the results only are given here with a few explanatory notes.

(a) Polar Diagram of $\theta_o/\epsilon(j\omega)$. (Nyquist Diagram).

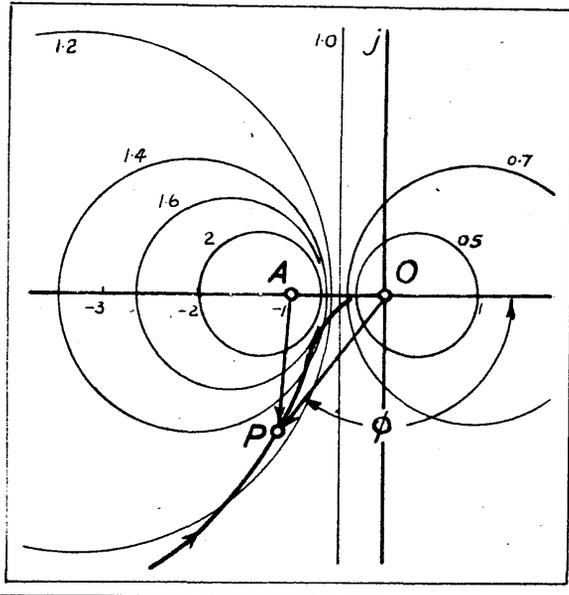


Fig. 21.

$$OP = |\theta_o/\epsilon(j\omega)|$$

$$\phi = \text{angle of } \theta_o/\epsilon(j\omega) \quad ; \quad \bar{AP} = \bar{AO} + \bar{OP} = 1 + \theta_o/\epsilon(j\omega) = \theta_i/\epsilon(j\omega)$$

$$\text{Ratio } OP/AP = |\theta_o/\theta_i(j\omega)| = Q$$

A line of constant Q is therefore the locus of points whose distance from two fixed points are in a constant ratio and is therefore a circle (Circle of Apollonius). The Q -contours form a family of coaxial circles of radii $\left| \frac{Q}{Q^2-1} \right|$ and centres $\left[\frac{-Q^2}{Q^2-1}, 0 \right]$.

(b) Polar Diagram of $\epsilon/\theta_0(j\omega)$. (Inverse Nyquist Diagram).

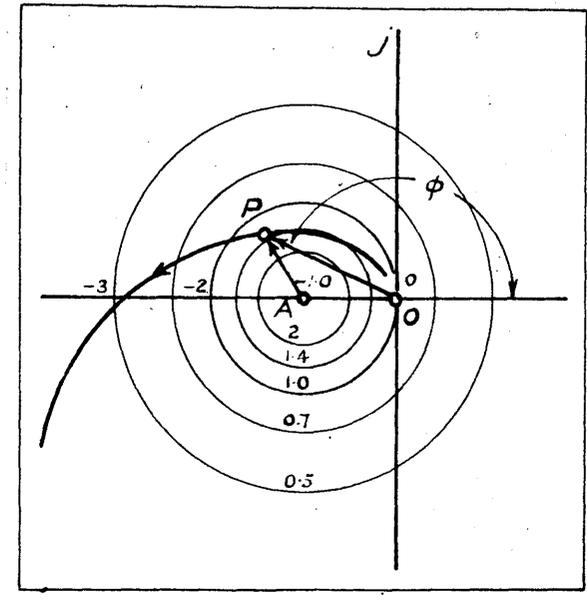


Fig. 22.

$$OP = |\epsilon/\theta_0(j\omega)|$$

$$\phi = \text{angle of } \epsilon/\theta_0(j\omega)$$

$$\overline{AP} = \overline{AO} + \overline{OP} = 1 + \epsilon/\theta_0(j\omega) = \theta_i/\theta_0(j\omega)$$

Hence length of $\overline{AP} = |\theta_i/\theta_0(j\omega)| = 1/Q$. The Q-contours in this case are a family of concentric circles of radii $1/Q$ and centres at $[-1, 0]$.

(c) Cartesian Diagram of $20 \log_{10} |\theta_{o/\epsilon}(j\omega)|$ against phase angle of $\theta_{o/\epsilon}(j\omega)$.
(Log-modulus - phase angle diagram).

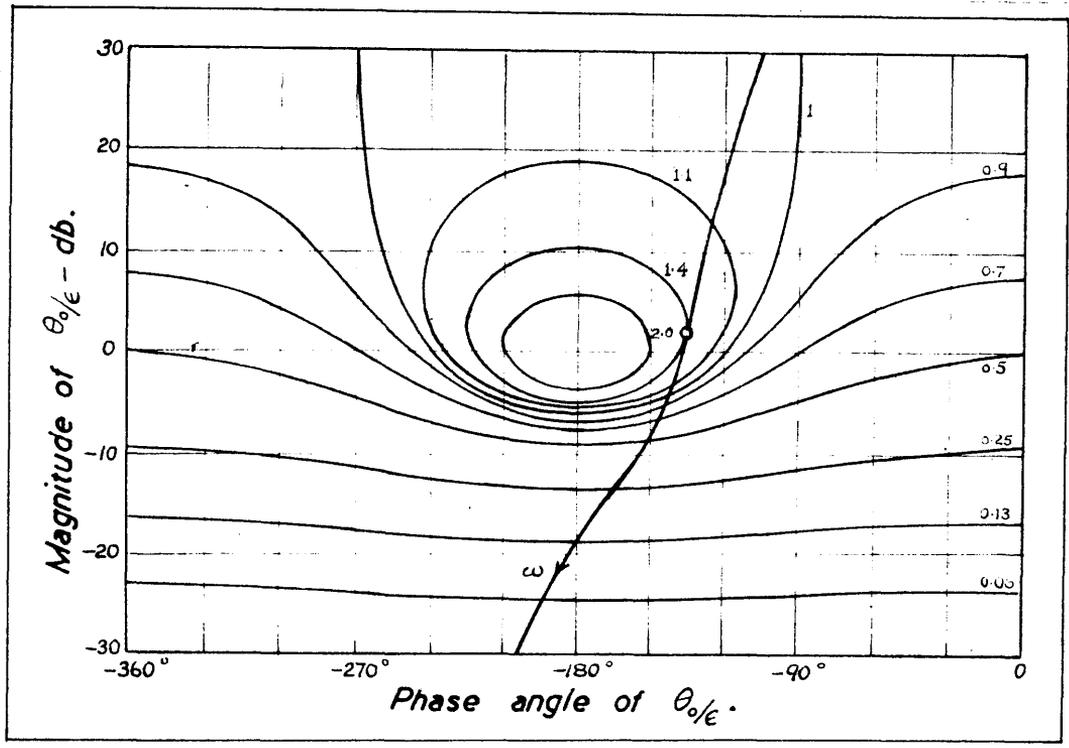


Fig. 23. Log-modulus phase-angle diagram.

The Q-contours may be derived from those in Fig. 21 by plotting the decibel magnitude of the vector OP as it traces out the Q-circle of that figure. More accurately, the magnitude can be calculated analytically and converted to decibels.

The use of the Q-contours on any of the diagrams is either to fix the gain factor so that a certain maximum Q will not be exceeded or, more frequently, to design the local 'dent' required in that part of the locus which approaches the critical point, once the gain factor has already been fixed by consideration of the steady-state errors. The design

procedure will enable a good first approximation to the final system adjustment to be fairly quickly obtained. It is evident, however, that not only the Q-value at the tangent point of the locus and the Q-contour is important, but also the nature of the tangency. If the two curves can be made to approach each other and touch over a range of frequencies, rather than just at one frequency, the system will possess a broader pass band and be more free from resonances. It may on the other hand become susceptible to unwanted inputs and the probability of these occurring must be considered. The use of Q-contours (or M-contours) has received much attention in America but the absence from the frequency-response diagram of any real information about the transient response (i.e. size of overshoot, number of overshoots, duration of decay of final overshoot etc.) leaves much to be desired in this technique.

The method, however, can be extended by the complex-frequency technique and, by using Campbell's³⁵ construction, the principal oscillatory mode may be found. With extra trouble, the additional roots of $1+Y(p)=0$ may be found as shown by Kusters and Moore⁴⁹. Having obtained the roots, however, and having calculated the coefficients of the various terms in the time response of the output, the magnitude of the overshoot is still not explicitly given; for this information the time-response of the output requires evaluation and plotting. An exception to this is the case when the principal mode term accounts for the whole of the overshoot by the time

maximum overshoot occurs. Campbell's method will then give fairly good results in analysing an experimental system. In Part II of the thesis this concept of the principal mode representation is carried further and a method of synthesising a system transfer function for a prescribed transient response is given.

2. Transient Response Basis. Whiteley's Standard Forms.

The classification of servo-mechanisms adopted by Whiteley in his paper is based upon firstly, the type of steady-state error which the system sustains and secondly, the order of the differential equation relating the input and output quantities. If a system has the differential equation

$$(a_0 D^r + a_1 D^{r-1} + \dots + a_{r-1} D + a_r) \theta_o = (b_0 D^s + b_1 D^{s-1} + \dots + b_{s-1} D + b_s) \theta_i \quad (8)$$

in which $s < r$ this yields the overall system transfer function

$$Q(p) = \frac{\theta_o(p)}{\theta_i} = \frac{b_0 p^s + b_1 p^{s-1} + \dots + b_{s-1} p + b_s}{a_0 p^r + a_1 p^{r-1} + \dots + a_{r-1} p + a_r} \quad (11)$$

The forms of $Q(p)$ for zero steady-state displacement, velocity and acceleration error can now be found using the Final Value theorem of Sec.2.6.

The required forms are in fact.

1. for zero steady-state displacement error

$$\frac{b_0 p^s + b_1 p^{s-1} + \dots + b_{s-1} p + a_r}{a_0 p^r + a_1 p^{r-1} + \dots + a_{r-1} p + a_r} \quad (33)$$

2. for zero steady-state velocity error

$$\frac{b_0 p^s + b_1 p^{s-1} + \dots + a_{r-1} p + a_r}{a_0 p^r + a_1 p^{r-1} + \dots + a_{r-1} p + a_r} \quad (34)$$

3. for zero steady-state acceleration error

$$\frac{b_0 p^s + b_1 p^{s-1} + \dots + a_{r-2} p^2 + a_{r-1} p + a_r}{a_0 p^r + a_1 p^{r-1} + \dots + a_{r-2} p^2 + a_{r-1} p + a_r} \quad (35)$$

The equality of the constant terms in the numerator and denominator thus ensures zero displacement error while agreement of the coefficients up to the term in p ensures zero velocity error and so on. A "Standard Form" can now be constructed for any one of the types (33) to (35), having a particular value of r (the order of the differential equation), by establishing once and for all, that set of coefficients which gives a satisfactory step-response. The basis of the forms given by Whiteley, reproduced in Table III is in most cases the limitation of the overshoot. Some of the forms give also a small subsequent undershoot.

In connection with the Table several points require to be noted.

- (a) The values of the coefficients are prescribed in terms of the fixed term in the denominator of $Q(p)$. This constant fixes the time-scale of the response. For example, the Standard Form (60) is

$$\frac{\theta_o(p)}{\theta_i} = \frac{18\omega_o^4 p + \omega_o^5}{p^5 + 9\omega_o p^4 + 29\omega_o^2 p^3 + 38\omega_o^3 p^2 + 18\omega_o^4 p + \omega_o^5} = Q(p)$$

If ω_o is increased to $a\omega_o$, the R.H.S. becomes

$$\frac{a^5 [18\omega_o^4 (p/a) + \omega_o^5]}{a^5 [(p/a)^5 + 9\omega_o (p/a)^4 + 29\omega_o^2 (p/a)^3 + 38\omega_o^3 (p/a)^2 + 18\omega_o^4 (p/a) + \omega_o^5]}$$

which is $Q(p/a)$. The response to a unit step-function is now

given by the inverse transform of $\frac{Q(p/a)}{p} = \frac{Q(p/a)}{a(p/a)}$, compared with

$Q(p)/p$ previously. The Scale Change Theorem¹¹ therefore gives

the output $\theta_o(at)$, i.e. an output a times as fast as $\theta_o(t)$,

but of the same shape. The term ω_o^{11} should, therefore, be

as large as possible.

TABLE III.

Class.	Type and General Form of	Max. Overshoot.	Standard Form - Coefficients of p.
A	Displacement and Zero- Displacement error systems. $\frac{\omega_0^n}{p^n + ap^{n-1} + \dots + \omega_0^n}$	5	$p^2 + 1.4\omega_0 + \omega_0^2$
		8	$p^3 + 2\omega_0 + 2\omega_0^2 + \omega_0^3$
		10	$p^4 + 2.6\omega_0 + 3.4\omega_0^2 + 2.6\omega_0^3 + \omega_0^4$
		Higher orders not checked: use binomial coefficients.	
B	Zero Velocity Error Systems. $\frac{zp + \omega_0^n}{p^n + ap^{n-1} + \dots + zp + \omega_0^n}$	10	$p^2 + 2.5\omega_0 + \omega_0^2$
		10	$p^3 + 5.1\omega_0 + 6.3\omega_0^2 + \omega_0^3$
		10	$p^4 + 7.2\omega_0 + 16\omega_0^2 + 12\omega_0^3 + \omega_0^4$
		10	$p^5 + 9\omega_0 + 29\omega_0^2 + 38\omega_0^3 + 18\omega_0^4 + \omega_0^5$ (60)
		10	$p^6 + 11\omega_0 + 43\omega_0^2 + 83\omega_0^3 + 73\omega_0^4 + 25\omega_0^5 + \omega_0^6$
C	Zero Acceleration Error Systems. $\frac{yp^2 + zp + \omega_0^n}{p^n + ap^{n-1} + \dots + yp^2 + zp + \omega_0^n}$	10	$p^3 + 6.7\omega_0 + 6.7\omega_0^2 + \omega_0^3$
		15	$p^4 + 7.9\omega_0 + 15\omega_0^2 + 7.9\omega_0^3 + \omega_0^4$
		20	$p^5 + 18\omega_0 + 69\omega_0^2 + 69\omega_0^3 + 18\omega_0^4 + \omega_0^5$
		20	$p^6 + 36\omega_0 + 25\omega_0^2 + 485\omega_0^3 + 25\omega_0^4 + 36\omega_0^5 + \omega_0^6$

- (b) It will be seen that the Standard Forms in the Table have their numerators limited to a constant, a linear factor and a quadratic factor for Classes A, B and C respectively. This is not always the case although most practical servo-mechanisms will fall within these categories. In the absence of Standard Forms having numerators in excess of the second degree, this coefficient in a proposed servo design will have to be neglected.
- (c) There is some choice in the Standard Forms available. A system may be analytically of Type 1, i.e. have a velocity error, but this may be so small that a zero-velocity error Form, may be used. Such a difference does not greatly matter, as the Standard Forms themselves do not necessarily represent the final adjustment of the system, but only a good first approximation. It will also be noted that Class A of the Table includes both Types 0 and 1 (in the notation of this work). Formally this is incorrect, but practically it is of little consequence.

In preparing these Standard Forms, Whiteley has allowed the equation $1 + \gamma(p) = 0$, to have only real roots and to obtain the overshoot as a function of only one parameter, various other restrictions have been laid down. These conditions do not enter into the application of the Standard Forms. This method of reducing the problem may be compared with that adopted in Part II, where the existence of at least one complex root-pair is essential to the theory.

In addition to these main design procedures, several other techniques exist. The root-locus method of Evans³⁷ is essentially a means of tracing the movement of the roots of $H(s)$ as the gain constant is varied. Again, specific information regarding the transient response is not easily obtained. Chestnut and Mayer³⁸ have computed a series of charts which enable an open-loop transfer function to be designed for any given value of the resonance peak of $\theta_o/\theta_i(j\omega)$ and for a given resonant frequency, or for a specified transient response. Their method may be criticised on two main points as follows.

- (i) The number of open-loop transfer function types is limited to four, all of which have a basic Type 1 control characteristic. The authors state that other control types may be approximated, as far as step-function response is concerned, by an equivalent Type 1 design. It is doubtful whether this will always be the case.
- (ii) The very large number of charts which are required. For the above four transfer-function types no fewer than 36 charts are given, each of which contains on the average 15 curves.

The above method does admittedly provide correlation between the frequency and transient response, for the forms which are considered. Even so, for the amount of information provided, the number of design charts is excessive.

Allied to the central design object of achieving specified frequency, and more important specified transient responses, are two further problems. These are (i) the correlation of the frequency and transient responses and (ii) the design of control systems in which noise is present in the input signal. The second-mentioned has been and still is the object of much research, and is closely linked with the design criterion of minimum r.m.s. error. Regarding the correlation of the frequency and transient response, a variety of methods are available for the prediction of the transient from the frequency response or vice-versa. In Part III of this thesis, an experimental study of a metadyne-controlled servo-mechanism is given, in which the transient response is (a) measured and (b) computed from the measured frequency response. Reference may be made to Chapter 11 for an explanation of three methods due to Campbell³⁵, Bedford and Fredendall³⁹, and Floyd⁴⁰, by which the conversion from frequency to transient response may be carried out. The two last-mentioned papers also deal with the converse problem.

In conclusion, it may be said that the methods mentioned so far, with the possible exception of that due to Chestnut and Mayer, do not permit the rapid design of servo-mechanism transfer functions from which the maximum overshoot and damping of the unit-step function response is easily predictable. Whiteley's Standard Forms, of course, do give the required transfer functions for specified maximum overshoots, but failure to realize the exact figures demanded by the Standard Forms leaves the question of

possible maximum overshoot unanswered. With the aim of meeting this deficiency, the theory of Part II is now presented and illustrated. The theory contained in this Part is a logical development of the idea of principal mode representation due to Campbell,³⁵ and follows a suggestion made in the conclusion of his paper.

PART II.

THE DESIGN OF LINEAR SERVO-MECHANISMS HAVING PRESCRIBED TRANSIENT
RESPONSES.

INTRODUCTION.

The brief review of design methods given in Chapter 5 has indicated that, so far, considerably more time and thought has been accorded to the frequency response method of investigation and design, than to the transient response method. This has been recently stated in a Progress Review by Whiteley* to whom is due the "Standard Form" Technique for obtaining specified transient responses. The method proposed in this Part has roughly the same object as Whiteley's work, that is, the production of system transfer functions which will give a definite known overshoot in the step-response. It is more flexible, however, than Whiteley's method since, as will be shown later, the overshoot is readily calculable in all cases.

The requisite theory is given in the following Chapter. Chapter 7 contains examples illustrating the method and shows how difficulty may arise in the inherently unstable Type 2 and Type 3 systems. This difficulty is resolved in Chapter 8 by using the transfer function $\epsilon/\theta_i(p)$, instead of the overall system transfer function $\theta_o/\theta_i(p)$. Numerical examples demonstrating this second method are then given, and final conclusions regarding both methods are drawn.

* A.L. Whiteley, Progress Review of Servo-Mechanisms, Proc.I.E.E. vol.98, I, p.289, Sept. 1951.

CHAPTER 6.

THE DESIGN OF $\theta_o/\theta_i(p)$ TRANSFER FUNCTIONS.

6.1. Description of Method.

Before entering into mathematical details, the term "prescribed transient responses" requires to be simply stated. By this is meant

- (a) the number of overshoots and undershoots; for a satisfactory response this is either one overshoot and no subsequent undershoot, or one overshoot with a small subsequent undershoot, possibly of 1 - 2% of the step-function input.
- (b) the size of the maximum overshoot; this will vary according to the application but not normally be allowed to exceed 15%.
- (c) the time at which maximum overshoot occurs, if this may be specified independently.

In the above specification of a transient response, it will be seen that no provision has been made for "dead-beat" response, i.e. response without overshooting. It is a necessity of the simple theory, that some overshoot should take place. There is not, in general, a great deal of difficulty in reducing this to as low a value as desired, and, furthermore, it is likely (as occurs in the results of Part III), that the practical system will in any case show less maximum overshoot than the design on paper indicates.* By using the full theory instead of a simple

* The initial formulation of the transfer function relating either the output or the error to the input will usually neglect physical friction, since the nature and amount of this is generally unknown, and its effect is stabilising in any case.

approximation it is possible to reduce the predicted overshoot to values of the order of a few per cent.

As pointed out in Sec. 2.5 however, there are fundamental limitations to the amount by which overshooting of a servo-mechanism may be reduced once its steady-state following properties are laid down. These are summarised below:-

Type Description	Type No.	Minimum Total of Overshoots and Undershoots in Step-Response.
Displacement-error	0	0
Zero displacement-error	1	0
Zero velocity-error	2	1
Zero acceleration-error	3	2

Since the majority of servo-mechanisms are of Type 1, with a very small velocity-error, or of Type 2, some overshooting is therefore inevitable from theoretical restrictions*. The deficiency of the simple theory in admitting only solutions giving an overshoot is therefore negligible.

6.2. Theoretical Considerations.

The theory is first given in terms of the $\theta_o/\theta_i(\rho)$ transfer function. Later considerations will show that the $e/\theta_i(\rho)$ form is generally easier to work with. This order has been so chosen, however, as the $\theta_o/\theta_i(\rho)$ analysis is not necessarily restricted to servo-mechanisms, but applies to the output-input relationships of linear systems in general. The theory makes use

*It is to be noted that linear systems are being dealt with. The effect of added non-linear factors may be to remove a small overshoot.

of the approximation to a total step-response by the term representing only the principal mode of oscillation, that is, that term having the least damping and, normally, the highest frequency. As far as the author is aware this seems to have been proposed first by Campbell, although others* have also used the idea. As a starting point it may be recalled from Sec. 5.2, that the overall transfer functions of Type 1, 2 and 3 servo-mechanisms are of the following form:-

$$\text{Type 1.} \quad \frac{\theta_o(p)}{\theta_i} = \frac{B_o p^s + B_1 p^{s-1} + \dots + B_{s-1} p + A_r}{A_o p^r + A_1 p^{r-1} + \dots + A_{r-1} p + A_r} \quad (33)$$

$$\text{Type 2.} \quad \frac{\theta_o(p)}{\theta_i} = \frac{B_o p^s + B_1 p^{s-1} + \dots + A_{r-1} p + A_r}{A_o p^r + A_1 p^{r-1} + \dots + A_{r-1} p + A_r} \quad (34)$$

$$\text{Type 3.} \quad \frac{\theta_o(p)}{\theta_i} = \frac{B_o p^s + B_1 p^{s-1} + \dots + A_{r-2} p^2 + A_{r-1} p + A_r}{A_o p^r + A_1 p^{r-1} + \dots + A_{r-2} p^2 + A_{r-1} p + A_r}, \quad \text{all } s < r \quad (35)$$

The Type 0 system can be taken along with the Type 1 system for the present purpose, although strictly speaking the constant terms in the numerator and denominator will only be very nearly equal instead of actually equal as (33) indicates. In (35) it has been stated that s is less than r . It is further true to say that for servo-mechanisms, s is less than $r-1$, in fact $s \leq r-2$. If at first sight cases arise where $s=r-1$, closer examination will show that an approximation has been made in deriving the transfer function. Here we will take it that $s \leq r-2$. This point is of minor importance in this Chapter but requires consideration in Chapter 8. If we use the Type 1 system in order to carry out the argument, we have the output transform in response to a unit-step function input

* Work on similar lines to Sec. 6.3 has been given by Mulligan⁴¹ without special reference to servo-mechanisms.

$$\theta_0(\rho) = \frac{B_0 \rho^s + B_1 \rho^{s-1} + \dots + B_{s-1} \rho + A_r}{A_0 \rho^r + A_1 \rho^{r-1} + \dots + A_{r-1} \rho + A_r} \cdot \frac{1}{\rho}$$

or in terms of the poles and zeros,

$$\theta_0(\rho) = \left[\frac{P_1 P_2 \dots P_r}{Z_1 Z_2 \dots Z_s} \right] \cdot \frac{(\rho + Z_1)(\rho + Z_2) \dots (\rho + Z_s)}{(\rho + P_1)(\rho + P_2) \dots (\rho + P_r)} \cdot \frac{1}{\rho} \quad (36)$$

in which $Z_1 Z_2 \dots$ etc. and $P_1 P_2 \dots$ etc. are minus the zeros and poles respectively and may be real or complex. As a consequence of the real coefficients

B_0, B_1, \dots etc., and A_0, A_1, \dots etc., however, complex poles and zeros must occur in conjugate pairs, and the multiplier $\left[\frac{P_1 P_2 \dots P_r}{Z_1 Z_2 \dots Z_s} \right]$ is wholly real. One

other restriction on the poles and zeros is to be laid down, namely, that they should all lie in the left-half of the complex plane. This must be so for the poles if stability is to be present, and for the zeros it will also be true if we exclude non-minimum phase components. Relation (36) therefore is the general form which all system transfer functions will take, the distinction between Types being as follows.

Types 0 and 1. Equal constant terms in numerator and denominator. This is already assured by the form of (36).

Type 2. Equal constant terms and equal terms in ρ . In terms of the poles and zeros this gives

$$\begin{aligned} (P_1 P_2 \dots P_r) \text{ (sum of products } Z_1 Z_2 \text{ etc. taken } (s-1) \text{ at a time)} &= \\ (Z_1 Z_2 \dots Z_s) \text{ (sum of products } P_1 P_2 \text{ etc. taken } (r-1) \text{ at a time)} & \end{aligned}$$

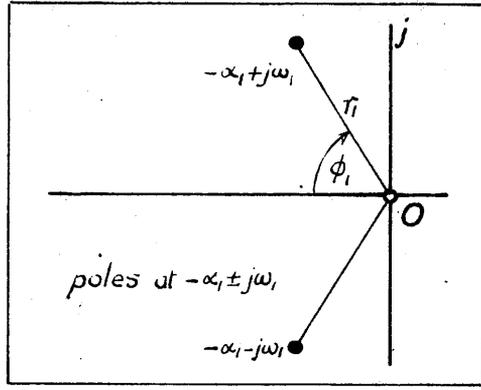
Type 3. Equal constant terms, equal terms in ρ and equal terms in ρ^2 . This gives the two conditions,

$$\begin{aligned} (P_1 P_2 \dots P_r) \text{ (sum of products } Z_1 Z_2 \text{ etc, taken } (s-2) \text{ at a time)} &= \\ (Z_1 Z_2 \dots Z_s) \text{ (sum of products } P_1 P_2 \text{ etc. taken } (r-2) \text{ at a time)} & \end{aligned}$$

The additional limitations occurring in Type 2 and Type 3 systems do not enter into the theory of the method, but only affect its application. Calculation of overshoot in Specific Cases of Relation (36).

1. One Complex Pole-Pair, No Zeros.

Fig. 24.



This is the simplest possible system. Let it be represented by

$$\theta_o(p) = r_1^2 \cdot \frac{1}{(p + \alpha_1 - j\omega_1)(p + \alpha_1 + j\omega_1)} \cdot \frac{1}{p}$$

The pole at $p=0$ gives $\frac{r_1^2}{(\alpha_1 - j\omega_1)(\alpha_1 + j\omega_1)} = 1$

The pole at $p = -\alpha_1 + j\omega_1$ gives $\frac{r_1^2}{2j\omega_1(-\alpha_1 + j\omega_1)} \cdot \varepsilon^{-\alpha_1 t} \cdot \varepsilon^{j\omega_1 t} = \frac{r_1^2 \varepsilon^{-\alpha_1 t}}{\omega_1 r_1 \varepsilon^{j(\pi - \phi_1)}} \cdot \frac{\varepsilon^{j\omega_1 t}}{2j}$

The pole at $p = -\alpha_1 - j\omega_1$ gives the conjugate of the above, i.e. $\frac{r_1^2 \varepsilon^{-\alpha_1 t}}{\omega_1 r_1 \varepsilon^{-j(\pi - \phi_1)}} \cdot \frac{\varepsilon^{-j\omega_1 t}}{-2j}$

The pole-pair $-\alpha_1 \pm j\omega_1$, therefore gives the term

$$\frac{r_1}{\omega_1} \varepsilon^{-\alpha_1 t} \sin(\omega_1 t - \overline{\pi - \phi_1}),$$

and the output is $\theta_o(t) = 1 + \frac{r_1}{\omega_1} \varepsilon^{-\alpha_1 t} \sin(\omega_1 t - \overline{\pi - \phi_1})$. (37)

The error quantity is $e(t) = \theta_i(t) - \theta_o(t)$

$$= -\frac{r_1}{\omega_1} \varepsilon^{-\alpha_1 t} \sin(\omega_1 t - \overline{\pi - \phi_1}).$$
 (38)

The overshoots of the output correspond to the minima of this expression. Thus

$$\begin{aligned} \frac{de(t)}{dt} &= -\frac{\zeta}{\omega_1} \left[-\alpha_1 \varepsilon^{-\alpha_1 t} \sin(\omega_1 t - \bar{\pi} - \phi_1) + \omega_1 \varepsilon^{-\alpha_1 t} \cos(\omega_1 t - \bar{\pi} - \phi_1) \right] \\ &= \frac{\zeta^2}{\omega_1} \varepsilon^{-\alpha_1 t} \left[\frac{\alpha_1}{\zeta} \sin(\omega_1 t - \bar{\pi} - \phi_1) - \frac{\omega_1}{\zeta} \cos(\omega_1 t - \bar{\pi} - \phi_1) \right] \\ &= \frac{\zeta^2}{\omega_1} \varepsilon^{-\alpha_1 t} \sin(\omega_1 t - \bar{\pi} - \phi_1 - \phi_1) \\ &= \frac{\zeta^2}{\omega_1} \varepsilon^{-\alpha_1 t} \sin(\omega_1 t - \bar{\pi}) = 0 \text{ for a minimum or a maximum.} \end{aligned}$$

When $t = \bar{\pi}/\omega_1$, the error is negative. This corresponds to the first overshoot the output, and it is therefore given in magnitude by

$$\frac{\zeta}{\omega_1} \varepsilon^{-\alpha_1 \bar{\pi}/\omega_1} \cdot \sin \phi_1 = \varepsilon^{-\alpha_1 \bar{\pi}/\omega_1} = h_1, \text{ say.} \quad (39)$$

The overshoot quite clearly becomes less as α_1/ω_1 becomes greater, i.e. the pole approaches the negative real axis. This simple case of only one complex pole will seldom if ever occur in practice. It has been taken as a starting point however, as it is the basis of the principal mode approximation. Before working out the next simplest case, it may be noted that the magnitude of the first undershoot is

$$\frac{\zeta}{\omega_1} \varepsilon^{-\alpha_1 2\bar{\pi}/\omega_1} \cdot \sin \phi_1 = \varepsilon^{-\alpha_1 \bar{\pi}/\omega_1} \cdot h_1.$$

The magnitude of each succeeding maximum overshoot or undershoot is given likewise by $\varepsilon^{-\alpha_1 \bar{\pi}/\omega_1}$ of its predecessor. This result holds for all systems in which the principal mode theory applies, irrespective of their complexity.

2. One Complex Pole-Pair, 1 Real Pole, No Zeros.

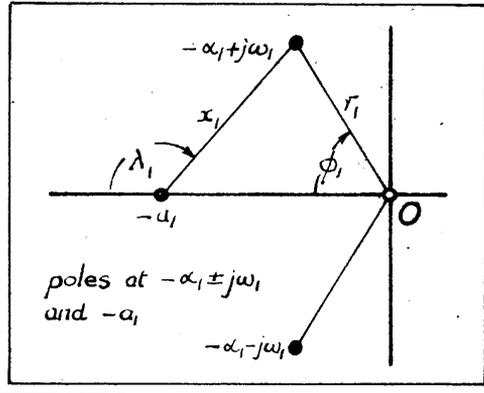


Fig. 25.

The overall system transfer function is given by

$$\frac{\theta_o(p)}{\theta_i} = a_1 r_1^2 \frac{1}{(p + a_1)(p + \alpha_1 - j\omega_1)(p + \alpha_1 + j\omega_1)}$$

Hence

$$\theta_o(p) = a_1 r_1^2 \frac{1}{(p + a_1)(p + \alpha_1 - j\omega_1)(p + \alpha_1 + j\omega_1)p}$$

The pole at $p=0$ gives $\frac{a_1 r_1^2}{a_1(\alpha_1 - j\omega_1)(\alpha_1 + j\omega_1)} = 1$

The pole at $p = -\alpha_1 + j\omega_1$ gives $\frac{a_1 r_1^2 \varepsilon^{-\alpha_1 t} \varepsilon^{j\omega_1 t}}{(-\alpha_1 + j\omega_1 + a_1) 2j\omega_1 (-\alpha_1 + j\omega_1)}$

Since $(-\alpha_1 + j\omega_1 + a_1)$ is the vector from the point $(-\alpha_1 + j \cdot 0)$ to the point $-\alpha_1 + j\omega_1$, this may be written as $x_1 \varepsilon^{j(\pi - \lambda_1)}$. The contribution due to this pole is

therefore $\frac{a_1 r_1^2 \varepsilon^{-\alpha_1 t} \varepsilon^{j\omega_1 t}}{x_1 \varepsilon^{j(\pi - \lambda_1)} \omega_1 r_1 \varepsilon^{j(\pi - \phi_1)} \cdot 2j}$

The pole at $p = -\alpha_1 - j\omega_1$ gives the conjugate, namely.

$$\frac{a_1 r_1^2 \varepsilon^{-\alpha_1 t} \varepsilon^{-j\omega_1 t}}{x_1 \varepsilon^{-j(\pi - \lambda_1)} \omega_1 r_1 \varepsilon^{-j(\pi - \phi_1)} \cdot -2j}$$

from which the complex pole-pair gives

$$\frac{a_1 r_1}{x_1 \omega_1} \cdot \varepsilon^{-\alpha_1 t} \sin(\omega_1 t - \pi - \phi_1 - \pi - \lambda_1)$$

The pole at $\rho = -a_1$ gives $\frac{a_1 \bar{\pi}^2 \xi^{-a_1 t}}{(-a_1 + \alpha_1 - j\omega_1)(-a_1 + \alpha_1 + j\omega_1) - \bar{\alpha}_1}$

Since $-a_1 + \alpha_1 - j\omega_1 = -x_1 \xi^{j(\bar{\pi} - \lambda_1)}$ and

$-a_1 + \alpha_1 + j\omega_1 = -x_1 \xi^{-j(\bar{\pi} - \lambda_1)}$, the term reduces to $-\frac{\bar{\pi}^2}{x_1^2} \xi^{-a_1 t}$.

The output is therefore $1 + \frac{a_1 \bar{\pi}}{x_1 \omega_1} \xi^{-\alpha_1 t} \sin(\omega_1 t - \bar{\pi} - \phi_1 - \bar{\pi} - \lambda_1) - \frac{\bar{\pi}^2}{x_1^2} \xi^{-a_1 t}$, (40)

and the error is $-\frac{a_1 \bar{\pi}}{x_1 \omega_1} \xi^{-\alpha_1 t} \sin(\omega_1 t - \bar{\pi} - \phi_1 - \bar{\pi} - \lambda_1) + \frac{\bar{\pi}^2}{x_1^2} \xi^{-a_1 t}$. (41)

It is clear that if the principal mode term is to represent the total error by the time the first minimum of (41) is reached, a_1 must certainly be greater than α_1 . Omitting such considerations meanwhile and assuming that the principal mode term does in fact represent the entire error at the time of maximum overshoot, the size of this overshoot can be worked out similarly to the previous example. We have

$$\begin{aligned} \frac{d\epsilon(t)}{dt} &= -\frac{a_1 \bar{\pi}}{x_1 \omega_1} \left[-\alpha_1 \xi^{-\alpha_1 t} \sin(\omega_1 t - \bar{\pi} - \phi_1 - \bar{\pi} - \lambda_1) + \omega_1 \xi^{-\alpha_1 t} \cos(\omega_1 t - \bar{\pi} - \phi_1 - \bar{\pi} - \lambda_1) \right] \\ &= \frac{a_1 \bar{\pi}^2}{x_1 \omega_1} \xi^{-\alpha_1 t} \left[\frac{\alpha_1}{\bar{\pi}} \sin(\omega_1 t - \bar{\pi} - \phi_1 - \bar{\pi} - \lambda_1) - \frac{\omega_1}{\bar{\pi}} \cos(\omega_1 t - \bar{\pi} - \phi_1 - \bar{\pi} - \lambda_1) \right] \\ &= \frac{a_1 \bar{\pi}^2}{x_1 \omega_1} \xi^{-\alpha_1 t} \sin(\omega_1 t - \bar{\pi} - \bar{\pi} - \lambda_1) = 0 \text{ for a min. or a max.} \end{aligned}$$

At $t = \frac{1}{\omega_1} (\bar{\pi} + \bar{\pi} - \lambda_1)$, the principal mode term has a minimum value whose magnitude is $\frac{a_1 \bar{\pi}}{x_1 \omega_1} \xi^{-\alpha_1 (\bar{\pi} + \bar{\pi} - \lambda_1) / \omega_1} \sin \phi_1$

$$= \frac{a_1}{x_1} \xi^{-\alpha_1 (\bar{\pi} + \bar{\pi} - \lambda_1) / \omega_1}$$

$$= \xi^{-\alpha_1 \bar{\pi} / \omega_1} \cdot \frac{1}{x_1 / a_1 \xi^{+\alpha_1 (\bar{\pi} - \lambda_1) / \omega_1}} \quad (42)$$

Comparison with (39) indicates that the effect of this added real pole has been to divide the overshoot in (39) by the amount $\chi/a_1 \varepsilon^{\alpha_1(\bar{n}-\lambda_1)/\omega_1}$. For points on the real axis this quantity is always greater than unity, which is the value it tends to assume as a_1 becomes very large, that is, the pole moves out to minus infinity.

Again, this is a very simple example but it nevertheless illustrates the essential step in obtaining the modification to the overshoot of a simple quadratic response, as a result of an added real pole, provided that the principal mode by itself constitutes that overshoot.

6.3. Generalisation of the Above Result.

The generalised overall transfer function may contain real poles, real zeros, pairs of complex poles, and pairs of complex zeros, all being present in any number provided that the degree of the denominator exceeds that of the numerator by two. There is also the possibility of repeated poles and repeated zeros, although it will be shown later that repeated poles are to be discouraged from the point of view of the principal mode approximation. The question of repeated poles and zeros is deferred meanwhile.

Before giving the general theory, the simple notation already used will be extended and defined. This is most easily understood from the diagram in Fig. 26.

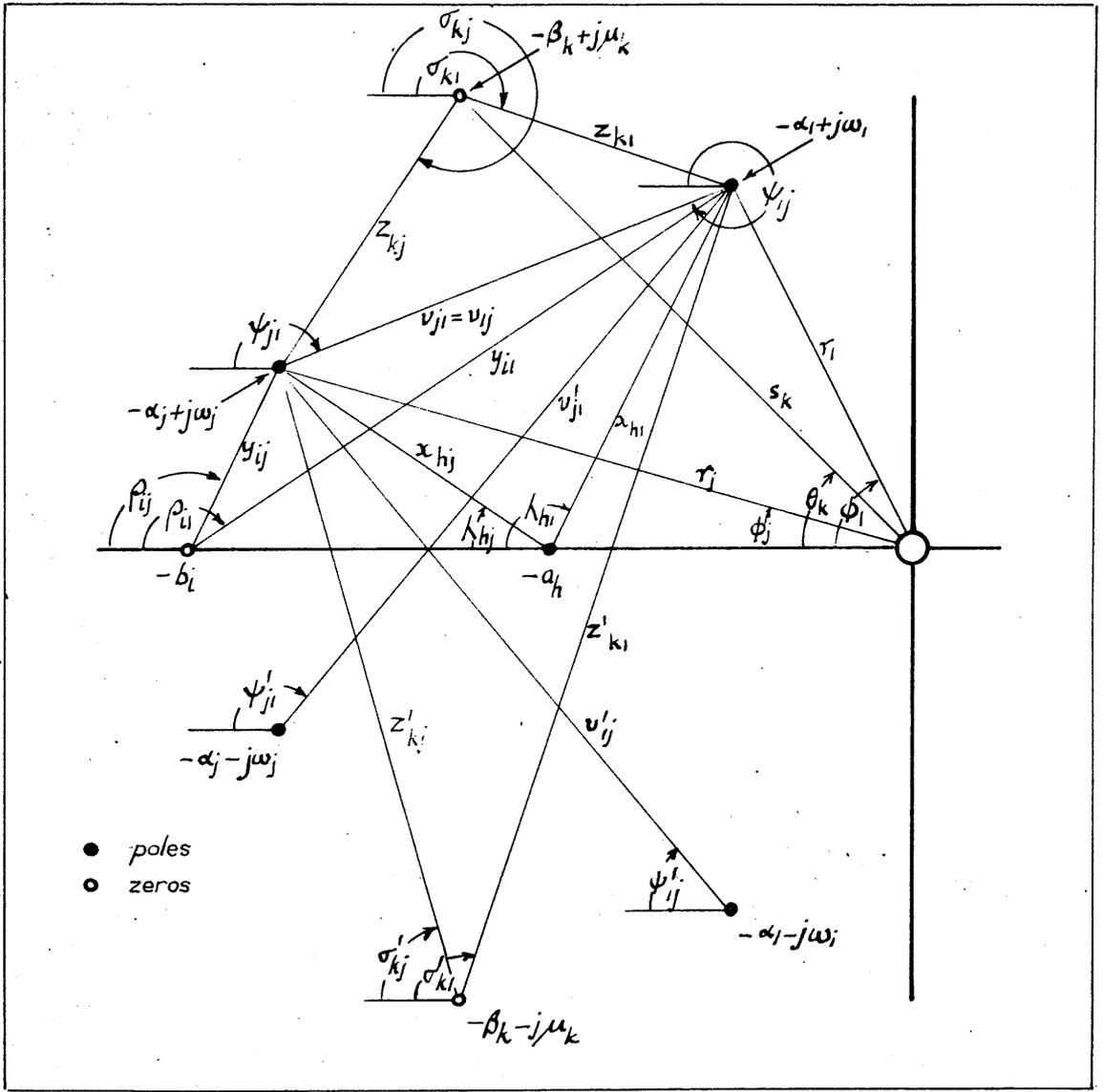


Fig. 26.

Meanings are attached to the symbols as follows:-

Positions of poles and zeros.

- $\alpha_1 \pm j\omega_1$ principal complex pole-pair
- $\alpha_j \pm j\omega_j$ j^{th} complex pole-pair
- a_h h^{th} real pole
- $\beta_k \pm j\omega_k$ k^{th} complex zero-pair
- b_i i^{th} real zero.

Distances

- r_1, r_j distances from origin to principal and j^{th} complex pole-pair
- s_k distance from origin to k^{th} complex zero-pair
- z_{k1}, z'_{k1} distances from k^{th} complex zero-pair $-\beta_k \pm j\omega_k$ to $-\alpha_1 + j\omega_1$
- z_{kj}, z'_{kj} " " k^{th} complex zero-pair $-\beta_k \pm j\omega_k$ to $-\alpha_j + j\omega_j$
- v_{j1}, v'_{j1} " " j^{th} complex pole-pair $-\alpha_j \pm j\omega_j$ to $-\alpha_1 + j\omega_1$
- x_{h1}, x_{hj} " " h^{th} real pole to $-\alpha_1 + j\omega_1$ and $-\alpha_j + j\omega_j$
- y_{i1}, y_{ij} " " i^{th} real zero to $-\alpha_1 + j\omega_1$ and $-\alpha_j + j\omega_j$

Angles

Angles are all measured in a clockwise direction from the negative real axis. In general we require the angle which is made with the negative real axis by the line from every pole and zero to the upper complex pole of every complex pole-pair. The angles are labelled as follows:-

- ψ_{j1}, ψ'_{j1} from j^{th} complex pole-pair $-\alpha_j \pm j\omega_j$ to $-\alpha_1 + j\omega_1$
- $\sigma_{k1}, \sigma'_{k1}$ from k^{th} complex zero-pair $-\beta_k \pm j\omega_k$ to $-\alpha_1 + j\omega_1$
- $\sigma_{kj}, \sigma'_{kj}$ from k^{th} complex zero-pair $-\beta_k \pm j\omega_k$ to $-\alpha_j + j\omega_j$
- $\lambda_{h1}, \lambda_{hj}$ from h^{th} real pole to $-\alpha_1 + j\omega_1$ and $-\alpha_j + j\omega_j$
- ρ_{i1}, ρ_{ij} from i^{th} real zero to $-\alpha_1 + j\omega_1$ and $-\alpha_j + j\omega_j$

The angles of the pole and zero vectors, $-\alpha_1 + j\omega_1, -\alpha_j + j\omega_j$ and $-\beta_k + j\omega_k$ are ϕ_1, ϕ_j , and θ_k respectively measured from the negative real axis, as shown on the diagram.

We are now able to consider the general overall system transfer function, in which there are m pairs of complex poles, f real poles,

n pairs of complex zeros and g real zeros, for the moment all being distinct, and with $2n+g \leq \left. \begin{matrix} * \\ 2m+f-2 \end{matrix} \right\}$. This gives the transfer function

$$\theta_o/\theta_i(\rho) = \frac{\prod_{h=1}^f a_h \prod_{j=1}^m r_j^2 \cdot \prod_{l=1}^g (\rho+b_l) \prod_{k=1}^n (\rho+\beta_k-j\mu_k \chi \rho+\beta_k+j\mu_k)}{\prod_{l=1}^g b_l \prod_{k=1}^m s_k^2 \cdot \prod_{h=1}^f (\rho+a_h) \prod_{j=1}^m (\rho+\alpha_j-j\omega_j \chi \rho+\alpha_j+j\omega_j)} \quad (43)$$

The output transform is the above expression with an added ρ in the denominator. It is therefore

$$\theta_o(\rho) = A \cdot \frac{\prod_{l=1}^g (\rho+b_l) \prod_{k=1}^n (\rho+\beta_k-j\mu_k \chi \rho+\beta_k+j\mu_k)}{\prod_{h=1}^f (\rho+a_h) \prod_{j=1}^m (\rho+\alpha_j-j\omega_j \chi \rho+\alpha_j+j\omega_j) \rho} \quad (44)$$

where A denotes the first factor of expression (43). The output time-function is evaluated similarly to the previous examples.

The pole at $\rho=0$, gives

$$A \cdot \frac{\prod_{l=1}^g b_l \prod_{k=1}^n (\beta_k-j\mu_k \chi \beta_k+j\mu_k)}{\prod_{h=1}^f a_h \prod_{j=1}^m (\alpha_j-j\omega_j \chi \alpha_j+j\omega_j)} = 1$$

The pole at $\rho=-\alpha_j+j\omega_j$, gives

$$A \cdot \frac{\prod_{l=1}^g (-\alpha_j+j\omega_j+b_l) \prod_{k=1}^n (-\alpha_j+j\omega_j+\beta_k-j\mu_k \chi -\alpha_j+j\omega_j+\beta_k+j\mu_k)}{\prod_{h=1}^f (-\alpha_j+j\omega_j+a_h) \prod_{\substack{q=1 \\ q \neq j}}^m (-\alpha_j+j\omega_j+\alpha_q-j\omega_q \chi -\alpha_j+j\omega_j+\alpha_q+j\omega_q)} \cdot \frac{\varepsilon^{-\alpha_j t} \varepsilon^{j\omega_j t}}{2j\omega_j(-\alpha_j+j\omega_j)}$$

From Fig. 26 we have

$$\begin{aligned} -\alpha_j+j\omega_j+b_l &= y_{ij} \varepsilon^{j(\pi-\beta_{lj})} & -\alpha_j+j\omega_j+a_h &= x_{hj} \varepsilon^{j(\pi-\lambda_{hj})} \\ -\alpha_j+j\omega_j+\beta_k-j\mu_k &= z_{kj} \varepsilon^{j(\pi-\sigma_{kj})} & -\alpha_j+j\omega_j+\alpha_q-j\omega_q &= u_{qj} \varepsilon^{j(\pi-\psi_{qj})} \\ -\alpha_j+j\omega_j+\beta_k+j\mu_k &= z'_{kj} \varepsilon^{j(\pi-\sigma'_{kj})} & -\alpha_j+j\omega_j+\alpha_q+j\omega_q &= u'_{qj} \varepsilon^{j(\pi-\psi'_{qj})} \end{aligned}$$

* The theory is unaffected by having $2n+g \leq \left. \begin{matrix} * \\ 2m+f-1 \end{matrix} \right\}$, but it is more correct to adopt the relation given.

Hence the contribution is

$$A. \frac{\prod_{l=1}^g y_{lj} \varepsilon^{j(\pi-\rho_{lj})} \prod_{k=1}^n z_{kj} z'_{kj} \varepsilon^{j(2\pi-\overline{\sigma_{kj}}+\sigma'_{kj})}}{\prod_{h=1}^f x_{hj} \varepsilon^{j(\pi-\lambda_{hj})} \prod_{q=1, q \neq j}^m u_{qj} u'_{qj} \varepsilon^{j(2\pi-\overline{\psi_{qj}}+\psi'_{qj})}} \cdot \frac{\varepsilon^{-\alpha_j t} \varepsilon^{j\omega_j t}}{\omega_j \eta_j \varepsilon^{j(\pi-\phi_j)} 2j}$$

The pole at $\rho = -\alpha_j - j\omega_j$ gives the conjugate of the above, namely

$$A. \frac{\prod_{l=1}^g y_{lj} \varepsilon^{-j(\pi-\rho_{lj})} \prod_{k=1}^n z_{kj} z'_{kj} \varepsilon^{-j(2\pi-\overline{\sigma_{kj}}+\sigma'_{kj})}}{\prod_{h=1}^f x_{hj} \varepsilon^{-j(\pi-\lambda_{hj})} \prod_{q=1, q \neq j}^m u_{qj} u'_{qj} \varepsilon^{-j(2\pi-\overline{\psi_{qj}}+\psi'_{qj})}} \cdot \frac{\varepsilon^{-\alpha_j t} \varepsilon^{j\omega_j t}}{\omega_j \eta_j \varepsilon^{-j(\pi-\phi_j)} -2j}$$

The pole-pair at $-\alpha_j \pm j\omega_j$ therefore gives

$$A. \frac{\prod_{l=1}^g y_{lj} \prod_{k=1}^n z_{kj} z'_{kj}}{\prod_{h=1}^f x_{hj} \prod_{q=1, q \neq j}^m u_{qj} u'_{qj}} \cdot \frac{\varepsilon^{-\alpha_j t}}{\omega_j \eta_j} \sin \left[\omega_j t - \overline{\pi-\phi_j} + \sum_{l=1}^g (\pi-\rho_{lj}) + \sum_{k=1}^n (2\pi-\overline{\sigma_{kj}}+\sigma'_{kj}) - \sum_{h=1}^f (\pi-\lambda_{hj}) - \sum_{q=1, q \neq j}^m (2\pi-\overline{\psi_{qj}}+\psi'_{qj}) \right] \quad (45)$$

The expression (45) may be denoted by $F_j(t)$; the contribution from the m complex poles is therefore $\sum_{j=1}^m F_j(t)$. As, however, we are interested in the term resulting from the principal pole-pair at $-\alpha_1 \pm j\omega_1$, we shall write this out in full. It is

$$F_1(t) = A. \frac{\prod_{l=1}^g y_{l1} \prod_{k=1}^n z_{k1} z'_{k1}}{\prod_{h=1}^f x_{h1} \prod_{j=2}^m u_{j1} u'_{j1}} \cdot \frac{\varepsilon^{-\alpha_1 t}}{\omega_1 \eta_1} \sin \left[\omega_1 t - \overline{\pi-\phi_1} + \sum_{l=1}^g (\pi-\rho_{l1}) + \sum_{k=1}^n (2\pi-\overline{\sigma_{k1}}+\sigma'_{k1}) - \sum_{h=1}^f (\pi-\lambda_{h1}) - \sum_{j=2}^m (2\pi-\overline{\psi_{j1}}+\psi'_{j1}) \right] \quad (46)$$

The real pole at $\rho = -a_h$ gives

$$A. \frac{\prod_{l=1}^g (b_l - a_h) \prod_{k=1}^n (-a_h + \beta_k - j\mu_k)(-a_h + \beta_k + j\mu_k) \varepsilon^{-a_h t}}{\prod_{l=1, l \neq h}^f (a_l - a_h) \prod_{j=1}^m (-a_h + \alpha_j - j\omega_j)(-a_h + \alpha_j + j\omega_j)(-a_h)} \\ = A. \frac{\prod_{l=1}^g (b_l - a_h) \prod_{k=1}^n (\beta_k - a_h^2 + \mu_k^2) \varepsilon^{-a_h t}}{\prod_{l=1, l \neq h}^f (a_l - a_h) \prod_{j=1}^m (\alpha_j - a_h^2 + \omega_j^2)(-a_h)} \quad (47)$$

If we denote (47) by $G_h(t)$, the total contribution from the real poles is $\sum_{h=1}^f G_h(t)$. The time expression for the output is therefore

$$\theta_o(t) = 1 + \sum_{j=1}^m F_j(t) + \sum_{h=1}^f G_h(t) \quad (48)$$

and for the error

$$e(t) = -\sum_{j=1}^m F_j(t) - \sum_{h=1}^f G_h(t) \quad (49)$$

In a similar manner to the previous examples we now differentiate the expression $-F_i(t)$, which, we assume, represents the total error by the time the first overshoot of the output (i.e. the first minimum of the error) is reached. This gives

$$\begin{aligned} \frac{dF_i(t)}{dt} &= \text{const.} \times \left[-\alpha_i \varepsilon^{-\alpha_i t} \sin \varepsilon t + \omega_i \varepsilon^{-\alpha_i t} \cos \varepsilon t \right] \\ &= \text{const.} \times \varepsilon^{-\alpha_i t} \sin \left[\omega_i t - \bar{\pi} + \sum_{l=1}^g (\bar{\pi} - \rho_{li}') + \sum_{k=1}^n (2\bar{\pi} - \bar{\sigma}_{ki} + \sigma_{ki}') - \sum_{h=1}^f (\bar{\pi} - \lambda_{hi}) - \sum_{j=2}^m (2\bar{\pi} - \bar{\psi}_{ji} + \psi_{ji}') \right] \end{aligned} \quad (50)$$

The expression (50) has its first zero at

$$t = \frac{1}{\omega_i} \left[\bar{\pi} + \sum_{h=1}^f (\bar{\pi} - \lambda_{hi}) + \sum_{j=2}^m (2\bar{\pi} - \bar{\psi}_{ji} + \psi_{ji}') - \sum_{l=1}^g (\bar{\pi} - \rho_{li}') - \sum_{k=1}^n (2\bar{\pi} - \bar{\sigma}_{ki} + \sigma_{ki}') \right] \quad (51)$$

Substituting this value of t in the expression for $-F_i(t)$, results in a minimum value for $-F_i(t)$, the magnitude of this being

$$A \cdot \frac{\prod_{l=1}^g y_{li} \prod_{k=1}^n z_{ki} z_{ki}'}{\prod_{h=1}^f a_{hi} \prod_{j=2}^m v_{ji} v_{ji}'} \cdot \frac{\varepsilon^{-\alpha_i t}}{\omega_i} \left[\bar{\pi} + \sum_{h=1}^f (\bar{\pi} - \lambda_{hi}) + \sum_{j=2}^m (2\bar{\pi} - \bar{\psi}_{ji} + \psi_{ji}') - \sum_{l=1}^g (\bar{\pi} - \rho_{li}') - \sum_{k=1}^n (2\bar{\pi} - \bar{\sigma}_{ki} + \sigma_{ki}') \right] \cdot \sin \phi_i \quad (52)$$

Since A is

$$\left[\prod_{h=1}^f a_{hi} \prod_{j=1}^m r_j^2 \right] / \left[\prod_{l=1}^g b_l \prod_{k=1}^n s_k^2 \right]$$

and $\sin \phi_i = \frac{\omega_i}{\bar{\pi}}$, the expression (52) which, we assume, represents the size of the overshoot is

$$\varepsilon \frac{-\alpha_l \bar{\omega}_l}{\bar{\omega}_l} \cdot \frac{\prod_{i=1}^g \frac{y_{li}}{b_i} \varepsilon \frac{\alpha_l (\bar{\omega}_l - \rho_{li})}{\bar{\omega}_l}}{\prod_{k=1}^n \frac{z_{ki} z'_{ki}}{s_k^2} \varepsilon \frac{\alpha_l (2\bar{\omega}_l - \bar{\sigma}_{ki} + \bar{\sigma}'_{ki})}{\bar{\omega}_l}} \cdot \frac{\prod_{h=1}^f \frac{x_{hl}}{a_h} \varepsilon \frac{\alpha_l (\bar{\omega}_l - \lambda_{hl})}{\bar{\omega}_l}}{\prod_{j=2}^m \frac{v_{ji} v'_{ji}}{\Gamma_j^2} \varepsilon \frac{\alpha_l (2\bar{\omega}_l - \bar{\psi}_{ji} + \bar{\psi}'_{ji})}{\bar{\omega}_l}} \quad (53)$$

The above expression gives therefore the modification to the overshoot $\varepsilon \frac{-\alpha_l \bar{\omega}_l}{\bar{\omega}_l}$ of the basic quadratic system. It is seen to be composed of similar factors, one for every pole and zero other than the principal pole-pair. Thus, as has already been shown in the second simple example, the modification introduced by a real pole is the divisor.

$$\frac{x_{hl}}{a_h} \varepsilon \frac{\alpha_l (\bar{\omega}_l - \lambda_{hl})}{\bar{\omega}_l} \quad (54)$$

In a similar fashion, a real zero introduces a multiplier

$$\frac{y_{li}}{b_i} \varepsilon \frac{\alpha_l (\bar{\omega}_l - \rho_{li})}{\bar{\omega}_l} \quad (55)$$

The complex pole-pair introduces the divisor

$$\frac{v_{ji} v'_{ji}}{\Gamma_j^2} \varepsilon \frac{\alpha_l (2\bar{\omega}_l - \bar{\psi}_{ji} + \bar{\psi}'_{ji})}{\bar{\omega}_l} \quad (56)$$

while the complex zero-pair introduces the multiplier

$$\frac{z_{ki} z'_{ki}}{s_k^2} \varepsilon \frac{\alpha_l (2\bar{\omega}_l - \bar{\sigma}_{ki} + \bar{\sigma}'_{ki})}{\bar{\omega}_l} \quad (57)$$

The forms for real poles and zeros are therefore the same, the only difference being that the former divides the overshoot and the latter multiplies it. The same may be said of complex poles and zeros, although their modifications are different from those of the real poles and zeros. It is understood however, that so far, these multipliers and divisors only hold for one particular position of the principal pole-pair, although the form will be maintained irrespective of this position. Furthermore, no

conditions have as yet been stated in order that the principal mode approximation should hold. Before examining this point, the effect of multiple poles and zeros is considered.

6.4. Effect of Multiple Poles and Zeros.

If we suppose that in the general expression (43) for the overall transfer function,

- (i) the real pole a_d is repeated d times,
- (ii) the real zero b_c is repeated c times,
- (iii) the complex pole-pair $-\alpha_s \pm j\omega_s$ is repeated δ times, and
- (iv) the complex zero-pair $-\beta_r \pm j\mu_r$ is repeated γ times,

the form of the output transform becomes

$$\theta_o(p) = \frac{a_d^d \prod_{h=1, h \neq d}^{f-d} a_h \cdot r_s^{2\delta} \prod_{j=1, j \neq \delta}^{m-\delta} r_j^2}{b_c^c \prod_{i=1, i \neq c}^{g-c} b_i \cdot s_r^{2\gamma} \prod_{k=1, k \neq \gamma}^{n-\gamma} s_k^2} \quad \times$$

$$\frac{(p+b_c)^c \prod_{i=1, i \neq c}^{g-c} (p+b_i) \cdot (p+\beta_r - j\mu_r)^\gamma (p+\beta_r + j\mu_r)^\gamma \prod_{k=1, k \neq \gamma}^{n-\gamma} (p+\beta_k - j\mu_k)(p+\beta_k + j\mu_k)}{(p+a_d)^d \prod_{h=1, h \neq d}^{f-d} (p+a_h) \cdot (p+\alpha_s - j\omega_s)^\delta (p+\alpha_s + j\omega_s)^\delta \prod_{j=1, j \neq \delta}^{m-\delta} (p+\alpha_j - j\omega_j)(p+\alpha_j + j\omega_j) p} \quad (58)$$

This is a cumbersome though quite regular expression and it is not intended to carry out the solution in the same detail as in the previous Section. The principal mode expression will be evaluated and certain results of the previous Section used to effect some shortcuts in the analysis. Let us write B for the first factor of the expression (58)

which is the quotient of the product of the poles and the product of the zeros. It is taken that the principal pole-pair $-\alpha_1 \pm j\omega_1$ is not a repeated pole, this fact being essential to the theory.

The pole at $-\alpha_1 + j\omega_1$, then gives

$$B \frac{(-\alpha_1 + j\omega_1 + b_c)^{g-c} \prod_{i=1, i \neq c}^{g-c} (-\alpha_1 + j\omega_1 + b_i)}{(-\alpha_1 + j\omega_1 + a_d)^{f-d} \prod_{h=1, h \neq d}^{f-d} (-\alpha_1 + j\omega_1 + a_h)} \cdot \frac{\epsilon^{-\alpha_1 t} \epsilon^{j\omega_1 t}}{2j\omega_1 (-\alpha_1 + j\omega_1)} x'$$

$$\frac{(-\alpha_1 + j\omega_1 + \beta_\gamma - j\mu_\gamma)^{\gamma} (-\alpha_1 + j\omega_1 + \beta_\gamma + j\mu_\gamma)^{\gamma} \prod_{k=1, k \neq \gamma}^{n-\gamma} (-\alpha_1 + j\omega_1 + \beta_k - j\mu_k) (-\alpha_1 + j\omega_1 + \beta_k + j\mu_k)}{(-\alpha_1 + j\omega_1 + \alpha_\delta - j\omega_\delta)^{\delta} (-\alpha_1 + j\omega_1 + \alpha_\delta + j\omega_\delta)^{\delta} \prod_{j=2, j \neq \delta}^{m-\delta} (-\alpha_1 + j\omega_1 + \alpha_j - j\omega_j) (-\alpha_1 + j\omega_1 + \alpha_j + j\omega_j)}$$

In a similar manner to the previous Section, this reduces to

$$B \frac{y_{c1}^c \epsilon^{j c (\pi - \beta_{c1})} \prod_{i=1, i \neq c}^{g-c} y_{i1} \epsilon^{j (\pi - \beta_{i1})} \cdot (z_{\gamma 1} z'_{\gamma 1})^{\gamma} j^{\gamma} (2\pi - \overline{\sigma_{\gamma 1} + \sigma'_{\gamma 1}})^{n-\gamma} \prod_{k=1, k \neq \gamma} z_{k1} z'_{k1} \epsilon^{j (2\pi - \overline{\sigma_{k1} + \sigma'_{k1}})}}{x_{d1}^d \epsilon^{j d (\pi - \lambda_{d1})} \prod_{h=1, h \neq d}^{f-d} x_{h1} \epsilon^{j (\pi - \lambda_{h1})} \cdot (u_{\delta 1} u'_{\delta 1})^{\delta} j^{\delta} (2\pi - \overline{\psi_{\delta 1} + \psi'_{\delta 1}})^{m-\delta} \prod_{j=2, j \neq \delta} u_{j1} u'_{j1} \epsilon^{j (2\pi - \overline{\psi_{j1} + \psi'_{j1}})}} \cdot \frac{\epsilon^{-\alpha_1 t} \epsilon^{j\omega_1 t}}{2j\omega_1 \epsilon^{j(\pi - \phi_1)}}$$

The pole-pair at $-\alpha_1 \pm j\omega_1$ therefore gives

$$B \frac{y_{c1}^c \prod_{l=1, l \neq c}^{g-c} y_{l1} \cdot (z_{\gamma 1} z'_{\gamma 1})^{\gamma} \prod_{k=1, k \neq \gamma}^{n-\gamma} z_{k1} z'_{k1}}{x_{d1}^d \prod_{h=1, h \neq d}^{f-d} x_{h1} \cdot (u_{\delta 1} u'_{\delta 1})^{\delta} \prod_{j=2, j \neq \delta}^{m-\delta} u_{j1} u'_{j1}} \cdot \frac{\epsilon^{-\alpha_1 t}}{\omega_1 \tau} \sin \left[\omega_1 t - (\pi - \phi_1) + c(\pi - \beta_{c1}) + \sum_{l=1, l \neq c}^{g-c} (\pi - \beta_{l1}) \right] \quad (59)$$

$$+ \gamma (2\pi - \overline{\sigma_{\gamma 1} + \sigma'_{\gamma 1}}) + \sum_{k=1, k \neq \gamma}^{n-\gamma} (2\pi - \overline{\sigma_{k1} + \sigma'_{k1}}) - d(\pi - \lambda_{d1}) - \sum_{h=1, h \neq d}^{f-d} (\pi - \lambda_{h1}) - \delta (2\pi - \overline{\psi_{\delta 1} + \psi'_{\delta 1}}) - \sum_{j=2, j \neq \delta}^{m-\delta} (2\pi - \overline{\psi_{j1} + \psi'_{j1}}) \Big].$$

This may be compared with the expression (46), from which it will be seen that analogous results to the previous Section will ensue, if the usual procedure for calculating the overshoot is carried out. In view of the similarity of the processes, only the final conclusion is quoted, namely, that for any pole or zero, real or complex, repeated c times say, the corresponding expressions (54) to (57) are also repeated c times. The multiplier for the zero b_c repeated c times is therefore

$$\frac{y_{c1}}{b_c} \propto \frac{\omega_1}{\omega_1} (\bar{\eta} - \rho_{c1})^c$$

and so on.

It is not intended formally to evaluate the terms resulting from the multiple poles, as the considerations of the next Section show that these in general should be avoided. There is however no inherent difficulty in obtaining these terms as can be found by consulting standard text-books on the subject .

6.5. General Conditions under which Principal Mode Approximation Holds.

Application of Foregoing Results to the Design of Transfer Functions.

The broad conditions to apply in order that the principal mode approximation should hold are obtained from the expressions (45) and (47) which give the terms due to complex poles other than the principal pole-pair, and those due to the real poles. Let us take it first of all that no complex poles other than the principal pole-pair are present. Then, in order that the terms represented by (47) be small, we should have,

- (i) a_h large, (ii) $(a_l - a_h), (\bar{\alpha}_l - a_h^2 + \omega_l^2)$ large, (iii) $(b_i - a_h), (\bar{\beta}_k - a_h^2 + \omega_k^2)$ small. The first condition is the most important, being obviously the

the placing of real poles well to the left of the principal pole. Condition (ii) merely states that all the poles should be well-separated and condition (iii) states that the zeros should be as near the poles as possible. This last condition is also evident, as in the limiting case of zeros coinciding with poles, the basic quadratic system will result by cancellation of these factors in the numerator and denominator of the transfer function.

If complex poles other than the principal pole-pair are present, similar considerations apply. In order that the zeros should be near the poles, it follows that it is advisable to insert complex zeros when there are subsidiary complex poles. It is realised that these are only general conditions, but in the numerical work done so far, no trouble has been experienced. Again, although horizontal separation of the poles is most important it is possible to bring up the nearest subsidiary pole fairly close to the principal pole if it is accompanied by a neighbouring zero. The condition that the poles should be well separated indicates that repeated poles are not generally permissible unless they are remote from the principal pole-pair.

Use of nearest real pole in addition to principal mode.

Under certain conditions where difficulty is experienced in obtaining a small enough overshoot by the use of the principal mode oscillation by itself, it may be advantageous to include that term which results from the next least-damped pole. As a real pole is easier to work with than a

complex pole-pair for the purposes of calculating the value of the term at any particular time, it is therefore best to make this next least-damped pole a real pole. Furthermore, it is clear that if this is to reduce the overshoot, the coefficient of this term in the time-expression for the error must be positive, since the first minimum of the error, corresponding to the overshoot in the output, is a negative quantity, see Fig. 27.

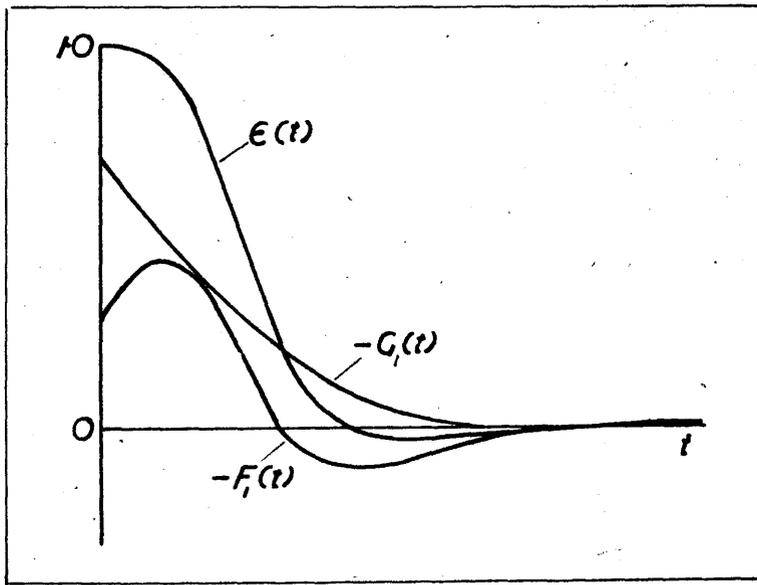


Fig. 27.

Thus in the expressions (48) or (49), if $G_1(t)$ is the term resulting from the nearest real pole, then $G_1(t)$ should be made negative. From the expression (47) giving the general form of the time-term resulting from a real pole, it is seen that the sign of $G_1(t)$ will depend on the relative numbers of sign-changes occurring in the factors $(b_i - a_i)$ and $(a_i - q_i)$. Hence we may state:-

- (i) the product of all the pole factors $(a_p - a_i)$ is positive since a_i is the pole furthest to the right; the denominator as a whole is therefore negative.
- (ii) the product of all the zero factors $(b_i - a_i)$ is positive if no zeros or if an even number of zeros lie to the right of a_i ;
- (iii) the product of all the zero factors $(b_i - a_i)$ is negative if an odd number of zeros lie to the right of a_i .

We reach the conclusion, therefore, that the term resulting from the nearest real pole will have the required negative sign if no zeros or an even number of zeros lie to the right of the nearest or least-damped real pole. This is a useful result to bear in mind. Since, however, zeros to the right of a_i are liable to seriously reduce the principal oscillatory mode, two at most may be permitted to lie in this region. It may make it clearer to state that a reduction in the principal oscillatory mode is quite welcome provided that the least damped real term is made negligible by the time the second stationary point of the principal oscillation is reached. For a satisfactory response the value of the principal mode oscillation at this point will have to be not greater than 1 - 2% say, of the input step. In other words the transient will be effectively finished.

Application to design.

The relation (53) gives the magnitude of the overshoot in the response to a unit-step. It is compounded of the basic overshoot $\varepsilon^{-\zeta_i \pi / \omega_i}$ multiplied by quantities $\frac{y_{U_i}}{b_i} \varepsilon^{\frac{\zeta_i'}{\omega_i} (\pi - \rho_i)}$, $\frac{z_{k_i} z_{k_i}'}{s k_i^2} \varepsilon^{\frac{\zeta_i'}{\omega_i} (2\pi - \sqrt{\zeta_i'^2 + \sigma_{k_i}^2})}$ for real and complex zeros respectively, and divided by

analogous quantities for real and complex poles. If therefore the position of the principal pole-pair is fixed, these quantities may be evaluated for other points on the left half of the complex plane. This may be done once and for all, and a design chart for this particular position of the principal pole-pair, so constructed. Similar charts may be prepared for other positions of the principal pole-pair. The work implied in such a series of charts may be partially reduced in the light of two of the considerations already mentioned which are required for the principal mode approximation, namely

- (i) poles to the left of the principal pole-pair, and
- (ii) zeros near poles.

To these we add one other restriction, which, in effect assures there will be no "kinks" or superposed high-frequency ripples on the response. It is that

- (iii) the imaginary part of any subsidiary complex pole should not exceed that of the principal pole.

With these conditions, the area of investigation in the complex plane is confined as shown in Fig. 28.

for the estimation of the time of the overshoot are given. We shall refer to the first three charts as overshoot charts and to the second three charts as angle charts.

1. Three positions of the principal pole-pair have been used, namely $-0.5 \pm j1$, $-1 \pm j1$ and $-2 \pm j1$. One overshoot chart and one angle chart is given for each of the above positions.
2. The logarithm to the base 10 of the quantities $\frac{y}{b} \varepsilon^{\frac{\alpha}{\omega} (\pi - \rho)}$ and $\frac{zz'}{s^2} \varepsilon^{\frac{\alpha}{\omega} (2\pi - \sqrt{\sigma} + \sigma)}$ has been given. This is therefore the increment in the logarithmic overshoot. Increments for zeros are taken with a positive sign, and those for poles with a negative sign, this sign being of course supplementary to the increment itself, which may be positive or negative.
3. The variation of the above quantities is shown as the real part of the pole or zero alters, each curve representing a fixed value of the imaginary part. This is best explained with reference to the diagram in Fig. 29.

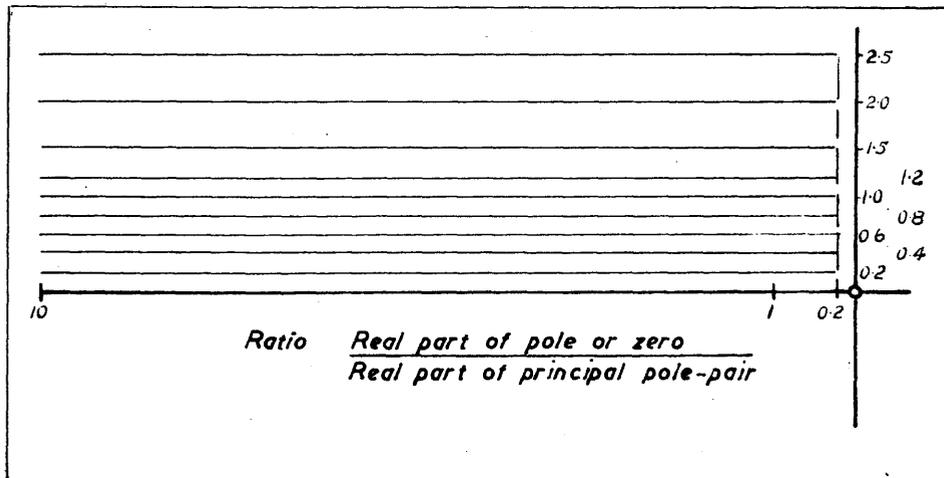


Fig. 29.

The fixed values of the imaginary part are 0, 0.2, 0.4, 0.6, 0.8, 1.0, 1.2, 1.5, 2.0 and 2.5 and the curves on the overshoot charts are therefore the previously mentioned increments, evaluated along the horizontal lines of Fig. 29. The abscissa of the curves is the ratio of the real part of the pole or zero to the real part of the principal pole. This is adopted for two reasons. First, it was desirable to use logarithmic paper in order to render the curves less steep, and hence obtain reasonable intersections with vertical lines. Such a logarithmic scale is not convenient for showing directly the real part of the pole or zero. Secondly, it is convenient to work in terms of the above ratio, from the viewpoint of the principal mode approximation.

4. The angle charts show the quantities $(\bar{\pi}-\rho)$ and $(2\bar{\pi}-\sigma+\sigma')$ for a real zero and a complex zero-pair respectively. These are, of course, the same for poles as for zeros. This information enables the final pole and zero positions to be fixed, if the speed of response of the system is specified. Thus from (51), the time of the first overshoot is

$$t = \frac{1}{\omega} \left[\bar{\pi} + \sum_{h=1}^f (\bar{\pi}-\lambda_{h1}) + \sum_{j=2}^m (2\bar{\pi}-\psi_{j1}+\psi'_{j1}) - \sum_{i=1}^g (\bar{\pi}-\rho_{i1}) - \sum_{k=1}^n (2\bar{\pi}-\sigma_{k1}+\sigma'_{k1}) \right]$$

that is $\bar{\pi} +$ (sum of angles for poles) $-$ (sum of angles for zeros).

Since the angles $(\bar{\pi}-\rho)$, $(2\bar{\pi}-\sigma_{k1}+\sigma'_{k1})$ etc. will remain invariant if the pole and zero pattern remains fixed, although changing in absolute magnitude, the time of the first overshoot is determined by the choice of ω . If this is increased k times and all other pole and zero distances similarly multiplied, the time for maximum overshoot to occur will be reduced by a

factor k . The procedure will be more evident if the examples in Chapter 7 are considered.

5. Range of curves.

The left-hand limit for the abscissa of the curves is 10. Beyond this point a maximum error of 0.025 in the logarithm of the factor will occur, if the contribution from such poles or zeros is neglected. For each zero or pole, this is equivalent to about 6% error respectively in the size of the overshoot^{*}, so that for practical purposes it is not serious. The error will depend on the relative numbers and spacing of poles and zeros beyond the left-hand limit, and if reasonable guesswork is employed, or if the poles and zeros are well out to the left, negligible error finally results. At the right-hand side, the curves are taken up to a value of the abscissa equal to 0.2, or to that value of abscissa at which the increment in log. overshoot becomes 2. In the numerical work so far performed it has not been necessary to approach even approximately either of those limits. Indeed it has already been said that the main working portion of the charts lies to the left of the abscissa value unity.

The same abscissa values limit the angle charts with an additional upper limit of 5 radians. This roughly corresponds to the upper limit of 2 in the overshoot charts. For poles and zeros beyond the abscissa limit

*. That is, an overshoot which is designed to be 10% will actually be 10.6% or 9.4%.

10, the angle $(\pi - \rho)$ is given with very little error by $1/b - \alpha_1$ radians, or it may be taken to be zero with, of course, rather more error.

The angle $(2\pi - \overline{\sigma} + \sigma')$ may be approximated by $2/b - \alpha_1$ radians for $\mu < 1$, and by zero for $\mu > 1$.

The use of the design charts is now demonstrated by a number of examples.

CHAPTER 7.

NUMERICAL WORK ILLUSTRATING THE USE OF THE $\frac{\theta}{\omega}$ DESIGN CHARTS

7.1. Notation.

The logarithm to the base 10 of the fractional (i.e. per unit) overshoot will be denoted by X . We therefore have, from (53),

$$X = -\frac{\alpha_1 \bar{\pi}}{\omega_1} \log_{10} \epsilon + \sum (\text{increments due to zeros}) - \sum (\text{increments due to poles}) \quad (60)$$

Let also $\sum \Delta_z = \sum (\text{increments due to zeros})$

$$\sum \Delta_p = \sum (\text{increments due to poles})$$

$\Delta_{b_i}, \Delta_{s_k} =$ increment due to i^{th} real zero and k^{th} complex zero-pair respectively.

$\Delta_{a_h}, \Delta_{r_j} =$ increment due to h^{th} real pole and j^{th} complex pole-pair respectively.

$$Y = \sum \Delta_z - \sum \Delta_p$$

$$L_{b_i}, L_{s_k} = (\bar{\pi} - \beta_{i1}), (2\bar{\pi} - \sqrt{\alpha_{k1}^2 + \sigma_{k1}^2}) \quad \text{respectively}$$

$$L_{a_h}, L_{r_j} = (\bar{\pi} - \lambda_{h1}), (2\bar{\pi} - \sqrt{\psi_{j1}^2 + \psi_{j1}^2}) \quad \text{respectively}$$

$$\sum L_z = \sum (\text{angles due to zeros})$$

$$\sum L_p = \sum (\text{angles due to poles})$$

The symbol Δ will be used to denote the increment in general from any pole or zero or complex pole- or zero-pair.

The relation (60) may then be written

$$X = -1.365 \frac{\alpha_1}{\omega_1} + Y \quad (61)$$

or
$$Y = X + 1.365, \quad (62)$$

which gives the required difference $\sum\Delta_z - \sum\Delta_p$, in terms of the desired overshoot and damping. It is useful to work out γ for overshoots of 5%, 10%, 15%, 20% and 25%, and for the three values of damping corresponding to the three positions of the principal pole-pair.

This is given in tabular form below.

TABLE IV.

Values of $\gamma = \sum\Delta_z - \sum\Delta_p$

$-\alpha, \pm j\omega,$	5%	10%	15%	20%	25%
$-0.5 \pm j1$	-0.618	-0.317	-0.140	-0.016	+0.081
$-1 \pm j1$	+0.064	+0.365	+0.542	+0.666	+0.763
$-2 \pm j1$	+1.429	+1.730	+1.907	+2.031	+2.128

We also have, for reference

TABLE V.

$-\alpha, \pm j\omega,$	r	r^2	$\phi,$	$\pi - \phi,$	$\sin \phi,$	Ratio undershoot overshoot.	Log Over- shoot.
$-0.5 \pm j1$	1.118	1.25	1.106	2.036	0.895	$\doteq 1/5$	γ -0.683
$-1 \pm j1$	1.414	2.0	0.785	2.357	0.707	1/23	γ -1.365
$-2 \pm j1$	2.236	5.0	0.463	2.679	0.447	1/530	γ -2.730

The examples which are now considered do not represent any particular systems but serve only to illustrate the application of the method. A practical problem due to Whiteley is given at the end of Chapter 9 and the circuit values are derived. This is done by the normal procedure of equating the unknown coefficients in the practical transfer function to those in the designed transfer function.

7.2. Type 1 Servo-Mechanisms.

Example 1. 1 complex pole-pair, 2 real poles, 2 real zeros.

Specification - 10% maximum overshoot at 0.5 second.

Subsequent undershoot not to exceed 1%.

Then
$$\frac{\theta_0}{\theta_i}(\rho) = \frac{a_1 a_2 \zeta^2}{b_1 b_2} \frac{(p+b_1)(p+b_2)}{(p+a_1)(p+a_2)(p+\alpha_1-j\omega_1)(p+\alpha_1+j\omega_1)} \quad (63)$$

With $-\alpha_1 \pm j\omega_1 = -1 \pm j$ the ratio $\frac{\text{undershoot}}{\text{overshoot}} \doteq \frac{1}{23}$, which will give

the required damping. From Table IV therefore,

$$Y = \sum \Delta_z - \sum \Delta_p = (\Delta_{b_1} + \Delta_{b_2}) - (\Delta_{a_1} + \Delta_{a_2}) = 0.365$$

If we take an arbitrary choice of $a_1 = 2\alpha_1$, $a_2 = 4\alpha_1$, we obtain from

the overshoot chart 2 and the angle chart 2, for the curves $\mu = 0$

Δ_{a_1}	0.193	L_{a_1}	0.79
Δ_{a_2}	0.037	L_{a_2}	0.32
$\sum \Delta_p$	0.230	$\sum L_p$	1.11

Hence $\sum \Delta_z = \sum \Delta_p + 0.365$. As a first trial this may be equally

apportioned to b_1 and b_2 , which gives $\Delta_{b_1} = 0.297, \Delta_{b_2} = 0.297$.

Thus $b_1 = 1.63\alpha_1$, $\Delta_{b_1} = 0.297$, $L_{b_1} = 1.03$

and similarly for b_2 . Since this places two zeros close to the nearest real pole at $-2\alpha_1$, we will expect the principal mode approximation to hold. But even if not, since we have an even number of zeros to the right of this pole, the resulting time term will be negative and will reduce the overshoot in any case.

The time of maximum overshoot is

$$t = \frac{1}{\omega_1} (\pi + \sum L_p - \sum L_z)$$

$$= \frac{1}{\omega_1} (3.142 + 1.11 - 2.06) = \frac{2.192}{\omega_1} \text{ sec.}$$

This is to equal 0.5 sec. Therefore $\omega_1 = \frac{2.192}{0.5} = 4.384$, and the required poles and zeros are

$$\begin{aligned} -\alpha_1 \pm j\omega_1 &= 4.384(-1 \pm j1), \quad \eta_1 = 4.384(1.414) \\ q_1 &= 4.384 \cdot 2 = 8.768 \\ q_2 &= 4.384 \cdot 4 = 17.536 \\ b_1 &= 4.384 \cdot 1.63 = 7.14 \\ b_2 &= 4.384 \cdot 1.63 = 7.14 \end{aligned}$$

The transfer function $\theta_o/\theta_i(p)$ is

$$\begin{aligned} \frac{\theta_o}{\theta_i}(p) &= \frac{(8.768)(17.536)(4.384 \cdot 1.414)^2}{(7.14)^2} \cdot \frac{(p+7.14)^2}{(p+8.768)(p+17.536)(p+4.384-j4.384)(p+4.384+j4.384)} \\ &= \frac{115.7(p+7.14)^2}{(p+8.768)(p+17.536)(p+4.384-j4.384)(p+4.384+j4.384)} \end{aligned} \quad (64)$$

The unit-step response given by (64) is

$$\theta_o(t) = 1 + 1.328 \varepsilon^{-4.384t} \sin(4.384t - 1.446) - 0.1029 \varepsilon^{-8.768t} + 0.422 \varepsilon^{-17.536t} \quad (65)$$

which is shown in Fig. 30.

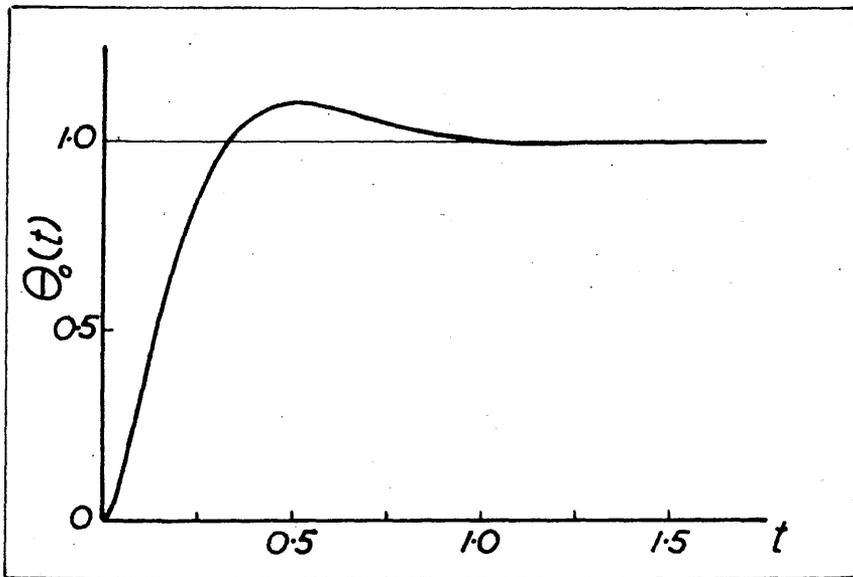


Fig. 30.

The value of the principal mode term at $t = 0.5$ sec. is 0.101, at which time the term $-0.1029\epsilon^{-0.8768t}$ is equal to -0.00129 . As the remaining term is negligible, it may be said that the overshoot is due entirely to the principal mode.

If the response (65) is expressed in terms of dimensionless time t' where $t' = 4.384t$, we have

$$\theta_o(t') = \frac{1 + 1.328\epsilon^{-t'} \sin(t' - 1.446) - 0.1029\epsilon^{-2t'} + 0.422\epsilon^{-4t'}}{\quad} \quad (66)$$

In this case maximum overshoot takes place at $t' = 4.384 \cdot 0.5 = 2.192$. As the true time scale of the response merely depends on multiplying the poles and zeros by some factor k , where

$$k = \frac{\pi + \sum L_p - \sum L_z}{\text{true time of maximum overshoot}}$$

it is simpler to work in terms of the dimensionless time $t' = kt$. This does not affect the size of the overshoot. Thus we have, in the present example

$$(i) \quad \theta_o/\theta_i(p) = \frac{2.4 \cdot (1.414)^2 (p+1.63)^2}{(1.63)^2 (p+2)(p+4)(p+1-j)(p+1+j)} \quad (67)$$

gives 10% overshoot at $t = 2.192$ sec.

$$(ii) \quad \theta_o/\theta_i(p) = \frac{2k \cdot 4k \cdot (1.414k)^2 (p+1.63k)^2}{(1.63k)^2 (p+2k)(p+4k)(p+k-jk)(p+k+jk)} \quad (68)$$

gives 10% overshoot at $t = \frac{2.192}{k}$ sec, that is at $t' = kt = 2.192$.

In the remaining examples, the time of the maximum overshoot will not be specified. Maximum overshoot will in fact always occur at $t' = \frac{\pi + \sum L_p - \sum L_z}{k}$

Example 2. 1 complex pole-pair, 2 real poles, 2 real zeros.

Specification - 10% maximum overshoot

Subsequent undershoot of about 2% is permitted.

This example is practically identical with the previous one and has been taken merely to show that a design with the same overshoot but different damping is possible. Cases may arise, however, when the design for a particular overshoot is limited by the choice of a particular position for the principal pole-pair and should this happen another chart should be used. Conclusions concerning this are stated at the end of this Chapter. Here we have

$$\frac{\theta_o(p)}{\theta_i} = \frac{a_1 a_2 \Gamma^2}{b_1 b_2} \cdot \frac{(p+b_1)(p+b_2)}{(p+a_1)(p+a_2)(p+\alpha_1-j\omega_1)(p+\alpha_1+j\omega_1)}$$

From Table IV, for a principal pole-pair at $-0.5 \pm j1$,

$$\gamma = \sum \Delta_z - \sum \Delta_p = -0.317$$

Since $\sum \Delta_z$ is positive for real zeros, then $\sum \Delta_p > 0.317$. This causes the poles, or at least one of the poles, to be relatively near the principal pole-pair. Suppose $a_1 = 1.8\alpha_1$, $a_2 = 3\alpha_1$, then from the overshoot chart 1 and angle chart 1

$$\begin{array}{ll} \Delta_{a_1} = 0.337 & L_{a_1} = 1.185 \\ \Delta_{a_2} = 0.145 & L_{a_2} = 0.78 \\ \hline \sum \Delta_p = 0.482 & \sum L_p = 1.965 \end{array}$$

Hence $\sum \Delta_z = 0.165$ and again taking equal zeros,

$$\frac{\Delta b_1 = \Delta b_2 = 0.082}{\sum \Delta_z = 0.164}, \quad b_1 = b_2 = 4\alpha_1, \quad \frac{L b_1 = L b_2 = 0.58}{\sum L_z = 1.16}$$

The required transfer function is

$$\begin{aligned} \frac{\theta_o(p)}{\theta_i} &= \frac{(0.9k)(1.5k)(1.118k)^2}{(2k)^2} \frac{(p+2k)^2}{(p+0.9k)(p+1.5k)(p+0.5k-jk)(p+0.5k+jk)} \\ &= \frac{0.4215k^2}{(p+0.9k)(p+1.5k)(p+0.5k-jk)(p+0.5k+jk)} \quad (69) \end{aligned}$$

The response to a unit step is

$$\theta_o(t') = 1 + 0.805 \varepsilon^{-0.5t'} \sin(t' - 2.835) - 0.814 \varepsilon^{-0.9t'} + 0.0585 \varepsilon^{-1.5t'} \quad (70)$$

Maximum overshoot occurs at $t' = \pi + 1.965 - 1.16 = 3.947$. The value of the principal mode term is then 0.1003, while the nearest real pole term is -0.0232 and the remaining term is negligible. The actual overshoot is therefore about 8%. The fact that the nearest real pole term is negative accords with the provision that no zeros or an even number of zeros should lie to the right of this pole. Fig. 31 below shows the response.

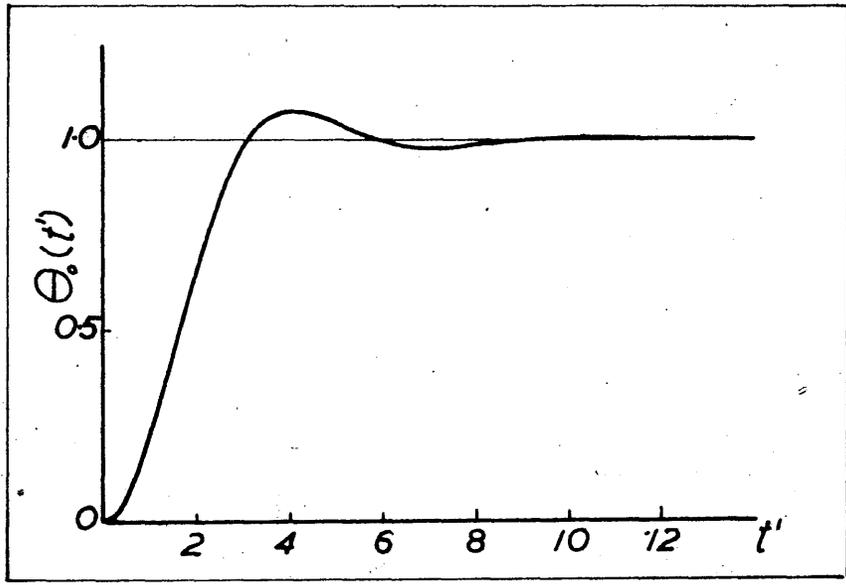


Fig. 31.

Example 3. 2 complex pole-pairs, 1 real pole, 1 complex zero-pair

Specification - 5% overshoot

Subsequent undershoot not to exceed 1%

$$\text{Here } \frac{\theta_0(\rho)}{\theta_i} = \frac{a_1 r_1^2 r_2^2}{s_1^2} \cdot \frac{(p + \beta_1 - j\mu_1)(p + \beta_1 + j\mu_1)}{(p + \alpha_1)(p + \alpha_1 - j\omega_1)(p + \alpha_1 + j\omega_1)(p + \alpha_2 - j\omega_2)(p + \alpha_2 + j\omega_2)}$$

A principal pole-pair at either $-0.5 \pm j1$ or $-1 \pm j1$ would give the required damping, the second case in fact rendering the undershoot negligible.

Taking this position, we have

$$\gamma = \sum \Delta_Z - \sum \Delta_P = 0.064$$

Let $a_1 = 4\alpha_1$, $\alpha_2 \pm j\omega_2 = 3\alpha_1 \pm j0.6\omega_1$. This is again quite arbitrary, the factors governing the choice simply being that a_1 and α_2 should both be relatively large compared with α_1 , and that the poles should not be too close to each other. From the charts,

$$\begin{array}{rcl} \Delta_{a_1} = 0.037 & L_{a_1} = 0.32 & \\ \Delta_{r_2} = 0.125 & L_{r_2} = 0.875 & \\ \hline \sum \Delta_P = 0.162 & \sum L_P = 1.195 & \end{array}$$

Hence $\sum \Delta_Z = 0.162 + 0.064 = 0.226$. From the overshoot chart 2, a zero-pair position of $-2.5\alpha_1 \pm j0.2\omega_1$, will give the necessary increment and is also in the vicinity of a complex pole-pair.

Therefore $-\beta_1 \pm j\mu_1 = -2.5\alpha_1 \pm j0.2\omega_1$, $L_{\beta_1} = 1.17$. This gives the required transfer function

$$\begin{aligned} \theta_o(p) &= \frac{(4k)(1.414k)^2(3.06k)^2}{(2.507k)^2} \frac{(p+2.5k-j0.2k)(p+2.5k+j0.2k)}{(p+4k)(p+k-jk)(p+k+jk)(p+3k-j0.6k)(p+3k+j0.6k)} \\ &= \frac{11.91k^3(p+2.5k-j0.2k)(p+2.5k+j0.2k)}{(p+4k)(p+k-jk)(p+k+jk)(p+3k-j0.6k)(p+3k+j0.6k)} \quad (71) \end{aligned}$$

The response to a unit-step input is

$$\theta_o(t') = 1 + 1.666 \varepsilon^{-t'} \sin(t' - 2.388) + 0.643 \varepsilon^{-3t'} \sin(0.6t' - 4.695) - 0.501 \varepsilon^{-4t'} \quad (72)$$

Maximum overshoot occurs at $t' = \sqrt{1 + 1.195} - 1.17 = 3.167$, at which time the principal mode term has the value 0.0492 and the remaining terms are then negligible. Fig. 32 below shows the response.

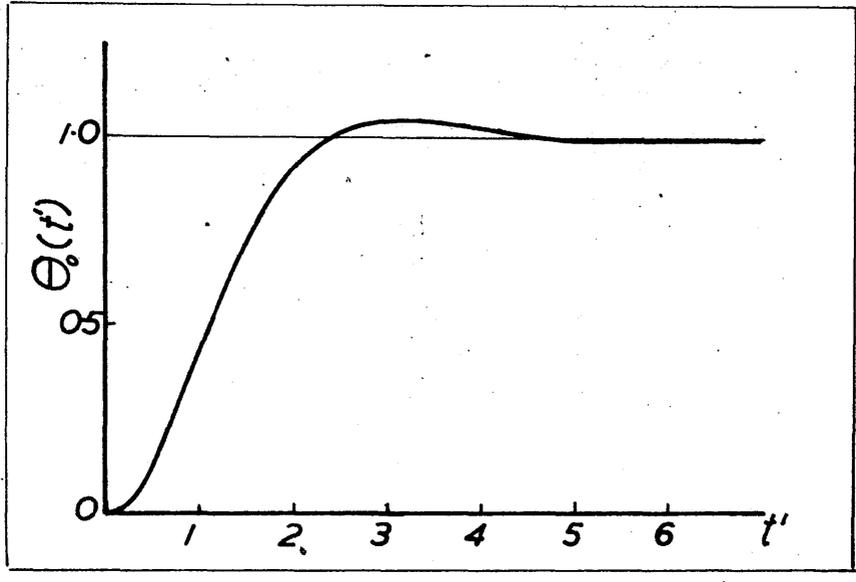


Fig. 32.

7.3. Type 2 Servo-Mechanisms.

As a result of the change to the basically unstable Type 2 system, we have the added condition stated in Sec. 6.2, namely that, for a transfer function of degrees s and r in the numerator and denominator respectively.

$$(P_1 P_2 \dots P_r) \cdot (\text{sum of products } Z_1 Z_2 \dots \text{ etc. taken } (s-1) \text{ at a time}) =$$

$$(Z_1 Z_2 \dots Z_s) \cdot (\text{sum of products } P_1 P_2 \dots \text{ etc. taken } (r-1) \text{ at a time}),$$

where $P_1, P_2 \dots$ etc. and $Z_1, Z_2 \dots$ etc. are minus the poles and zeros.

This complex relation very much restricts the possible pole and zero positions. For a transfer function of the same form as Example 3 of the previous Section, the condition becomes

$$(a_1 r_1^2 r_2^2) 2\beta_1 = s_1^2 (r_1^2 r_2^2 + 2\alpha_1 \alpha_2 r_1^2 + 2\alpha_1 \alpha_2 r_2^2) \tag{73}$$

or

$$\alpha_1 = \frac{r_1^2 r_2^2 s_1^2}{2\beta_1} \frac{1}{r_1^2 r_2^2 - (2\alpha_2 r_1^2 + 2\alpha_1 r_2^2) \frac{s_1^2}{2\beta_1}}$$

Since $a_1 > 0$ we must have $s_1^2/2\beta_1 < \frac{\gamma_1^2 \gamma_2^2}{2\alpha_2 \gamma_1^2 + 2\alpha_1 \gamma_2^2}$. It is convenient before proceeding further to make a Table of $s_1^2/2\beta_1$ for a variety of positions of the zero. This is given below

TABLE VI
Values of $s_1^2/2\beta_1 = (\beta^2 + \mu^2)/2\beta$

μ	$\beta = 0.5$	1.0	1.5	2.0	2.5	3.0	large
0.2	0.29	0.52	0.763	1.01	1.258	1.507	$\beta/2$
0.4	0.41	0.58	0.803	1.04	1.282	1.525	"
0.6	0.61	0.68	0.87	1.09	1.322	1.56	"
0.8	0.89	0.82	0.963	1.16	1.378	1.607	"
1.0	1.25	1.0	1.083	1.25	1.45	1.667	"
1.2	1.69	1.22	1.23	1.36	1.538	1.74	"
1.5	2.5	1.625	1.5	1.562	1.70	1.875	"
2.0	4.25	2.5	2.083	2.0	2.05	2.167	"
2.5	6.5	3.625	2.833	2.562	2.5	2.542	"

From a general point of view it can be seen from the Table that $s_1^2/2\beta_1$ will be lowest when both β_1 and μ_1 are small. Reference to the charts will show however that the contribution to the logarithmic overshoot increases very rapidly under those conditions. Herein lies the limitation due to the change of Type. In view of the rather complex relation produced, very little headway can be made unless by numerical trial and error processes. It is precisely on account of such difficulties that a design in terms of

of the poles and zeros of the error to input transfer function is preferred (see Chapters 8 and 9). A solution in terms of $\theta_o/\theta_i(p)$ is, however, carried out below.

Example 1. 2 complex pole-pairs, 1 real pole, 1 complex zero-pair.

Specification - 5% overshoot

Subsequent undershoot not to exceed 1%.

This is the same form of $\theta_o/\theta_i(p)$ and the same specification as in the previous (Type 1) example. It represents a good performance for a Type 2 system. Whether or not these requirements could be relaxed would, in a practical system, depend upon the size of the system and upon the resulting values of the circuit components. This and similar questions of a practical nature are not at present under consideration.

$$\text{We have } \frac{\theta_o(p)}{\theta_i} = \frac{a_1^2 r_1^2 r_2^2}{s_j^2} \frac{(\rho + \beta_1 - j\mu_1)(\rho + \beta_1 + j\mu_1)}{(\rho + \alpha_1)(\rho + \alpha_1 - j\omega_1)(\rho + \alpha_1 + j\omega_1)(\rho + \alpha_2 - j\omega_2)(\rho + \alpha_2 + j\omega_2)} \quad (74)$$

$$\text{with } a_1 = \frac{r_1^2 r_2^2 s_1^2 / 2\beta_1}{r_1^2 r_2^2 - (2\alpha_2 r_1^2 + 2\alpha_1 r_2^2) \frac{s_1^2}{2\beta_1}} \quad (75)$$

$$\text{and therefore } \frac{s_1^2}{2\beta_1} < \frac{r_1^2 r_2^2}{2\alpha_2 r_1^2 + 2\alpha_1 r_2^2} \quad (76)$$

By way of comparing the effect of the change of Type, suppose the same complex poles are adopted as in the previous example, that is

$$-\alpha_1 \pm j\omega_1 = -1 \pm j1$$

$$-\alpha_2 \pm j\omega_2 = -3 \pm j0.6, \quad \Delta r_2 = 0.125$$

As before, we require $\gamma = \Sigma \Delta_z - \Sigma \Delta_p = 0.064$. Substituting in (76)

gives $\frac{s_1^2}{2\beta_1} < 0.609$ and from Table VI, a value of $-\beta_1 \pm j\mu_1$ in the neighbourhood of $-1 \pm j0.2$ is indicated. Suppose we take

$$-\beta_1 \pm j\mu_1 = -1.1 \pm j0.2$$

giving $\frac{s_1^2}{2\beta_1} = 0.568$ and from the overshoot chart 2, $\Delta_{s_1} = 1.175$. From (75)

$$a_1 = 8.38$$

and hence

$$\Delta_{a_1} = 0.007$$

$$\text{Therefore } \gamma = 1.175 - 0.007 - 0.125 = 1.043$$

which is much too large. This in fact will give a logarithmic overshoot of $1.043 - 1.365 = -0.322$, which corresponds to about 48%. Practically all of this is due to the increment from the zero-pair at $-1 \pm j0.2$, which is fixed by the relation (76). If the overshoot is to be reduced, must be decreased by moving the zero-pair to the left. This entails an increase in $\frac{s_1^2}{2\beta_1}$, and hence in the R.H.S. of (76). It is possible to increase this if either or both complex-pairs are moved to the left.

If, therefore, we take $-\alpha_1 \pm j\omega_1 = -2 \pm j1$

then $\frac{s_1^2}{2\beta_1} < \frac{s_2^2}{10\alpha_2 + 4\beta_2^2} \rightarrow 1.25$ as α_2 increases indefinitely.

$$\text{Also } \gamma = \sum \Delta_z - \sum \Delta_p = 1.429$$

From Table VI and the overshoot chart 3, it will be seen that Δ_{s_1} , itself will approach the value of γ , as soon as $\frac{s_1^2}{2\beta_1}$ decreases even a small amount below the value 1.25. In turn, this means that both a_1 and α_2 are bound to be large, since relation (76) will barely be satisfied. Let us therefore account γ wholly to the zero and neglect $\sum \Delta_p$. From the conditions

$$\Delta_{s_1} = 1.429, \quad \frac{s_1^2}{2\beta_1} < 1.25$$

the best position of the zero is about $-1.18\alpha_1 + j0.2$, for which $L_{s_1} = 2.4$.

This gives $\frac{s_1^2}{2\beta_1} = 1.189$, and from (75), after a few trial values, we obtain

$$a_1 = 130, \quad -\alpha_2 \pm j\omega_2 = -60 \pm j1$$

The value of ω_2 is immaterial from the overshoot point of view, provided that $\alpha_1 \gg \omega_2$. The required transfer function is therefore

$$\frac{\theta_o(p)}{\theta_i} = \frac{(130k)(2.236k)^2(60k)^2}{(2.365k)^2} \cdot \frac{(p+2.36k-j0.2k)(p+2.36k+j0.2k)}{(p+130)(p+2k-jk)(p+2k+jk)(p+60k-jk)(p+60k+jk)}$$

$$= \frac{4.165 \times 10^5 k^3 (p + 2.36k - j0.2k)(p + 2.36k + j0.2k)}{(p + 130k)(p + 2k - jk)(p + 2k + jk)(p + 60k - jk)(p + 60k + jk)} \quad (77)$$

and the response is

$$\theta_o(t') = 1 + 0.475 \varepsilon^{-2t'} \sin(t' - 0.249) - 98.2 \varepsilon^{-60t'} \sin(t' - 0.0022) - 0.65 \varepsilon^{-130t'} \quad (78)$$

shown in Fig. 33.

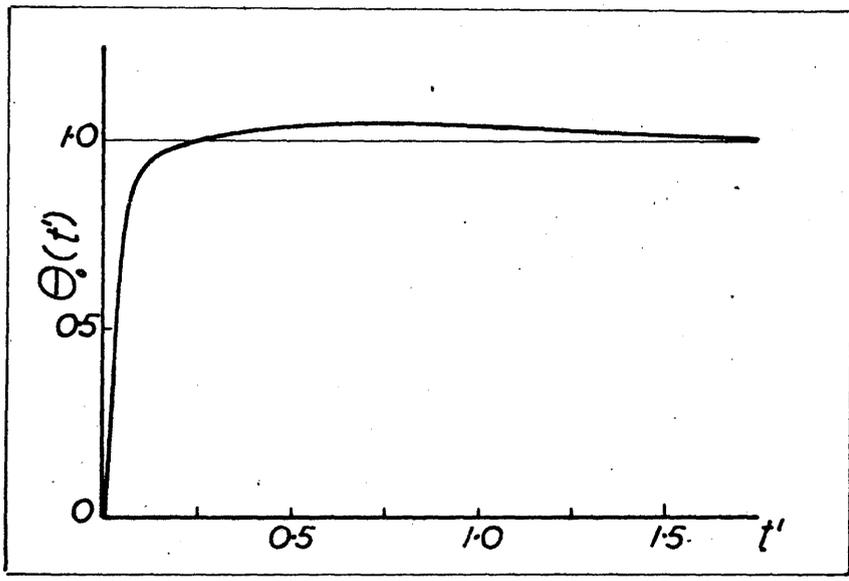


Fig. 33.

The maximum overshoot occurs at $t' = \pi + \sum L_p - \sum L_z = 0.742$, and is due entirely to the principal mode which has the value 0.051. No particular merit is claimed for the above transfer function; in fact, it would probably be better to obtain a solution using the negative term due to the real pole and allow the principal mode to contribute about 10%, instead of the present 5.1%. The design, however, does show that

- (a) low overshoots with Type 2 systems demand poles with very large damping,

- (b) the principal mode term has an initial value not greatly exceeding its value at the time of maximum overshoot,
- (c) the complex zeros and subsidiary complex poles approach the negative real axis, and may be considered as real.

A Type 2 example is now given in which a larger overshoot is permitted and only real zeros and subsidiary poles are considered. As a consequence of these more reasonable conditions, it will be seen that the subsidiary poles become considerably smaller than in (77)

Example 2. 1 complex pole-pair, 3 real poles, 2 real zeros

Specification - 15% overshoot

Subsequent undershoot not to exceed 1%.

$$\text{Then } \frac{\theta_o(p)}{\theta_i} = \frac{a_1 a_2 a_3 r_1^2}{b_1 b_2} \cdot \frac{(p + b_1)(p + b_2)}{(p + \alpha_1)(p + \alpha_2)(p + \alpha_3)(p + \alpha_1 - j\omega_1)(p + \alpha_1 + j\omega_1)} \quad (79)$$

A principal pole-pair at $-1 \pm j1$ will give the required damping. From

Table IV
$$Y = \sum \Delta_z - \sum \Delta_p = 0.542$$

As in the previous example, let Y be made up wholly from $\sum \Delta_z$, and for simplicity let us take $b_1 = b_2$. . .

Thus $\Delta_{b_1} = \Delta_{b_2} = 0.271$ and from the overshoot chart 2, $b_1 = b_2 = 1.7\alpha_1 = 1.7$.

The additional condition for a Type 2 system now gives

$$a_1 a_2 a_3 r_1^2 (b_1 + b_2) = b_1 b_2 [2\alpha_1 a_1 a_2 a_3 + r_1^2 (a_1 a_2 + a_1 a_3 + a_2 a_3)] \quad (80)$$

Further the least value which a_1 can have if Δ_{α_1} may be neglected is about $1/0.271$. In any case, as there will be two zeros to the right of α_1 , the contribution of this term, if any, to the overshoot will oppose that due to the principal mode. It will therefore be quite safe to fix a_1 at the

value $10k_1 = 10$. Substitution in (80) now gives the relation between a_2 and a_3 ,

$$a_2 = \frac{20a_3}{1.55a_3 - 20} \quad (81)$$

from which we see that a_3 must exceed $\frac{20}{1.55} \doteq 13$, if a_2 is to be positive. A number of values of a_2 and a_3 are possible. A reasonable spacing of the poles is produced however if $a_2 = 20$, $a_3 = 36.4$. We thus obtain the required transfer function

$$\frac{\theta_o(p)}{\theta_i} = \frac{(10k \times 20k \times 36.4k \times 1.414k)^2}{(1.7k)^2} \cdot \frac{(p+1.7k)^2}{(p+10k)(p+20k)(p+36.4k)(p+k-jk)(p+k+jk)}$$

$$\frac{5040k^3 (p+1.7k)^2}{(p+10k)(p+20k)(p+36.4k)(p+k-jk)(p+k+jk)} \quad (82)$$

This gives the response

$$\theta_o(t') = 1 + 0.869 \varepsilon^{-t'} \sin(t' - 0.627) - 1.604 \varepsilon^{-10t'} + 1.427 \varepsilon^{-20t'} - 0.307 \varepsilon^{-36.4t'} \quad (83)$$

which is shown below.

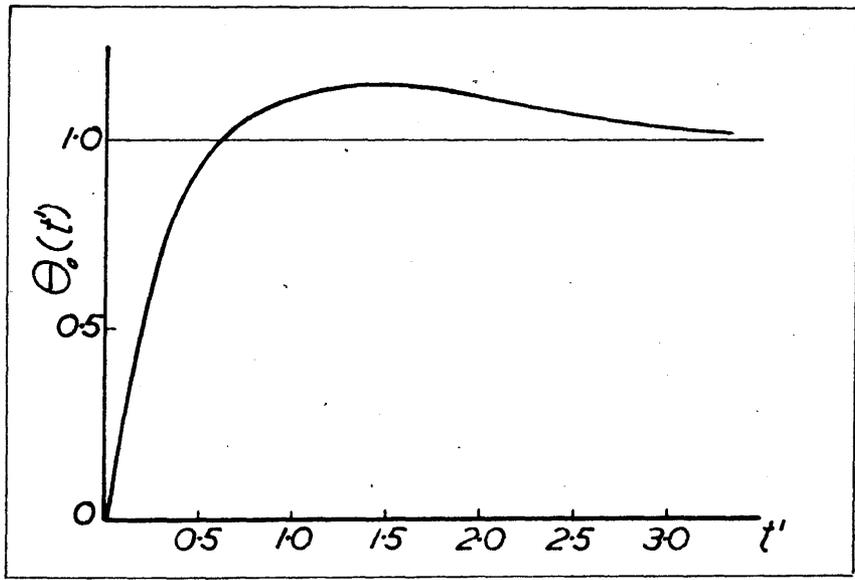


Fig. 34.

The maximum overshoot occurs at $t' = \pi + \Sigma L_p - \Sigma L_z$. From the angle chart 2, $L_{b1} = 0.97$, and ΣL_z is therefore 1.94. The approximation $1/(a_1 - \alpha_1)$ may be used to calculate ΣL_p . In this manner, $\pi + \Sigma L_p - \Sigma L_z = 1.452$ at which time the principal mode has the value 0.1496. The other terms are then negligible.

Sufficient examples have now been given to show that the design objectives are realizable with an accuracy ample for the purpose. All the above calculations have been performed with an ordinary slide-rule giving logarithms to three figures. Similarly three figures at the most can be extracted from any chart reading. Relatively coarse approximations may also be employed without greatly affecting the final result. One disconcerting feature, however, is the extra trouble experienced with Type 2 systems, although the previous example does show that a reasonable specification can be met without undue labour. It is nevertheless sufficient warning that the design of a Type 3 system by this method is not going to be easy, as two very complex equations containing products of poles taken $(r-1)$ and $(r-2)$ at a time, require to be satisfied simultaneously. In such cases, a better approach is afforded by designing in terms of the poles and zeros of the $e_{\theta_i}^{\prime}(p)$ transfer function. This method is therefore now considered.

CHAPTER 8.

THE DESIGN OF $e/\theta_i(p)$ TRANSFER FUNCTIONS.

8.1. Introductory. Transfer Function Forms.

The theory given in Chapter 6 is now developed in terms of the transfer function $e/\theta_i(p)$. In this case the distinctions between the various types of control become more apparent. This depends upon a minor difference in the forms of the transfer functions. Thus, from the basic servo

equations $\frac{\theta_o}{e} = Y(p)$, $\frac{e}{\theta_i}(p) = \frac{1}{1 + Y(p)}$

$$\frac{\theta_o(p)}{\theta_i} = \frac{Y(p)}{1 + Y(p)}$$

we have

$$\frac{e}{\theta_i}(p) = \frac{\theta_i/\theta_o(p) - 1}{\theta_i/\theta_o(p)}$$

Using the relations (33) to (35) of Sec. 6.2, gives

$$\begin{aligned} \text{Type 1. } \frac{e}{\theta_i}(p) &= \frac{(A_0 p^r + A_1 p^{r-1} + \dots + A_{r-1} p + A_r) - (\beta_0 p^s + \beta_1 p^{s-1} + \dots + \beta_{s-1} p + A_r)}{A_0 p^r + A_1 p^{r-1} + \dots + A_{r-1} p + A_r} \\ &= \frac{A_0 p^r + A_1 p^{r-1} + \dots + (A_{r-s} - \beta_0) p^s + \dots + (A_{r-1} - \beta_{s-1}) p}{A_0 p^r + A_1 p^{r-1} + \dots + A_{r-1} p + A_r} \end{aligned} \quad (84)$$

$$= \frac{p(p + U_1 X p + U_2) \dots (p + U_{r-1})}{(p + V_1 X p + V_2) \dots (p + V_r)} \quad (85)$$

where $U_1, U_2 \dots$ etc., and $V_1, V_2 \dots$ etc., are minus the zeros and poles of the function, and

$$U_1 U_2 \dots U_{r-1} = (A_{r-1} - \beta_{s-1}) / A_0$$

$$V_1 V_2 \dots V_r = A_r / A_0$$

As usual, the zeros and poles may be real or complex and in the latter case, they occur as conjugate pairs. Since the poles and zeros are to be confined to the left-hand plane, $U_1, U_2 \dots V_1, V_2 \dots$ etc., will all be positive real

quantities or complex with positive real parts. The analogous expressions for Type 2 and Type 3 controls are,*

Type 2.
$$\frac{\epsilon_i}{\theta_i}(p) = \frac{A_0 p^r + A_1 p^{r-1} + \dots + (A_{r-5} - B_0) p^5 + \dots + (A_{r-2} - B_{5-2}) p^2}{A_0 p^r + A_1 p^{r-1} + \dots + A_{r-1} p + A_r} \quad (86)$$

$$= \frac{\rho^2 (p + U_1)(p + U_2) \dots (p + U_{r-2})}{(p + V_1)(p + V_2) \dots (p + V_r)} \quad (87)$$

where $U_1 U_2 \dots U_{r-2} = (A_{r-2} - B_{5-2}) / A_0$
 $V_1 V_2 \dots V_r = A_r / A_0$

Type 3.
$$\frac{\epsilon_i}{\theta_i}(p) = \frac{A_0 p^r + A_1 p^{r-1} + \dots + (A_{r-5} - B_0) p^5 + \dots + (A_{r-3} - B_{5-3}) p^3}{A_0 p^r + A_1 p^{r-1} + \dots + A_r} \quad (88)$$

$$= \frac{\rho^3 (p + U_1)(p + U_2) \dots (p + U_{r-3})}{(p + V_1)(p + V_2) \dots (p + V_r)} \quad (89)$$

where $U_1 U_2 \dots U_{r-3} = (A_{r-3} - B_{5-3}) / A_0$
 $V_1 V_2 \dots V_r = A_r / A_0$

The transforms of the error quantity when these systems are subjected to unit step inputs, are given by the expressions (85), (87) and (89) with an additional ρ in the denominator. The error transforms for the control Types 1, 2 and 3, all of the r^{th} order, are therefore distinguished as follows,

* The Type 0 control may again be taken along with Type 1, although strictly speaking it yields a transfer function form

$$\frac{\epsilon_i}{\theta_i}(p) = \frac{(p + U_1)(p + U_2) \dots (p + U_r)}{(p + V_1)(p + V_2) \dots (p + V_r)} \quad \text{where}$$

$U_1 U_2 \dots U_r = (A_r - B_5) / A_0$ and $V_1 V_2 \dots V_r = A_r / A_0$. If, however, the gain of the system is high, then $A_r \doteq B_5$ and relation (85) may be used.

Type 1. r poles; $r-1$ zeros of which none occur at the origin; (90)

Type 2. r poles; $r-1$ zeros of which one occurs at the origin; (91)

Type 3. r poles; $r-1$ zeros of which two occur at the origin. (92)

It will be shown that these differences are more easily handled in design, than the complex expressions of poles and zeros which resulted from stating the equality of certain coefficients in the numerator and denominator of the $\theta_o/\theta_i(\rho)$ transfer function. A further advantage of the method is pointed out later.

Certain restrictions, however, do arise. From (84), (86) and (88), it will be seen that the coefficients of $\rho^r, \rho^{r-1} \dots \rho^{s+1}$ in the numerator and denominator are equal. The forms (85), (87) and (89) automatically ensure that the coefficients of highest powers agree, but for the next and lower powers of ρ , additional relations result. These are,

$$\text{coefficient of } \rho^{r-1} (A_1/A_o) \quad U_1 + U_2 + \dots + U_{r-1} = V_1 + V_2 + \dots + V_r \quad (93)$$

$$\text{coefficient of } \rho^{r-2} (A_2/A_o) \quad \Sigma'_2 U = \Sigma'_2 V \quad (94)$$

where $\Sigma'_2 U$ is the sum of all possible products of $U_1, U_2 \dots U_{r-1}$, taken two at a time, and similarly for $\Sigma'_2 V$.

$$\text{coefficient of } \rho^{r-3} (A_3/A_o) \quad \Sigma'_3 U = \Sigma'_3 V \quad (95)$$

where $\Sigma'_3 U$ is the sum of all possible products of $U_1, U_2 \dots U_{r-1}$, taken three at a time, and similarly for $\Sigma'_3 V$.

It is therefore clear that unless $s = r-2$, or at least $s = r-3$, the theory built upon the $e/\theta_i(\rho)$ transform gives little improvement over the method based upon the $\theta_o/\theta_i(\rho)$ transform. A second point to note is that

the number of zeros to be dealt with increases. This is not, however, a source of trouble. A brief review of theory analogous to that in Chapter 6 is now given.

8.2. General Solution of Principal Mode Approximation to Overshoot in Unit-Step Response, in terms of Poles and Zeros of $\epsilon/\theta_i(p)$ Transfer Function.

As the results for Types 2 and 3 may be obtained from particular cases of the expression (85), in which one and two, respectively, of the zeros are allowed to approach and finally co-incide with the origin, it will be sufficient to carry out the general analysis in terms of a Type 1 error transform. Using the notation given in Sec. 6.3, this may be written,

$$\epsilon(p) = \frac{\prod_{i=1}^g (p + b_i) \prod_{k=1}^n (p + \beta_k - j\mu_k)(p + \beta_k + j\mu_k)}{\prod_{h=1}^f (p + \alpha_h) \prod_{j=1}^m (p + \alpha_j - j\omega_j)(p + \alpha_j + j\omega_j)} \quad (96)$$

where $g + 2n = f + 2m - 1$

The pole at $p = -\alpha_j + j\omega_j$ gives

$$\begin{aligned} & \frac{\prod_{i=1}^g (-\alpha_j + j\omega_j + b_i) \prod_{k=1}^n (-\alpha_j + j\omega_j + \beta_k - j\mu_k)(-\alpha_j + j\omega_j + \beta_k + j\mu_k)}{\prod_{h=1}^f (-\alpha_j + j\omega_j + \alpha_h) \prod_{q=1, q \neq j}^m (-\alpha_j + j\omega_j + \alpha_q - j\omega_q)(-\alpha_j + j\omega_j + \alpha_q + j\omega_q)} \cdot \frac{\epsilon^{-\alpha_j t} \epsilon^{j\omega_j t}}{\epsilon} \\ &= \frac{\prod_{i=1}^g y_{ij} \epsilon^{j(\pi - \beta_{ij})} \prod_{k=1}^n z_{kj} z'_{kj} \epsilon^{j(2\pi - \sigma_{kj} + \sigma'_{kj})}}{\prod_{h=1}^f x_{hj} \epsilon^{j(\pi - \lambda_{hj})} \prod_{q=1, q \neq j}^m u_{qj} u'_{qj} \epsilon^{j(2\pi - \psi_{qj} + \psi'_{qj})}} \cdot \frac{\epsilon^{-\alpha_j t} \epsilon^{j\omega_j t}}{2j\omega_j} \end{aligned}$$

The pole at $p = -\alpha_j - j\omega_j$ gives the conjugate of this, so that the pole-pair

at $-\alpha_j \pm j\omega_j$ gives

$$\frac{\prod_{i=1}^g y_{ij} \prod_{k=1}^n z_{kj} z'_{kj}}{\prod_{h=1}^f x_{hj} \prod_{q=1, q \neq j}^m u_{qj} u'_{qj}} \cdot \frac{\varepsilon^{-\alpha_j t}}{\omega_j} \quad (97)$$

$$\sin \left[\omega_j t + \sum_{l=1}^g (\bar{\pi} - \beta_{lj}) + \sum_{k=1}^n (2\bar{\pi} - \overline{\sigma_{kj} + \sigma'_{kj}}) - \sum_{h=1}^f (\bar{\pi} - \lambda_{hj}) - \sum_{q=1, q \neq j}^m (2\bar{\pi} - \overline{\psi_{qj} + \psi'_{qj}}) \right]$$

This expression we shall denote by $M_j(t)$. The contribution from the m complex poles is therefore $\sum_{j=1}^m M_j(t)$. In particular, the principal mode term is

$$M_1(t) = \frac{\prod_{i=1}^g y_{i1} \prod_{k=1}^n z_{k1} z'_{k1}}{\prod_{h=1}^f x_{h1} \prod_{j=2}^m u_{j1} u'_{j1}} \cdot \frac{\varepsilon^{-\alpha_1 t}}{\omega_1} \quad (98)$$

$$\sin \left[\omega_1 t + \sum_{l=1}^g (\bar{\pi} - \beta_{l1}) + \sum_{k=1}^n (2\bar{\pi} - \overline{\sigma_{k1} + \sigma'_{k1}}) - \sum_{h=1}^f (\bar{\pi} - \lambda_{h1}) - \sum_{j=2}^m (2\bar{\pi} - \overline{\psi_{j1} + \psi'_{j1}}) \right]$$

In like manner, the contribution from the real poles will be denoted by $\sum_{h=1}^f N_h(t)$, where $N_h(t)$ is the term resulting from the pole $-a_h$, and given by

$$\frac{\prod_{i=1}^g (b_i - a_h) \prod_{k=1}^n (\beta_k - a_h^2 + \mu_k^2)}{\prod_{l=1, l \neq h}^f (a_l - a_h) \prod_{j=1}^m (\alpha_j - a_h^2 + \omega_j^2)} \cdot \varepsilon^{-a_h t} \quad (99)$$

The error is therefore

$$e(t) = \sum_{j=1}^m M_j(t) + \sum_{h=1}^f N_h(t) \quad (100)$$

Assuming that the principal mode approximation holds, (98) is now differentiated in order to calculate the maximum value of the first overshoot.

This occurs when

$$\omega_1 t + \sum_{l=1}^g (\bar{\pi} - \rho_{li}) + \sum_{k=1}^n (2\bar{\pi} - \overline{\sigma_{ki}} + \sigma'_{ki}) - \sum_{h=1}^f (\bar{\pi} - \lambda_{hi}) - \sum_{j=2}^m (2\bar{\pi} - \overline{\psi_{ji}} + \psi'_{ji}) - \phi_1 = \bar{\pi},$$

that is, when

$$t = \frac{1}{\omega_1} \left[\bar{\pi} + \phi_1 + \sum_{h=1}^f (\bar{\pi} - \lambda_{hi}) + \sum_{j=2}^m (2\bar{\pi} - \overline{\psi_{ji}} + \psi'_{ji}) - \sum_{l=1}^g (\bar{\pi} - \rho_{li}) - \sum_{k=1}^n (2\bar{\pi} - \overline{\sigma_{ki}} + \sigma'_{ki}) \right] \quad (101)$$

The condition (101), substituted in (98), gives $M_1(t)$ the required negative sign, (minimum values of the error quantity corresponding to overshoots of the output). The expression for t is the time of the first minimum of (98) and therefore the time at which the first overshoot has its maximum value.

The size of this maximum is given by

$$\frac{\prod_{l=1}^g y_{li} \prod_{k=1}^n z_{ki} z'_{ki}}{\prod_{h=1}^f x_{hi} \prod_{j=2}^m u_{ji} v'_{ji}} \cdot \frac{\omega_1}{\bar{\sigma}_1} \left[\bar{\pi} + \phi_1 + \sum_{h=1}^f (\bar{\pi} - \lambda_{hi}) + \sum_{j=2}^m (2\bar{\pi} - \overline{\psi_{ji}} + \psi'_{ji}) - \sum_{l=1}^g (\bar{\pi} - \rho_{li}) - \sum_{k=1}^n (2\bar{\pi} - \overline{\sigma_{ki}} + \sigma'_{ki}) \right] \quad (102)$$

$$= \frac{\omega_1}{\bar{\sigma}_1} \cdot \frac{\prod_{l=1}^g y_{li} \sum_{i=1}^g \frac{\alpha_i}{\omega_i} (\bar{\pi} - \rho_{li}) \prod_{k=1}^n z_{ki} z'_{ki} \sum_{i=1}^g \frac{\alpha_i}{\omega_i} (2\bar{\pi} - \overline{\sigma_{ki}} + \sigma'_{ki})}{\prod_{h=1}^f x_{hi} \sum_{i=1}^g \frac{\alpha_i}{\omega_i} (\bar{\pi} - \lambda_{hi}) \prod_{j=2}^m u_{ji} v'_{ji} \sum_{i=1}^g \frac{\alpha_i}{\omega_i} (2\bar{\pi} - \overline{\psi_{ji}} + \psi'_{ji})} \quad (103)$$

This is an entirely analogous expression to relation (53) of Chapter 6. In this case, however, the generalised overshoot is given in terms of factors, $y_{ii} \varepsilon^{\frac{\alpha_i}{\omega_i} (\bar{\pi} - \rho_i)}$ etc., which depend on the poles and zeros of the $\varepsilon/\theta_i(\rho)$ transform. These factors do not differ greatly in form from those resulting in the $\theta_0/\theta_i(\rho)$ analysis. The quantities are compared below:-

<u>Pole or Zero Type.</u>	<u>Form in θ_0/θ_i analysis</u>	<u>Form in ε/θ_i analysis.</u>	
Real pole	$x/\alpha \cdot \varepsilon^{\frac{\alpha_i}{\omega_i} (\bar{\pi} - \lambda)}$	$x \cdot \varepsilon^{\frac{\alpha_i}{\omega_i} (\bar{\pi} - \lambda)}$	(104)

Complex pole-pair	$uv'/r \cdot \varepsilon^{\frac{\alpha_i}{\omega_i} (2\bar{\pi} - \bar{\psi} + \bar{\psi}')$	$uv' \cdot \varepsilon^{\frac{\alpha_i}{\omega_i} (2\bar{\pi} - \bar{\psi} + \bar{\psi}')}$	(105)
-------------------	--	---	-------

Real zero	$y/b \cdot \varepsilon^{\frac{\alpha_i}{\omega_i} (\bar{\pi} - \rho)}$	$y \varepsilon^{\frac{\alpha_i}{\omega_i} (\bar{\pi} - \rho)}$	(106)
-----------	--	--	-------

Complex zero-pair	$zz'/s \cdot \varepsilon^{\frac{\alpha_i}{\omega_i} (2\bar{\pi} - \bar{\sigma} + \bar{\sigma}')$	$zz' \varepsilon^{\frac{\alpha_i}{\omega_i} (2\bar{\pi} - \bar{\sigma} + \bar{\sigma}')}$	(107)
-------------------	--	---	-------

Charts giving the logarithms of the quantities in the right-hand column have been computed. Calculations with these charts are based, for Type 1 servo-mechanisms, upon the expression (103). The overshoots in the case of Type 2 and Type 3 servo-mechanisms are given by the expressions (108) and (110), deduced below.

Type 2 servo-mechanism.

In the expression (103), let the real zero b_g tend to the origin. Then

$$y_{g_i} \rightarrow r_i \quad \text{and} \quad (\bar{\pi} - \rho_{g_i}) \rightarrow (\bar{\pi} - \phi_i)$$

The principal mode approximation to the overshoot is therefore

$$\varepsilon^{\frac{\alpha_i}{\omega_i} 2\phi_i} \cdot \frac{\prod_{l=1}^{g-1} y_{li} \varepsilon^{\frac{\alpha_i}{\omega_i} (\bar{\pi} - \rho_{li})} \prod_{k=1}^n z_{ki} z'_{ki} \varepsilon^{\frac{\alpha_i}{\omega_i} (2\bar{\pi} - \bar{\sigma}_{ki} + \bar{\sigma}'_{ki})}}{\prod_{h=1}^f x_{hi} \varepsilon^{\frac{\alpha_i}{\omega_i} (\bar{\pi} - \lambda_{hi})} \prod_{j=2}^m v_{ji} v'_{ji} \varepsilon^{\frac{\alpha_i}{\omega_i} (2\bar{\pi} - \bar{\psi}_{ji} + \bar{\psi}'_{ji})}} \quad (108)$$

Also from the expression (101), the time for maximum overshoot is obtained by letting $(\bar{\pi} - \rho_{g_i}) \rightarrow (\bar{\pi} - \phi_i)$. This gives

$$t = \frac{1}{\omega_1} \left[2\phi_1 + \sum_{h=1}^f (\bar{\pi} - \lambda_{h1}) + \sum_{j=2}^m (2\bar{\pi} - \psi_{j1} + \psi'_{j1}) - \sum_{l=1}^{q-1} (\bar{\pi} - \beta_{l1}) - \sum_{k=1}^n (2\bar{\pi} - \sqrt{\sigma_{k1} + \sigma'_{k1}}) \right] \quad (109)$$

Type 3 servo-mechanism.

In a similar manner to the above, let two real zeros tend to the origin, that is,

$$y_{g1}, y_{(g-1)1} \rightarrow r_1$$

and

$$(\bar{\pi} - \beta_{g1}), (\bar{\pi} - \beta_{(g-1)1}) \rightarrow \bar{\pi} - \phi_1$$

Substitution in (103) gives an overshoot of magnitude

$$r_1 \varepsilon^{\frac{\alpha_1}{\omega_1} (3\phi_1 - \bar{\pi})} \cdot \frac{\varepsilon^{\frac{g-2}{\omega_1} \sum_{j=2}^m (\bar{\pi} - \beta_{j1})} \prod_{k=1}^n z_{k1} z'_{k1} \varepsilon^{\frac{\alpha_1}{\omega_1} (2\bar{\pi} - \sqrt{\sigma_{k1} + \sigma'_{k1}})}}{\varepsilon^{\frac{\alpha_1}{\omega_1} \sum_{h=1}^f (\bar{\pi} - \lambda_{h1})} \prod_{j=2}^m v_{j1} v'_{j1} \varepsilon^{\frac{\alpha_1}{\omega_1} (2\bar{\pi} - \psi_{j1} + \psi'_{j1})}} \quad (110)$$

Likewise from (101), the time at which this occurs is

$$t = \frac{1}{\omega_1} \left[3\phi_1 - \bar{\pi} + \sum_{h=1}^f (\bar{\pi} - \lambda_{h1}) + \sum_{j=2}^m (2\bar{\pi} - \psi_{j1} + \psi'_{j1}) - \sum_{l=1}^{g-2} (\bar{\pi} - \beta_{l1}) - \sum_{k=1}^n (2\bar{\pi} - \sqrt{\sigma_{k1} + \sigma'_{k1}}) \right] \quad (111)$$

With small values of ϕ_1 , this expression may become negative. In this case, the first minimum of the principal mode term occurs $2\bar{\pi}/\omega_1$ seconds later than that given by (111). There is no difficulty in calculating the size of the minimum, according to the methods previously given, but it must be checked if this minimum represents the first overshoot in the output. The principal mode term will now have a negative initial value and must pass through a positive maximum before the required minimum is reached. Interaction with a positive real term may produce a minimum value of the error quantity prior to this, and hence render the principal mode approximation quite invalid. A design is not impossible under these circumstances but it would be largely a matter of trial and error and is better avoided.

Before applying the above results to numerical designs, brief conclusions regarding the use of the least-damped real pole are given.

8.3. Use of Least-Damped Real Pole in addition to Principal Mode Term.

From the relation (99), the term resulting from the nearest or least-damped real pole, situated at $-a_1$, is

$$N_1(t) = \frac{\prod_{i=1}^g (b_i - a_1) \prod_{k=1}^n (\beta_k - a_1^2 + \mu_k^2)}{\prod_{h=2}^f (a_h - a_1) \prod_{j=1}^m (\alpha_j - a_1^2 + \omega_j^2)} \cdot \xi^{-a_1 t} \quad (112)$$

This applies for a Type 1 servo-mechanism. In the case of a Type 2 system, the term becomes

$$N_1(t) = -a_1 \frac{\prod_{i=1}^{g-1} (b_i - a_1) \prod_{k=1}^n (\beta_k - a_1^2 + \mu_k^2)}{\prod_{h=2}^f (a_h - a_1) \prod_{j=1}^m (\alpha_j - a_1^2 + \omega_j^2)} \cdot \xi^{-a_1 t} \quad (113)$$

and for Type 3,

$$N_1(t) = a_1^2 \frac{\prod_{i=1}^{g-1} (b_i - a_1) \prod_{k=1}^n (\beta_k - a_1^2 + \mu_k^2)}{\prod_{h=2}^f (a_h - a_1) \prod_{j=1}^m (\alpha_j - a_1^2 + \omega_j^2)} \cdot \xi^{-a_1 t} \quad (114)$$

If now the term $N_1(t)$ is to help in reducing the overshoot by subtracting from the principal mode value at that point, the sign of $N_1(t)$ should be positive as shown in Fig. 27. But $a_1 < a_h$ and therefore $\prod_{h=2}^f (a_h - a_1) > 0$. Also $\prod_{j=1}^m (\alpha_j - a_1^2 + \omega_j^2) > 0$ and $\prod_{k=1}^n (\beta_k - a_1^2 + \mu_k^2) > 0$. We have therefore the following conditions in order that $N_1(t)$ should be positive -

- Type 1. No zeros or an even number of zeros to the right of a_1 ;
- Type 2. An odd number of zeros to the right of a_1 ;
- Type 3. As Type 1.

It may not always be possible to use this technique, for example, in Type 2 systems having few zeros and poles, and it is impossible to apply it in a Type 2 system with only one zero (distinct from the origin) and only one pole. This follows from condition (93), in which $u_1 = b_1$ and $\sum V = a_1 + 2\alpha_1$, and hence b_1 is necessarily greater than a_1 .

8.4. Notes on the Design Charts.

1. The overshoot charts show the logarithm to the base 10 of the quantities $y \varepsilon^{\frac{\alpha_1 (\pi - \rho)}{\omega_1}}$ and $z z' \varepsilon^{\frac{\alpha_1 (2\pi - \sigma + \sigma')}{\omega_1}}$ for real zeros or poles and complex zero- or pole-pairs respectively. As these charts simply correspond to the $\theta_{\theta_i}(\rho)$ overshoot charts already described, similar remarks pertaining to the abscissae and range of the charts apply.
2. For the calculation of the time of the maximum overshoot in accordance with equations (101), (109), and (111), the angles $(\pi - \rho)$, $(2\pi - \sigma + \sigma')$ etc., may be taken from the $\theta_{\theta_i}(\rho)$ angle charts already provided.
3. For poles and zeros beyond the abscissa limit 10, the quantities $\log_{10} y \varepsilon^{\frac{\alpha_1 (\pi - \rho)}{\omega_1}}$ and $\log_{10} z z' \varepsilon^{\frac{\alpha_1 (2\pi - \sigma + \sigma')}{\omega_1}}$ are approximated by $\log_{10} b$ and $\log_{10} s^2$ respectively. The approximations for $(\pi - \rho)$ and $(2\pi - \sigma + \sigma')$ are the same as those adopted previously (Sec. 6.5, Note 5).

Whereas at first sight, it would seem that these approximations, however necessary, constitute a drawback to the method, in fact the very opposite holds - especially in the $\theta_{\theta_i}(\rho)$ form of the design. It is a fact that in this method the inherently unstable Type 2 and 3 systems normally require all their subsidiary poles and zeros to be remotely situated from

the principal pole-pair and thus the use of the charts is completely unnecessary. Furthermore, as it is immaterial whether remote poles or zeros are real or complex, it is easier to make these wholly real. A design may therefore be worked out entirely by the use of a slide-rule, as logarithms to three figures are amply sufficient. A very real advantage is thus obtained where it is most needed - in the design of satisfactory $e/\theta_i(p)$ transfer functions for the basically unstable Type 2 and Type 3 servo-mechanisms. Such designs are demonstrated in the following Chapter.

CHAPTER 9.

USE OF THE $\epsilon/\theta_i(p)$ DESIGN CHARTS. COMPARISON OF $\theta/\theta_i(p)$ AND $\epsilon/\theta_i(p)$ METHODS.

9.1. Introductory. Tables of Constants.

It has been pointed out in Sec. 7.3 that Type 2 and Type 3 transfer functions are best designed with only real subsidiary poles and real zeros. This also holds for the present method and in the following designs only such poles and zeros are considered. Such a procedure simplifies the use of the charts, by confining the readings to the curves, $\mu = 0$, and avoiding the awkward "bunching" of the curves pertaining to complex poles and zeros. Complete charts have been given, however, as it may happen that a complex zero-pair is desired.

Before considering the examples, it is useful to prepare a Table, similar to Table IV, which gives $Y = \Sigma \Delta_z - \Sigma \Delta_p$ for a number of overshoots each with the three degrees of damping. This is based upon the expressions (103), (108) and (110) for Types 1, 2 and 3 servo-mechanisms respectively.

In logarithmic form, these are

$$\text{Type 1.} \quad X = \log_{10} \text{overshoot} = -\log_{10} r_1 - 0.4343 \frac{\alpha_1 (\pi + \phi_1)}{\omega_1} + Y \quad (115)$$

$$\text{Type 2.} \quad X = -0.8686 \frac{\alpha_1 \phi_1}{\omega_1} + Y \quad (116)$$

$$\text{Type 3.} \quad X = \log_{10} r_1 - 0.4343 \frac{\alpha_1 (3\phi_1 - \pi)}{\omega_1} + Y \quad (117)$$

We have therefore the following Table giving the value of $Y = \Sigma \Delta_z - \Sigma \Delta_p$.

TABLE VII

Values of $Y = \sum \Delta_z - \sum \Delta_p$

Type	$-d_1 \pm j\omega_1$	Overshoot					Log. Overshoot X
		5%	10%	15%	20%	25%	
1	$-0.5 \pm j1$	-0.33	-0.029	+0.148	+0.272	+0.369	Y -0.971
	$-1 \pm j1$	+0.556	+0.857	+1.034	+1.158	+1.255	Y -1.857
	$-2 \pm j1$	+2.179	+2.48	+2.657	+2.781	+2.878	Y -3.48
2	$-0.5 \pm j1$	-0.82	-0.519	-0.342	-0.218	-0.121	Y -0.481
	$-1 \pm j1$	-0.619	-0.318	-0.141	-0.017	+0.080	Y -0.682
	$-2 \pm j1$	-0.497	-0.196	-0.019	+0.105	+0.202	Y -0.804
3	$-0.5 \pm j1$	-1.311	-1.01	-0.833	-0.709	-0.612	Y +0.01
	$-1 \pm j1$	-1.793	-1.492	-1.315	+1.191	-1.094	Y +0.492
	$-1 \pm j1^*$	+0.939	+1.24	+1.417	+1.541	+1.638	Y -2.24
	$-2 \pm j1$	-3.174	-2.873	-2.696	-2.572	-2.475	Y +1.873
	$-2 \pm j1^*$	+2.285	+2.586	+2.763	+2.887	+2.984	Y -3.586

* Entries here are calculated from

$$X = \log_{10} r_1 - 0.4343 \frac{\alpha_1}{\omega_1} (3\phi_1 + \pi) + Y$$

i.e. the case of maximum overshoot occurring $2\pi/\omega_1$ seconds later than that given by (111). hr

9.2. Examples.

Example 1. Type 2, 1 complex pole-pair, 3 real poles, 3 real zeros.

Specification - 15% maximum overshoot

Subsequent undershoot of the order of 2% is

permitted.

Then

$$\frac{\epsilon_i(\rho)}{\theta_i(\rho)} = \frac{\rho(\rho + b_1)(\rho + b_2)(\rho + b_3)}{(\rho + a_1)(\rho + a_2)(\rho + a_3)(\rho + \alpha_1 - j\omega_1)(\rho + \alpha_1 + j\omega_1)} \quad (118)$$

For this example, let us take the case when the degree of the denominator in $\theta_0/\theta_i(\rho)$ exceeds that of the numerator by 2. This gives the condition (93) of Sec. 8.1 as follows,

$$b_1 + b_2 + b_3 = a_1 + a_2 + a_3 + 2\alpha_1 \quad (119)$$

Since the position $-\pm j\omega_1$ for the principal pole-pair will give the required damping, the required γ from Table VII is

$$\gamma = \sum \Delta_z - \sum \Delta_p = -0.141$$

A number of designs are tried.

(a) Suppose $a_1 = 10, a_2 = 15, a_3 = 20, b_1 = 5$. γ is then given as follows.

$$a_1 = 10 \quad \Delta_{a_1} = 1.0 \quad (\text{chart 2 or approximated by})$$

$$a_2 = 15 \quad \Delta_{a_2} = 1.176 \quad (\text{approximated by})$$

$$a_3 = 20 \quad \Delta_{a_3} = 1.301 \quad (\quad " \quad " \quad)$$

$$\sum a = 45 \quad \sum \Delta_p = 3.477 \therefore b_2 + b_3 = 42 \text{ and try } b_2 = 12, b_3 = 30. \text{ Thus}$$

$$\Delta_{b_1} = 0.72, \Delta_{b_2} = 1.079, \Delta_{b_3} = 1.477, \sum \Delta_z = 3.276$$

This gives $\gamma = -0.201$, \log_{10} overshoot = $-0.201 - 0.682 = -0.883$

and therefore about 13% overshoot.

(b) Suppose $a_1 = 5, a_2 = 10, a_3 = 20, b_1 = 4$.

Thus

$$a_1 = 5 \quad \Delta_{a_1} = 0.72$$

$$a_2 = 10 \quad \Delta_{a_2} = 1.0$$

$$a_3 = 20 \quad \Delta_{a_3} = 1.301$$

$$\sum a = 35 \quad \sum \Delta_p = 3.021$$

$$\therefore b_2 + b_3 = 33, \text{ and try } b_2 = 6, b_3 = 27.$$

Then

$$b_1 = 4 \quad \Delta_{b_1} = 0.64$$

$$b_2 = 6 \quad \Delta_{b_2} = 0.79$$

$$b_3 = 27 \quad \Delta_{b_3} = 1.431$$

$$\sum \Delta_z = 2.861$$

$$\text{and } \gamma = -0.161$$

$$X = \log_{10} \text{ overshoot} = -0.161 - 0.682 = -0.843, \text{ i.e. } 14.34\%$$

It will be near enough to adopt these pole positions. To complete the data, we have

$$L_{a_1} = 0.24, L_{a_2} = 0.11, L_{a_3} = 0.053$$

$$L_{b_1} = 0.32, L_{b_2} = 0.2, L_{b_3} = 0.037$$

$$\underline{\Sigma L_p = 0.403}$$

$$\underline{\Sigma L_z = 0.557}$$

Hence maximum overshoot will occur at

$$t^1 = 2\phi_1 + \Sigma \Delta p - \Sigma \Delta z = 1.417$$

The required transfer function is

$$\frac{\epsilon_i(p)}{\theta_i} = \frac{p^2 (p + 4k \chi p + 6k \chi p + 27k)}{(p + 5k \chi p + 10k \chi p + 20k \chi p + k - jk \chi p + k + jk)} \quad (120)$$

which gives the error

$$\epsilon(t^1) = \underline{0.8345 \epsilon^{-t^1} \sin(t^1 + 2.511) + 0.0863 \epsilon^{-5t^1} + 0.995 \epsilon^{-10t^1} - 0.577 \epsilon^{-20t^1}}; \quad (121)$$

the computed overshoot is 14.33%.

The above design is only one of a number which are possible with fairly wide variations in the poles and zeros. For instance the transfer function

$$\frac{\epsilon_i(p)}{\theta_i} = \frac{p^2 (p + 5k \chi p + 33.5k)^2}{(p + 10k \chi p + 20k \chi p + 40k \chi p + k - jk \chi p + k + jk)}$$

will also give about 15% overshoot. Such a wide choice of designs results primarily from having a relatively large number of poles and zeros. Conversely, a limited number of poles and zeros restricts their possible locations.

Example 2. Type 3, 1 complex pole-pair, 3 real poles, 2 real zeros.

Specification - 10% maximum overshoot

Subsequent undershoot of order of 2% permitted.

Then

$$\frac{\epsilon_i(p)}{\theta_i} = \frac{p^3 (p + b_1 \chi p + b_2)}{(p + a_1 \chi p + a_2 \chi p + a_3 \chi p + \alpha_1 - j\omega_1 \chi p + \alpha_1 + j\omega_1)} \quad (122)$$

Further let us take the case when the degree of the denominator in exceeds that of the numerator by 3. This gives the extra conditions

$$b_1 + b_2 = a_1 + a_2 + a_3 + 2\alpha_1 \quad (123)$$

$$b_1 b_2 = a_1 a_2 + a_1 a_3 + a_2 a_3 + 2\alpha_1 (a_1 + a_2 + a_3) + \eta^2 \quad (124)$$

In contrast to the previous design, this one is limited. Since a principal pole-pair at either $-1 \pm j$ or $-2 \pm j$ produces the condition of a principal mode maximum before the required minimum (assuming that all poles and zeros will be remote and hence $\sum L_p - \sum L_z \doteq 0$, the position $-0.5 \pm j$ will be taken. From Table VII therefore,

A number of designs are now tried.

(a) Suppose $a_1 = 10, a_2 = 20, b_1 = 40$. Conditions (123) and (124) give

$$\left. \begin{aligned} 40 + b_2 &= 31 + a_3 \\ 40 b_2 &= 231.25 + 31 a_3 \end{aligned} \right\}$$

from which $a_3 = 65.7, b_2 = 56.7$. The logarithmic increments

are	a_1	10	Δa_1	1.0	[$\log_{10} a_1$]	b_1	40	Δb_1	1.602	[$\log_{10} b_1$]
	a_2	20	Δa_2	1.301	etc.	b_2	56.7	Δb_2	1.754	etc.
	a_3	65.7	Δa_3	1.818						
			$\sum \Delta p$	4.119				$\sum \Delta z$	3.356	

$$Y = \sum \Delta z - \sum \Delta p = -0.763$$

$$X = -0.763 + 0.01 = -0.753 \quad \text{and overshoot} = 17.6\%$$

(b) Suppose $a_1 = 10, a_2 = 40, b_1 = 60$.

$$\left. \begin{aligned} \text{Then} \quad 60 + b_2 &= 51 + a_3 \\ 60 b_2 &= 451.25 + 51 a_3 \end{aligned} \right\}$$

from which $a_3 = 110, b_2 = 101$. Hence,

a_1	10	Δa_1	1.0	b_1	60	Δb_1	1.778
a_2	40	Δa_2	1.602	b_2	101	Δb_2	2.005
a_3	110	Δa_3	2.041				
		$\sum \Delta p$	4.643			$\sum \Delta z$	3.783

$$Y = \sum \Delta z - \sum \Delta p = -0.860$$

$$X = -0.860 + 0.01 = 0.85, \quad \text{and overshoot} = 14.1\%$$

In this example a number of other trials were necessary. These showed firstly, that the overshoot decreased as the sum $a_1 + a_2$ increased, more especially if a_1 and a_2 were not too far apart, and secondly, that the choice of b_2 did not greatly affect the overshoot for given values of a_1 and a_2 . The final design is

(c)	a_1	15	Δa_1	1.176	$L a_1$	0.069	b_1	60	Δb_1	1.778	$L b_1$	0.0168
	a_2	30	Δa_2	1.477	$L a_2$	0.0339	b_2	81.4	Δb_2	1.911	$L b_2$	0.0123
	a_3	95.4	Δa_3	1.980	$L a_3$	0.0105						

$$\text{Hence } Y = -0.944, \quad X = -0.944 + 0.01 = -0.934;$$

this gives a principal mode value of 11.64% at $t' = (3\phi_1 - \bar{\eta} + \sum L p - \sum L z) = 0.26$

In this design the term due to the nearest real pole will be positive since there are no zeros to the right of a_1 . The transfer function is

$$\frac{\epsilon(p)}{\epsilon_i} = \frac{p^3(p + 60k)(p + 81.4k)}{(p + 15k)(p + 30k)(p + 95.4k)(p + 0.5k - jk)(p + 0.5k + jk)} \quad (125)$$

and the error is

$$\epsilon(t') = \frac{0.148 \epsilon^{-0.5t'} \sin(t' - 3.988) + 2.643 \epsilon^{-15t'} - 1.625 \epsilon^{-30t'} + 0.0952 \epsilon^{-95.4t'}}{\quad} \quad (126)$$

Fig. 35 below indicates the conditions.

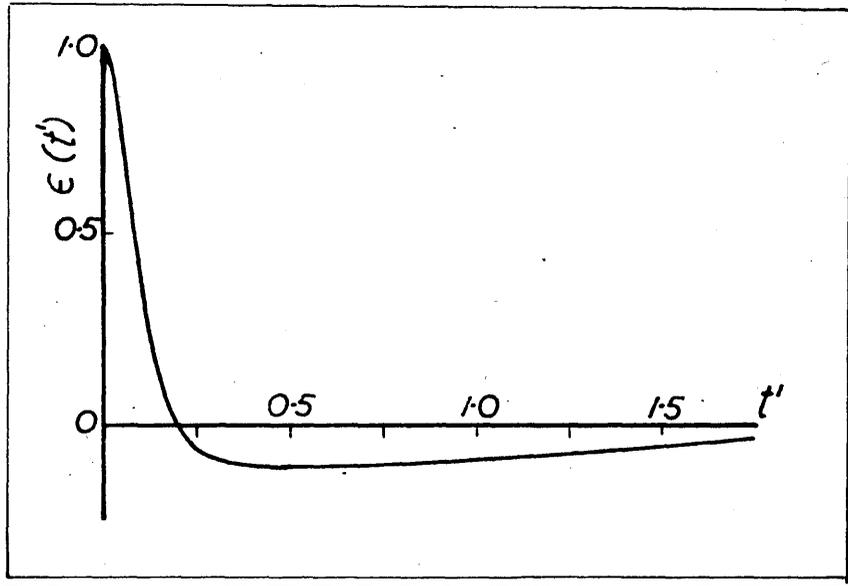


Fig. 35.

The value of the principal mode at $t' = 0.26$ is -0.1163 . The good agreement between the predicted and computed values, indicates that the approximations $\Delta_{a_i} = \log_{10} a_i$ etc., $L_{a_i} = 1/(a_i - \alpha_i)$ etc. are quite reasonable. It should be noted that although the term due to the nearest real pole causes the overshoot at $t' = 0.26$ to be about 6%, the maximum overshoot still is about 11% and occurs later. This is due to the slow decay of the principal mode term relative to the nearest real pole. A real pole with less damping than this would be required to materially alter the overshoot.

Example 3. Type 3, 1 complex pole-pair, 2 real poles, 1 real zero.

This example illustrates the derivation of the circuit values in a practical problem.*

* See reference 21, Appendix 2.

The open-loop transfer function of an induction motor-controlled servo-mechanism is given by

$$\frac{\theta_o(p)}{\epsilon} = \frac{3500}{p^2} \quad (127)$$

to a first approximation. It may be stabilised by the use of the input network shown below.

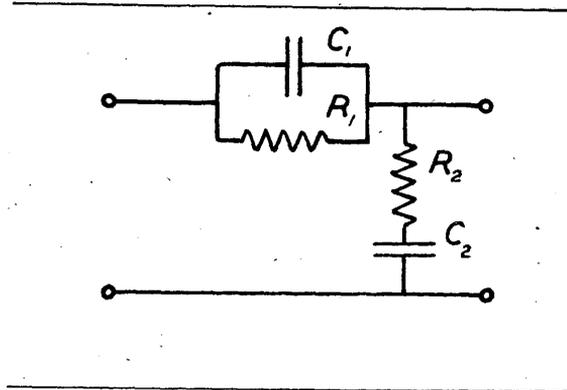


Fig. 36.

This has the transfer function,

$$\frac{e_o(p)}{e_i} = \frac{p^2 + p\left[\frac{1}{C_1 R_1} + \frac{1}{C_2 R_2}\right] + \frac{1}{C_1 R_1 C_2 R_2}}{p^2 + p\left[\frac{1}{C_1 R_1} + \frac{1}{C_2 R_2} + \frac{1}{C_1 R_2}\right] + \frac{1}{C_1 R_1 C_2 R_2}} \quad (128)$$

giving the resultant loop transfer function

$$\frac{\theta_o(p)}{\epsilon} = \frac{p^2 + p\left[\frac{1}{C_1 R_1} + \frac{1}{C_2 R_2}\right] + \frac{1}{C_1 R_1 C_2 R_2}}{p^2 + p\left[\frac{1}{C_1 R_1} + \frac{1}{C_2 R_2} + \frac{1}{C_1 R_2}\right] + \frac{1}{C_1 R_1 C_2 R_2}} \cdot \frac{3500}{p^2} \quad (129)$$

The $\frac{\epsilon_o(p)}{\theta_i}$ transfer function is therefore,

$$\frac{\epsilon_o(p)}{\theta_i} = \frac{p^2(p^2 + p\left[\frac{1}{C_1 R_1} + \frac{1}{C_2 R_2} + \frac{1}{C_1 R_2}\right] + \frac{1}{C_1 R_1 C_2 R_2})}{p^4 + p^3\left(\frac{1}{C_1 R_1} + \frac{1}{C_2 R_2} + \frac{1}{C_1 R_2}\right) + p^2\left(\frac{1}{C_1 R_1 C_2 R_2} + 3500\right) + p\left(\frac{1}{C_1 R_1} + \frac{1}{C_2 R_2}\right)3500 + \frac{3500}{C_1 R_1 C_2 R_2}} \quad (130)$$

As the steady-state acceleration error will be very small it is better to treat this as a Type 3 system. In this case we have

$$\frac{\epsilon_o(p)}{\theta_i} = \frac{p^3(p + \left[\frac{1}{C_1 R_1} + \frac{1}{C_2 R_2} + \frac{1}{C_1 R_2}\right])}{p^4 + p^3\left(\frac{1}{C_1 R_1} + \frac{1}{C_2 R_2} + \frac{1}{C_1 R_2}\right) + p^2 3500 + p\left(\frac{1}{C_1 R_1} + \frac{1}{C_2 R_2}\right)3500 + \frac{3500}{C_1 R_1 C_2 R_2}} \quad (131)$$

$$= \frac{p^3(p+b_1)}{(p+a_1)(p+a_2)(p+\alpha_1-j\omega_1)(p+\alpha_1+j\omega_1)} \quad (132)$$

where $b_1 = a_1 + a_2 + 2\alpha_1$.

As in the previous example, take the principal pole-pair position $-0.5 \pm j$.

For a 10% overshoot we have, from Table VII

$$Y = -1.01$$

Further since $b_1 > a_1$ or a_2 , the term due to the nearest real pole will be positive. A few trials are made.

(a) Suppose $a_1 = 10, a_2 = 20$

$$\text{Then } \Delta a_1 = 1, \Delta a_2 = 1.301, \Sigma \Delta p = 2.301$$

$$\underline{b_1 = 31, \Delta b_1 = 1.491, \Sigma \Delta z = 1.491}$$

$$\text{Hence } Y = -0.81 \text{ and } X = \log_{10} \text{ overshoot} = -0.81 + 0.01 = -0.8,$$

that is, about 16% overshoot.

(b) Suppose $a_1 = 15, a_2 = 30$

$$\text{Then } \Delta a_1 = 1.176, \Delta a_2 = 1.477, \Sigma \Delta p = 2.653$$

$$\underline{b_1 = 46, \Delta b_1 = 1.663, \Sigma \Delta z = 1.663}$$

and $Y = -0.99$. This is near enough the required value.

Completing the data, we have $L_{a_1} = 0.069, L_{a_2} = 0.0339, L_{b_1} = 0.022$. The principal mode minimum will therefore occur at $t^1 = 3\phi_1 + \Sigma L_p - (\pi + \Sigma L_z)$, i.e. at $3(1.106) + 0.103 - (3.142 + 0.022) = 0.257$. The required transfer function

$$\text{is } \frac{\epsilon(p)}{\theta_i} = \frac{p^3(p+46k)}{(p+15k)(p+30k)(p+0.5k-jk)(p+0.5k+jk)} \quad (134)$$

$$= \frac{p^4 + 46kp^3}{p^4 + 46kp^3 + 496.3k^2p^2 + 506k^3p + 562.5k^4} \quad (135)$$

Equating the coefficients of (131) to those of (135)

$$\frac{3500}{C_1R_1C_2R_2} = 562.5k^4 \quad (136)$$

$$\left(\frac{1}{C_1R_1} + \frac{1}{C_2R_2}\right) 3500 = 506k^3 \quad (137)$$

$$3500 = 496.3k^2 \quad (138)$$

$$\left(\frac{1}{c_1 R_1} + \frac{1}{c_2 R_2} + \frac{1}{c_1 R_2} \right) = 46k \quad (139)$$

Solving, we have $k = 2.655$, $k^2 = 7.06$, $k^3 = 18.75$, $k^4 = 49.8$.

The resistance R_2 would not normally be less than about 1000Ω to avoid undue loading and we shall suppose it is fixed at that value. The equations then give $C_1 = 8.35\mu F$, $C_1 R_1 C_2 R_2 = 0.1248$. If $C_2 = 10\mu F$, then $R_1 = 1.5 M\Omega$.

The error is given by

$$e(t) = 0.133\epsilon^{-1.328t} \sin(2.655t + 3.991) + 2.2\epsilon^{-39.85t} - 1.103\epsilon^{-79.7t} \quad (140)$$

The principal mode has the value -0.1047 at $t = 0.257/k = 0.0968$, at which time the term $2.2\epsilon^{-39.85t}$ is equal to 0.0464 . The maximum overshoot will still be of the order of 10%, but will occur slightly later than the above value of t indicates.

9.3. Comparison of the $\frac{\theta_o}{\theta_i}(\rho)$ and $\frac{\epsilon_o}{\epsilon_i}(\rho)$ Design Methods. Conclusions.

The following points may be made regarding the two design procedures.

1. The $\frac{\theta_o}{\theta_i}(\rho)$ method is easily applicable to Type 0 and Type 1 systems irrespective of their order; no restriction other than the required $\Sigma\Delta_z - \Sigma\Delta_p$ limits the pole and zero positions. For Type 2 systems, the additional relationship expressing the equality of the coefficients of powers of ρ in the numerator and denominator, causes some trouble in the case of high order systems (i.e. 6 or more). Type 3 designs are complicated by two additional relationships compared with Type 1 designs.
2. The $\frac{\epsilon_o}{\epsilon_i}(\rho)$ method may be equally easily applied to all Types. In all Types however, if the degree of the denominator of $\frac{\theta_o}{\theta_i}(\rho)$ exceeds that of the numerator of $\frac{\theta_o}{\theta_i}(\rho)$ by two, we have the further condition

$$\text{sum of zeros} = \text{sum of poles.}$$

If the degree of the denominator of $\theta_o/\theta_i(p)$ exceeds that of the numerator by three, two extra conditions arise, namely

$$\begin{array}{l} \text{sum of zeros} = \text{sum of poles} \\ \left. \begin{array}{l} \text{sum of products of zeros taken two} \\ \text{at a time} \end{array} \right\} = \left\{ \begin{array}{l} \text{sum of products of poles taken} \\ \text{two at a time.} \end{array} \right. \end{array}$$

If the degree of the denominator of $\theta_o/\theta_i(p)$ exceeds that of the numerator by more than three, three conditions occur.

It follows from points 1 and 2 therefore that a Type 3 system coming within this last category will cause trouble in design. The difficulty may only be resolved by judicious trial and error.

3. For systems requiring remotely situated poles, the $\theta_o/\theta_i(p)$ transfer function has its zeros near the principal pole-pair position. Under similar circumstances, the $\zeta_{\theta_i}(p)$ transfer function has its zeros also remotely situated. All the subsidiary poles and zeros of the $\zeta_{\theta_i}(p)$ transfer function being thus situated beyond the range of the charts, these may be dispensed with. This is a considerable advantage.

Conclusions.

The following remarks apply to both methods.

1. The principal mode approximation to the overshoot is possible except in the case of Type 2 and Type 3 systems of low order (i.e. 3 or 4) and having a low value of maximum overshoot. In such cases there are not sufficient subsidiary poles and zeros to enable all relationships to be simultaneously satisfied. By using the term due to the nearest real pole, however, the difficulty of the small overshoot may usually

be resolved. It is possible, nevertheless, that even with this technique a case may arise when a given overshoot is unattainable. This indicates that the response requires to be composed of wholly real terms. Under these conditions it may be possible to use one of Whiteley's Standard Forms.

2. The principal pole-pair position to be selected will depend on the required overshoot and damping. Except for Type 3 systems, in which the least damped position at $-0.5 \pm j1$ has to be accepted, there is usually a choice of two out of the three principal pole-pair positions. It is difficult to make a statement on this point because of the dependence of $\Sigma \Delta_z - \Sigma \Delta_p$ on overshoot, damping, and the relative numbers of subsidiary poles and zeros in any particular transfer function. In general, however, the design will be easier, (i) if $\Sigma \Delta_p - \Sigma \Delta_z$ does not depart too widely from zero for equal numbers of subsidiary poles and zeros (distinct from the origin), and (ii) if $\Sigma \Delta_z - \Sigma \Delta_p$ is a small positive or negative quantity according as the number of zeros is greater than or is less than the number of subsidiary poles.

Finally, it may be stated that the writer's preference lies with the $e/\theta_i(\rho)$ design method, firstly, on account of the segregation of Types and secondly, since the charts are in most cases unnecessary.

PART III.

EXPERIMENTAL INVESTIGATION OF TRANSIENT AND FREQUENCY RESPONSES
OF METADYNE SERVO-MECHANISM.

INTRODUCTION.

The servo-mechanism whose transient and frequency responses are described in the following four Chapters is the Admiralty Metadyne Power Control Apparatus for a 2-Pdr. R.P.50 "M" Mark 7 Mounting. This equipment was obtained as a result of a generous offer by the Admiralty and consisted of the complete control gear with the exception of the transmitting magslips and those items of switchgear situated on the mounting proper. The equipment also comprised several components and numerous circuits all incidental to the basic servo-mechanism sequence. These are the h.f. alternator set forming the power supply for the thermionic amplifier, control panel and automatic voltage regulator for the h.f. output, automatic starters for the h.f. set, and for the metadyne motor-generator set. Control potentiometers for joystick manipulation of the mounting were also included, and a d.c. contactor panel for effecting change-over from this operation to fully automatic (servo) operation, i.e. remote control from the transmitting magslips. As the author had not at the time been dealing with such remote position control apparatus, a decision was made to set up the equipment as near as possible to the original scheme for which the gear was designed. At the same time, it was attempted to separate the training and elevating motions as far as possible and to make available certain important points in the circuit for laboratory testing and so on. This required the construction of an auxiliary switchgear panel and a few modifications to the d.c. contactor panel. The training moment-of-inertia of the mounting was simulated by a

correctly-proportioned inertia which was direct-coupled to the gun-driving motor armature. A gear-box linking the fine and coarse magflip transmitters was also constructed. The description of the last two items is given in the Chapter following. As the operation of the h.f. alternator-set, starting arrangements, d.c. contactor panel and auxiliary switchgear panel, is not relevant to the main servo-mechanism sequence, the description of these components and circuits is omitted. The circuit diagram of Fig. 37 therefore represents only the essential connections, the actual circuit diagram being considerably more complex. The preliminary construction and setting-up, testing and measurement was carried out during the Summers of 1948, -49, and -50. The final tests which are given in Chapters 11 and 12 were obtained during the Summer of 1951.

Nature of Investigation.

The object of the investigation is to determine to what extent quantitative prediction of the transient response of a practical "linear" servo-mechanism is possible from a knowledge of the frequency response. By a practical "linear" system is understood a system whose operation is made as linear as possible. The results will decide, it is hoped, just how far the highly-developed frequency-response design techniques can be of use when applied to a practical servo-mechanism. At the outset, it is certain that this limit will depend on the complexity of the system and upon the extent of non-linear operation which takes place under normal amplitude signals. The system tested is moderately complex and normally non-linear in its operation. This will be the case with most high-power, high-accuracy systems.

Result of Investigation.

For the system tested the investigation showed that the predicted overshoot exceeded the measured overshoot by 16% of the input step, with a smaller discrepancy for a more oscillatory response. The measured time of the maximum overshoot exceeded the predicted time by 40%, with again better agreement for a more oscillatory response. It has been attempted to extend these conclusions in Sec. 13.4, subject to the provisions stated therein.

CHAPTER 10.

DESCRIPTION OF APPARATUS AND METHODS OF TESTING.

For convenience in following out the description of the apparatus and methods of testing, the order in which this is presented is given. This is as follows.

10.1. Description of Apparatus.

Fig. 37. Circuit diagram of servo-mechanism and apparatus associated with sinusoidal measurements.

Fig. 38. Thermionic amplifier circuit (simplified).
Note on operation.

Fig. 39. Photograph of general layout of apparatus.

Fig. 40. Photograph of details of resetting.

Fig. 41. Photograph of details of transmitting magslips and gearbox, transient response paper recorder, and sinusoidal input-motion generator.

Layout and mechanical construction of main components.

10.2. Description of Methods of Testing.

Transient response measurement

Fig. 42. Circuit diagram of frequency response measuring gear.
Frequency response measurement.

Factors limiting the linear operation of the system.

Throughout Figs. 37 to 42, a consistent numbering system for the components has been employed. The key to this is given on Fig. 37 and is repeated for convenience opposite Fig. 39.

Circuit Diagram of Servo-Mechanism and Apparatus associated with Sinusoidal Measurements.

1. Dial indicating angle of input (i.e. fine transmitting magslip) shaft.
2. Dial indicating angle of output (i.e. fine resetting magslip) shaft.
3. Thermionic amplifier.
4. Metadyne generator.
5. Motor.
6. Load inertia.
7. Fine transmitting magslip.
8. Coarse transmitting magslip.
9. Transmitting resolver for phase measurement.
10. Drive motor for sine input motion.
11. Sine wave drive to input shaft.
12. Resolver giving output shaft angle for phase measurement.
13. Paper recorder.
14. Pen recording output motion.
15. Resetting magslip.
16. Resetting gear-box.
17. H.F. alternator set.
18. Transmitting gear-box.
19. Free running pulley on coarse magslip shaft.
20. Motor shaft.
21. Metadyne generator set driving motor.
22. Receiving resolver for phase measurement.
23. Variac supplying 9.
24. Variac supplying 12.
25. Phase correcting circuit for 50-c/s excitation to 12.

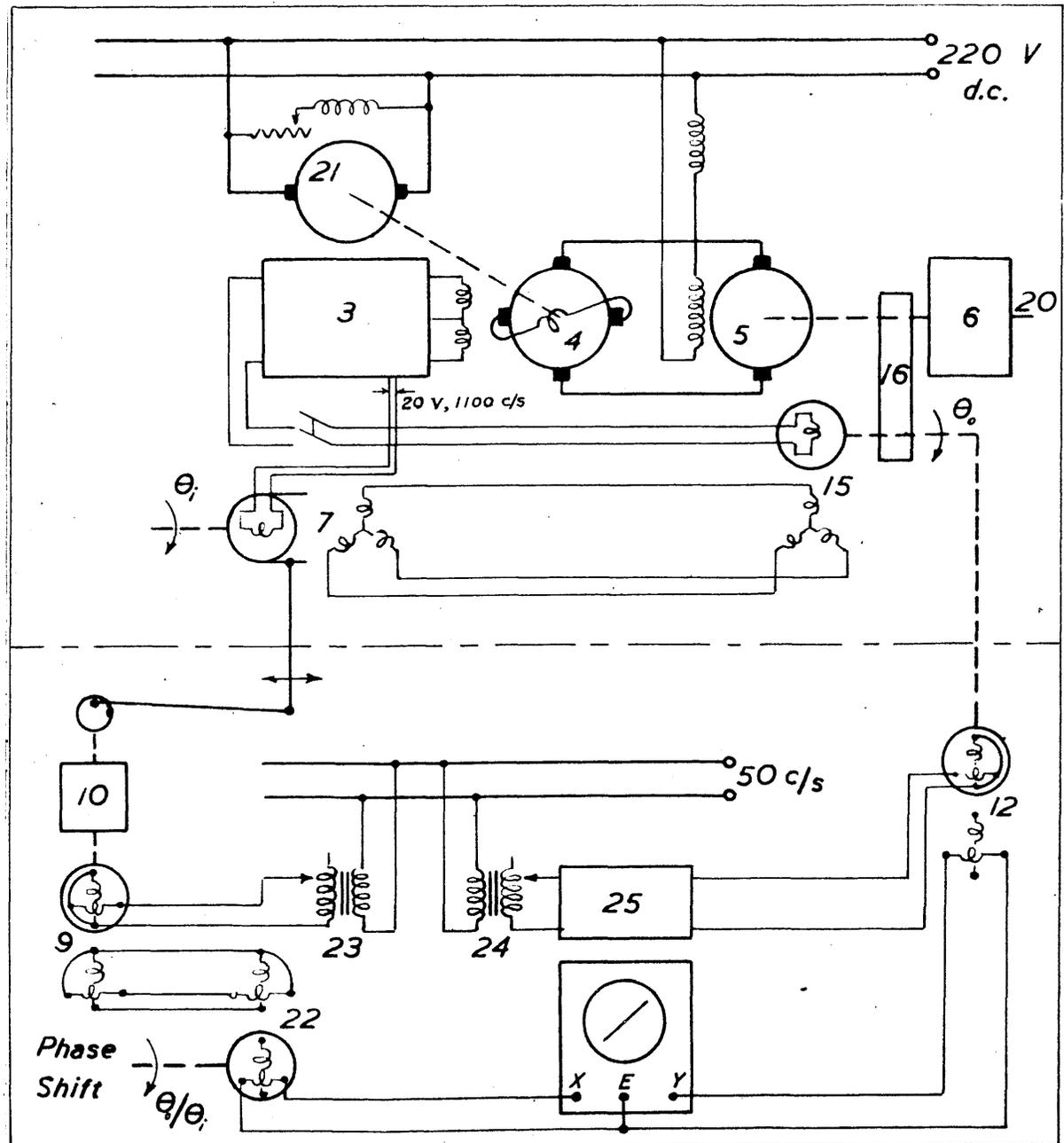


Fig. 37

10.1. Description of Apparatus.

Description of servo-mechanism.

The circuits above the chain-dotted line in Fig. 37 comprise the servo-mechanism proper, as distinct from the frequency response measurement circuits which are below the chain-dotted line. The load inertia 6 is a simple brake pulley which has been proportioned to simulate the training moment-of-inertia of the mounting referred to the motor shaft 20. Details of the actual training motion were as follows.*

Training moment-of-inertia	8,800 slug-ft ²
Gun-driving motor armature inertia referred to mounting	7,150 "
Gear ratio motor shaft to mounting	293 1/3 to 1

$$\text{Total m.i. referred to motor shaft} = \frac{15950 \times 32.2}{(293 \frac{1}{3})^2} \text{ lb-ft}^2 = 5.96 \text{ lb-ft}^2$$

The gun-driving motor armature inertia was measured by the method of bifilar suspension and found to be 2.07 lb.-ft². An added inertia was therefore constructed to bring the total m.i. referred to the motor shaft up to 6 lb-ft². Error indication is obtained by the magflip transmitter-resetter system of which only the fine transmission is shown in the diagram. In practice this is geared up 36 times from the speed of the mounting with a coarse magflip link to take over for larger misalignments. Provision has been made for this in the transmitting gear-box described later. As all the tests described here, however, required only fine magflip signals, the coarse transmission has been omitted from the circuit diagram. The gear-

* Figures supplied by Secretary of Admiralty for Naval Ordnance Department.

train from motor shaft to fine resetter, indicated by item 16 on Fig. 37, was fixed at 25 to 3. The main details of the experimental set-up are therefore:

Total moment-of-inertia referred to motor shaft	6.06 lb-ft ²
Gear ratio motor shaft to fine resetting magslip	25 to 3
Gear ratio fine to coarse transmission	36 to 1
(Effective gear ratio motor shaft to mounting = $\frac{25}{3} \times 36 = 300$ to 1)	

A high-frequency alternator provides a 200-V, 1100-c/s power input to the thermionic amplifier 3, which gives by a separate winding on the high-tension transformer, a 20-V, 1100-c/s supply for the transmitting magslip rotors. The transmitting magslips are Admiralty Pattern 10428, supplied by Muirhead & Co., while the resetting magslips received with the equipment are Admiralty Pattern 10429.

The angular error between the shaft of the fine transmitting magslip 7, and that of the fine resetting magslip 15, is measured in magnitude by the r.m.s. value of the voltage induced in the resetting magslip rotor winding, the output being 0.6 /degree error. A phase change of 180° in this voltage indicates the error has altered from positive to negative or vice versa.*

* No output voltage is obtained from the resetting magslip rotor when its magnetic axis is at right angles to the pulsating field produced by the stator windings. The direction of this field depends in turn upon the positions of the transmitting magslip rotor and stator. In the laboratory set-up described herein, it is convenient, for this condition, to fix pointers on both the transmitting and resetting magslip shafts in the vertically upward direction and to set the indicating dials 1 and 2 to read zero. If the transmitting magslip shaft turns clockwise through 20° say, the load inertia is driven clockwise by $20 \times 25/30$ and turns the resetter shaft clockwise through 20°. Similarly for anti-clockwise displacements of the transmitter.

This point is explained in connection with the description of the first stage of the amplifier. The thermionic amplifier output supplies the main variator windings of the metadyne-generator 4. There are arranged in differential or push-pull form as shown in Fig. 37. A d.c. shunt motor 21, forms the drive for the metadyne generator which supplies the armature power of a separately excited d.c. gun-driving motor 5. This motor is fitted with a brake at the end of the shaft remote from the load, with provision for lifting the brake shoes magnetically by flow of the normal shunt-field current through a special brake winding; or mechanically for inspection purposes by means of a small lever on the brake structure itself. Throughout the tests the brake was permanently lifted.

The ratings of these machines are as follows:-

Metadyne generator set driving motor	- 220 V, 5.6 h.p., 2800 r/m
Metadyne generator	35 - 70 V, 1.2 kW, 2800 r/m
Gun-driving motor	28 - 220V, 40A, 660 r/m

The figures for the first two machines refer to the continuous power output which they are capable of supplying. The current rating of the gun-driving motor is also continuous. Having more direct bearing on the operation of the system, the peak output power of the metadyne generator is about 2.5 kW, which is accompanied by a peak input power to the driving motor of about 5 kW. The supplies are

metadyne-generator set driving-motor	- 220 V, d.c.
high-frequency alternator set driving-motor	- 100 V, d.c.

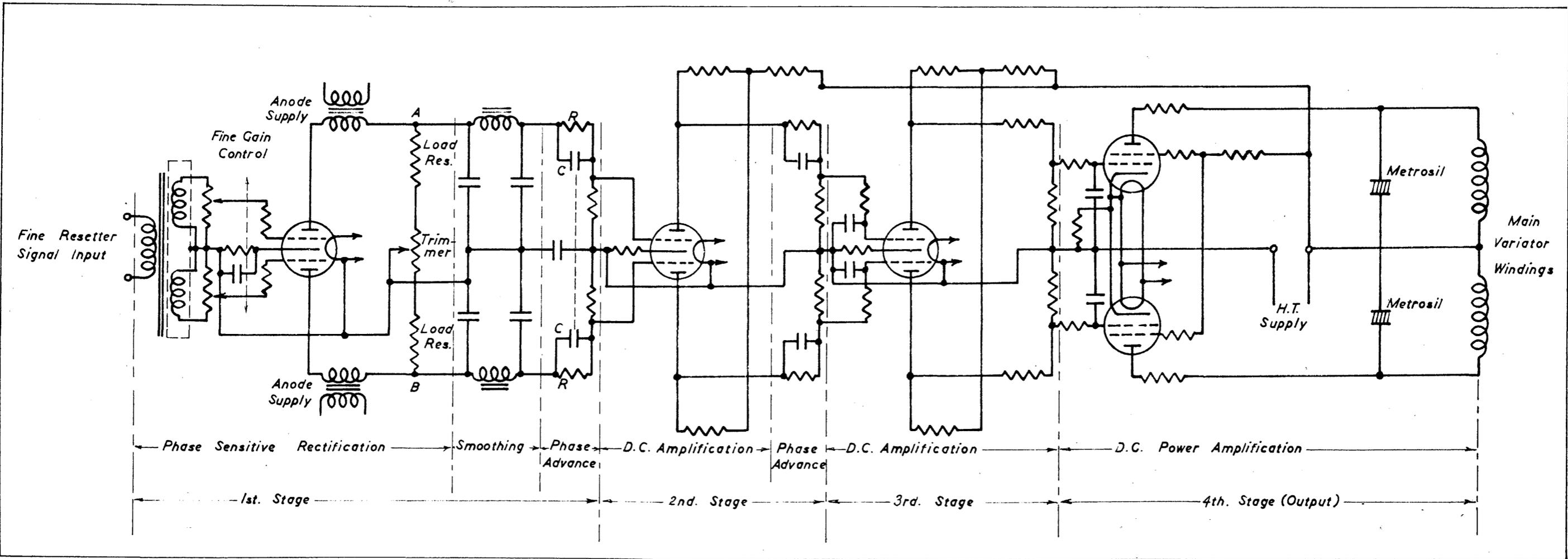


Fig. 38 Simplified Thermionic Amplifier Circuit.

The first is obtained from another d.c. machine and the second from a 3-phase, mercury-arc rectifier.

Thermionic amplifier - note on operation.

The circuit diagram of Fig. 38 has been abstracted from a complete diagram which was supplied with the equipment. It is included for the mere explanation of the functions of the various stages. Technical details of the design have not been altered in any way from the original and they are omitted in this description. The amplifier has four stages, namely,

1. Phase-sensitive rectification, smoothing, and phase-advancing.
2. Voltage amplification and phase-advancing.
3. Voltage amplification.
4. Output stage producing current in variator windings proportional to stage input voltage.

Phase sensitive rectification.

In the absence of an a.c. signal input to the fine grid transformer, the anodes of the first valve carry equal currents during the positive half-cycle of their alternating supply voltages, which are in phase with each other. No net voltage drop appears over the two load resistors, i.e. across the points A and B. Due, however, to the centre tapped secondary of the grid transformer, a signal input will supply the grids in anti-phase. Since the magstrip excitation is taken from a further secondary winding on the h.t. transformer supplying also the first-stage anodes, one grid voltage will be in phase with its anode voltage, while the other grid voltage will be in anti-phase to its anode voltage. Unequal currents are therefore taken

in each half of the valve and a net output voltage appears across AB. If the angular error between input and output changes its sign, the signal input to the grid transformer sustains a 180° phase-change. The anode previously taking the smaller current will now take the larger one and a change in the polarity of the net output voltage across AB results. This output voltage is smoothed and phase-advanced before being passed to the grids of the second valve. The phase-advancing circuit operates on changing d.c. signals, and has a number of condensers C of differing values, any one of which may be selected by a four-position switch.

Balanced D.C. Amplification.

The remaining stages constitute a straightforward balanced d.c. amplifier, with a fixed amount of phase-advance provided in the second stage. "Metrosil" non-linear discharge resistors are fitted across both halves of the variator windings in order to limit peak transient voltages which may arise across them. The standing current in each half of the variator winding is nominally 40 mA, with a maximum swing of ± 40 mA due to the input signal. In practice it is necessary to balance the whole forward sequence by adjustment of the trimming resistor in the first stage. This is altered until no metadyne output current flows when the system is set in the position of zero error. For the purposes of linear operation, the whole ± 40 mA swing cannot be used. The limit of linearity in the gain of the forward sequence is reached, in fact, at about ± 25 mA swing.

General Layout of Apparatus.

1. Dial indicating angle of input (i.e. fine transmitting magslip) shaft.
2. Dial indicating angle of output (i.e. fine resetting magslip) shaft.
3. Thermionic amplifier.
4. Metadyne generator.
5. Motor.
6. Load inertia.
7. Fine transmitting magslip.
8. Coarse transmitting magslip.
9. Transmitting resolver for phase measurement.
10. Drive motor for sine input motion.
11. Sine wave drive to input shaft.
12. Resolver giving output shaft angle for phase measurement.
13. Paper recorder.
14. Pen recording output motion.
15. Resetting magslip.
16. Resetting gear-box.
17. H.F. alternator set.
18. Transmitting gear-box.
19. Free running pulley on coarse magslip shaft.
20. Motor shaft.
21. Metadyne generator set driving motor.
22. Receiving resolver for phase measurement.
23. Variac supplying 9.
24. Variac supplying 12.
25. Phase correcting circuit for 50-c/s excitation to 12.

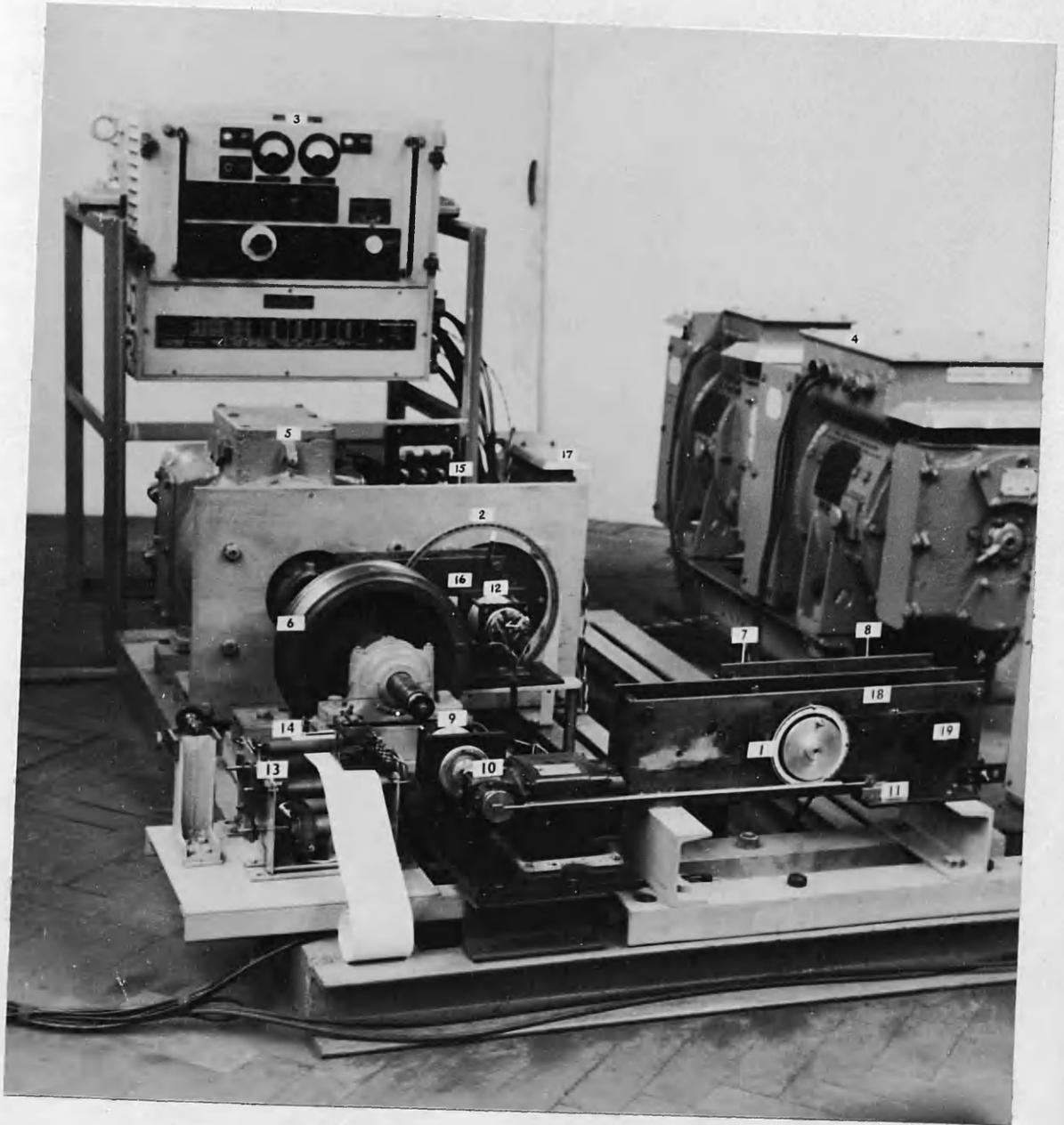


Fig. 39. General layout of apparatus.

Layout.

Referring to the general view given in Fig. 39, the thermionic amplifier 3 is supported in an angle-iron framework in the rear of the photograph with the high-frequency alternator set 17 underneath. The metadyne motor-generator set 4 on the right, comprises twin-metadyne generators and a driving motor assembled as one unit. The starters for both these sets are not shown in the photograph. The other apparatus, excluding instruments, is mounted on a $5\frac{1}{2}'$ x 3' bedplate made up from channel and Tee-irons. Directly in front of the thermionic amplifier is the gun-driving motor 5, with the load inertia 6. To the right of the servo-motor and wholly concealed in this view are the resetting magslips in their gear-box 15. In the front of the photograph, the paper recorder 13 is to the left; the drive-motor 10, for the slider-crank sinusoidal input-motion-generator 11 is in the centre while the transmitting magslips 7 and 8, with their gear-box 18, are on the right.

Mechanical construction of main components.

The motor 5 is bolted down on cross-channels. The normal bolts which held the front end-shield are replaced by four double-ended bolts which perform the same task and support also a rectangular steel plate. This plate is also bolted in three places to the composite forward motor cross-channel and has two holes bored in it, one through which the motor shaft projects and the other into which the resetter gear box is spigoted.

The plate serves as a reference plane and establishes a rigid connection between the motor shaft and the shaft of the fine mag slip resetter, between which the 25 to 3 gear-train is placed. This gear-box, 16, is built separately and later fixed in position on the reference plate. The load inertia has been made in the form of a brake pulley* whose m.i. has been finally adjusted by screwing on a brass ring. The load is keyed on to a $1\frac{1}{2}$ " diameter shaft supported in Hoffmann self-aligning ball bearings to reduce friction as far as possible. This shaft is solidly-coupled to the motor shaft.

Gear-boxes (see Figs. 40 and 41).

A similar construction has been employed for the 25 to 3 gear-train 16, as for the 36 to 1 gear-train 18, which is between the fine and coarse transmitting mag slips 7 and 8. Both are made of bright mild steel plates clamped parallel with each other and separated by accurately made tubular distance pieces. The gear-box 16 is made up of two plates $\frac{1}{4}$ " thick while the gear-box 18, which supports also the transmitting mag slips, is composed of three plates $5/16$ " thick. The gear-wheel shafts are of $\frac{1}{4}$ " diameter silver steel throughout and run in brass bushes which are pressed into the steel plates. Between every pair of fixed bushes, movable bushes are provided which allow $1/10$ " movement of the bush in any direction at right angles to the bore. The centres are initially marked out as accurately as possible. In this manner, the backlash can be reduced to a very small

* A suitable casting happened to be available. Steady loading of the motor may possibly be required in future.

amount and it will be possible to take up a limited amount of wear. The gear-trains themselves are composed of Bond's Standard 40 D.P. Spur Gear Wheels.

Two circular scales, 1 and 2, graduated in intervals of one degree are used to measure the angles of the transmitting and resetting magslips shafts. As mentioned previously, the pointers mounted on the shafts read zero when they point vertically upward. The transmitter and resetter dial indications will hereafter be referred to as the input and output angles of the system, with clockwise rotations designated positive.

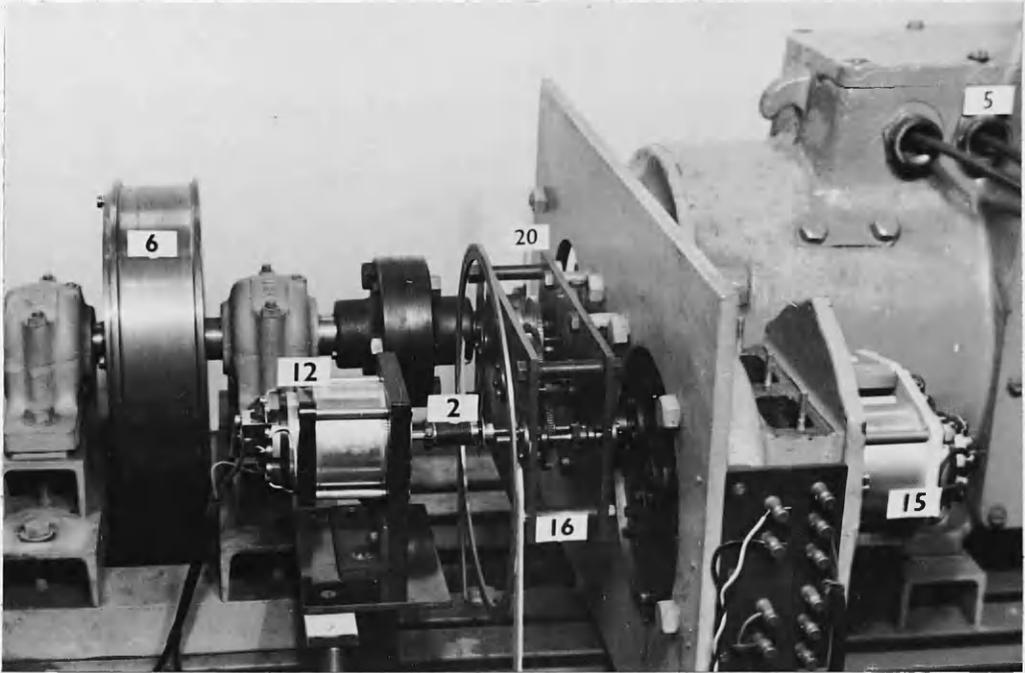


Fig. 40. Details of resetting.

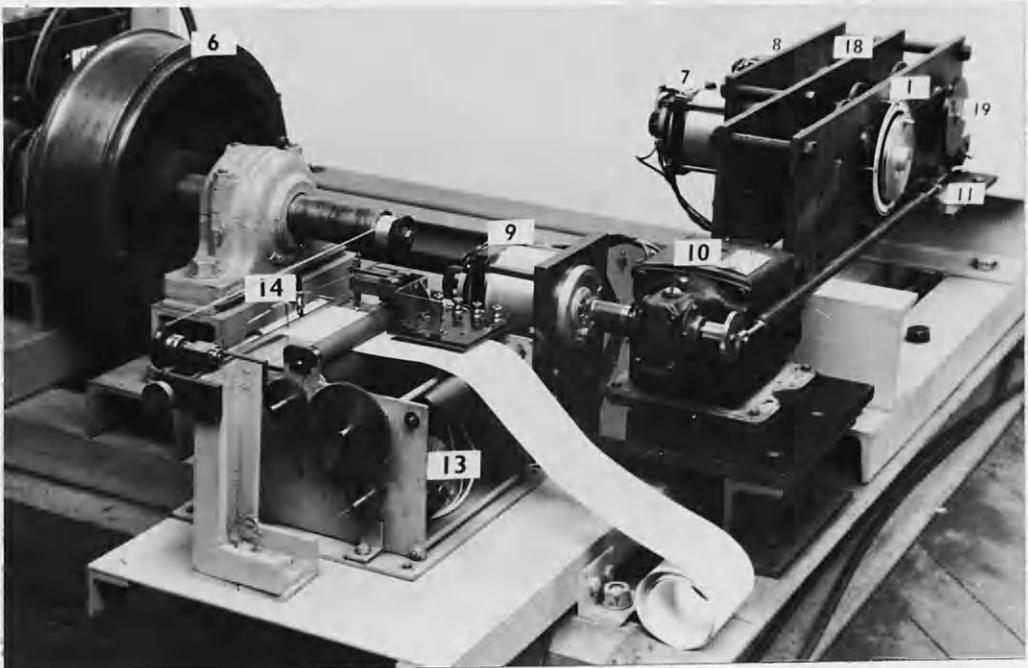


Fig. 41. Details of paper recorder and transmitting gearbox.

10.2. Description of Methods of Testing.

Transient Response Measurement.

A discussion of the factors which limited the movement of the output in order that its response might be considered linear is given in Sec.10.3. Here it is sufficient to note that this limited range of motion allowed the possibility of recording on paper by a pen directly moved by the output shaft or any shaft geared to it. The advantages of simplicity, accuracy, ease of inspection and measurement of traces compared with oscillographic methods seemed to outweigh the possible trouble in constructing a reliable recorder with a suitable time-marking device, and hence this method was adopted.

The simple recorder 13, shown in Fig. 41, consists of a series of rollers running between two upright plates. The rollers which pull the paper through are in the centre of the recorder and are rubber-covered. The lower of these is driven through a gear-train from a small 24-V d.c. motor fitted with a worm-gear, and mounted within the recorder on its base-plate. The upper roller runs in floating bushes which are pressed downwards by springs. Paper which is fed in between the rollers is thus pulled through by a friction grip. Before entering the rollers, the paper passes over a brass plate forming a suitable writing surface and is kept flat prior to being written on by a brass strip mounted transversely to the paper and close to the writing surface. This strip contains a narrow groove extending the full width of the recorder and forms a guiding slot for a short length of 8 B.A. screwed rod which is rigidly fixed to the pen-holder. The screwed rod also provides a means of regulating the

writing pressure. BIRO pens are used for both the output and time-marking traces. The output pen is moved by a flexible steel tape passing round a Tufnol pulley on the motor shaft and round a similar one on the far side of the recorder from the motor shaft. A limited amount of travel in the horizontal direction is provided at the driven pulley in order to tension the steel tape. The friction at the sharp edges of the tape is ample to overcome the slight resistance of the pen motion and no trouble has been experienced with slip. The same may be said of the oscillatory drive to the fine transmitting magslip described under frequency response measurement, where the torque conditions resisting the drive are much more arduous. It may be of interest to state that the steel tape used was magnetic recording tape. The scale of the recording was fixed at $1'' = 10^\circ$ fine magslip rotation.

The time-marking pen is fixed to the armature of a small relay. Flow of current through the coil causes the pen to be lifted off the paper on which it is normally held by a light spring. The circuit of the relay coil is closed twice per second by a contact driven by a geared synchronous motor supplied at mains frequency. The time-marking trace is therefore a series of short lines or dashes, the ends of which indicate half-second intervals. It is necessary to assume that the paper speed does not alter between these points. Reference to the actual records shows that this is quite justifiable.

Frequency Response Measurement.

The lower part of Fig. 37 shows the additional circuits and apparatus required for frequency response measurement. The diagram is repeated be-

low for convenience.

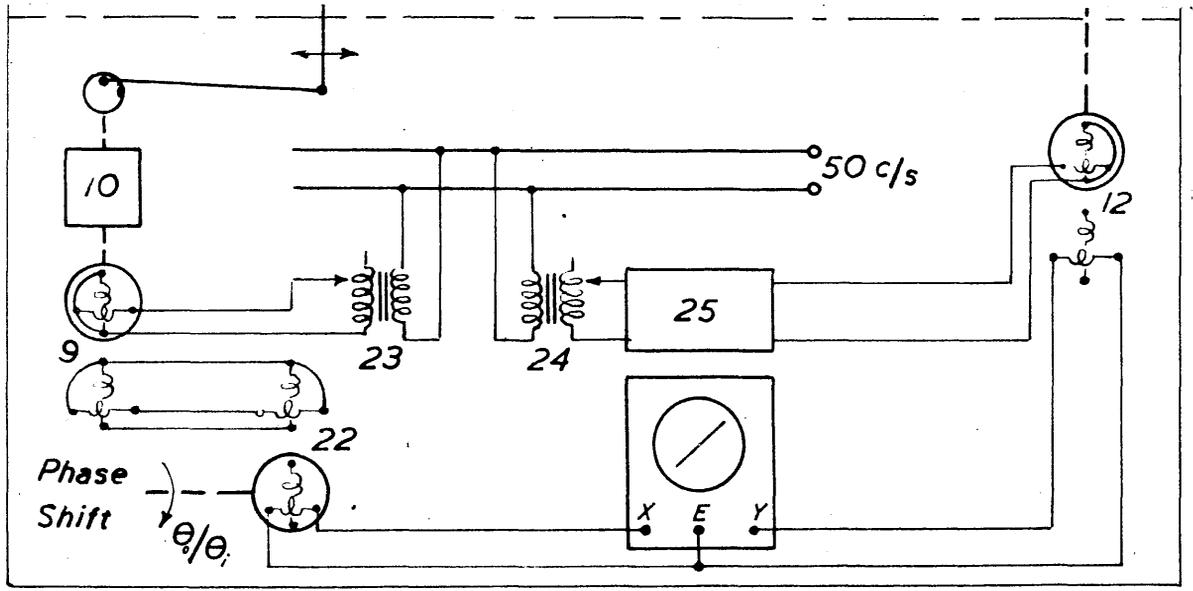


Fig. 42.

The shaft of the fine transmitting mag-slip 7 is moved sinusoidally by the slider-crank and steel-tape drive shown in the photograph of Fig. 39 and in the detailed view of Fig. 41. For the range of frequencies required this method proved quite satisfactory. The ratio of connecting-rod length to crank length is about 35 for input motions of $\pm 10^\circ$ amplitude and about 70 for input motions of $\pm 5^\circ$ amplitude so that negligible error is introduced in assuming the movement of the input shaft to be a true sinusoid. The crank-pin screws into the face of the half-coupling on the drive-motor shaft and a set of tapped holes for input motions of $\pm 5^\circ$, $\pm 10^\circ$, $\pm 20^\circ$ is provided. The part of the crank pin which rotates in the connecting rod is turned eccentrically to the part which screws into the half-coupling, and by this means, the crank length may be adjusted to give

the exact movement required to the input shaft. The amplitude of oscillation of the output shaft, i.e. fine resetter shaft, is observed on the dial shown. Frequency is obtained by timing a given number of oscillations with a stop watch.

The angle of lag of the output with respect to the input shaft is measured by means of a cathode-ray tube and using a null method of detection. Three resolver-type magslips^{*} were used. The resolver 12, directly driven by the output shaft, has its rotor excited by a 50-c/s supply, and provides at its stator winding terminals an alternating voltage proportional to the output shaft angle[†], and one whose phase with respect to the 50-c/s supply, changes by 180° as the output angle goes from positive to negative. The Variac 24 provides a means of altering the rotor excitation in order to maintain approximately constant amplitude trace on the CRO with varying amplitudes of the output oscillation. A CR phase-shifting network 25, is also required for reasons given later.

* No particular significance attaches to the use of resolver-type magslips. The resolver 12, coupled to the output shaft was originally employed because of the large signal voltage it was possible to obtain, but later a fixed setting of the oscilloscope amplifier was used. The resolvers 9 and 22 were used simply on account of their availability.

† More correctly, proportional to the sine of the output shaft angle. The maximum amplitude occurring in the tests was just under 16° so that the greatest error introduced is about 1.1%.

The resolvers 9 and 22 provide at the rotor terminals of the latter, a 50-c/s alternating voltage whose amplitude varies sinusoidally at the frequency of rotation of the resolver rotor 9. This rotor is directly coupled to the output shaft of the drive motor 10 responsible for the sinusoidal input shaft movement. The stators of 9 and 22 are connected in a manner similar to ordinary mag-slip stators and function likewise. The rotor of 22 can be turned freely by hand to any position indicated by a suitable dial and pointer system, and according to its angular displacement, so the phase of the amplitude variation of its output voltage will be altered. This output voltage, in short, is an a.c. signal having sinusoidal amplitude variation at the frequency of the input sine motion and whose phase with respect to the input shaft motion can be adjusted by the angular displacement of the rotor. The Variac 23 enables the actual voltage obtained to be set at a convenient value for observation on the C.R.O. The pattern which results is explained below.

Before making any phase measurements, it is necessary to line-up the voltages on the vertical and horizontal deflecting plates. Thus when θ_i, θ_o and the plane-shift dial indication are all made zero, the stators of 12 and 22 may be rotated until a stationary spot is obtained on the C.R.O. screen. If now θ_i and θ_o are given equal or unequal displacements in the same direction a trace will appear on the oscilloscope, and provided the 50-c/s voltages on the X and Y plates are in phase, this will be a straight line. The tilt of the line, of course, depends on the excitations

supplied by the Variacs 23 and 24. Phase shift between these 50-c/s signals occurred. This may be partly assigned to the difference in the circuits of the two channels and to the employment of an amplifier in the Y direction. It was removed by the CR phase-correcting circuit 25. As a fixed setting of the oscilloscope amplifier was used, the network 25 need only be adjusted at the start and a straight-line trace is obtained thereafter. All the above details are performed with θ_i and θ_o stationary.

The movement of the trace when θ_i and θ_o are both varying sinusoidally at the same frequency can best be explained with reference to Fig. 43.

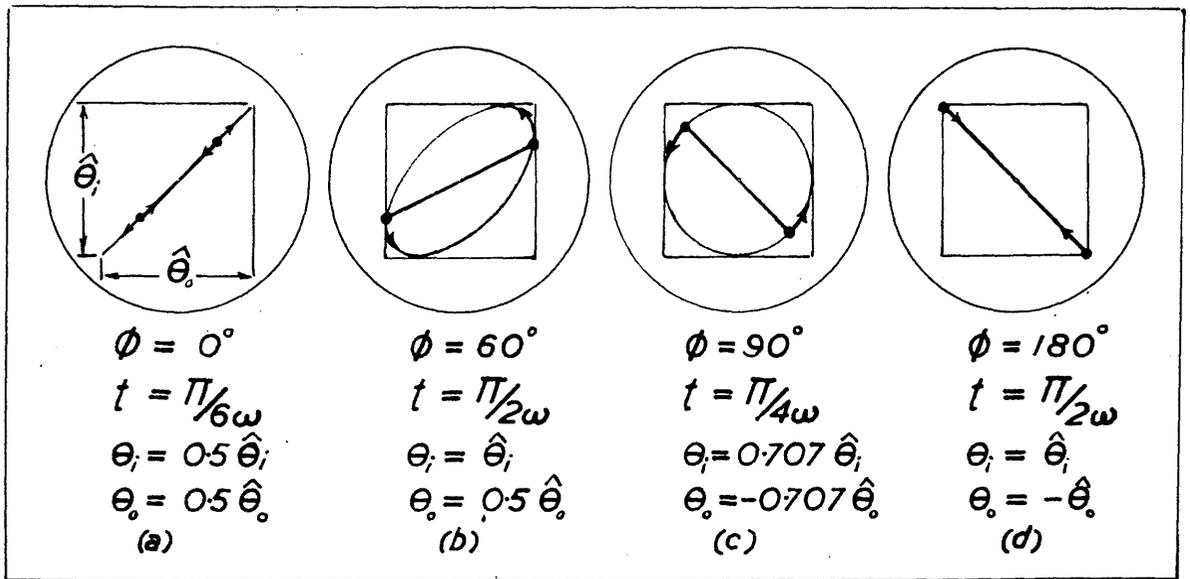


Fig. 43. C.R.O. traces obtained with various phase shifts of $\theta_o/\theta_i(j\omega)$.

The quantities indicated immediately below each diagram are (i) the phase-shift of $\theta_o/\theta_i(j\omega)$, (ii) the instant of time concerned, the input being given by $\theta_i = \hat{\theta}_i \sin \omega t$, (iii) the instantaneous value of θ_i , and (iv) the instantaneous value of θ_o . The pattern obtained for zero phase-

shift of $\theta_o/\theta_i(j\omega)$ is indicated in Fig. 43a. As the sinusoidal amplitude variations are in phase, the straight line trace merely alters its length without altering its direction. The frequency of this alteration in length is twice that of the input signal frequency. In Fig. 43b the conditions for 60° phase-lag of θ_o with respect to θ_i' are given. The extremities of the line follow the elliptical course shown and the line itself undergoes periodic variation of its length while it rotates anti-clockwise. A 90° phase-shift appears as Fig. 43c, in which the extremities follow a circle. The 180° phase-shift pattern takes the form of Fig. 43d. The procedure for phase-measurement is therefore to rotate the rotor of the resolver 22 (thus altering the phase of the amplitude variation with respect to the input) until the zero phase-shift pattern, Fig. 43a is obtained. The rotation indicated on the dial provided is the required angle. For small amplitude variations of θ_o , the vertical deflection was increased by altering the Variac 24. Hence it was possible to obtain the in-phase trace as a 45° line for all experimental points and obtain the greatest precision. In practice the range of dial adjustment within which no detectable change appeared in the trace was not greater than 2° at the lowest frequency point taken nor greater than 4° at the highest frequency point. The mid-point of this zone was taken as the phase-angle.

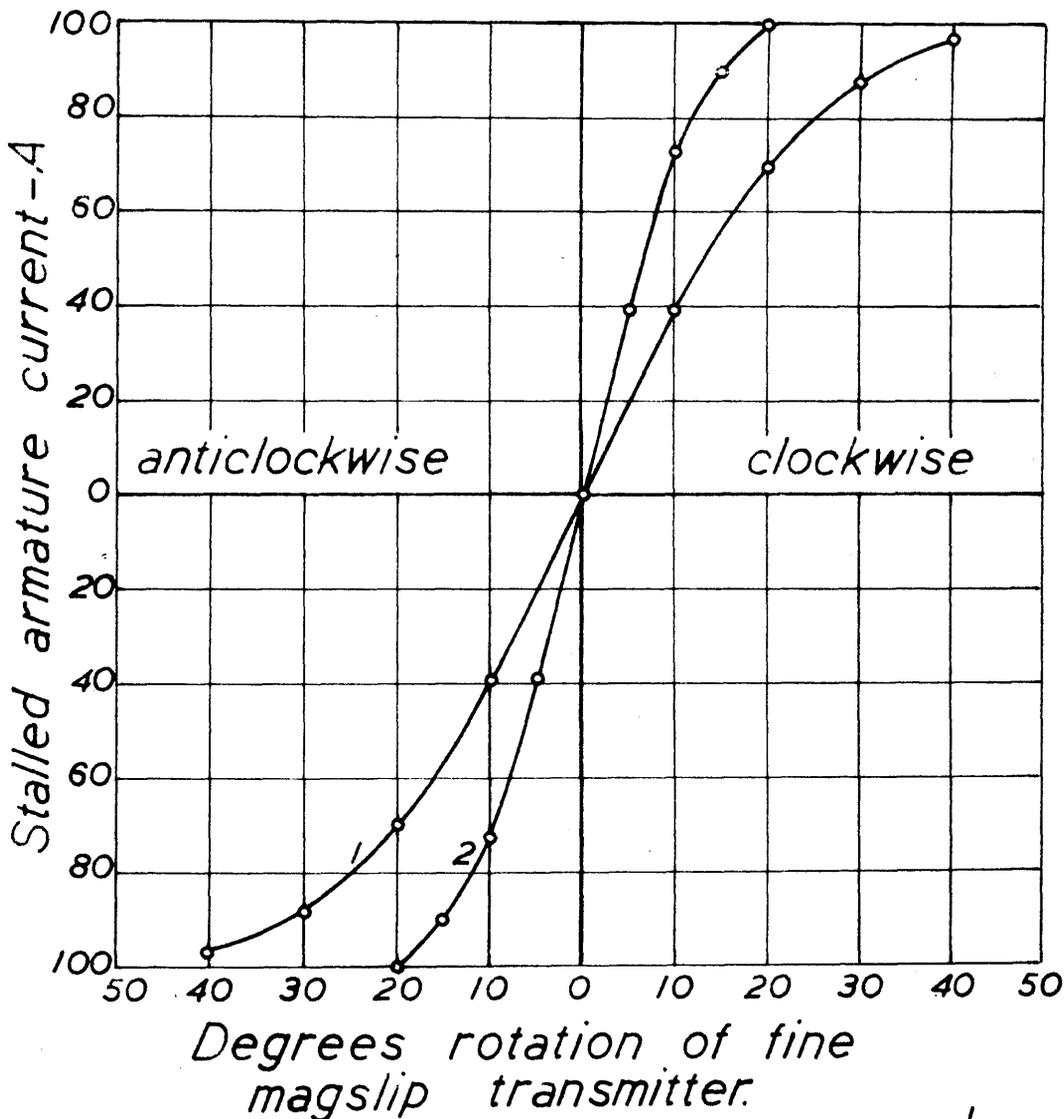
Extent of Linear Operation of the System.

The two main factors which restrict the range of linear movements of the system are (a) static error and (b) curvature of the control "stiffness" characteristic, that is, the relationship of motor torque to error.

The maximum static error between the transmitting and resetting mag-slip shafts is made up from (i) the inherent static inaccuracy of the mag-slip link itself, which usually is of the order of $\frac{1}{4}^{\circ}$ to $\frac{1}{2}^{\circ}$ and (ii) the maximum deviation which can occur without the developed torque of the motor exceeding the stiction torque. The second of these error components is inversely proportional to the control stiffness. For the two values of control stiffness which are used in the tests, the total static errors at the fine mag-slip shaft were approximately 1.5° and 0.8° for the lower and higher gain settings respectively.

Curvature of the overall stiffness characteristic is illustrated by Fig. 44, which gives the variation of stalled motor armature current with degrees rotation of the fine mag-slip transmitter. The control stiffness in lb-ft/degree error for each of the two gain settings is obtained by multiplying the slope of the curves by 0.294, which is the motor developed torque per armature ampere with normal field excitation. This torque constant was determined by previous tests. The overall sensitivity curves of Fig. 44 are given for two gain settings of the thermionic amplifier, with the larger gain very nearly equal to twice the smaller. Linearity of the characteristics may be reasonably assumed up to 10° and 20° error, for the higher and lower gain settings respectively.

The above factors i.e. static error and curvature of the control characteristic severely limit the range of movement within which the operation may be considered linear. The following are signal amplitudes used throughout the tests.



1. gain setting K_1 -stiffness $1.15 \text{ lb-ft}/^\circ$
 2 gain setting $K_2 = 2K_1$ -stiffness $2.3 \text{ lb-ft}/^\circ$

Fig.44 Overall Sensitivity.

Response Number.	(i) Gain Setting (ii) Control Stiffness, lb-ft/° error (iii) CR Value of 1st stage phase-advance.	Magnitude of Input Step in Transient Response.	Amplitude of Sinusoidal Input- Variation.
I	K_1 1.15 lb-ft/° 0.11 sec.	20°	± 10°
II	$2K_1$ 2.3 lb-ft/° 0.11 sec.	9°	± 5°
III	$2K_1$ 2.3 lb-ft/° 0.077 sec.	9°	± 5°

Three transient and three frequency responses were taken. Response II differs from I only in that twice the control stiffness of I was used. Response III differs from II only in that a less stabilising CR value in the 1st stage phase-advancing circuit was used. The responses have not been chosen widely different as the object was to examine if reasonably small changes in the frequency response of a practical system could give accurate prediction of the transient response. Further, responses which show a number of oscillations have not been measured as these are not acceptable in practice.

CHAPTER 11.

EXPERIMENTAL FREQUENCY AND STEP RESPONSES.

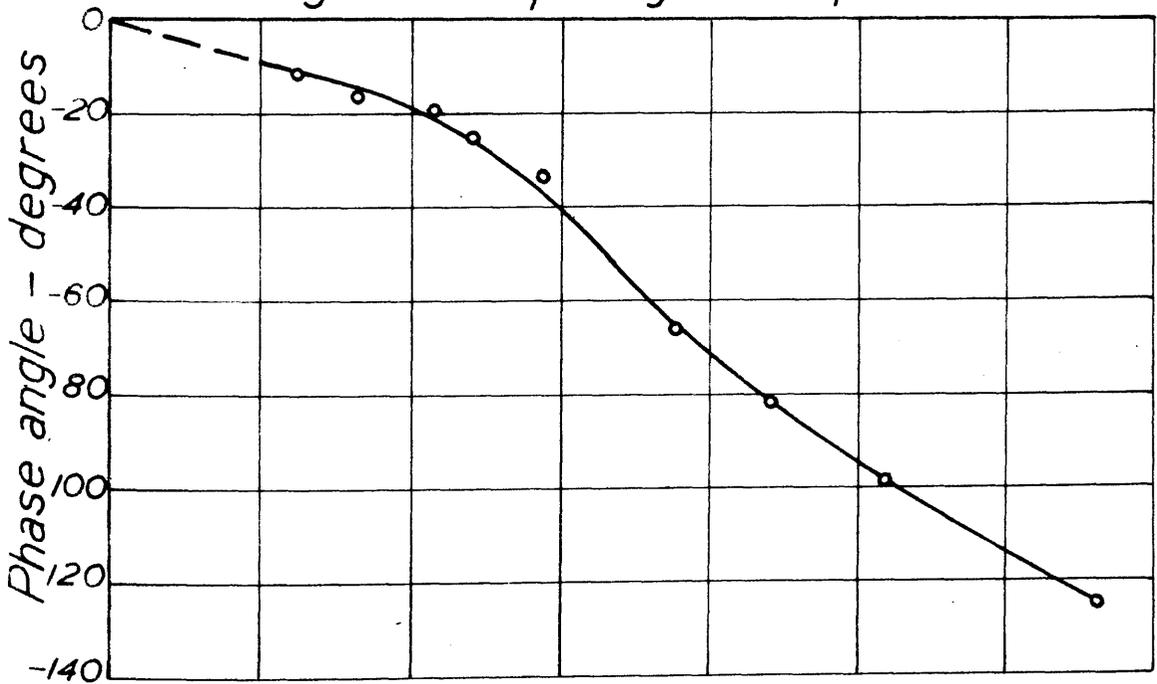
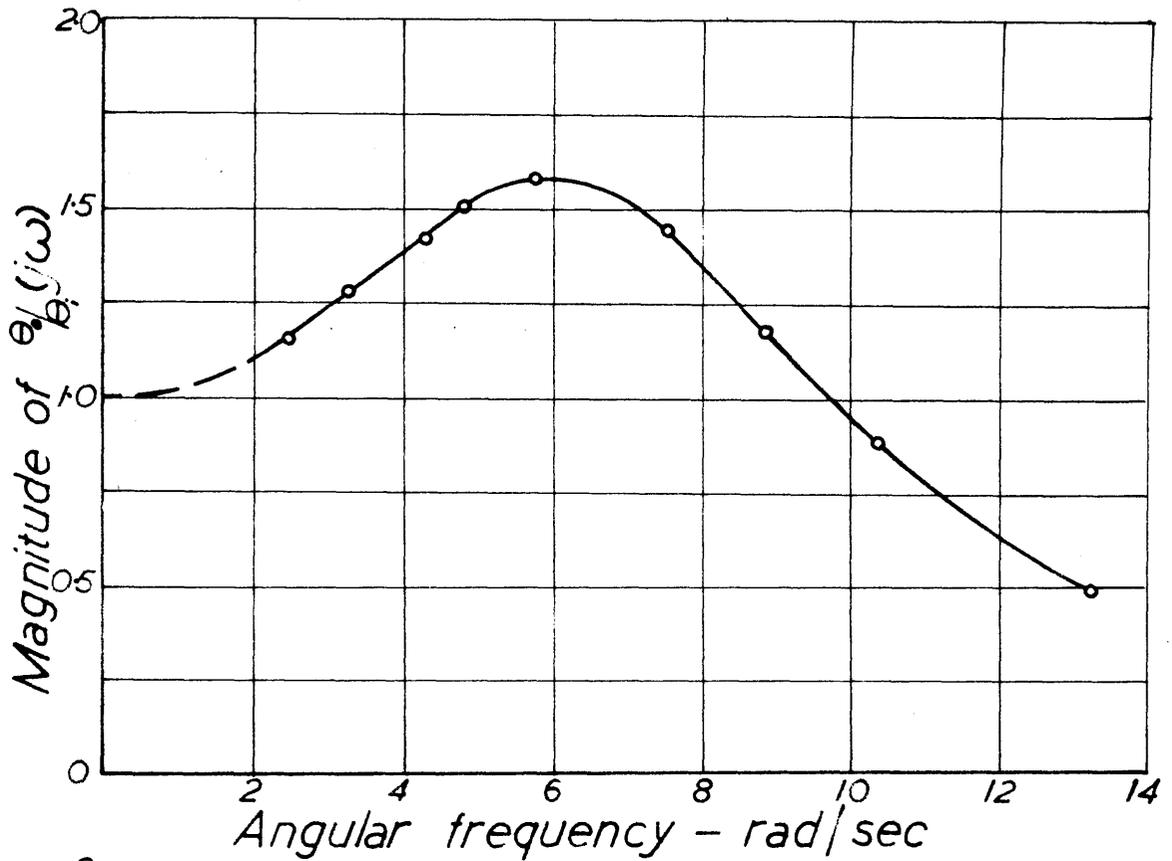
STEP RESPONSES CALCULATED FROM FREQUENCY RESPONSES.

11.1. Frequency Responses.

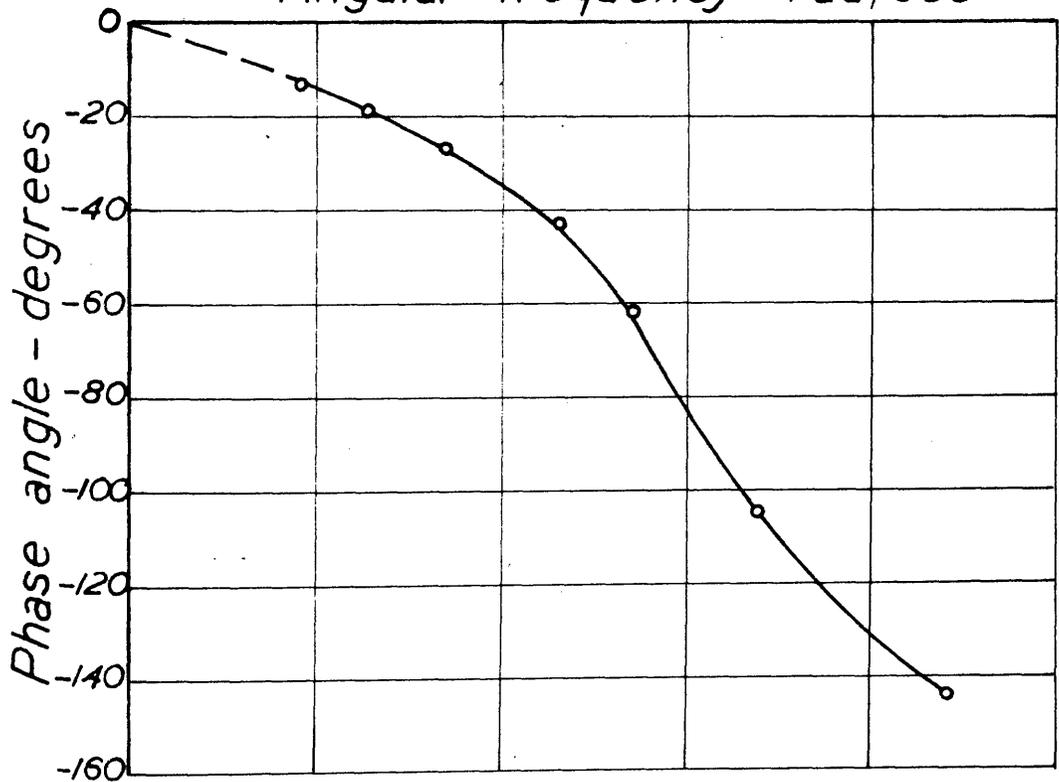
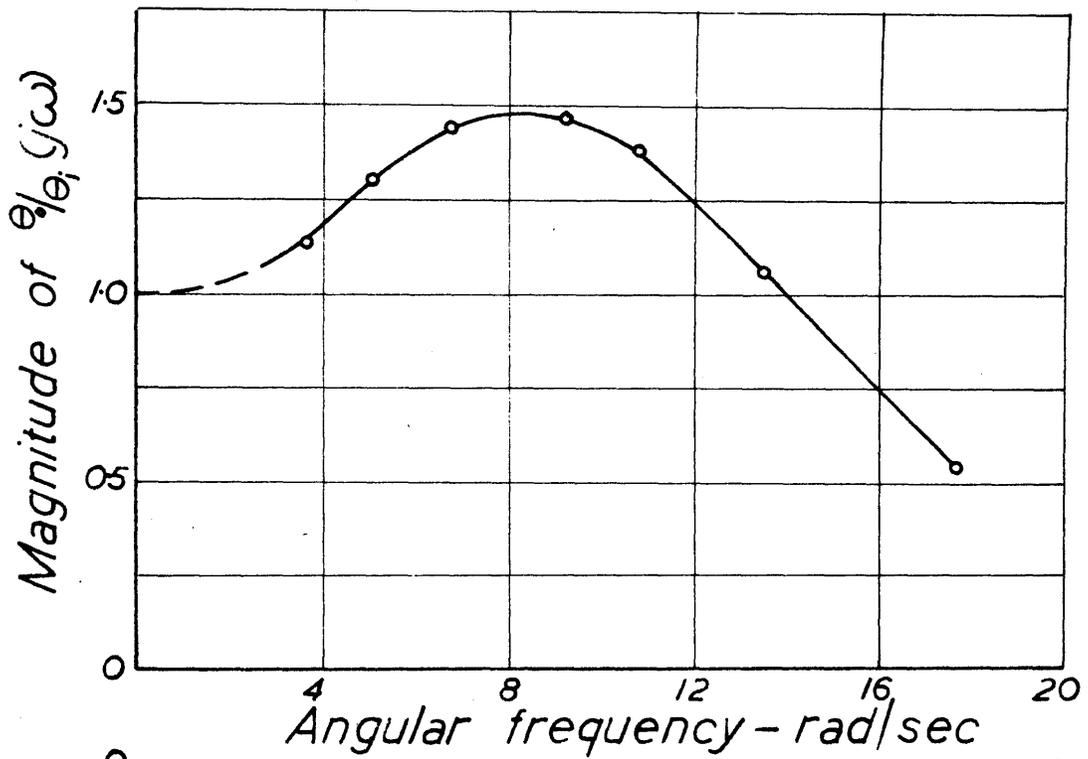
On the three following pages are shown the frequency responses I, II and III, taken under the conditions stated. The first experimental point for these responses was taken at the lowest frequency for which the movement of the output could be considered sinusoidal. Waveforms of the actual output motion at this and other points throughout the experimental frequency range are given later, and show reasonable sinusoidal motion of the output. This point is discussed in relation to the oscillograms of the motor armature current which are shown in Sec. 12.2(b).

The highest frequency point at which readings were observed was, for all three responses, the frequency at which the response became approximately 0.5. This restriction was set by the total dead zone then becoming about 20% of the amplitude of the output oscillation. It was evident that readings beyond this point would have no value.

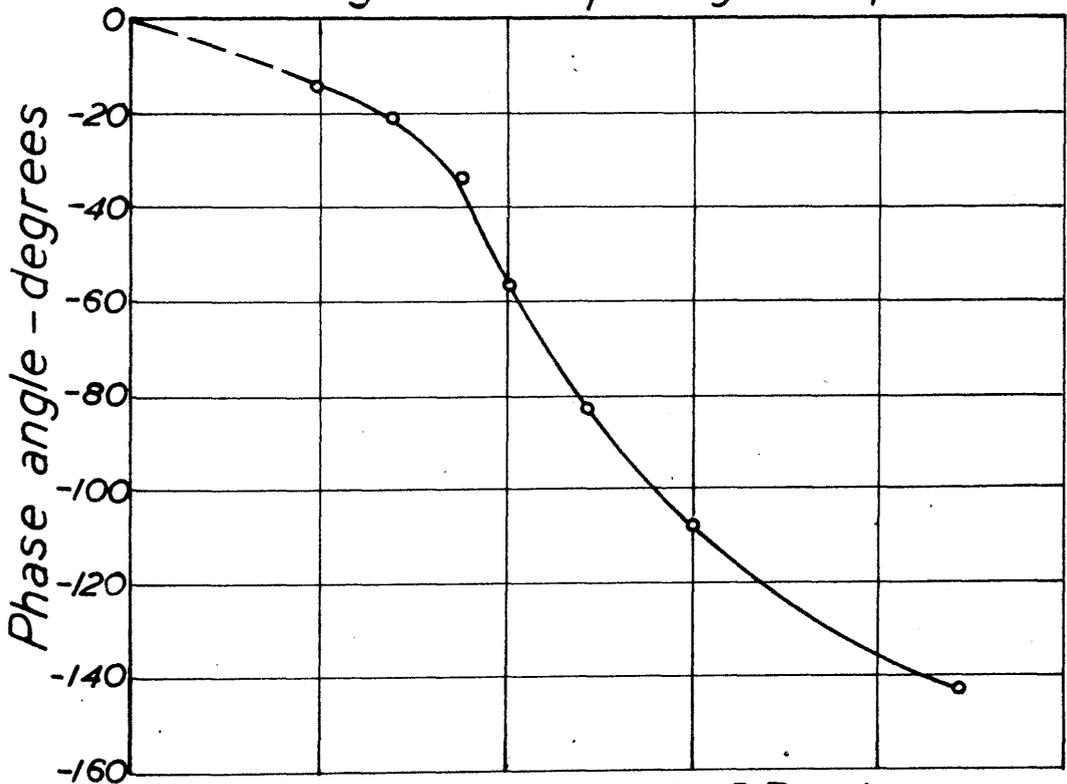
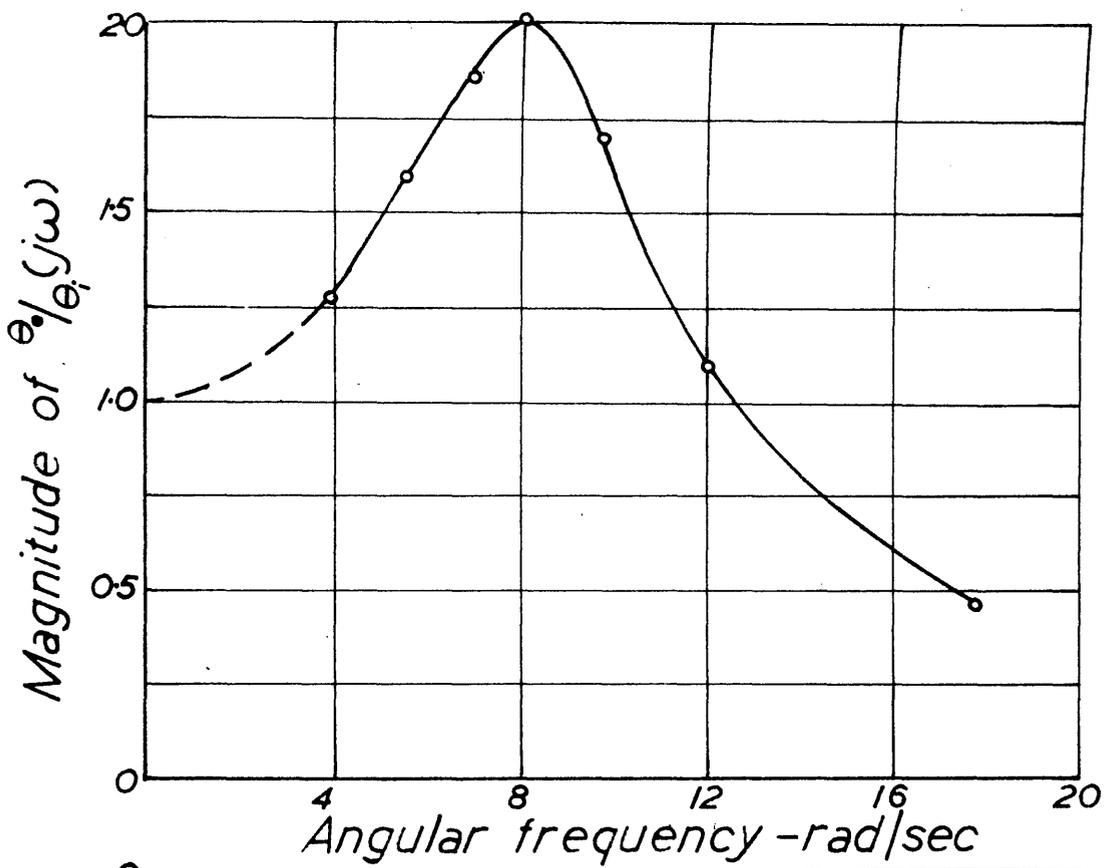
The amplitude response curves have been produced to give $|\frac{\theta_o}{\theta_i}(j\omega)|=1$ at very low frequencies and a similar assumption of zero phase-shift has also been made. Practically, of course, this is not the case, due to stiction. A recording of the output motion at a low frequency is given in Sec. 12.2(b).



stiffness 1.15 lb-ft/°, CR value 0.11 sec
 Fig. 45 Frequency Response I.



stiffness 23lb-ft/l°, CR value 0.11 sec
 Fig. 46 Frequency Response II.



stiffness 23lb-ft/l°, CR value 0.077 sec

Fig. 47 Frequency Response III.

11.2. Step Responses.

Figures 48, 49 and 50 are step-response records taken on the paper recorder described.

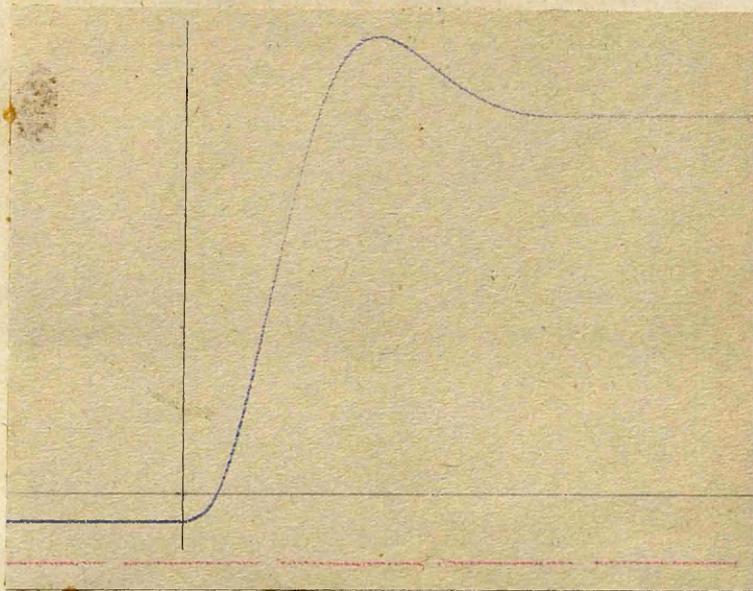


Fig. 48. Experimental Transient Response I.

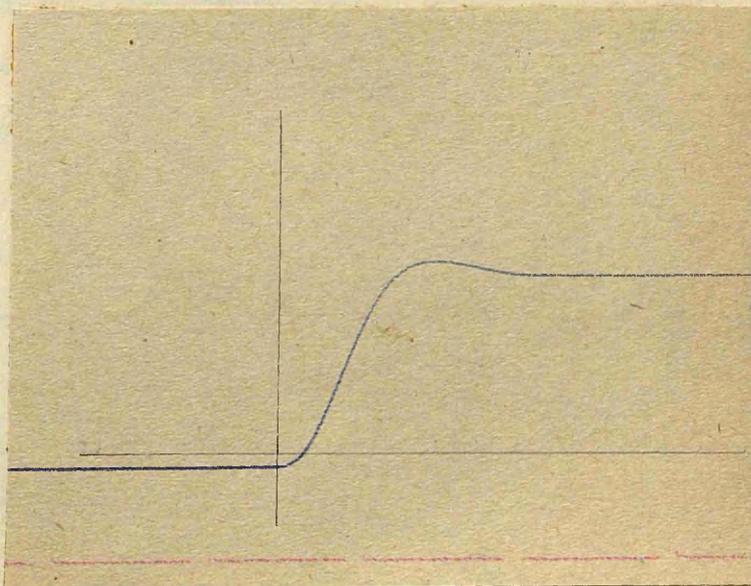


Fig. 49. Experimental Transient Response II.

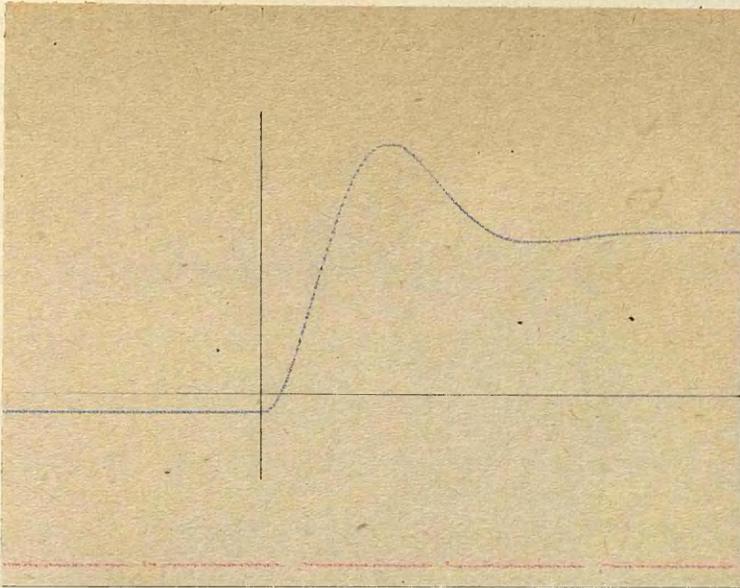


Fig. 50. Experimental Transient Response III.

The step was applied to the system by disconnecting the resetter error voltage from the amplifier input, turning the transmitting mag slip off by the desired amount and re-connecting the resetter output to the amplifier, the entire system of course being operative meanwhile. An attempt was made to eliminate the effect which static error has on the rate of response of the system by taking up the static error before inserting the required offset for the step response. The detailed procedure is illustrated in considering the step response I, Fig. 48, for which condition the static error was largest and approximately 1.5° . Details are

- (a) turn input i.e. transmitting mag slip shaft slowly to any angle α°
output, i.e. resetting mag slip shaft follows to approximately $(\alpha-1.5)^\circ$
- (b) disconnect resetting mag slip output from amplifier and turn input to $(\alpha+20)^\circ$, where 20° is magnitude of input step function.
- (c) start paper drive in recorder.

- (d) start time-marking device
- (e) re-connect resetting magstrip output to amplifier input, observe response and final output angle
- (f) stop paper drive and time-marker.

The re-closing of the switch in (e) coincides with the instant the time-marking pen returns to the paper. The $t=0$ position on the response is then given by measuring from the time-mark a distance equal to the spacing between the two pens. Some error is incurred in this manner but in practice this position co-incident very closely with the point at which the first departure of the output from zero took place.

For the purposes of comparison with the calculated step responses, the detail of each response has been abstracted from the appropriate record and enlarged in the diagram of Figs. 51, 52 and 53.

Before giving these on a unit step response basis, the effect of taking up the static error prior to offsetting the input is taken into account. This is discussed in relation to the experimental response I, Fig. 51, having an input step of 20° . The output in this case showed ultimate co-incident with the input, but it will not always be so, as the output may come to rest anywhere within $\pm 1.5^\circ$ approximately, of the input. In the first place, therefore, only values prior to the point D have any meaning as far as linearity is concerned. In the second place, by the time point C is reached, the output has run through $(20 + 1.5)^\circ$. The correct time for cross-over is more nearly at the point B. The correct overshoot is nevertheless obtained by the height of the shaded part at any

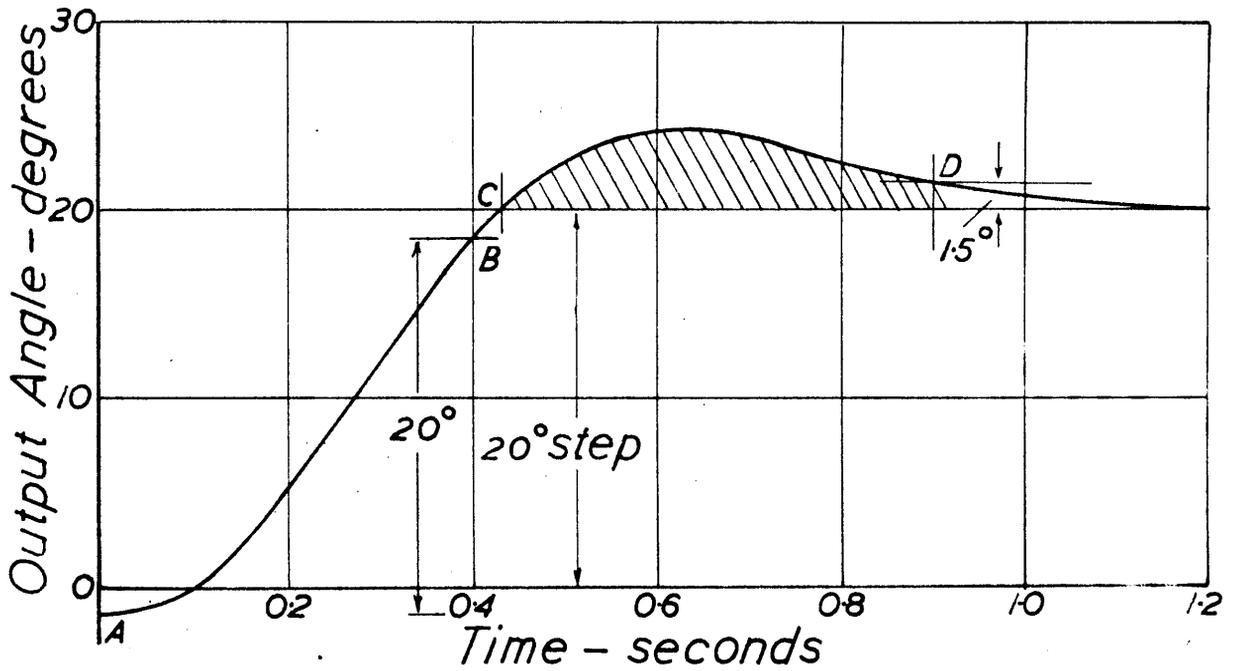


Fig.51 Transient Response I.

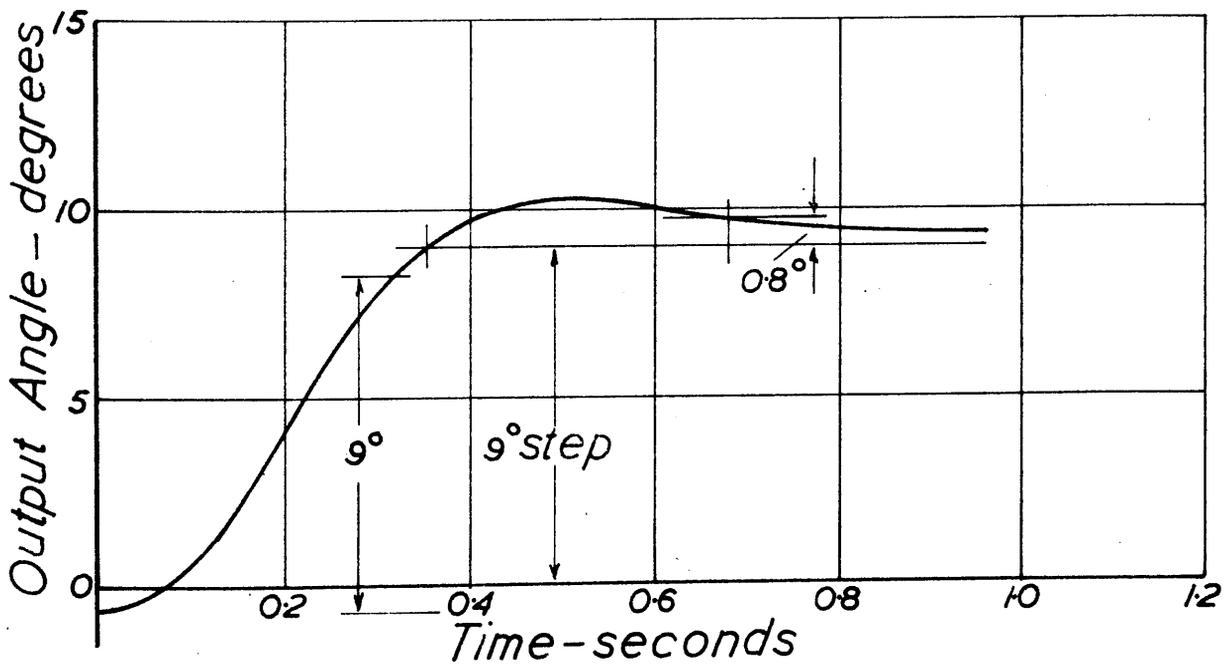


Fig.52 Transient Response II.

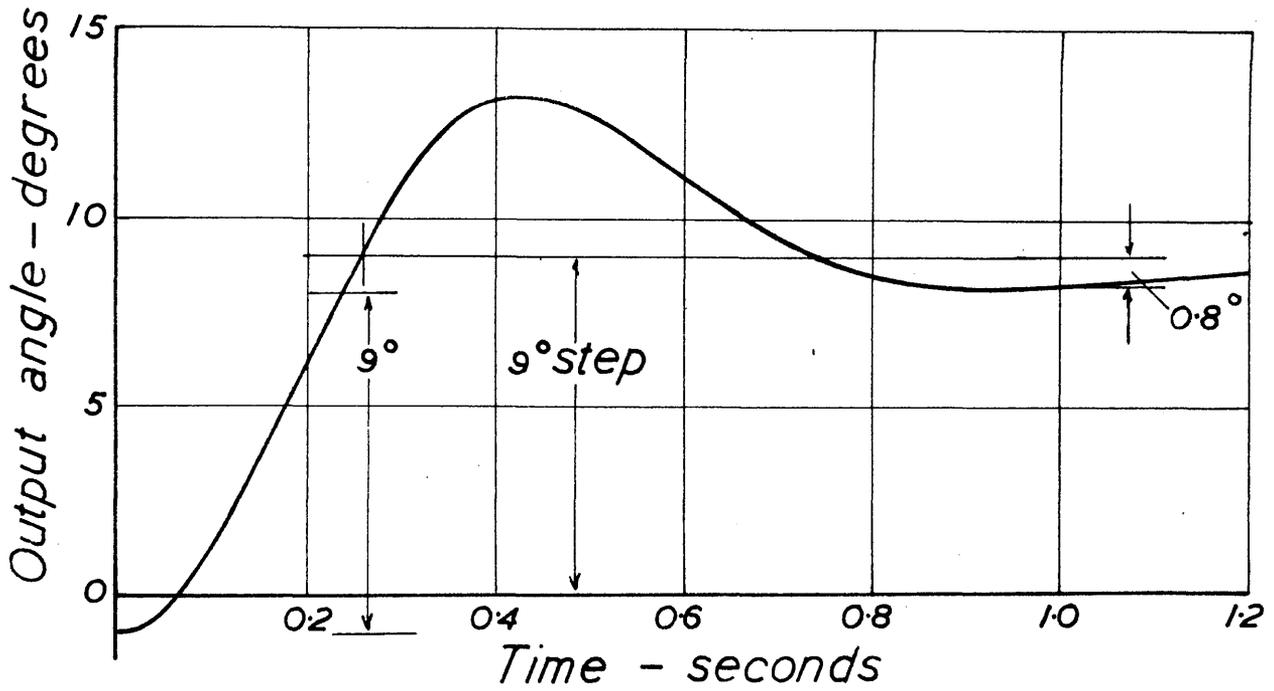


Fig. 53 Transient Response III.

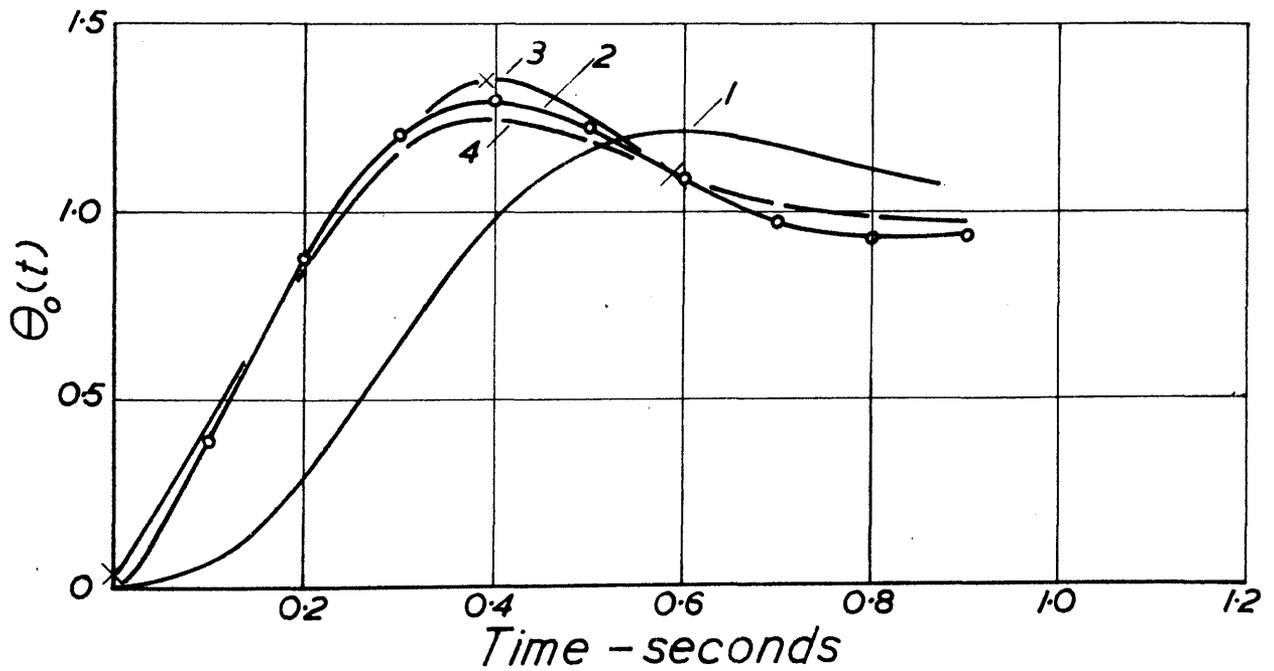


Fig. 54 Transient Response I.

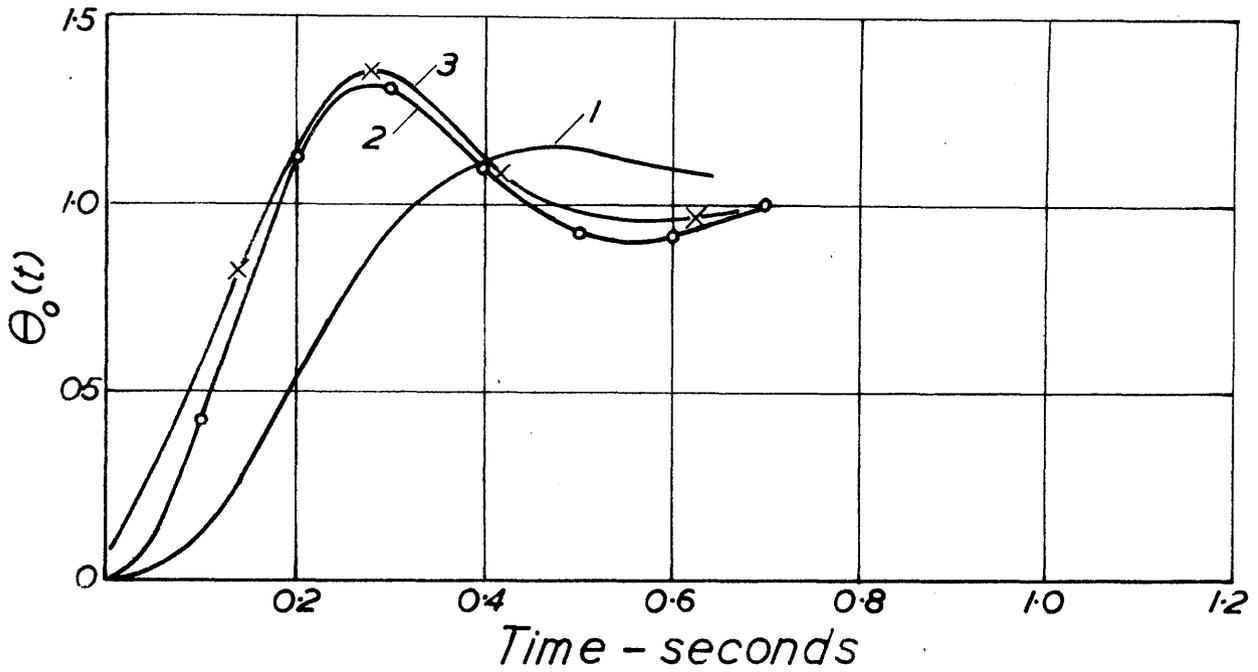


Fig. 55 Transient Response II.

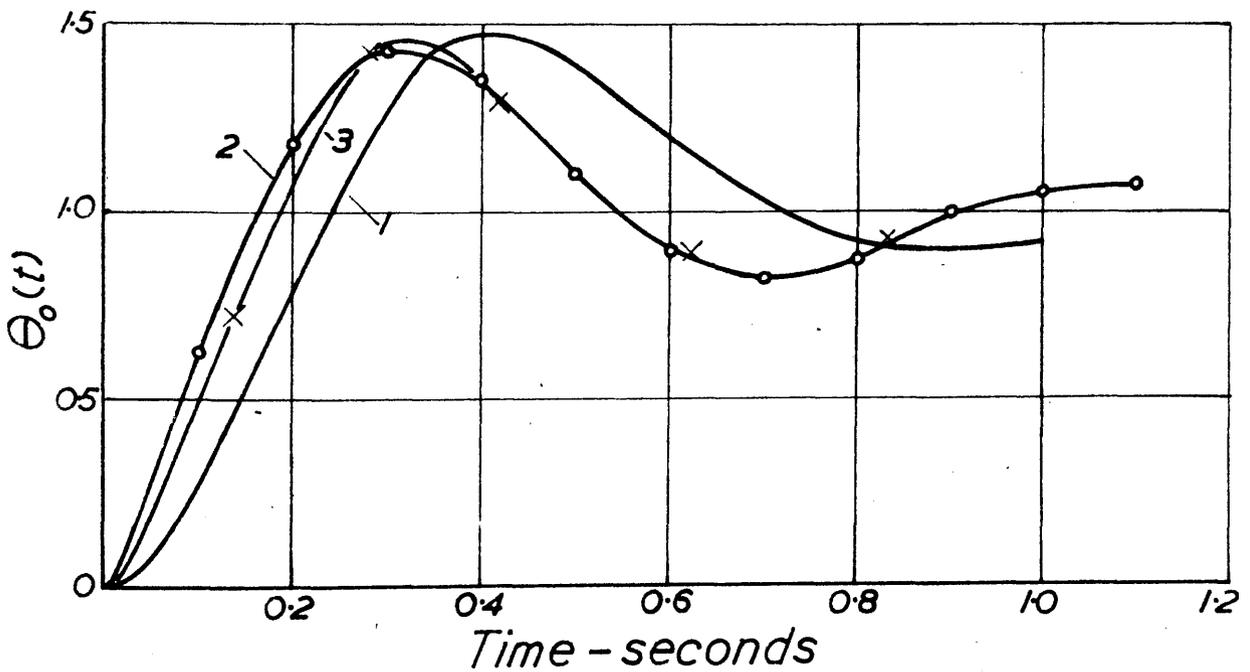


Fig. 56 Transient Response III.

instant up to the point D. This shaded part therefore has been transposed to the point B before evaluating the response on a unit step basis, as shown by curve 1 in Fig. 54. The overall effect is a slight decrease (0.03 seconds) in the time to reach cross-over and to reach maximum overshoot, compared with the recorded value. Even with these restrictions it is still not possible to assert that a truly linear response has been measured, the most important defect still remaining being the increase of friction to the stiction value momentarily at the point of maximum overshoot. This will tend to reduce the maximum overshoot compared with the linear value. The response curves 1 of Figs. 54, 55 and 56 represent, however, the nearest approach to linearity that a practical system is likely to give.

11.3. Calculation of Step Response from Frequency Response.

The step responses for each of the frequency responses I, II and III have been calculated (a) by Campbell's³⁵ method and (b) by the method of Bedford and Fredendall.³⁹ In addition Floyd's⁴⁰ method has also been used for response I. The methods are given below without proof, response I being used for illustrative calculations. The results for this response are given in Fig. 54, and for responses II and III in Figs. 55 and 56 respectively.

Campbell's Method.

Fig. 57 gives the vector response $\frac{\theta_i}{\theta_o}(j\omega)$. From it the Nyquist locus $\frac{\theta_i}{\theta_o}(j\omega)$ has been constructed as shown by the heavy line of Fig. 58, on which the angular frequency points are marked. By erecting a set of

curvilinear squares on this locus, the value of $p = \alpha_1 + j\omega_1$, for which $Y(p) = \frac{\theta_0}{\epsilon}(\rho) = -1$ is found. This is the principal oscillatory root of the characteristic equation

$$1 + Y(p) = 0$$

and from it, the error of the system when excited by a unit step in the input displacement, may be approximately found. The error is in fact

$$e(t) = \frac{2\epsilon}{|p_1| |Y'(p)|} \cos(\omega_1 t - \delta - \beta) \quad (141)$$

where $\alpha_1 + j\omega_1 = |p_1| \angle \delta$ is the value of p satisfying $1 + Y(p) = 0$ and $|Y'(p)|$ is the magnitude of the derivative $\frac{dY(p)}{dp}$ at $p = -\alpha_1 + j\omega_1$. For regular functions such as $Y(p)$, we have

$$\frac{dY(p)}{dp} = \frac{\partial Y}{\partial \alpha} = \frac{1}{i} \frac{\partial Y}{\partial \omega} \quad (142)$$

so that $Y'(p) = \frac{dY(p)}{dp}$ may be evaluated either along a contour of constant ω or along one of constant α . Taking the first method, we obtain $Y'(p) = \text{vector } \overline{AB}$, that is, the change in $Y(p)$ for unit increase in α , measured along the tangent to the ω contour through $[-1, 0]$.

Thus

$$Y'(p) = |Y'(p)| \angle \beta = AB \angle \beta$$

Details abstracted from the diagram of Fig. 58 are, therefore

$$\begin{aligned} \alpha_1 + j\omega_1 &= -3 + j7.1 \\ |p_1| \angle \delta &= 7.71 \angle 112.9^\circ \\ |Y'(p)| \angle \beta &= 0.25 \angle -110^\circ \end{aligned}$$

Hence

$$\begin{aligned} e(t) &= \frac{2\epsilon}{7.71 \times 0.25} \cos(7.1t - 112.9^\circ + 110^\circ) \\ \theta_0(t) &= \underline{1 - 1.037\epsilon^{-3t} \cos(407t - 2.9^\circ)} \quad (143) \end{aligned}$$

This response is shown by curve 2 of Fig. 54. Similarly curves 2 of Figs. 55 and 56 are the outputs calculated by Campbell's method from

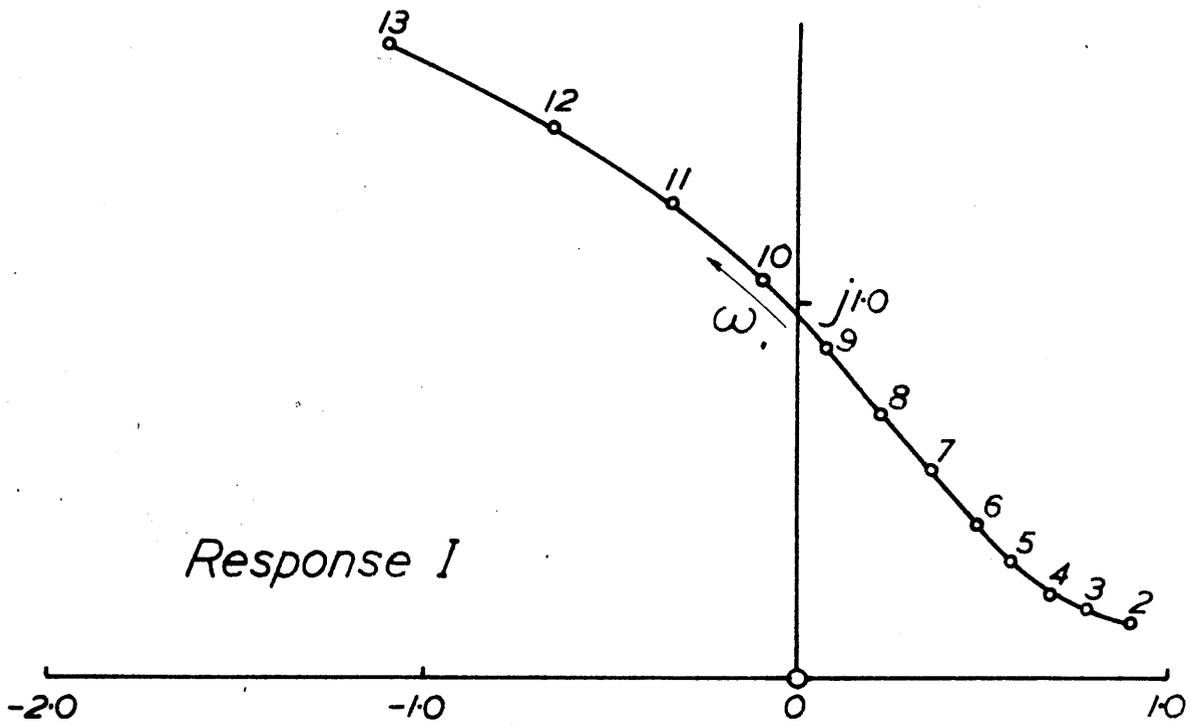


Fig.57 $\theta_i/\theta_o(j\omega)$ Locus.

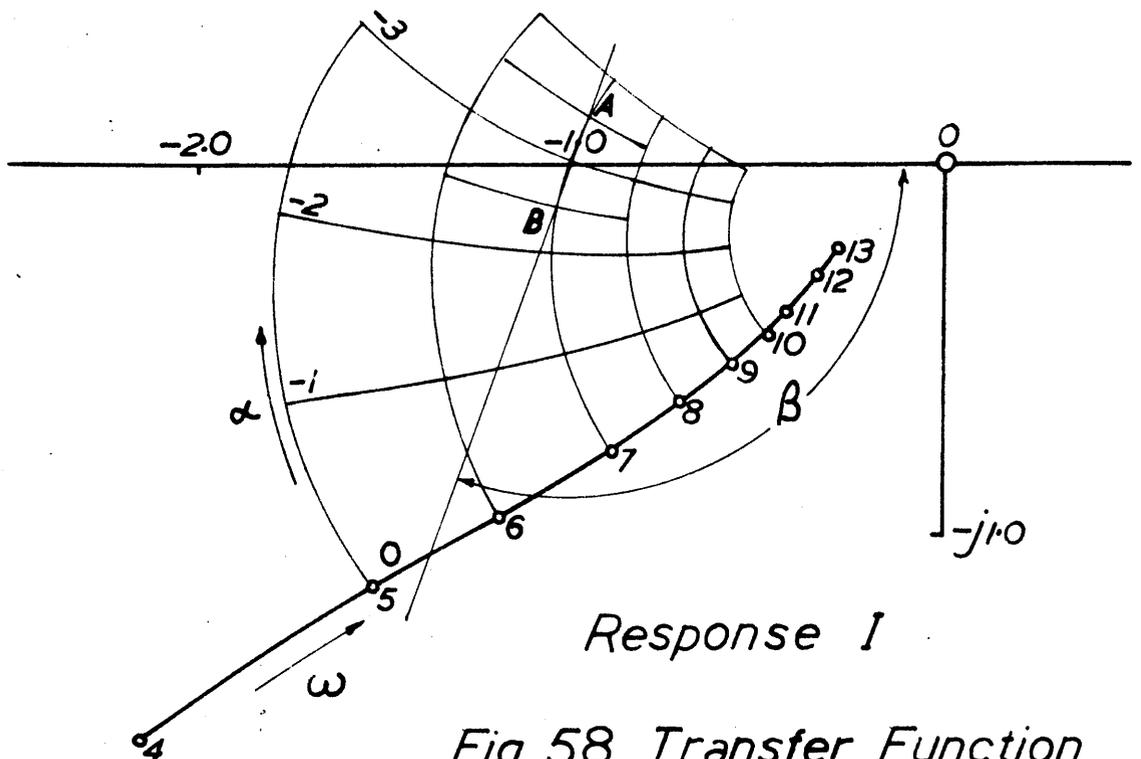


Fig.58 Transfer Function $\theta_i/\theta_o(p)$.

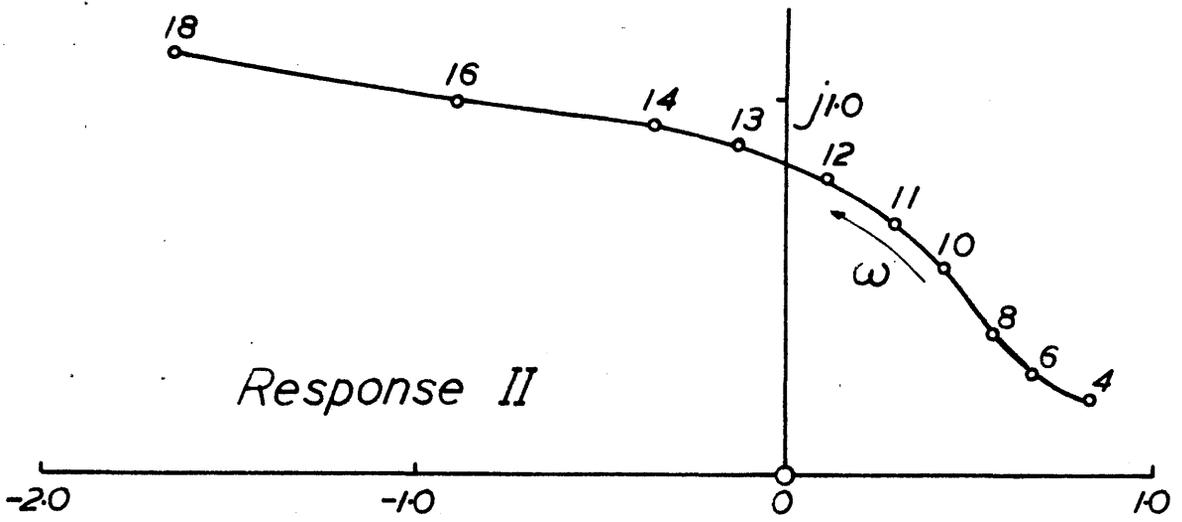


Fig. 59 $\theta_i/\epsilon(j\omega)$ Locus.

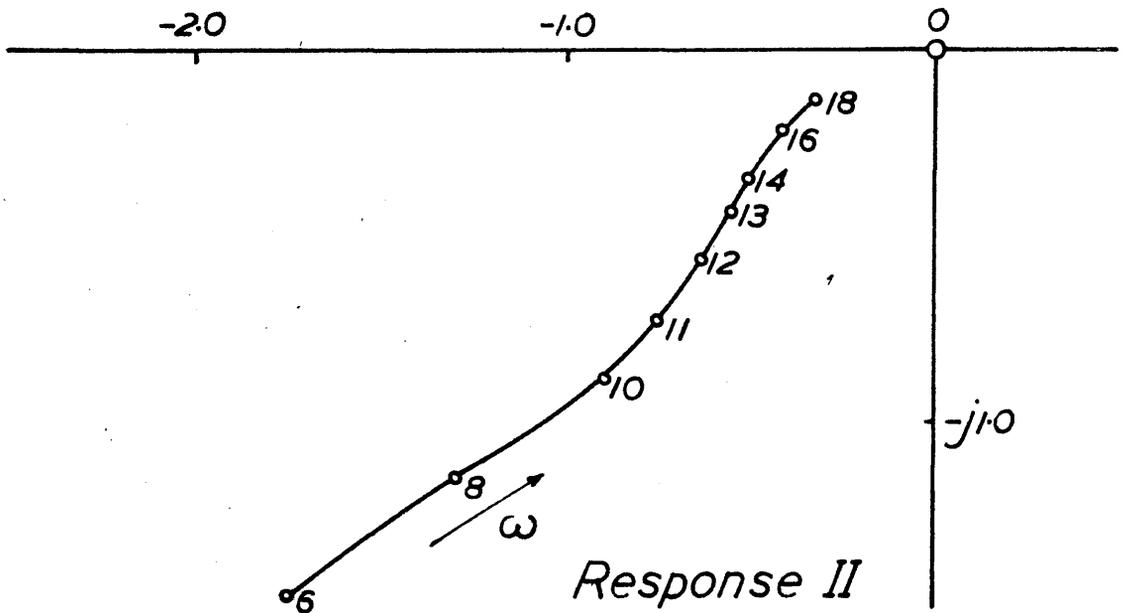


Fig. 60 Transfer Function $\theta_i/\epsilon(p)$.

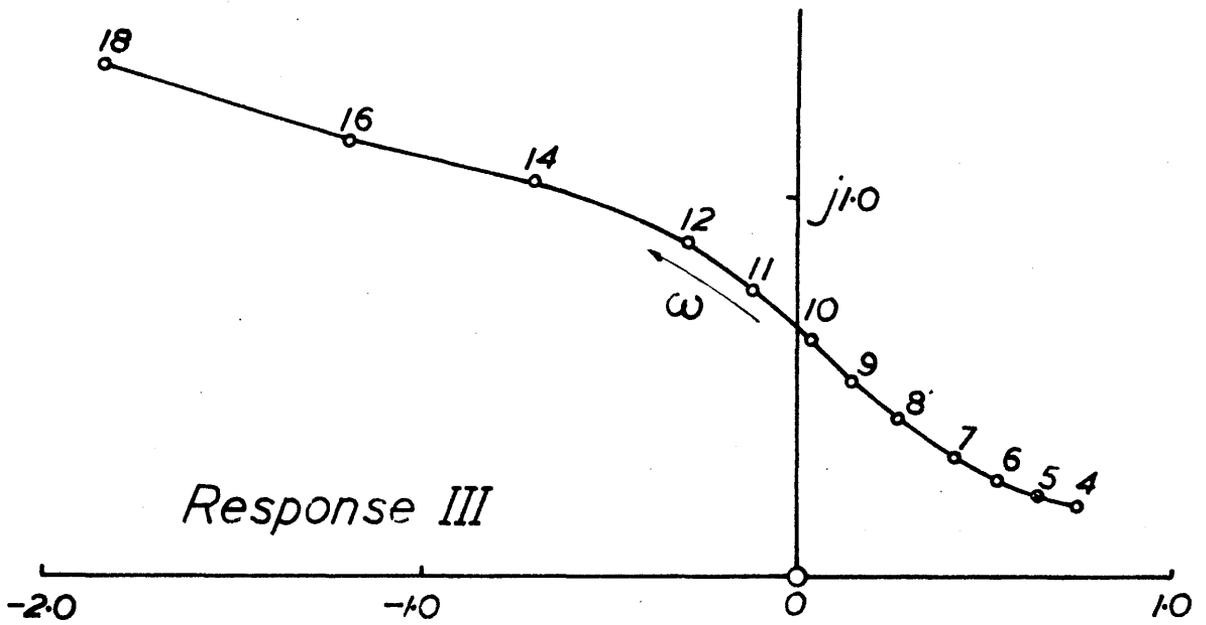


Fig. 61 $\theta_i/\theta_o(j\omega)$ Locus.

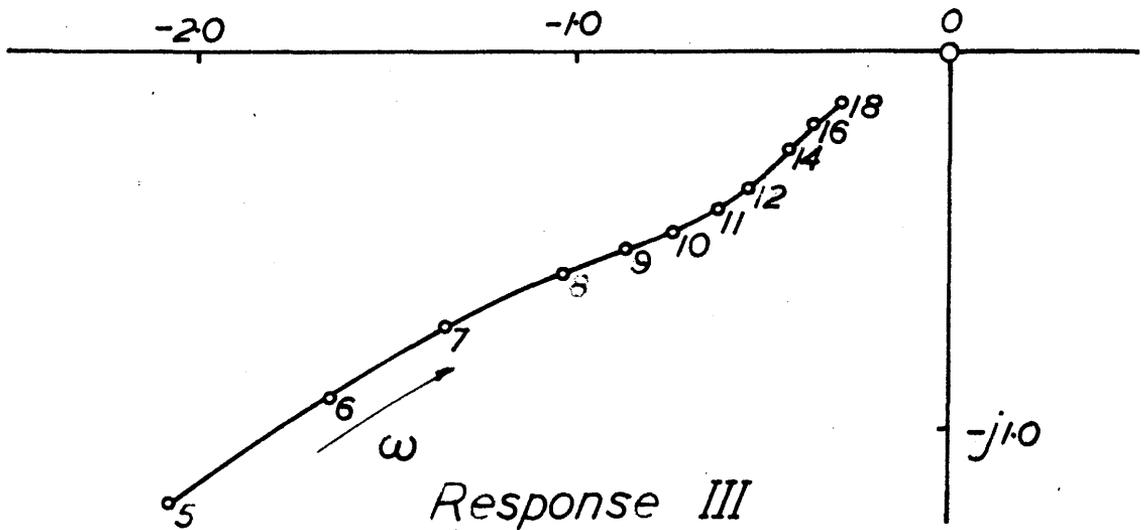


Fig. 62 Transfer Function $\theta_i/\epsilon(p)$.

the frequency responses II and III. These are respectively

$$\theta_o(t) = 1 - 1.123 \varepsilon^{-4.3t} \cos(681t - 31)^\circ \quad (144)$$

and
$$\theta_o(t) = 1 - 0.953 \varepsilon^{-2.3t} \cos(473t - 13)^\circ \quad (145)$$

the details of which have been computed from the diagrams of Figs. 60 and 62.

Method of Bedford and Fredendall .

For the purposes of comparison with Campbell's method, the harmonic synthesis procedure of Bedford and Fredendall was carried out for all three responses. The method depends upon the Fourier Series representation of a rectangular wave of amplitude unity, namely

$$\frac{4}{\pi} \left(\sin \theta + \frac{1}{3} \sin 3\theta + \frac{1}{5} \sin 5\theta + \dots \right) \quad (146)$$

On adding a constant height of unity to this, the periodic wave shown in Fig. 63 is obtained.

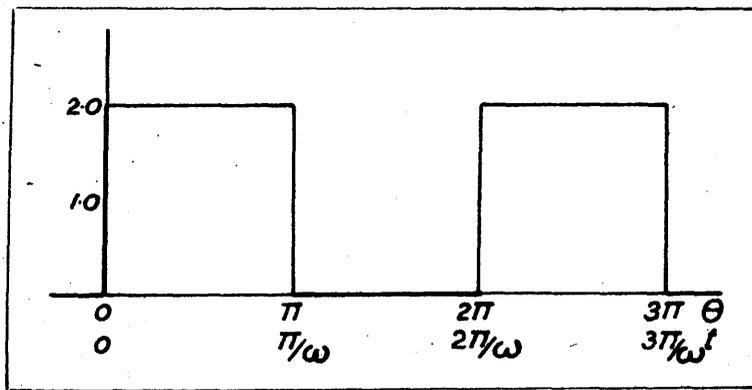


Fig. 63.

Taking half the height of this variation and letting the period increase indefinitely will result in a unit step applied at $t = 0$. For practical purposes, however, it is not necessary to increase the period indefinitely.

Provided the half-period is sufficiently long to allow the transient decay of the response to be completed, then the response of the output to any one of the rectangular blocks will be the same as the response to a unit step.

The method consists therefore of

- (a) deciding upon a fundamental half-period to satisfy the above condition;
- (b) resolving the rectangular wave into a Fourier Series; this will be

$$\frac{1}{2} \left[1 + \frac{4}{\pi} \left(\sin \theta + \frac{1}{3} \sin 3\theta + \frac{1}{5} \sin 5\theta + \dots \right) \right], \quad \theta = \omega t \quad (147)$$

- (c) inspection of the frequency response $\frac{\theta_o}{\theta_i}$ to obtain the phase-shift and gain at each of the frequencies $\omega, 3\omega, 5\omega \dots$; let these be ϕ_n and R_n for any frequency $n\omega$, where n is odd;
- (d) synthesis of the output response, namely

$$\begin{aligned} & \frac{1}{2} \left[1 + \frac{4}{\pi} \left(R_1 \sin(\theta + \phi_1) + \frac{R_3}{3} \sin(3\theta + \phi_3) + \dots + \frac{R_n}{n} \sin(n\theta + \phi_n) \right) \right] \\ & = \frac{1}{2} \left[1 + \frac{4}{\pi} \left(a_1 \sin \theta + \frac{a_3}{3} \sin 3\theta + \dots + \frac{a_n}{n} \sin n\theta \dots \right. \right. \\ & \quad \left. \left. + b_1 \cos \theta + \frac{b_3}{3} \cos 3\theta + \dots + \frac{b_n}{n} \cos n\theta \right) \right] \quad (148) \end{aligned}$$

where $a_n = R_n \cos \phi_n$ = Real part of $\theta_o/\theta_i(j\omega)$
 $b_n = R_n \sin \phi_n$ = $\frac{1}{j}$ (Imaginary part of $\theta_o/\theta_i(j\omega)$)
 and $\theta_o/\theta_i(j\omega) = 1$, at $\omega = 0$.

The method is straightforward and with the aid of Tables* of $\frac{1}{n} \sin n\theta$ and $\frac{1}{n} \cos n\theta$, the computation of (148) can be done fairly quickly.

* See Appendix III.

The composition of a rectangular wave using 11 and 15 harmonics is shown in Fig. 64. Its features are (a) a finite rate of rise, (b) increasing positive values before $t = 0$, and (c) about 8.5% maximum overshoot. To some extent (a) and (c) cancel in comparison with the true rectangular form. With regard to (b), we shall expect at $t = 0$ a small positive value of the response of any system to such an input. In applying the method, a 15-harmonic composition has been adopted for all the responses, but the fundamental periods chosen differ. For response I, the period is 7 secs. and for responses II and III, it is 5 secs. Taking response I, of Fig. 45, we obtain,

n	1	3	5	7	9	11	13	15
ω	0.897	2.69	4.49	6.28	8.07	9.87	11.67	13.46
R_n	1.01	1.2	1.47	1.57	1.33	.95	.67	.46
ϕ_n	-4°	-12°	-22°	-45°	-73°	-95°	-111°	-127°

The calculation is given in full in Appendix IV. The response is shown in Fig. 65. Similar results for responses II and III are given in Figs. 66 and 67. Comparison of these with the measured transient responses is shown by curves 3 of Figs. 54, 55 and 56 all respectively.

It will be seen that the response of Fig. 65 is not zero before $t = 0$ (it is in fact 0.04). A calculation up to 21 harmonics, using the same fundamental, was made for comparison. For this, the frequency responses of Fig. 45 were extrapolated by eye. This gave

$$\begin{array}{lll}
 R_{17} = 0.3 & R_{19} = 0.18 & R_{21} = 0.1 \\
 \phi_{17} = -138^\circ & \phi_{19} = -147^\circ & \phi_{21} = -154^\circ
 \end{array}$$

so that the response is effectively zero thereafter. The calculation

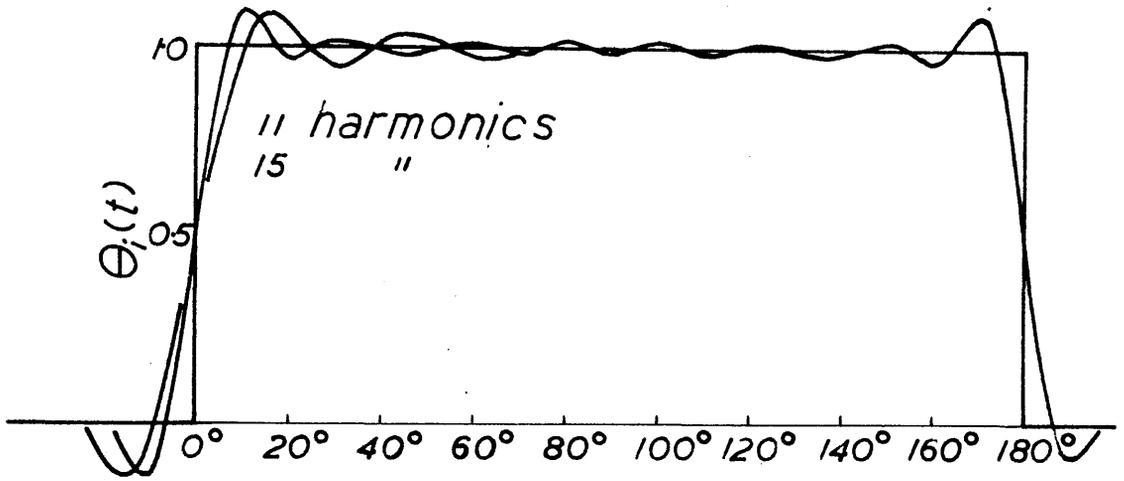


Fig. 64

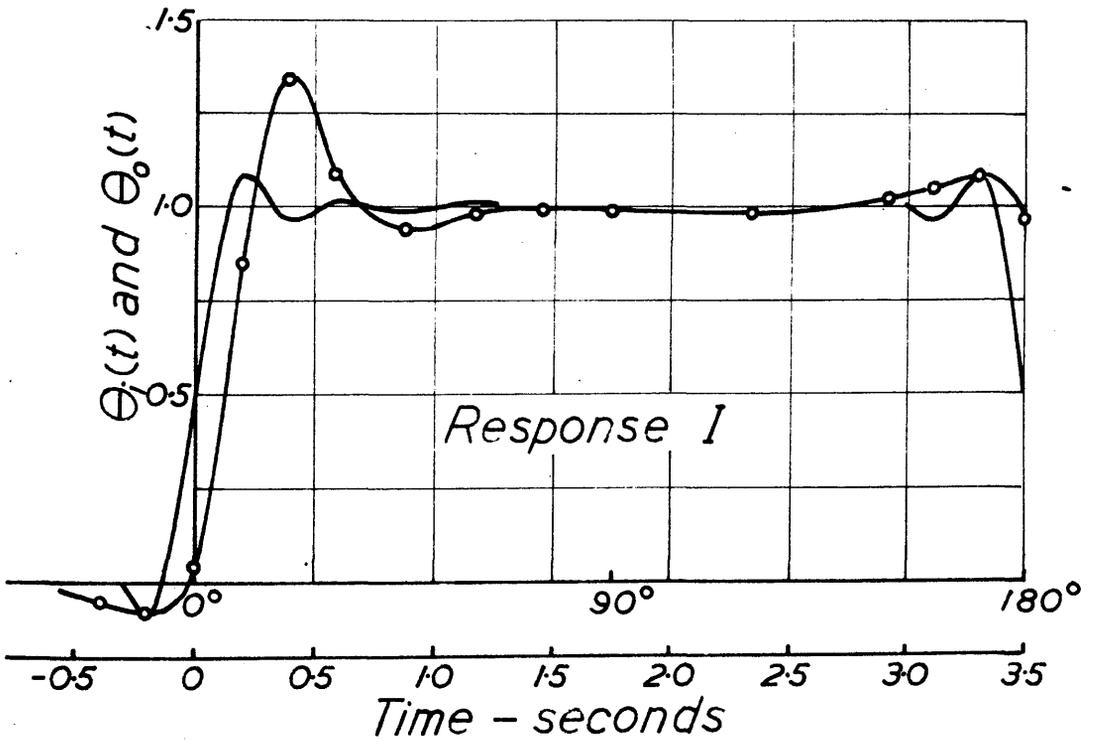


Fig. 65

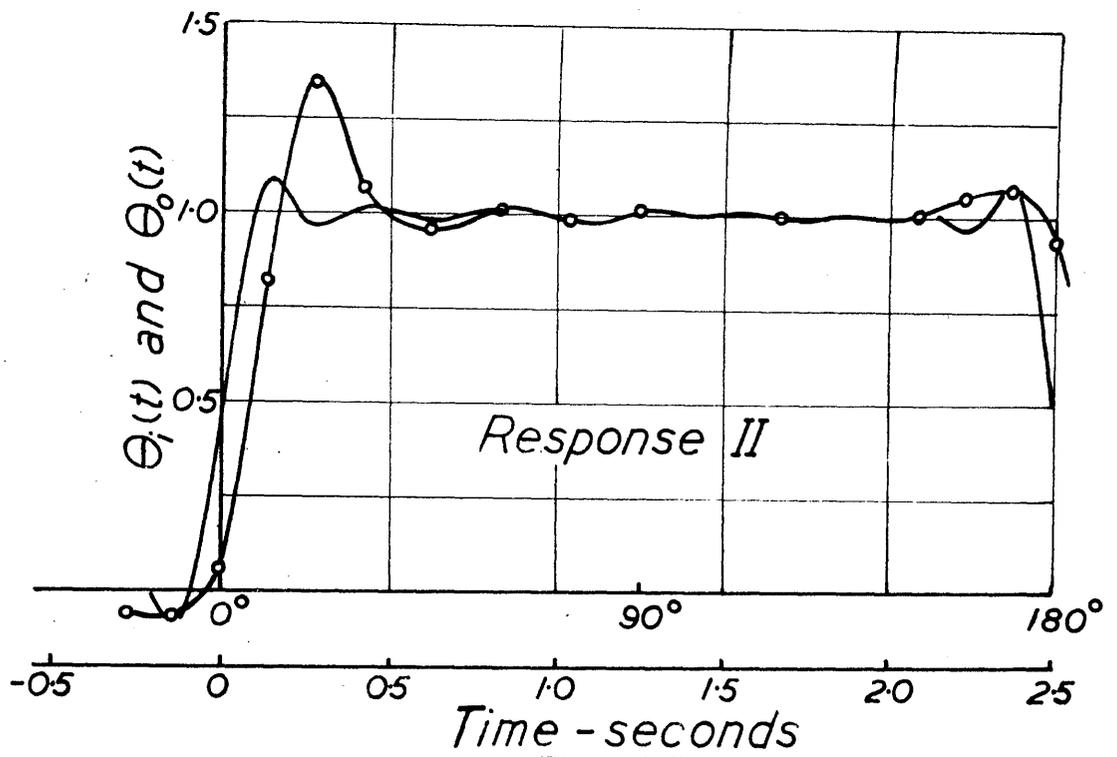


Fig. 66

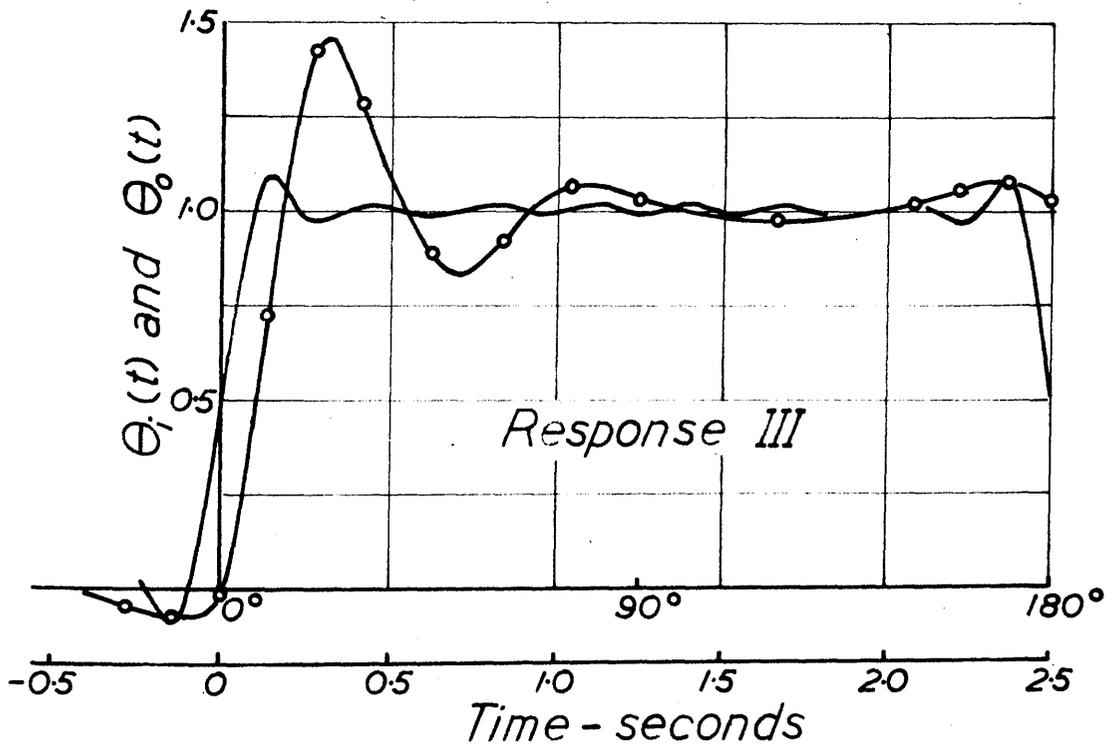


Fig. 67

gives the figure of $\theta_0 = 0.03$ at $t = 0$. This shows only a slight reduction of the $t = 0$ value compared with the 15-harmonic calculation. It is clear that many more harmonics need to be present in the composite input wave before actual physical conditions resembling a unit-step can be approached. The deficiency of the experimental responses' not being known above a finite frequency has not greatly affected the result. In fact it is probable that more exact measurements at the low frequencies (where the response is unity or slightly above unity) would more directly influence the initial value. At $t = 0$, the output contribution is wholly due to the b_n values and it is therefore possible by expressing the result in terms of the a_n values alone to make the initial output zero.*

Floyd's Method .

Floyd's method is essentially a means of performing an approximate inverse Fourier Transformation upon the system frequency response function $\theta_o/\theta_i(j\omega)$. It has been shown in Sec. 2.4 of Part I that the inverse Laplace Transformation of $\theta_o/\theta_i(p)$ is in fact the response of the system to a unit impulse. In the notation of Sec. 2.4, $f_j(t)$ = response to a unit impulse $\delta(t)$ applied at $t = 0$, and $\frac{\theta_o(p)}{\theta_i(p)} = Q(p)$ = system frequency response function.

* The series $\frac{4}{\pi} \sum \frac{a_n}{n} \sin n\theta$ may be used over the ranges $0 < \theta < 90^\circ$, $180^\circ < \theta < 270^\circ$, i.e. the quarter periods commencing at $t = 0, t = \pi/\omega$. As the sine expansion is symmetrical about 90° and 270° , the correct response is not obtained outside these ranges. Here it is preferred to keep a response which at least is physical, although not the exact representation of the unit-step response

Hence

$$f_s(t) = \frac{1}{2\pi j} \int_{c-j\infty}^{c+j\infty} Q(p) \varepsilon^{pt} dp \quad (149)$$

Since all the poles of $Q(p)$ are in the left-half plane, this may be written as a real frequency integral

$$f_s(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} Q(j\omega) \varepsilon^{j\omega t} d\omega \quad (150)$$

$$= \frac{1}{2\pi} \int_{-\infty}^{+\infty} \theta_o/\theta_i(j\omega) \varepsilon^{j\omega t} d\omega \quad (151)$$

The above may be reduced to

$$f_s(t) = \frac{2}{\pi} \int_0^{\infty} (\text{Re. } \frac{\theta_o(j\omega)}{\theta_i}) \cos \omega t d\omega \quad (152)$$

which is the basis of the approximate method due to Floyd. The approximation consists in representing the real part of the $\theta_o/\theta_i(j\omega)$ response by a series of straight lines, as shown in Fig. 68. Considering now the area between the curve and the ω axis, the straight line approximation enables this to be made up of a number of trapezoids. For instance, up to the frequency at which the curve crosses the ω axis, it is composed roughly of three trapezoids, AdeOA, BedAB and one of negative area, AabBA. These are shown separately below, trapezoid numbers 3, 2 and 1 respectively. Trapezoids 4 and 5 which complete the area required are also shown. Each of these trapezoids is identified by three quantities, which, shown for trapezoid 3, are r_3 , ω_3 and Δ_3 , where r_3 = height of trapezoid (positive or negative)

$$\omega_3 = \frac{\omega_d + \omega_e}{2}$$

$$\Delta_3 = \omega_e - \omega_d$$

The integral (152) for this trapezoid, when evaluated, gives

$$\frac{2}{\pi} \omega_3 r_3 \left(\frac{\sin \omega_3 t}{\omega_3 t} \right) \left(\frac{\sin \Delta_3 t}{\Delta_3 t} \right) \quad (153)$$

corresponding results holding for the other trapezoids. Summing therefore,

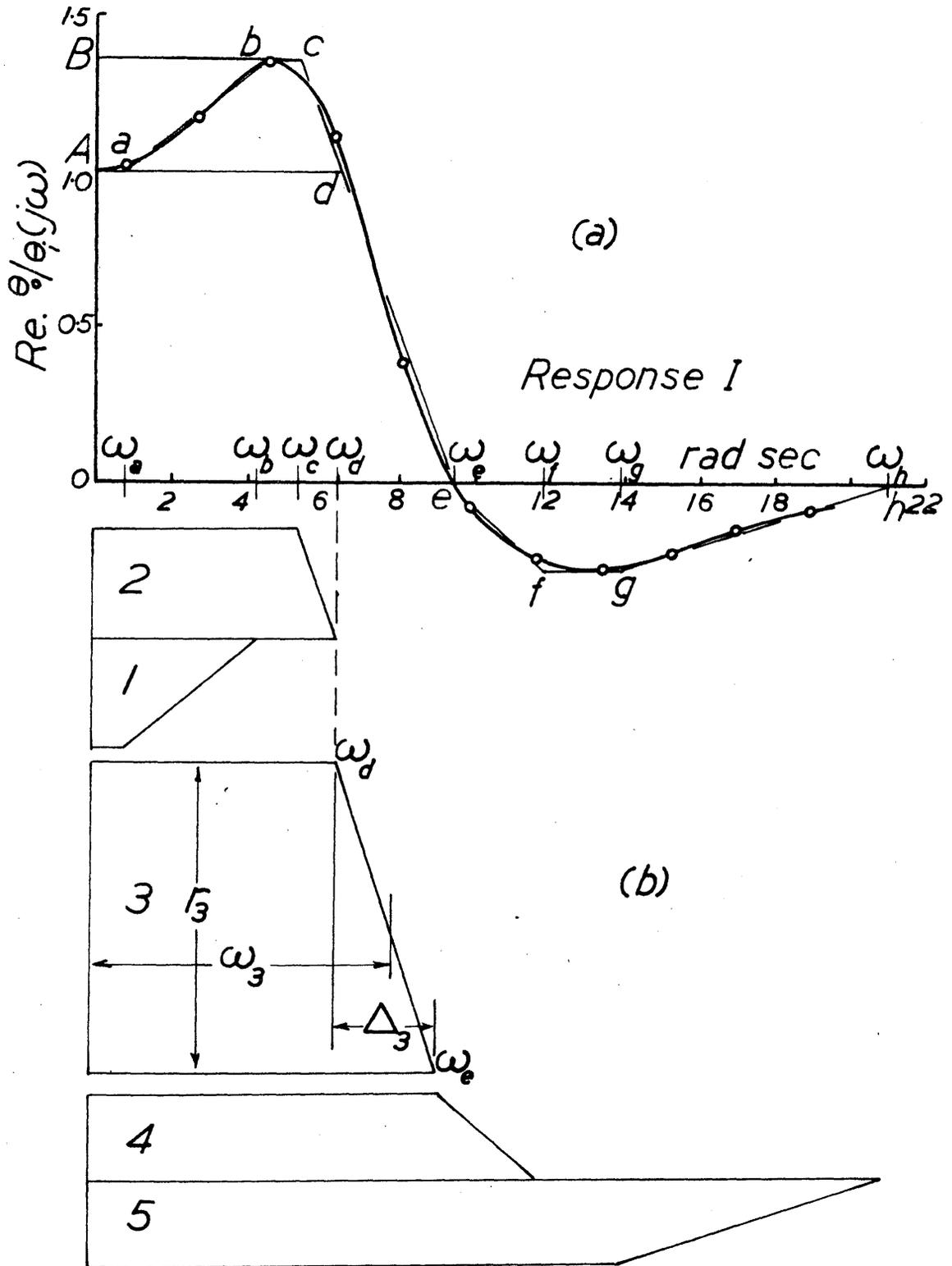


Fig.68 Diagrams illustrating Floyd's method.

the total due to all the trapezoids gives the impulse response,

$$r_s(t) = \sum_{h=1}^5 \frac{z}{\pi} \omega_n r_n \left(\frac{\sin \omega_n t}{\omega_n t} \right) \left(\frac{\sin \Delta_n t}{\Delta_n t} \right) \quad (154)$$

This is computed using a Table of $\frac{\sin x}{x}$.

The numerical values for the response of Fig. 68 are

n	1	2	3	4	5
r	-0.35	0.35	1.0	0.28	-0.28
ω	2.55	5.9	7.95	10.65	17.5
Δ	3.5	1.0	3.1	2.3	7

Result (154) is therefore

$$r_s(t) = \frac{z}{\pi} \left[-0.892 \left(\frac{\sin 2.55t}{2.55t} \right) \left(\frac{\sin 3.5t}{3.5t} \right) + 2.065 \left(\frac{\sin 5.9t}{5.9t} \right) \left(\frac{\sin 1.0t}{1.0t} \right) + 7.95 \left(\frac{\sin 7.95t}{7.95t} \right) \left(\frac{\sin 3.1t}{3.1t} \right) \right. \\ \left. + 2.98 \left(\frac{\sin 10.65t}{10.65t} \right) \left(\frac{\sin 2.3t}{2.3t} \right) - 4.9 \left(\frac{\sin 17.5t}{17.5t} \right) \left(\frac{\sin 7t}{7t} \right) \right] \quad (155)$$

This is computed and then integrated numerically by the time-series operator $\frac{\int [1, 1]}{z [1, -1]}$, where δ is the spacing of the ordinates defining the curve. $\delta = 0.05$ sec has been taken for the integration, and results in the step response shown by curve 4 of Fig. 54. The agreement with the other two calculated responses is quite good. For this reason it was considered unnecessary to carry out Floyd's method for the responses II and III, and since consistent agreement between the two previous methods had already been obtained for all three responses.

The discrepancy between the measured step responses and the calculated responses forms the discussion of the following Chapter.

* See Tustin, A. A Method of Analysing the Behaviour of Linear Systems in Terms of Time Series. Proc. Inst. Elect. Engrs. vol. 94, IIA, 1, p.130. 1947.

CHAPTER 12.

DISCUSSION OF RESULTS.

12.1. Comparison of Calculated with Measured Responses.

The measured π -step responses are given by the curves 1 of Figs. 54, 55 and 56, and the abstraction of these results from the transient records has already been dealt with. In the same diagrams are also given the calculated step responses, the curves numbered 2 being those obtained by Campbell's method and the curves numbered 3 being those resulting from the method of Bedford and Fredendall. Finally curve 4 of Fig. 54, gives the step response calculated by Floyd's method.

Examination of these curves shows the following general features,

- (i) the calculated responses in each case are substantially the same in respect of the time and size of the maximum overshoot.
- (ii) in all cases the time at which maximum overshoot is reached in the calculated responses is considerably less than the time for maximum overshoot in the measured responses. The discrepancy however becomes smaller as the degree of oscillation in the response becomes greater. Thus we have

Response No.	Measured Max.Overshoot.	Time of Meas. Max.Overshoot.	Time of Calc. Max.Overshoot.	Difference as % Meas. Value.
II	15%	0.47 sec.	0.28	40.4
I	21%	0.59	0.4	32.2
III	47%	0.41	0.31	24.4

(iii) there is reasonable agreement between the calculated and measured values of maximum overshoot except when the latter has its smallest value. Actual values are

Response No.	Measured Max.Overshoot.	Calculated * Max.Overshoot.	Difference compared with Meas. Value.
II	15%	31%	+ 16%
I	21	30	+ 9
III	47	43	- 4

* (Campbell's method).

Briefly, therefore, it may be said that the agreement is best in regard to both time and size of the maximum overshoot, when the degree of oscillation in the response is greatest.

12.2. Discussion of discrepancy in results.

The lack of agreement in the time scale of the responses is the most evident feature. The reason for this and the excessive calculated values of the maximum overshoot in the response II, may be investigated under three headings.

- (a) Accuracy of methods used in converting from frequency to step response.
 - (b) Extent of linear operation under frequency response conditions.
 - (c) Reproducibility of system step and frequency responses, and order of accuracy of experimental measurements.
- (a) Accuracy of methods of calculation.

In Campbell's method the approximation made is to assume that the response will not differ greatly from that of a second-order system having

only one complex root-pair identical with the principal complex root of the system under investigation. It is therefore clear that this method will give good results if the system, albeit of high order, has in fact one complex root-pair from which all other roots are remotely situated. The method also assumes the response $\frac{\theta_o}{\theta_i}(j\omega) \rightarrow 1$ as ω tends to an indefinitely low value. Campbell's procedure will therefore give good results for oscillatory systems but will give poor results for systems damped by large real roots. The graphical accuracy of the method may be checked by performing a calculation in a simple second-order system, having only one complex root-pair. This has been done elsewhere and establishes an order of accuracy of 2 - 3%.

The Fourier synthesis method on the other hand depends on a measured characteristic throughout its whole range and will depend therefore on the shape of this irrespective of what root and zero pattern produces it. By taking an analytical example, the normal accuracy of the method can be shown to be again about 2 - 3%.

From the above therefore we may conclude that, although the approximations made in calculation may contribute errors of 2 - 3% under the best conditions, a major discrepancy of 25 - 40% is not accountable to them.

(b) Extent of linear operation of system under frequency response conditions.

Waveforms of output angle and motor armature current.

It has already been stated that experimental frequency response

measurements were made down to the lowest frequency at which the output motion was apparently a sinusoid. For frequencies lower than this, the output shows a definite stationary time at the peak amplitude of its movement. The stationary period is of course due to a finite error being required to allow a driving torque equal to the stiction torque to be developed. As the frequency is lowered so the stationary period increases. Fig. 69 below gives a recording of the output motion at a frequency of about 0.12 c/s, with a sinusoidal input movement of $\pm 10^\circ$ amplitude, and a control stiffness equal to that of response I.

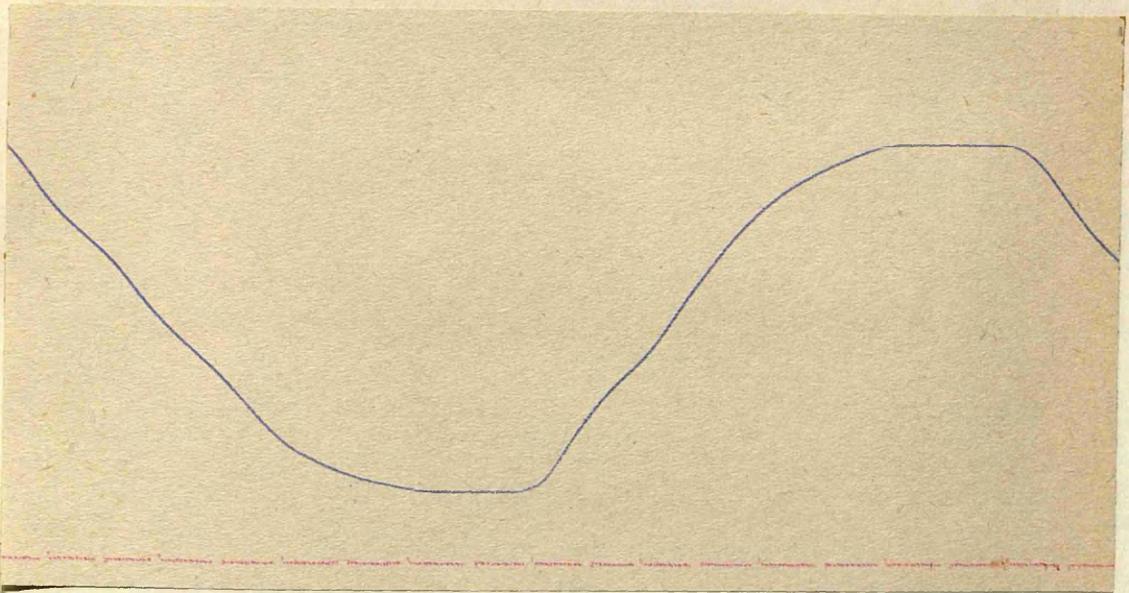


Fig. 69.

The stationary period for the output is about 0.8 sec. and the amplitude of the output movement is only $\pm 9.3^\circ$ *. A large increase

* The amplitude scale for the paper recordings is $1^\circ = 0.1''$ and time-marking dashes occur every 0.5 sec. (see Sec. 10.2).

in the control stiffness causes jerkiness to be apparent when the output leaves its stationary position, as in Fig. 70, but at the same time, the stationary period decreases and the output amplitude more nearly approaches the input amplitude.

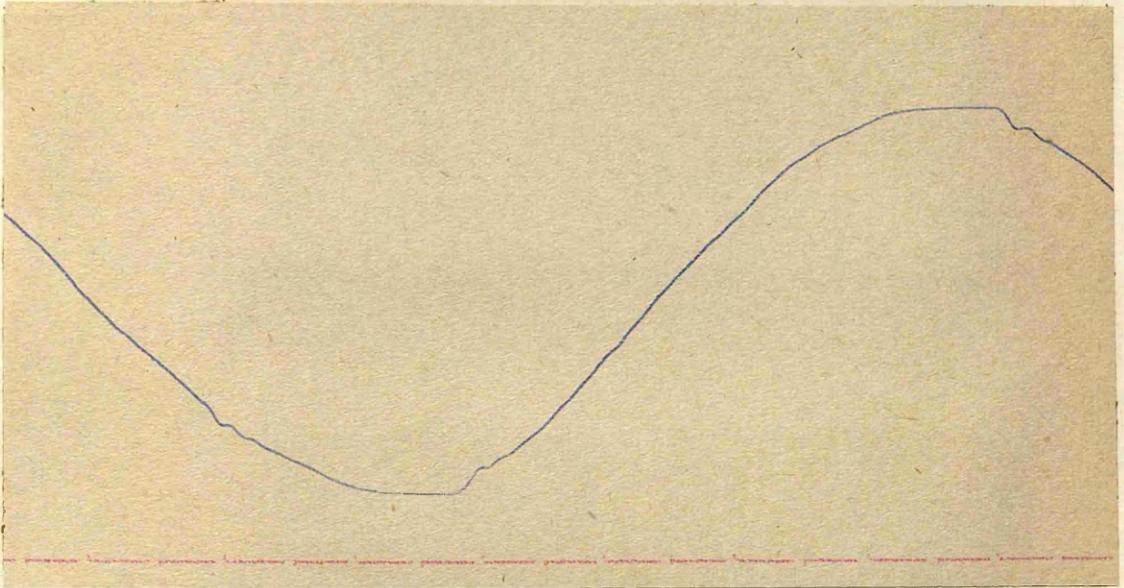


Fig. 70.

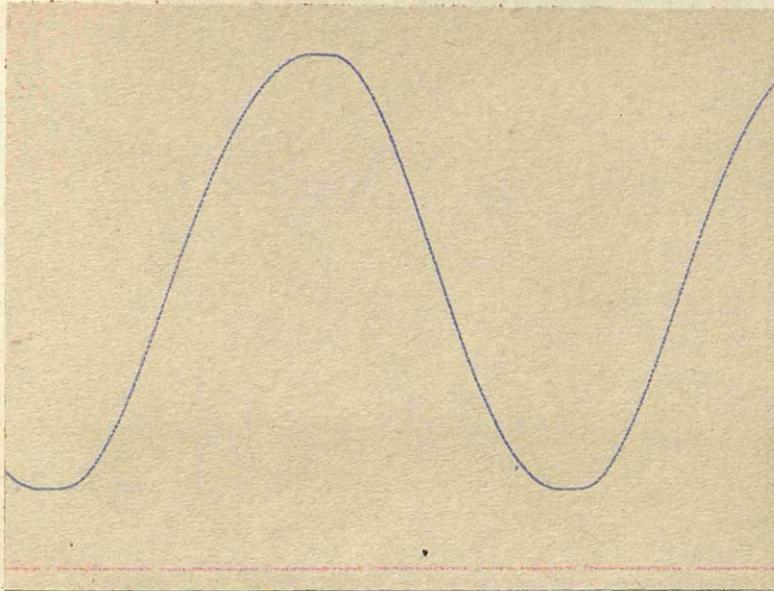
The frequency in Fig. 69 is considerably less than the lowest points at which the frequency responses I, II and III were commenced. Response I, having control stiffness 1.15 lb-ft/o, started at approximately 0.4 c/s and responses II and III, having control stiffness 2.3 lb-ft/o, started at about 0.6 c/s. Records were taken of the output motion and of the motor armature current for several points throughout the range of experimental frequency response measurement. The frequency points, for each of the two values of control stiffness, are as follows.

1. Lowest experimental point.
2. Lowest frequency at which the motor armature current waveform has a reasonable fundamental sine component.
3. Frequency for which $\left| \frac{\theta_o}{\theta_i}(j\omega) \right|$ was a maximum, i.e. the resonant frequency.
4. Frequency at which $\left| \frac{\theta_o}{\theta_i}(j\omega) \right|$ again became unity.
5. Frequency at which $\left| \frac{\theta_o}{\theta_i}(j\omega) \right|$ decreased to 0.5.

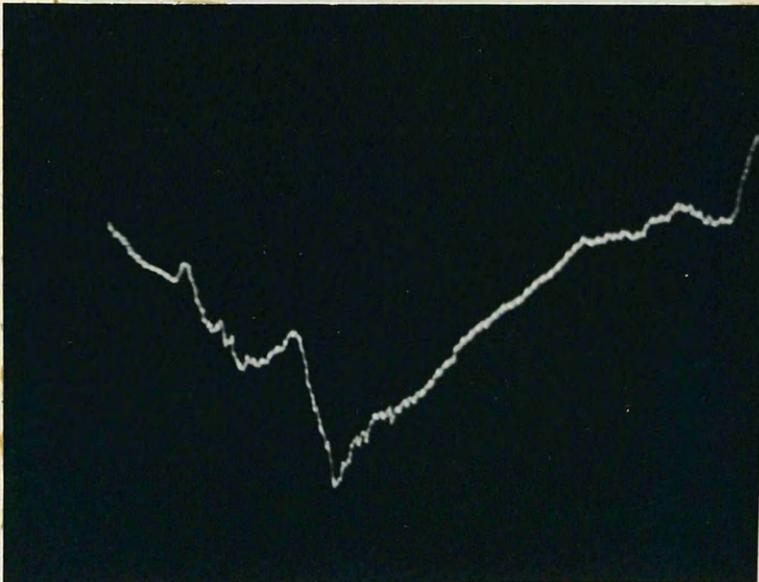
The output motion records Figs. 71a to 75a refer to the frequencies 1 to 5 just described with a control stiffness of 1.15 lb-ft/°.

Oscillograms Figs. 71b to 75b give the motor armature current variation for these points. Following this group are records Figs. 76a to 80a, and oscillograms Figs. 76b to 80b, which represent the same points 1 to 5 (but not the same frequencies) for a control stiffness of 2.3 lb-ft/°.

Response I. Figs. 71 to 75. Control stiffness 1.15 lb-ft/°. Input amplitude $\pm 10^\circ$



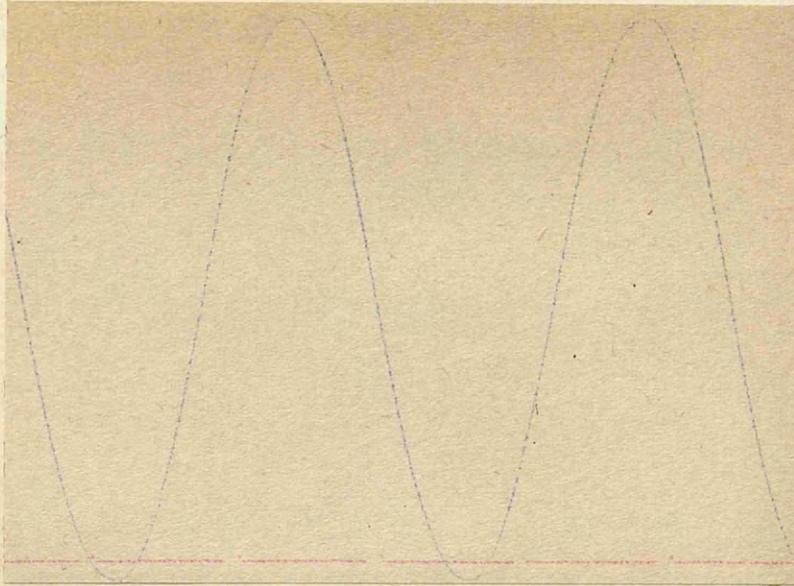
(a) Output.



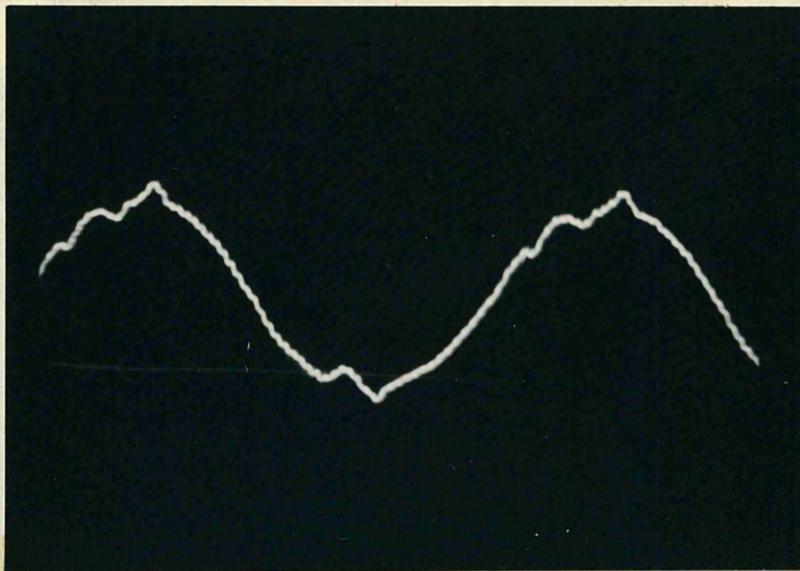
(b) Motor armature current. (approx. 0.5 cycle)

Fig. 71. 'Point 1. Approx. lowest experimental reading.

$$\omega = 2.62, \left| \frac{\theta_o}{\theta_i}(j\omega) \right| = 1.14.$$



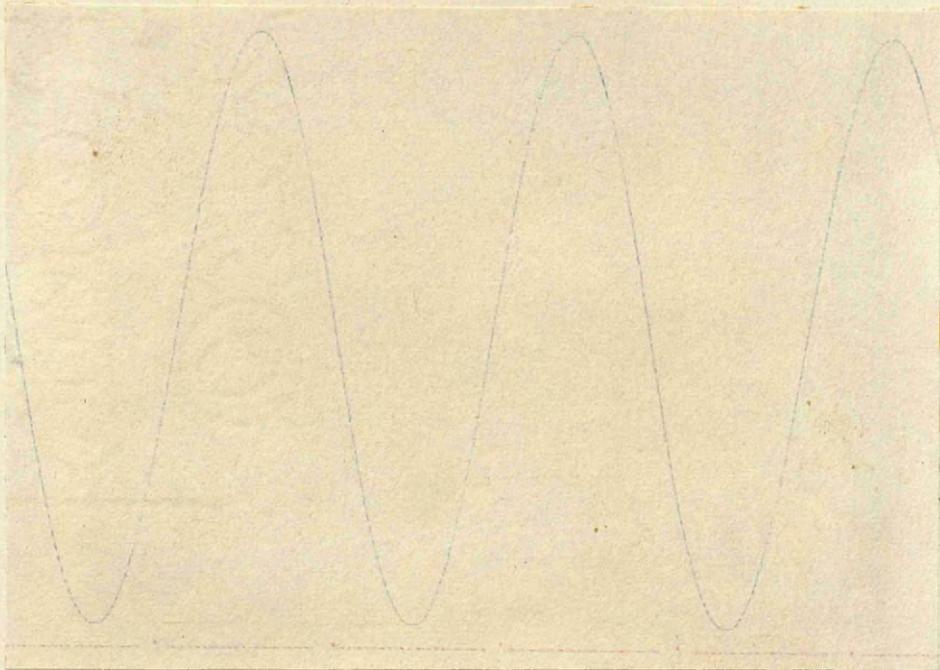
(a) Output.



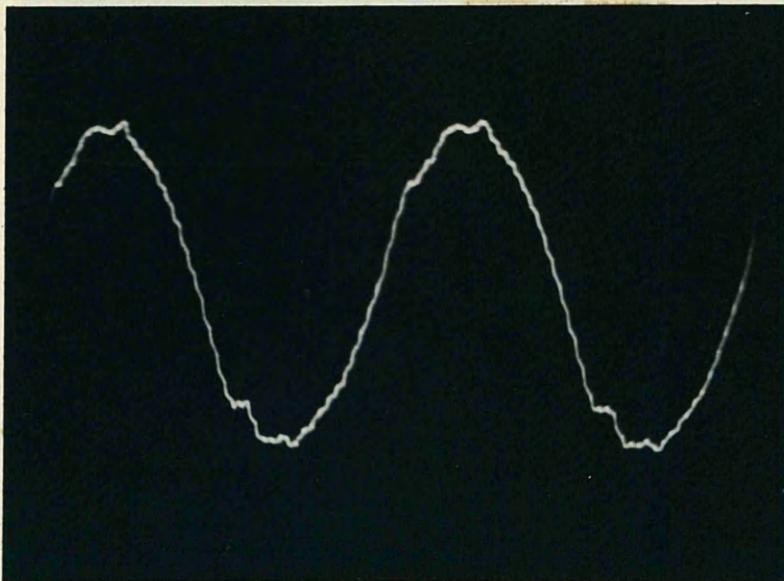
(b) Motor armature current.

Fig. 72. Point 2. Lowest frequency for reasonable fundamental sine component in motor armature current waveform.

$$\omega = 5.19, \left| \frac{\theta_o(j\omega)}{\theta_i} \right| = 1.49.$$



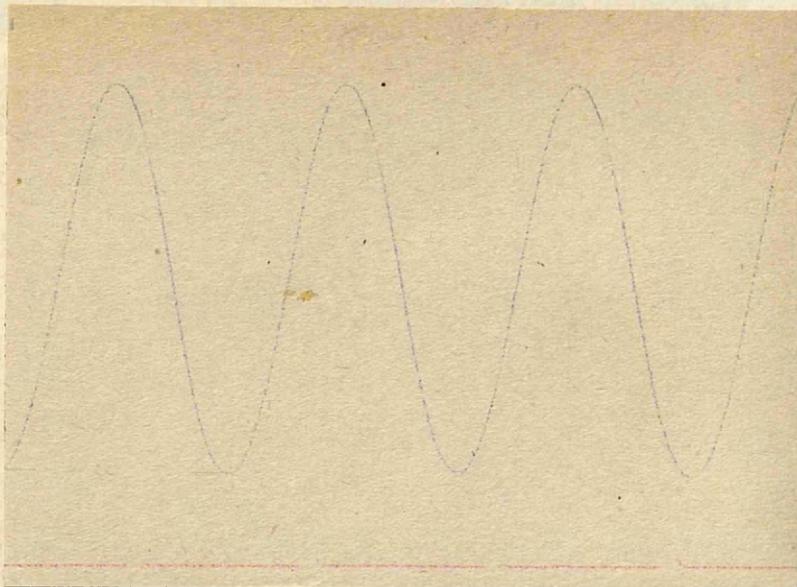
(a) Output.



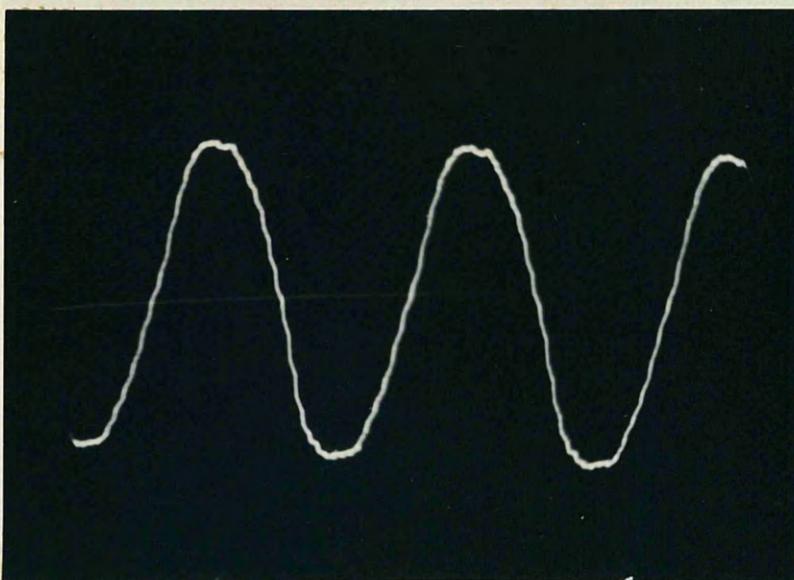
(b) Motor armature current.

Fig. 73. Point 3. Approx. resonant frequency.

$$\omega = 6.88, \left| \frac{\theta_o}{\theta_i}(j\omega) \right| = 1.55.$$



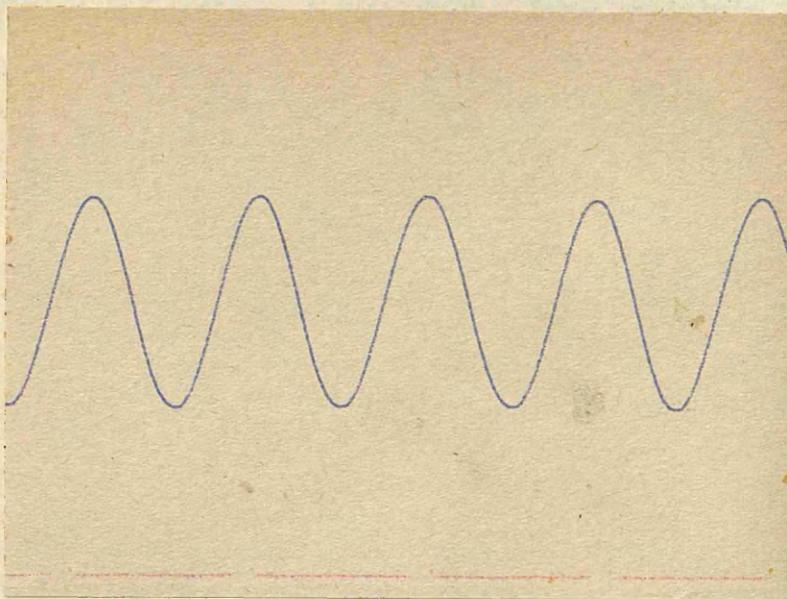
(a) Output.



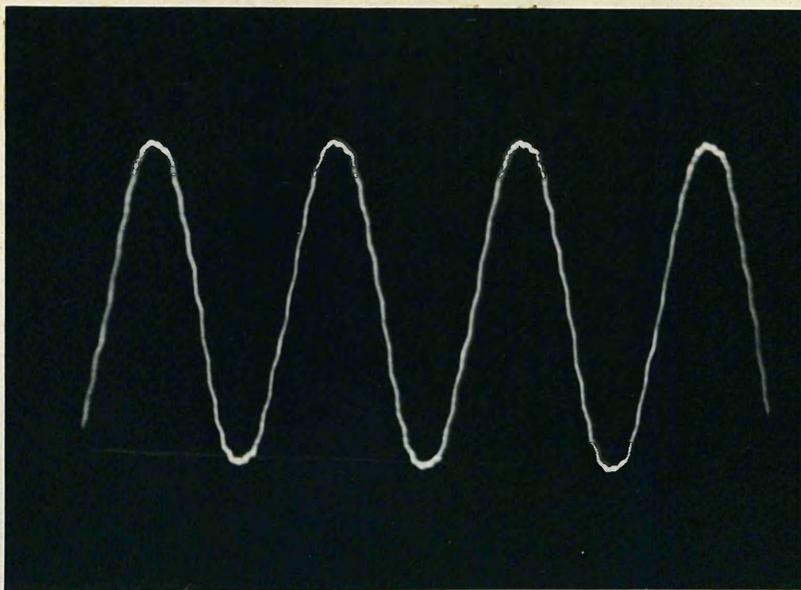
(b) Motor armature current.

Fig. 74. Point 4. Frequency at which $\left| \frac{\theta_o}{\theta_i}(j\omega) \right| \doteq 1.0$.

$$\omega = 9.72, \quad \left| \frac{\theta_o}{\theta_i}(j\omega) \right| = 1.02$$



(a) Output.

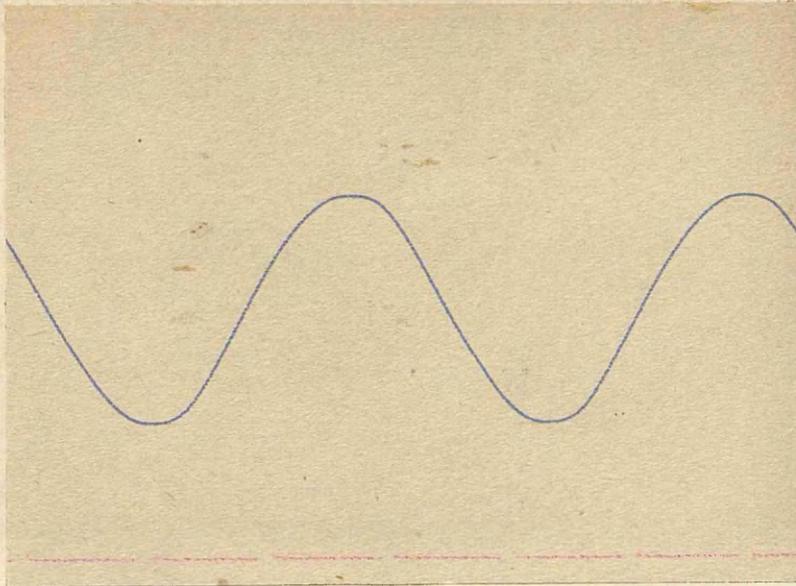


(b) Motor armature current.

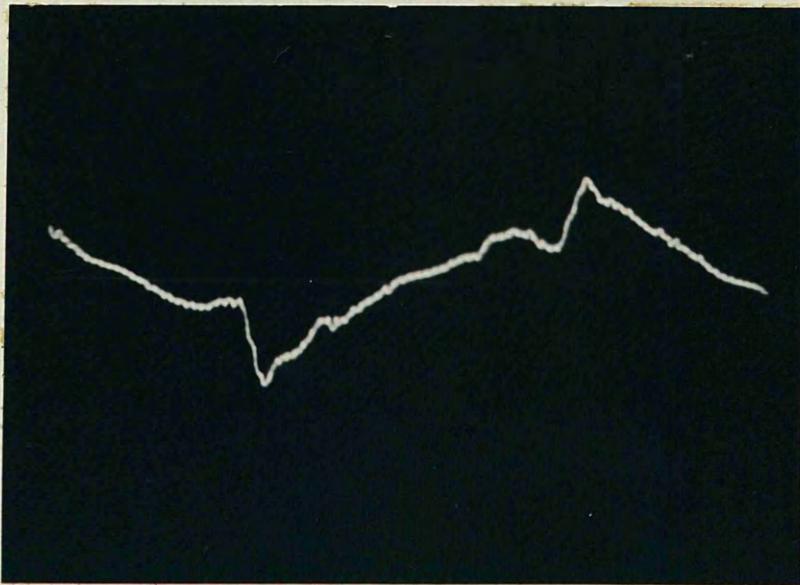
Fig. 75. Point 5. Frequency at which $\left| \frac{\theta_o}{\theta_i}(j\omega) \right| \doteq 0.5$.

$$\omega = 13.46, \quad \left| \frac{\theta_o}{\theta_i}(j\omega) \right| = 0.55.$$

Response II. Control stiffness 2.3 lb-ft/° . Input amplitude $\pm 5^\circ$.



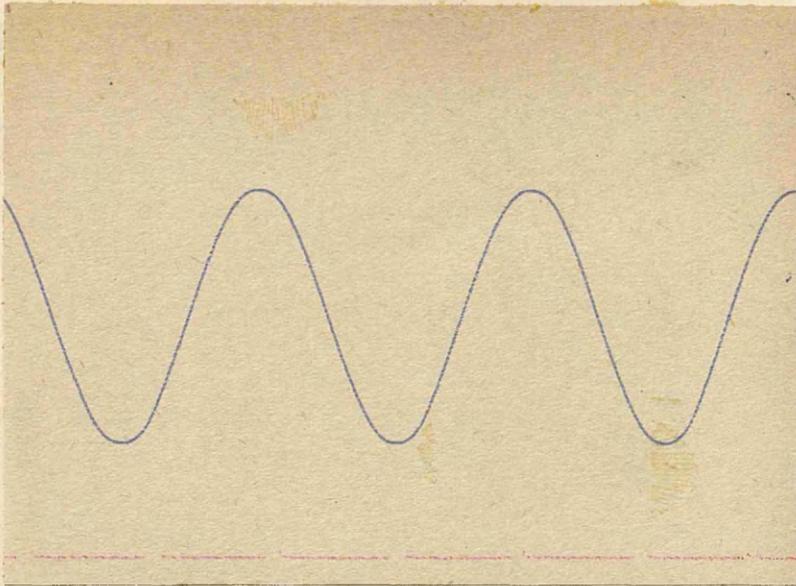
(a) Output.



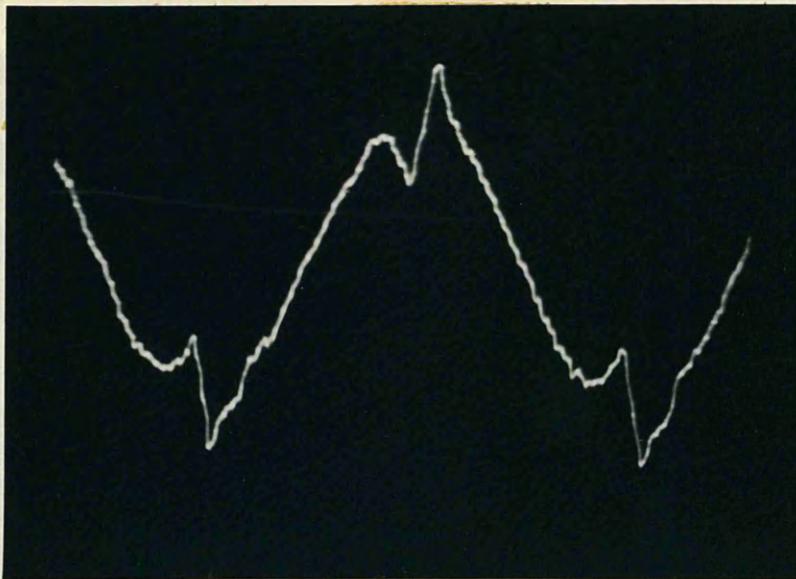
(b) Motor armature current.

Fig. 76. Point 1. Approx. lowest experimental reading.

$$\omega = 3.81, \quad \left| \frac{\theta_o(j\omega)}{\theta_i} \right| = 1.2.$$



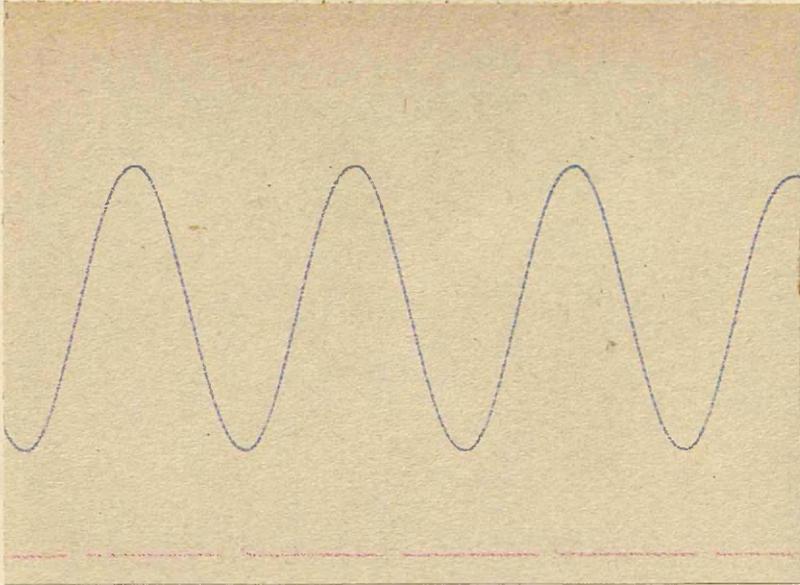
(a) Output.



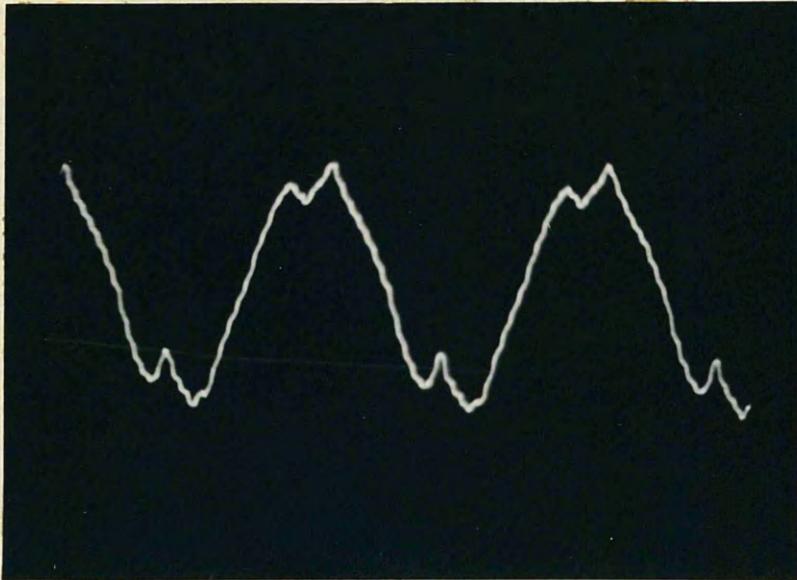
(b) Motor armature current.

Fig. 77. Point 2. Lowest frequency for reasonable fundamental sine component in motor armature current wave-form.

$$\omega = 5.66, \quad \left| \frac{e_o(j\omega)}{e_i} \right| = 1.35.$$



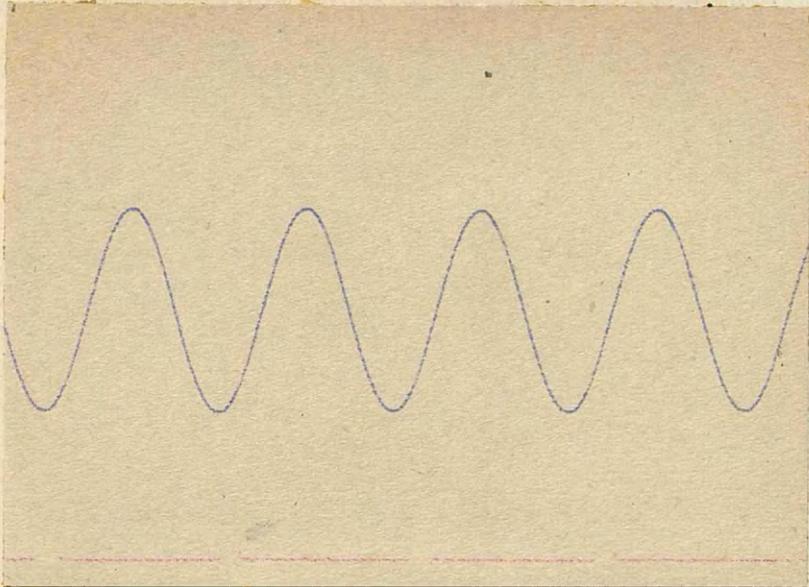
(a) Output.



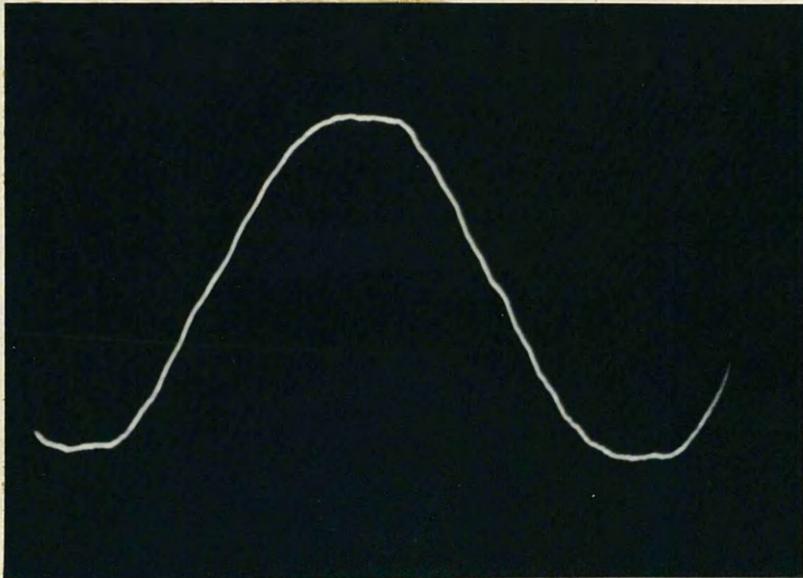
(b) Motor armature current.

Fig. 78. Point 3. Approx. resonant frequency.

$$\omega = 9.03, \quad \left| \frac{\theta_o}{\theta_i}(j\omega) \right| = 1.51.$$

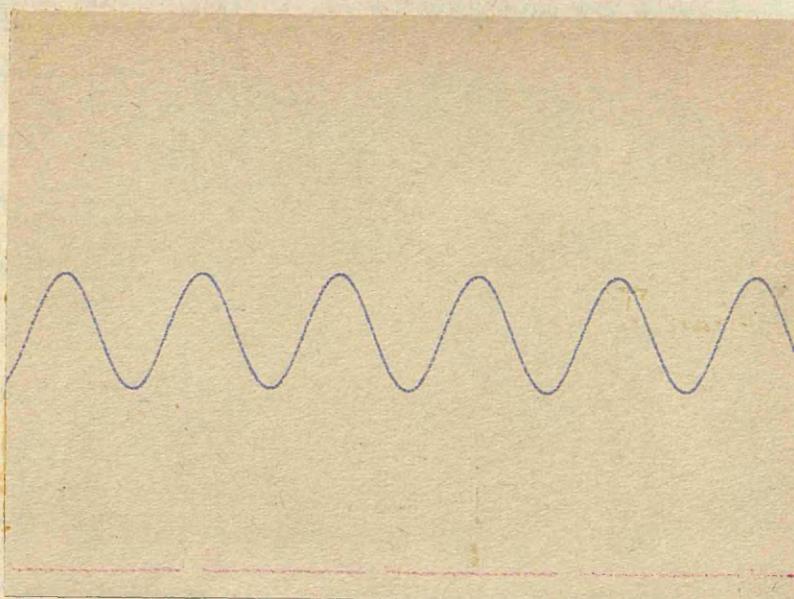


(a) Output.

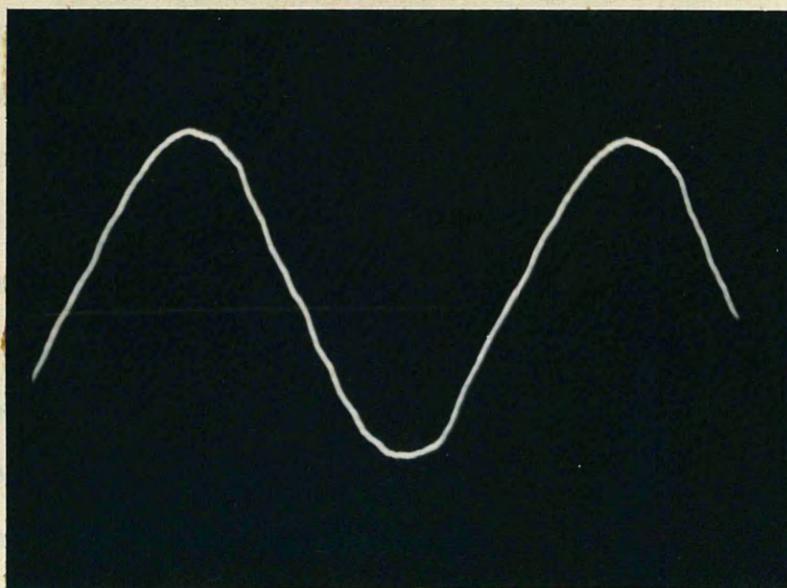


(b) Motor armature current.

Fig. 79. Point 4. Frequency at which $\left| \frac{\theta_o}{\theta_i}(j\omega) \right| \doteq 1$.
 $\omega = 13.37$, $\left| \frac{\theta_o}{\theta_i}(j\omega) \right| = 1.08$.



(a) Output.



(b) Motor armature current.

Fig. 80. Point 5. Frequency at which $\left| \frac{\theta_o}{\theta_i}(j\omega) \right| \doteq 0.5$.

$$\omega = 18.6, \quad \left| \frac{\theta_o}{\theta_i}(j\omega) \right| = 0.6.$$

The following observations may be made from the foregoing records and oscillograms.*

- (i) The output motion, as far as the eye can judge, is sinusoidal except at the lowest experimental point. At these points the waveform is not far off sinusoidal but shows flattening of the peaks. The development of these waveforms, i.e. Figs. 71a and 76a from the low frequency waveform Fig. 69 is apparent. This effect has already been discussed.
- (ii) The waveforms of motor armature current show quite clearly that near-sinusoidal operation of the system only results once approximately resonant-frequency has been reached. Above this frequency the waveforms are reasonably sinusoidal. The development of this sinusoidal waveform from the somewhat ragged but not entirely random variation of Fig. 71b proceeds in an orderly fashion.
- (iii) The waveforms for a control stiffness of 1.15 lb-ft/^o show good correlation with those for a control stiffness of 2.3 lb-ft/^o, where more "peakiness" results. At the lower frequencies, in the region represented by Figs. 72b and 77b, a reasonable prediction of the waveform is possible, from a consideration of the friction characteristic of the load, as shown below.

* Enlargements of film negatives taken on a Cossor Model 1089 Oscilloscope with single-sweep time base.

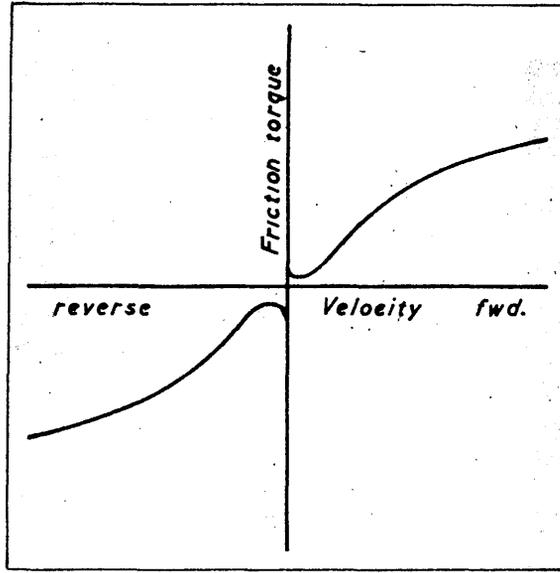


Fig. 81. Friction/speed characteristic of servo-mechanism load.

Consider the total driving torque required to move an inertia load sinusoidally against the above friction torque. Since the motor is operating with a fixed field current, the total torque waveform will also represent that of the motor armature current. The conditions are given in Fig. 82.

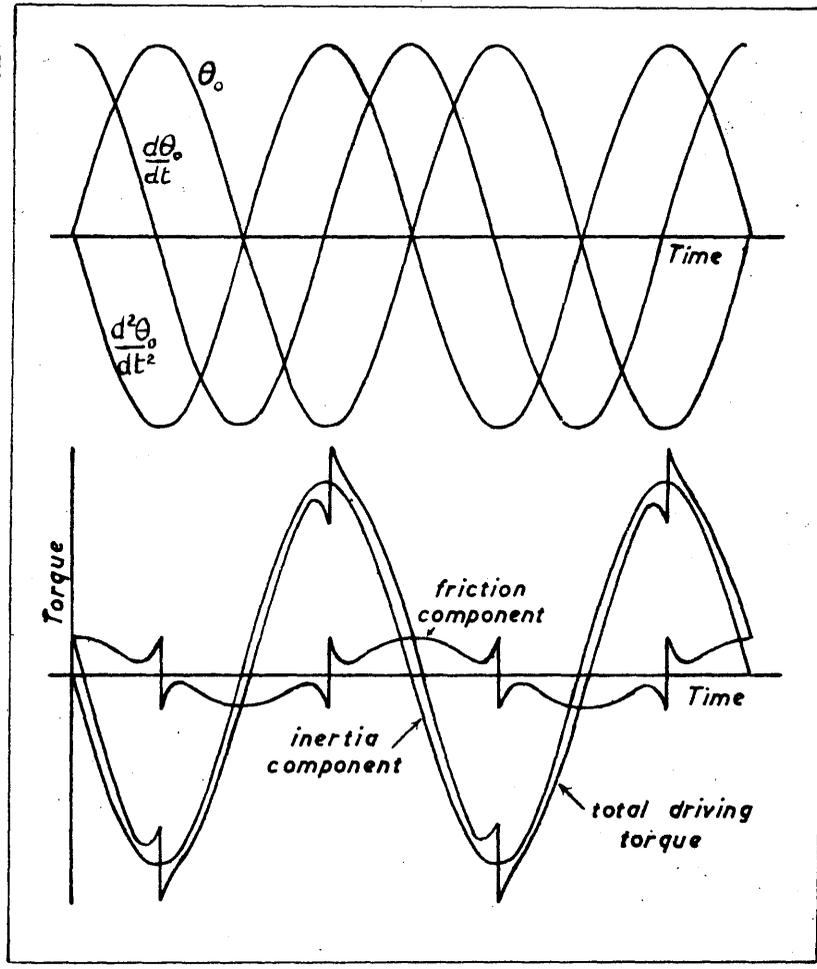


Fig. 82. Non-Sinusoidal torque for sinusoidal angular motion.

The upper diagram shows the sinusoidal displacement, velocity and acceleration of the load inertia. From these curves the acceleration and friction components of the driving torque may be drawn in, as the lower figure shows. The sum of the two represents the total torque and hence the armature current waveform. Reference to the oscillograms in Fig. 72b and 77b indicates that the actual waveform of motor current

possesses the predicted shape, any variation being only a question of the relative amounts of friction and acceleration components.

Likewise, the improvement in waveform as the frequency is raised is due to the acceleration torque increasing while the stiction torque remains essentially the same. Further the friction torque required at any velocity also becomes smaller in relation to the acceleration torque.* The result is obtained, therefore, that sinusoidal operation of the system is only taking place above a frequency which is approximately equal to the resonant frequency. For any system, however, it is possible to make the acceleration torque higher in relation to the stiction torque by running the input and output through larger amplitudes at any particular frequency, and in this manner sinusoidal operation may be extended to lower frequencies. Further discussion of this point is given in Sec. 13.2, which offers criticism of the experimental technique.

Linear range of control.

The linear range of the overall sensitivity characteristic, see Fig. 44, has already placed limits on the maximum input steps which could be used for transient response measurements. A similar restriction applies to the frequency response. The $\epsilon(j\omega)$ loci of Figs. 58, 60, and 62, enable this to be checked. The maximum errors for the responses I, II and III are, in

* The ratio $\frac{\text{max. velocity}}{\text{max. acceleration}}$ being inversely proportional to the frequency.

fact, 14.7° , 8.3° , and 9.3° which are all within the linear range of their respective sensitivity characteristics. Thus far the dynamic sensitivity characteristics of the system have not been mentioned. These relate to the motor torque per unit error for any given motor speed. More frequently, the information is given in terms of the torque/speed curves of the motor for given values of the excitation of the generator supplying the motor armature power. These curves can in turn be prepared from the output current/output voltage characteristics of the generator, in this case the metadyne generator. The result of steady load tests on the metadyne generator to determine these characteristics is shown in Fig. 83. In an ideal machine possessing no residual magnetism these would be symmetrical about the origin. The possible effect of curvature of the dynamic characteristics can be deduced from these curves. From the frequency response results, the maximum velocity attained by the motor was about 16 rad/sec. This corresponds to motor voltages of the order of 15V and therefore it is apparent that only a small portion of the curves of Fig. 83 on either side of zero voltage, was actually used. From this we may safely conclude that the effect of curved dynamic characteristics introduces negligible error.

Interaction with power-supply system.

This was not detectable during transient response measurements. A maximum variation of 5V in 220V occurred in the metadyne-generator driving-motor voltage, in the course of sinusoidal measurements.

The overall effect of stiction, non-linear friction, saturation and curvature of the dynamic control characteristics can be found from

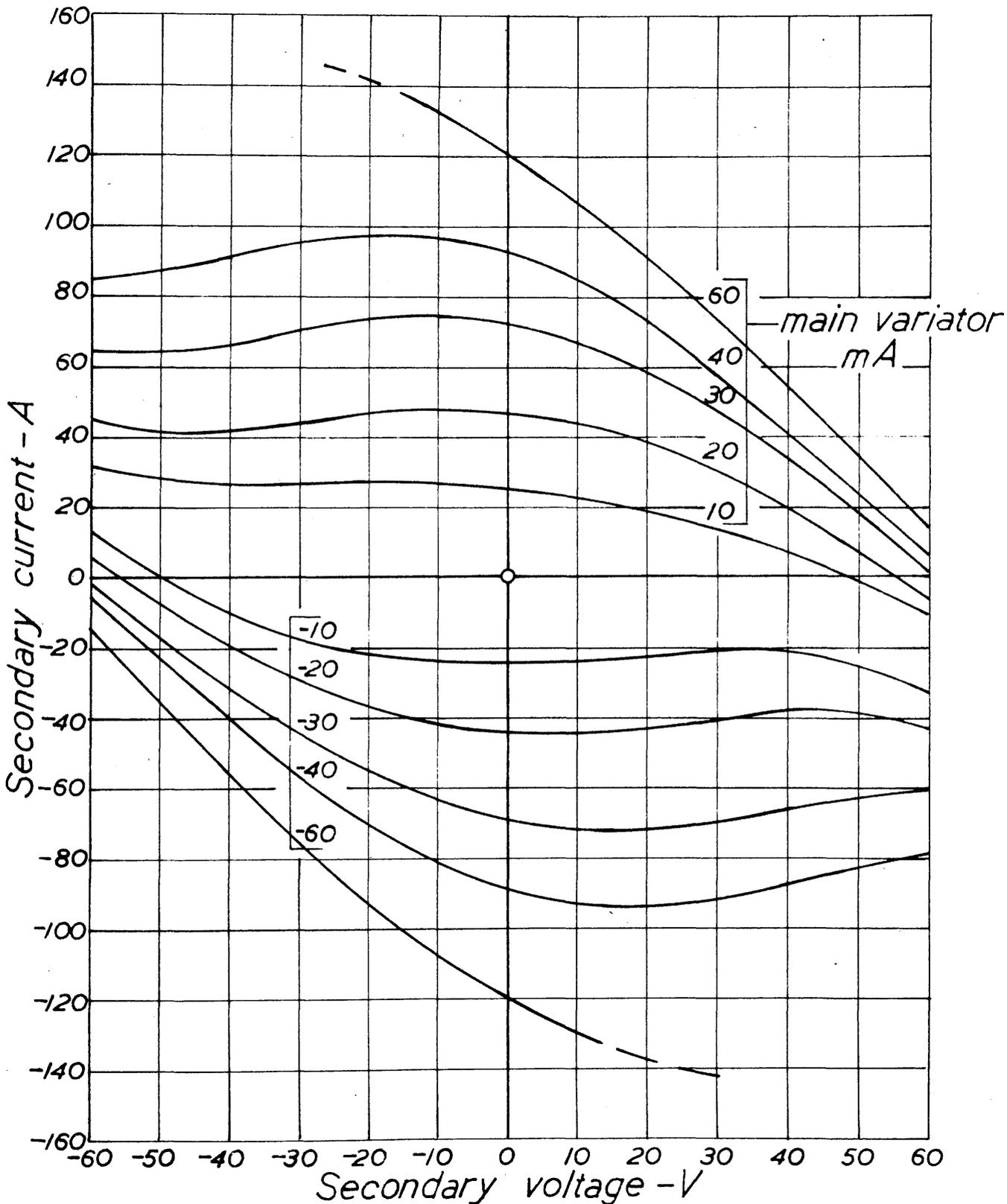


Fig. 83 Metadyne generator secondary current-secondary voltage characteristics

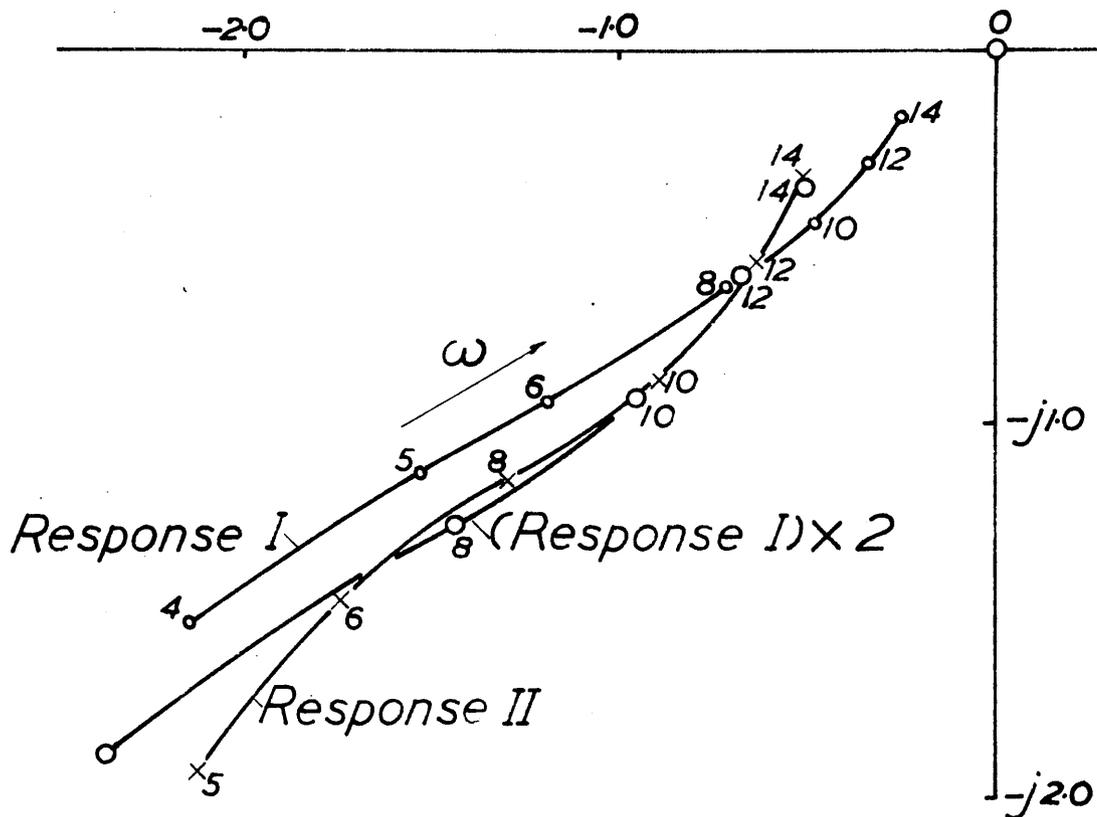


Fig. 84 Check of linear operation by comparison of $\theta/e(j\omega)$ loci.

consideration of two $\frac{G}{s}(j\omega)$ loci having known control stiffnesses. Thus since response II only differed from response I in having twice the control stiffness, the $\frac{G}{s}(j\omega)$ locus of I may be expanded radially in the Nyquist plane by a factor of 2, and compared with the $\frac{G}{s}(j\omega)$ locus of II. Fig. 84 demonstrates this and supplements the conclusion already drawn from considering the wave-forms, namely, that very nearly linear conditions obtain from $\omega = 8$ rad/sec. upwards.

(c) Reproducibility of system . Order of accuracy of experimental measurements.

In a complex system such as the one tested, the question of consistent results immediately arises. This has been checked thoroughly and it has been established that, given similar conditions prior to testing, e.g. the same running-in periods under steady sinusoidal motion, it is possible to produce transient responses which are practically identical. The maximum error in transient recording amounts in fact to no more than the very slight backlash between the motor shaft and the fine resetter shaft. This is of the order of 3° at the motor shaft and represents a possible discrepancy between the recorded angle and the actual angle of the fine resetter shaft, of $9/25^\circ$. For a 9° step, this represents a 4% error, and for a 20° step, about 2.5% error.

Not quite the same accuracy applies to the frequency responses, considerable error being possible at the higher frequencies with input motions of 5° . The error in observing the dial indication being about $\frac{1}{4}^\circ$, the least accuracy

is 10%, which occurs at the highest experimental point of responses II and III. The accuracy with larger output movements increases proportionally. Reference to the response curves shows it to vary from 5% to 2% for all but the above mentioned points. The uncertainty in measuring the phase-shift of $\frac{\theta_o}{\theta_i}(j\omega)$ has already been stated in Sec.10.2. This measurement gave very consistent results.

A criticism of the experimental technique in frequency response measurement is given in the next and concluding Chapter.

CHAPTER 13.

CONCLUSION. FURTHER WORK.

13.1. Conclusions from Experimental Results.

A high-accuracy, high-power servo-mechanism has been subjected to step inputs under conditions representing the nearest approach possible to linear operation. The system has also been subjected to steady sinusoidal inputs over a finite frequency range limited at the lower end by non-sinusoidal output movement, and at the higher end, by inaccurate error voltage indication, due to the small amplitude of output movement. Approximately linear operation took place over the top 70 - 75% of this frequency range.

The measured and calculated step responses show some measure of correlation but are most noticeably different in their times for maximum overshoot, with least discrepancy occurring when the degree of oscillation is greatest. The bulk of this error is due to assuming in the analysis that true sinusoidal conditions hold from zero frequency upwards. The assumed frequency response characteristics in the region of zero frequency, have in fact conferred higher values of the effective amplitude and smaller values of the effective phase-shift of $\frac{\theta_o(j\omega)}{\theta_i}$ than actually occurs. The term effective is used since we know quite definitely that **non-sinusoidal** conditions exist; (the effective amplitude and phase-shift might, however, be considered as the amplitude and phase-shift resulting if only the fundamental component of the output were taken). These

assumed characteristics result in the calculated step-response being faster than in practice. We are able to attribute the greater part of the discrepancy to this assumption, since measurements made in the frequency range where the output motion, but not the motor armature current, was sinusoidal, are reasonably accurate. The above assumption also accounts for the better agreement for an oscillatory response than for a well-damped one. For instance, in the harmonic analysis method this results from the increase in the resonance peak in comparison with the remainder of the response. Since true sinusoidal conditions have in fact been obtained at resonant frequencies the calculated response will be more nearly correct. Campbell's method will, of course, give the best results when the response is fairly oscillatory, due to the assumption made in the method that the principal mode itself accounts for the total overshoot.

13.2. Criticism of Technique of Frequency Response Measurement.

In the method employed in the tests, a fixed amplitude of input oscillation was used for any particular curve being determined. This resulted in the accelerating torque required being comparable with the friction / stiction torque at low frequencies. If the input amplitude is raised however as the frequency is lowered, then it has been shown that better sinusoidal conditions result. Brown and Campbell⁴ give a method which will accomplish this. Trouble will, of course, arise if the increased velocities carry the motor operation into a curved region of its torque/speed characteristics, or into a region in which they are not equi-spaced. This happens fairly soon in the case of a metadyne-generator type servo such as

the one tested. The method used is, nevertheless, criticised on this point.

13.3. Analytical Techniques.

All methods give essentially the same predicted step response, as the underlying assumption of sinusoidal conditions from zero frequency upwards is made in each case. It is encouraging, however, that in this case where nothing is known of the type of roots of the system characteristic equation, Campbell's very approximate method gives agreement with the other two. From this it may be expected that the principal mode theory of Part II of this thesis will provide reasonable results in practice given approximate linear conditions. It represents the same type of approximation as Campbell's method, but one which does take into account the remaining roots of the system characteristic equation.

13.4. Generalisation of Conclusions.

The object of the investigation was to determine to what extent quantitative prediction of a system transient response was possible from knowledge of a measured frequency response. The results obtained for the present system have already been given and the error in prediction has been discussed.

The question now arises of whether a general statement may be made on the evidence of the above investigation. While it would be unreasonable to claim a certain limit of prediction for all systems, some indication is possible of the maximum error in prediction which is likely to occur if correct experimental technique is employed. This requires (a) reasonable

sine-wave conditions throughout the whole system to as low a frequency as possible, and (b) measurements of high accuracy at low frequencies and at the resonant frequencies. Loss of accuracy at frequencies greater than that at which the response decreases to 0.5 say, may be tolerated.

Under these conditions the maximum overshoot ought to be predictable to within 10% of the input step and with greater accuracy for greater degrees of oscillation. The time for maximum overshoot should be calculable to within 25% of the true value and will certainly be so for oscillatory responses. The results of the investigation do not warrant any closer limits being set. It is, however, in the case of relatively complex high-accuracy systems of considerable power output that these maximum divergences of calculated and actual step responses will occur. Medium- and low-power servo-mechanisms, excepting the inherently non-linear "on-off" types, will almost certainly give better agreement. Similarly a frequency response design on paper is not likely to afford a closer guide to the transient response which will occur in practice, than the above figures indicate. It is nevertheless reassuring that the maximum overshoot in practice will be less than the linear prediction theory suggests.* On the other hand, the response-time in practice will exceed the value predicted by calculation.

13.5. Further Work.

The greater part of the work relating to servo-mechanisms which has

* The measured response III admittedly shows a greater overshoot than the predicted response. As however the 4% difference could very well result from the inaccuracy in measurement and calculation, the two overshoots are for practical purposes the same.

been done in the past has assumed linear conditions of operation, and in proportion to this amount of work, few papers have dealt with non-linear effects.^{42,43,46} Recently, however, particulars relating to specific non-linear systems and also reviews of the non-linear oscillation problem as a whole have been appearing.* It is clear at the outset that very little generalisation of the non-linear problem can be obtained. The problem has been and is being tackled from a more practical viewpoint, by considering separately each type of non-linearity as it occurs, e.g. servos with torque limitation, backlash, non-linear friction and curvature of motor torque/speed characteristics. From any one of these investigations the generalisation of results is in itself difficult and much work still remains to be done to sort out the conclusions reached by different experiments on the subject. The qualitative effect is relatively easy to establish but non-dimensional presentation of quantitative results is hindered by lack of uniformity of experimental conditions. The clearing-up and unification of the results for each type of non-linearity on its own is a practical problem of immediate importance.

Further, more consideration should be given to fundamentally non-linear systems of the "on-off" type. The same may be also said of a.c. operated systems the full development of which has been delayed on account of technical difficulties of control and also on account of the added complexity of theory with a carrier-wave present in the sequence.

* See references 44, 45, 50-57.

In addition to these essentially practical problems, there is still room for further theoretical developments. Even in the linear theory, the question of multi-loop systems having more than one input, intended or otherwise, can be explored. Related to this is the problem of interaction with the power-supply and consideration of the whole system in terms of energy flow. The technique of tensor analysis has still to be usefully applied to such multi-loop systems.

The interesting problem of optimisation and the bearing of information theory upon practical servo-mechanism designs also offers much scope to theoretical workers. Just exactly to what extent practical servo-mechanism designs may hope to profit from this aspect is not obvious. At present this refinement seems less important than the production of reliable results regarding the basic non-linearities which occur in every practical system. Together with a study of "on-off" systems, this is, in the author's opinion, the most pressing objective at the present time. It is the author's intention to proceed to this work.

BIBLIOGRAPHY.

The following bibliography is not intended to be exhaustive. This has been decided on in view of the amount of literature on the subject and since there are already fairly complete bibliographies existing in certain text-books. These books have been included in the bibliography and also the Proceedings of several Conferences on Automatic Control generally. Apart from these works, the list contains only those references which occur in the text of the thesis.

BOOKS

1. BODE, H.W. Network Analysis and Feedback Amplifier Design.
Van Nostrand. 1945.
2. JAMES, H.M., NICHOLS, N.B., and PHILLIPS, R.S.
Theory of Servo-mechanisms.
McGraw-Hill. 1947.
3. SMITH, E.S. Automatic Control Engineering.
McGraw-Hill. 1944
4. BROWN, G.S. and CAMPBELL, D.P.
Principles of Servomechanisms.
Wiley. 1948.
5. CHESTNUT, H. and MAYER, R.W.
Servomechanisms and Regulating System Design. Vol.1.
Wiley. 1951.
6. PORTER, A. An Introduction to Servo-Mechanisms.
Methuen. 1950.
7. MACMILLAN, R.H. An Introduction to the Theory of Control in Mechanical Engineering.
Cambridge University Press. 1951.
8. LEONHARD, A. Die selbsttatige Regelung.
Springer. 1949.

9. OPPELT, W. Grundgesetze der Regelung.
Wolfenbutteler Verlagsanstalt. 1947
10. OPPELT, W. Stetige Regelvorgänge.
Wissenschaftliche Verlagsanstalt K.G. Hanover. 1949.

References 3 to 7 contain bibliographies, in the main, of the British and American contributions to the subject. Items 8 to 10 list German work.

11. GARDNER, M.F. and BARNES, J.L.
Transients in Linear Systems.
Wiley. 1942.
12. CARSLAW, H.S. and JAEGER, J.C.
Operational Methods in Applied Mathematics.
Clarendon Press. 1941.
13. JAEGER, J.C. An Introduction to the Laplace Transformation.
Methuen. 1949.
14. GUILLEMIN, E.A. Communication Networks. Vols. I and II.
Wiley. 1931.

PROCEEDINGS OF CONFERENCES ON AUTOMATIC CONTROL.

15. Society of Chemical Industry, Chemical Engineering Group.
Proc. Soc. Chem. Ind. 18. 1936.
16. Joint Conference of Society of Chemical Industry, Institute of
Physics and Institution of Chemical Engineers.
Proc. Soc. Chem. Ind. 27. 1945.
17. Convention of Institution of Electrical Engineers.
Proceedings, 94, IIA, 1 and 2. May 1947.

PAPERS AND ARTICLES.

18. FOSTER, E.W. and LUDBROOK, L.C.
Some Industrial Electronic Servo and Regulator Systems.
J. Inst. Elect. Engrs. vol. 94, IIA, 1, p.100. 1947.
19. LIU, Y.J. Charts for Verifying the Stability of Servo-mechanisms.
M.I.T. Dept. of Elec. Eng. 1941.
20. EVANS, L.W. Solution of the Cubic Equation and the Cubic Charts.
M.I.T. Dept. of Elec. Eng. 1943.
21. WHITELEY, A.L. Theory of Servo-mechanisms with Particular Reference
to Stabilisation.
J. Inst. Elect. Engrs. vol. 93, II, p.353. 1946.

22. ROUTH, E.J. On the Stability of Motion (Adams Prize Essay).
Macmillan. 1877.
23. HURWITZ, A. Ueber die Bedingungen, unter welchen eine Gleichung
nur Wurzeln mit negativen reellen Theilen besitzt.
Math. Ann. vol. 46, p.273. 1895.
24. NYQUIST, H. Regeneration Theory.
Bell Syst. Tech. J. vol. 11, p. 126. 1932.
25. LÜTHI, A. Fading Conditions for Regulator Equations of Any
Order.
Escher Wyss News vol. 15/16. 1942/1943.
26. LEONHARD, A. Neues Verfahren zur Stabilitätsuntersuchung.
Arch. Elekt. vol. 38, p. 17. 1944.
27. LEONHARD, A. Stabilitätskriterium insbesondere von Regelkreisen
bei vorgeschriebener Stabilitätsgute.
Arch. Elekt. vol. 39, p. 100. 1948.
28. DEMONTVIGNIER, M. and LEFÈVRE, P.
Une nouvelle methode harmonique d'etude de la
stabilite des systemes lineaires.
Rev. Gen. Elect. vol. 58, p.263. 1949.
29. GÖRK, E. Stabilitätskriterium.
Arch. Elekt. Ubertrag. vol.4, p. 89. 1950.
30. VAZSONI, A. A Generalisation of Nyquist's Stability Criteria.
J. App. Phys. vol. 20, p. 863. 1949.
31. GROSSMANN, H.H. Elementare Begründung und Verscharfung des
Hurwitzschen Stabilitätskriteriums.
Schweizer Arch. angew. Wiss. Tech. vol.14, p.242.1948.
32. SHERMAN, S. A Note on Stability Calculations and Time Lag.
Quart.App. Math. vol. 5, p.92, 1947.
33. MINORSKY, N. Self-excited Oscillations in Dynamical Systems
possessing Retarded Actions.
Trans. Amer. Soc. Mech. Engrs. (J. App. Mech.),
vol. 63. 1941.

34. ANSOFF, H.I. and KRUMHANSL, J.A.
A General Stability Criterion for Linear Oscillating Systems with Constant Time Lag.
Quart. App. Math. vol. 6, p. 337. 1948.
35. CAMPBELL, W.W. Use of Nyquist Diagram.
Min. of Supply Servo Lib. Ref. G74. 1945.
36. PROFOS, P. A New Method for the Treatment of Regulation Problems.
Sulzer Tech. Review. No. 2, p.1. 1945.
37. EVANS, W.R. Control System Synthesis by Root Locus Method.
Trans. Amer. Inst. Elect. Engrs. vol. 69, p. 66, 1950.
38. CHESTNUT, H. and MAYER, R.W.
Comparison of Steady State and Transient Performance of Servo-mechanisms.
Trans. Amer. Inst. Elect. Engrs. vol. 68, p.765, 1949.
39. BEDFORD, A.V. and FREDENDALL, G.L.
Transient Response of Television Amplifiers.
Proc. Inst. Radio Engrs. N.Y. vol. 30, p. 440. 1942.
40. FLOYD, G.F. Method for Approximating Transient Response from Frequency Response.
See reference 4, chap. 11.
41. MULLIGAN, J.R. The Effect of Pole and Zero Locations on the Transient Response of Linear Dynamic Systems.
Proc. Inst. Radio Engrs. N.Y. vol. 37, p.516. 1949.
42. TUSTIN, A. The Effects of Backlash and Speed Dependent Friction on the Stability of Closed-Cycle Control Systems.
J. Inst. Elect. Engrs. vol. 94, IIA, 1, p. 143. 1947.
43. TUSTIN, A. A Method of Analysing the Effect of Certain Kinds of Non-Linearity in Closed-Cycle Control Systems.
J. Inst. Elect. Engrs. vol. 94, IIA, 1, p. 152. 1947.
44. MICHEL, J.G.L. and PORTER, A.
The Effect of Friction on the Behaviour of Servo-mechanisms at Creep Speeds.
Proc. Inst. Elect. Engrs. vol.98, II, p. 297. 1951
45. HSIA, Miss Pei-Su. A Graphical Analysis for Non-Linear Systems.
Proc. Inst. Elect. Engrs. (to be published).
46. WILLIAMS, F.C. and RITSON, F.J.U.
Electronic Servo Simulators.
J. Inst. Elect. Engrs. vol. 94, IIA, 1, p. 112, 1947.

The following are some of the papers presented at the Conference on Automatic Control organised in July 1951 by the Department of Scientific and Industrial Research. The bound Proceedings has, as yet, not been published.

47. DZUNG, L.S. The Stability Criterion.
48. LEONHARD, A. Relative Damping as a Criterion for Stability.
49. KUSTERS, N.L. and MOORE, W.J.M.
A Generalisation of the Frequency Response Method
for the Study of Feedback Control Systems.
50. SCOTT, W.E. An Introduction to the Analysis of Non-Linear
Closed Cycle Control Systems.
51. CHERRY, E.C. Some Theorems in Generalised Mechanics relevant to
Non-Linear Systems.
52. LOEB, J. A General Linearising Process for Non-Linear
Control Systems.
- 53.UTTLEY, A.M. and HAMMOND, P.H.
The Stabilising of On-Off Control Servo-mechanisms.
54. MINORSKY, N. Review of Non-Linear Phenomena, and Their Potential
Possibilities in Control Systems.
55. JONES, R.W. Stability Criteria for Certain Non-Linear Systems.
56. HARMER, J.D. The Jerking Motion Caused by Static Friction in
Position Control Systems.
57. LIVERSIDGE, J.H. Frequency Analysis Applied to Devices with
Backlash and Resilience.

APPENDIX I.

THE EXPONENTIAL FOURIER SERIES FOR A PERIODIC FUNCTION.

The trigonometrical form of the Fourier Series for a function

$p(t)$, periodic in time $T = \frac{2\pi}{\omega}$, is

$$p(t) = \frac{a_0}{T} + \frac{2}{T} \sum_{n=1}^{\infty} (a_n \cos n\omega t + b_n \sin n\omega t) \quad (1a)$$

where

$$a_0 = \int_{-T/2}^{T/2} p(t) dt, \quad (2)$$

$$a_n = \int_{-T/2}^{T/2} p(t) \cos n\omega t dt, \quad (3)$$

$$b_n = \int_{-T/2}^{T/2} p(t) \sin n\omega t dt. \quad (4)$$

Putting $\cos n\omega t = \frac{1}{2} (\varepsilon^{jn\omega t} + \varepsilon^{-jn\omega t})$, $\sin n\omega t = \frac{1}{2j} (\varepsilon^{jn\omega t} - \varepsilon^{-jn\omega t})$ in (1a)

we have

$$p(t) = \frac{a_0}{T} + \frac{1}{T} \sum_{n=1}^{\infty} [(a_n - jb_n) \varepsilon^{jn\omega t} + (a_n + jb_n) \varepsilon^{-jn\omega t}]. \quad (1b)$$

From (3) and (4)

$$a_n - jb_n = \int_{-T/2}^{T/2} p(t) \varepsilon^{-jn\omega t} dt = P(n\omega) \quad (5)$$

$$a_n + jb_n = \int_{-T/2}^{T/2} p(t) \varepsilon^{jn\omega t} dt = P(-n\omega) \quad (6)$$

Putting $n=0$ in (5) or (6), also

$$P(n\omega)_{n=0} = \int_{-T/2}^{T/2} p(t) dt = a_0 \quad (7)$$

Thus, substituting in (1b)

$$p(t) = \frac{1}{T} \sum_{n=-\infty}^{n=+\infty} P(n\omega) \varepsilon^{jn\omega t}. \quad (1c)$$

This is the exponential form of the Fourier Series for a period function.

The coefficient $P(n\omega)$ is given by

$$P(n\omega) = \int_{-T/2}^{T/2} p(t) \varepsilon^{-jn\omega t} dt \quad (5)$$

The n^{th} frequency component in (1a) is

$$\frac{2}{T} (a_n \cos n\omega t + b_n \sin n\omega t) = \frac{2}{T} r_n \cos (n\omega t - \phi_n) \quad , \quad \text{where}$$

$r_n^2 = a_n^2 + b_n^2 = |P(n\omega)|^2$ and $\phi_n = \tan^{-1} \frac{b_n}{a_n}$. Hence the coefficients $P(n\omega)$ represent the relative amplitudes and phases of the frequency components, the actual amplitudes being $\frac{2}{T} |P(n\omega)|$. $|P(n\omega)|$ and ϕ_n define two real line spectra, the amplitude and phase spectra. Both are contained in the complex coefficient $P(n\omega)$, which we note from (5) and (6) also has the property $P(n\omega) = \text{conjugate. } P(-n\omega)$.

Extension of Range of Periodic Function.

By allowing the periodic time T to increase indefinitely, it becomes possible to represent a non-periodic function $f(t)$. This process requires that the fundamental period ω become infinitesimal and the frequency component $n\omega$ become the continuous variable ω . Thus,

writing (1c) as

$$p(t) = \frac{1}{2\pi} \sum_{n=-\infty}^{n=+\infty} P(n\omega) \varepsilon^{jn\omega t} \cdot \omega,$$

in the limit $\left. \begin{array}{l} T \rightarrow \infty \\ n\omega \rightarrow \omega \\ \omega \rightarrow d\omega \end{array} \right\}$, this becomes

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(\omega) \varepsilon^{j\omega t} d\omega, \quad (8)$$

where, from (5),

$$F(\omega) = \int_{-\infty}^{\infty} f(t) \varepsilon^{-j\omega t} dt. \quad (9)$$

Equations (8) and (9) express the inverse and direct Fourier Transformation. Strictly speaking several conditions are required to be satisfied by $f(t)$ in order that its Fourier Transform $F(\omega)$ should exist. The most important of these is that

should be finite. $\int_{-\infty}^{\infty} |f(t)| dt$

APPENDIX II.

TABLES TO ASSIST IN THE CALCULATION OF NYQUIST DIAGRAMS AT REAL AND COMPLEX FREQUENCIES.

Each of the Tables I to V on the following pages refers to one value of relative damping, defined by the angle β as explained in Sec. 5.1; the values of β corresponding to Tables I to V are 0° , 30° , 45° , 60° and 75° .

Table I gives the factor $(1+jx)^{\pm 1}$ in polar form. Tables II to V show the values of (a) $(1+x(-\sigma+j))^{\pm 1}$ in polar form and (b) $x(-\sigma+j)$, $x^2(-\sigma+j)^2$ in co-ordinate form; the range of x is $0.10 < x < 10$. With the aid of these Tables, factors such as $(1+pT)^{\pm 1}$, $(1+ap+bp^2)^{\pm 1}$ may be evaluated for values of $p = u(-\sigma+j)$, $\sigma = \tan \beta$.

TABLE I. $\beta = 0^\circ$.

x	$(1+jx)^{\pm 1} = R^{\pm 1} \angle \pm \phi^\circ$		
	R	R^{-1}	ϕ°
0.1	1.005	0.995	5.7
0.2	1.02	0.980	11.3
0.3	1.04	0.962	16.8
0.4	1.08	0.926	21.7
0.5	1.12	0.893	26.5
0.6	1.17	0.855	30.8
0.8	1.28	0.781	38.6
1.0	1.41	0.707	45.0
1.2	1.56	0.641	50.4
1.6	1.89	0.529	57.8
2.0	2.24	0.446	63.3
3.0	3.16	0.316	71.8
4.0	4.12	0.243	76.4
5.0	5.10	0.196	78.5
6.0	6.08	0.164	80.8
8.0	8.06	0.124	82.8
10.0	10.05	0.100	84.3

TABLE II. $\beta = 30^\circ$

x	$(1+x[-\sigma+j])^{\pm 1} = R^{\pm 1} \angle \pm \phi^\circ$			$x(-\sigma+j)$		$x^2(-\sigma+j)^2$	
	R	R^{-1}	ϕ°	Real	Imag.	Real	Imag.
0.1	0.947	1.06	6.1	-0.058	0.1	-0.007	-0.012
0.2	0.913	1.10	12.7	-0.115	0.200	-0.027	-0.046
0.3	0.88	1.14	19.9	-0.173	0.3	-0.061	-0.104
0.4	0.866	1.15	27.5	-0.231	0.4	-0.107	-0.184
0.5	0.869	1.15	35.2	-0.288	0.5	-0.167	-0.288
0.6	0.888	1.13	42.5	-0.346	0.6	-0.240	-0.414
0.8	0.964	1.04	56.1	-0.462	0.8	-0.427	-0.736
1.0	1.09	0.917	66.5	-0.577	1.0	-0.667	-1.15
1.2	1.24	0.806	75.5	-0.692	1.2	-0.961	-1.66
1.6	1.6	0.625	90.	-0.924	1.6	-1.71	-2.94
2.0	2.00	0.500	94.3	-1.15	2.0	-2.67	-4.60
2.5	2.54	0.394	100.	-1.44	2.5	-4.17	-7.19
3.0	3.09	0.324	103.7	-1.73	3.00	-6.05	-10.4
4.0	4.21	0.238	108.1	-2.31	4.00	-10.7	-18.4
5.0	5.35	0.187	110.7	-2.88	5.0	-16.7	-28.8
6.0	6.40	0.156	112.6	-3.46	6.0	-24.0	-41.4
8.0	8.78	0.114	114.3	-4.62	8.0	-42.7	-73.6
10.0	11.09	0.090	115.5	-5.77	10.0	-66.7	-115.0

TABLE III. $\beta = 45^\circ$

x	$(1 + x[-\sigma + j])^{\pm 1} = R^{\pm 1} \angle \pm \phi^\circ$			$x(-\sigma + j)$		$x^2(-\sigma + j)^2$	
	R	R^{-1}	ϕ°	Real	Imag.	Real	Imag.
0.1	0.906	1.10	6.3	-0.1	0.1	0.	-0.02
0.2	0.825	1.21	14.0	-0.2	0.2	0	-0.08
0.3	0.762	1.31	23.2	-0.3	0.3	0	-0.18
0.4	0.721	1.39	33.7	-0.4	0.4	0	-0.32
0.5	0.707	1.41	45.0	-0.5	0.5	0	-0.50
0.6	0.721	1.39	56.4	-0.6	0.6	0	-0.72
0.8	0.825	1.21	75.9	-0.8	0.8	0	-1.28
1.0	1.00	1.00	90.0	-1.0	1.0	0	-2.00
1.2	1.22	0.82	99.4	-1.2	1.2	0	-2.88
1.6	1.71	0.585	110.5	-1.6	1.6	0	-5.12
2.0	2.24	0.446	116.5	-2.0	2.0	0	-8.0
2.5	3.08	0.325	119.1	-2.5	2.5	0	-12.5
3.0	3.61	0.277	123.7	-3.0	3.0	0	-18.
4.0	5.00	0.200	126.9	-4.0	4.0	0	-32
5.0	6.40	0.156	128.7	-5.0	5.0	0	-50
6.0	7.81	0.128	129.8	-6.0	6.0	0	-72
8.0	10.63	0.094	131.1	-8.0	8.0	0	-128
10.0	13.45	0.074	132.0	-10.0	10.0	0	-200

TABLE IV. $\beta = 60^\circ$

x	$(1+x[-\sigma+j])^{\pm 1} = R^{\pm 1} \angle \pm \phi^\circ$			$x(-\sigma+j)$		$x^2(-\sigma+j)^2$	
	R	R^{-1}	ϕ°	Real	Imag.	Real	Imag.
0.1	0.833	1.20	6.9	-0.173	0.1	0.02	-0.035
0.2	0.684	1.46	17.0	-0.346	0.2	0.08	-0.138
0.3	0.566	1.77	32.0	-0.520	0.3	0.18	-0.312
0.4	0.505	1.98	52.4	-0.693	0.4	0.32	-0.554
0.5	0.518	1.93	74.8	-0.866	0.5	0.50	-0.867
0.6	0.602	1.66	93.8	-1.04	0.6	0.72	-1.25
0.8	0.89	1.12	116.0	-1.39	0.8	1.28	-2.22
1.0	1.24	0.806	126.2	-1.73	1.0	2.00	-3.46
1.2	1.62	0.617	131.9	-2.08	1.2	2.88	-4.98
1.6	2.39	0.418	137.8	-2.77	1.6	5.12	-8.87
2.0	3.18	0.314	140.7	-3.46	2.0	8.0	-13.9
2.5	4.16	0.24	143.1	-4.33	2.5	12.5	-21.6
3.0	5.16	0.194	144.5	-5.20	3.0	18.0	-31.2
4.0	7.16	0.14	146.0	-6.93	4.0	32.0	-55.4
5.0	9.15	0.109	146.8	-8.66	5.0	50.0	-86.7
6.0	11.2	0.090	147.6	-10.4	6.0	72.0	-125
8.0	15.1	0.066	148.3	-13.9	8.0	128.0	-222
10.0	19.1	0.052	148.9	-17.3	10.0	200	-346

TABLE V. $\beta = 75^\circ$

x	$(1 + x[-\sigma + j]) = R^{ \pm\phi^\circ}$			$x(-\sigma + j)$		$x^2(-\sigma + j)^2$	
	R	R^{-1}	ϕ°	Real	Imag.	Real	Imag.
0.1	0.635	1.57	9.1	-0.373	0.1	1.293	-0.746
0.2	0.324	3.09	38.1	-0.746	0.2	2.59	-1.49
0.3	0.322	3.11	111.8	-1.12	0.3	3.88	-2.23
0.4	0.632	1.58	140.8	-1.49	0.4	5.18	-2.98
0.5	1.00	1.00	150.5	-1.87	0.5	6.47	-3.73
0.6	1.38	0.725	153.9	-2.24	0.6	7.76	-4.46
0.8	2.14	0.467	157.7	-2.98	0.8	10.36	-5.96
1.0	2.91	0.344	159.7	-3.73	1.0	12.9	-7.46
1.2	3.67	0.272	161.5	-4.48	1.2	15.5	-8.95
1.6	5.22	0.192	162.4	-5.96	1.6	20.7	-11.9
2.0	6.76	0.148	162.9	-7.46	2.0	25.9	-14.9
2.5	8.70	0.115	163.3	-9.33	2.5	32.4	-18.7
3.0	10.6	0.094	163.5	-11.2	3.0	38.8	-22.4
4.0	14.5	0.069	163.5	-14.9	4.0	51.7	-29.8
5.0	18.4	0.054	164.2	-18.7	5.0	64.6	-37.3
6.0	22.2	0.045	164.6	-22.4	6.0	77.6	-44.6
8.0	29.9	0.033	164.6	-29.8	8.0	104	-59.6
10.0	37.7	0.026	164.6	-37.3	10.0	129	-74.6

APPENDIX III.

TABLES OF $\frac{1}{n} \sin n\theta$ AND $\frac{1}{n} \cos n\theta$ FOR n ODD UP TO $n = 21$

(Intervals of 10° in θ for $0^\circ < \theta < 30^\circ$, and 15° in θ for $30^\circ < \theta < 90^\circ$)

n	0	10°	20°	30°	45°	60°	75°	90°
1	0	0.174	0.342	0.500	0.707	0.866	0.966	1.000
3	0	0.167	0.289	0.333	0.236	0	-0.236	-0.333
5	0	0.153	0.197	0.100	-0.141	-0.173	0.052	0.200
7	0	0.134	0.092	-0.071	-0.101	0.124	0.037	-0.143
9	0	0.111	0	-0.111	0.079	0	-0.079	0.111
11	0	0.085	-0.058	-0.045	0.064	-0.079	0.088	-0.091
13	0	0.059	-0.076	0.038	-0.054	0.067	-0.074	0.077
15	0	0.033	-0.058	0.067	-0.047	0	0.047	-0.067
17	0	0.010	-0.020	0.029	0.041	-0.051	-0.015	0.059
19	0	-0.009	0.018	-0.026	0.037	0.046	-0.013	-0.052
21	0	-0.024	0.041	-0.047	-0.034	0	0.034	0.047

For $90^\circ < \theta < 180^\circ$, $\frac{1}{n} \sin n(90+\phi) = \frac{1}{n} \sin n(90-\phi)$.

Table of $\frac{1}{n} \sin n\theta$

n	0	10°	20°	30°	45°	60°	75°	90°
1	1.000	0.985	0.940	0.866	0.707	0.500	0.259	0
3	0.333	0.289	0.167	0	-0.236	-0.333	-0.236	0
5	0.200	0.129	-0.035	-0.173	-0.141	0.100	0.193	0
7	0.143	0.049	-0.109	-0.124	0.101	0.071	-0.138	0
9	0.111	0	-0.111	0	0.079	-0.111	0.079	0
11	0.091	-0.031	-0.070	0.079	-0.064	0.045	-0.023	0
13	0.077	-0.049	-0.013	0.067	-0.054	0.038	-0.020	0
15	0.067	-0.058	0.033	0	0.047	-0.067	0.047	0
17	0.059	-0.058	0.055	-0.051	0.041	0.029	-0.057	0
19	0.053	-0.052	0.049	-0.045	-0.037	0.027	0.051	0
21	0.048	-0.041	0.024	0	-0.034	-0.047	-0.034	0

For $90^\circ < \theta < 180^\circ$, $\frac{1}{n} \cos n(90 + \phi) = -\frac{1}{n} \cos n(90 - \phi)$.

Table of $\frac{1}{n} \cos n\theta$

APPENDIX IV.

HARMONIC ANALYSIS METHOD OF OBTAINING STEP RESPONSE FROM FREQUENCY RESPONSE.

Fundamental period = 7 sec. $\omega = 0.897$

From response I, Fig. 45,

n	1	3	5	7	9	11	13	15	17	19	21
ω	0.897	2.69	4.49	6.28	8.07	9.87	11.67	13.46	15.27	17.05	18.85
R	1.01	1.20	1.47	1.57	1.33	0.95	0.67	0.46	0.30	0.18	0.10
ϕ	-4°	-12°	-22°	-45°	-73°	-95°	-111°	-127°	-138°	-147°	-154°
a_n	1.007	1.173	1.362	1.11	0.589	-0.083	-0.24	-0.277	-0.223	-0.151	-0.090
b_n	-0.0705	-0.249	-0.551	-1.11	-1.272	-0.946	-0.626	-0.367	-0.201	-0.098	-0.044

17th to 21st harmonic by extrapolation.

$$a_n = R_n \cos \phi_n, \quad b_n = R_n \sin \phi_n.$$

t	0	0.19	0.39	0.58	0.87	1.17	1.46	1.75
θ	0	10°	20°	30°	45°	60°	75°	90°
a_n								
1.007	0	0.1752	0.3444	0.5035	0.7120	0.8720	0.9730	1.007
1.173	0	0.1960	0.3390	0.3910	0.2770	0	-0.277	-0.3910
1.362	0	0.2083	0.2683	0.1362	-0.1920	-0.2355	0.0708	0.2724
1.110	0	0.1487	0.1021	-0.0788	-0.1121	0.1377	0.0405	-0.1587
0.389	0	0.0432	0	-0.0432	-0.0307	0	-0.0307	0.0432
-0.083	0	-0.0070	0.0048	0.0037	-0.0053	0.0065	-0.0073	0.0075
-0.240	0	-0.0142	0.0182	-0.0091	0.0130	-0.0161	0.0178	-0.0185
-0.277	0	-0.0091	0.0161	-0.0186	0.0130	0	-0.0130	-0.0186
-0.223	0	-0.0022	0.0045	-0.0065	-0.0091	0.0114	0.0033	-0.0132
-0.151	0	0.0014	-0.0027	-0.0039	-0.0056	-0.0069	-0.0020	0.0078
-0.090	0	0.0022	-0.0037	-0.0042	0.0031	0	-0.0031	-0.0042

$\frac{a_n \sin n\theta}{n}$ terms

t	0	0.19	0.39	0.58	0.87	1.17	1.46	1.75
θ	0	10°	20°	30°	45°	60°	75°	90°

b_n								
-0.0705	-0.0705	-0.0694	-0.0663	-0.0611	-0.0499	-0.0353	-0.0183	0
-0.249	-0.0829	-0.0720	0	0	0.0587	0.0829	0.0587	0
-0.551	-0.1102	-0.0711	0.0193	0.0954	0.0777	-0.0551	-0.1064	0
-1.11	-0.1587	-0.0544	0.1210	0.1376	-0.1121	-0.0787	0.1530	0
-1.272	-0.1412	0	0.1412	0	-0.1005	0.1412	-0.1005	0
-0.946	-0.0861	0.0294	0.0662	-0.0747	0.0605	-0.0426	0.0218	0
-0.626	-0.0482	0.0307	0.0081	-0.0420	0.0338	-0.0238	0.0125	0
-0.367	-0.0246	0.0213	-0.0121	0	-0.0173	0.0246	-0.0173	0
-0.201	-0.0119	0.0117	-0.0111	0.0103	-0.0082	-0.0058	0.0115	0
-0.098	-0.0052	0.0051	-0.0048	0.0044	0.0036	-0.0026	-0.0050	0
-0.044	-0.0021	0.0018	-0.0011	0	0.0015	0.0021	0.0015	0

$\frac{b_n \cos n\theta}{n}$ terms

**The Institution of Engineers and Shipbuilders
in Scotland**

Servo Control Problems

Additional Paper.

A. J. O. CRUICKSHANK, B.Sc.

**PUBLISHED BY THE INSTITUTION
ELMBANK CRESCENT
GLASGOW
1961**

Servo Control Problems

BY

A. J. O. CRUICKSHANK, B.Sc.

The responsibility for the statements and opinions expressed in papers and discussions rests with the individual authors; the Institution as a body merely places them on record.

*Reprinted from the
Transactions of the Institution of Engineers
and Shipbuilders in Scotland*

1951

Paper No. 1141

SERVO CONTROL PROBLEMS

By A. J. O. CRUICKSHANK,* B.Sc.

23rd January, 1951

SYNOPSIS

The paper describes the general aims and problems of servo-mechanism operation. After an explanation of the action of the control, the various input signals met with are discussed. A short review of electric and hydraulic servo components is then given, with a note of the control effect produced by the various power units. Servo-mechanism performance is then considered and the question of stability and steady-state following are treated. In the first case, the action of time lags is pointed out with the aid of their steady-state frequency response characteristics. Non-linear effects such as backlash and variable friction are also described. The paper concludes with a summary of the methods used in practice for stabilization and reduction of steady-state error.

INTRODUCTION

The primary purpose of the class of control systems known generally as servo-systems is the control of the position, velocity or other attribute of the output member of a piece of apparatus, in such a manner that the magnitude of this quantity is in accordance with the dictates of some earlier, essentially time-varying quantity. The controlled quantity or output quantity as it is usually termed, might, for example, be the angular position

* Of Glasgow University.

in the training direction of an anti-aircraft gun. This would be required to move and follow as accurately as possible some other quantity, known as the input quantity, in this example the angular position of a light shaft in a remotely situated computing mechanism.

The characteristic of such control systems is that the action of the control is dependent upon the error or difference between the instantaneous input and output values ; it means, therefore, a continuous comparison of the output quantity with its desired value, as represented by the input, in order to determine the magnitude and sense of the error. This error quantity then, through the medium of power-amplification, is ultimately responsible for driving the output back into coincidence with the input and in so-doing reducing itself to zero. From the point of view of the error, such a system is self-zeroing. That some type of power amplification, for example, electric or hydraulic, is necessary, follows from the fact that very considerable forces are normally required for the rapid movements of the output member, subjected in addition perhaps, to extraneous disturbances. It is not possible to derive this power from the error quantity directly since such power would not normally be sufficient, and in any event, this procedure would adversely affect the true value of the input quantity itself. It is, therefore, axiomatic that servo-systems are power-amplifying.

Consider, for example, a ship steering-gear which, though not normally regarded as such, is a simple form of hydraulic servo-mechanism for the purpose of controlling the rudder angle in accordance with the angle of the helm. The latter is, therefore, the input quantity. Its value is continuously compared with the output quantity of the system, namely, the rudder angle, usually by means of the differential action of a floating lever. One end of this lever is actuated by the tiller movement and the other end by the movement of a receiver telemotor operated from a transmitting unit controlled by the angle of the helm. The resulting movement of the differential lever, representing the difference between the two angles, sets the stroke of a variable-delivery pump delivering high-pressure oil to a double-acting ram which moves the rudder in the correct direction. The basic elements of a servo-control are thus present, namely, an error-determining device forming the error quantity,

which, after hydraulic power amplification, is the ultimate cause of moving the rudder into the position of correspondence. The essential operation of servo-control is illustrated in Fig. 1, wherein is shown the continuous feedback of information about the output magnitude in order to compare it with the input magnitude and so determine the error, ϵ . The error quantity is here defined as $\epsilon = \theta_1 - \theta_0$.

In electrical engineering terms, the error can be thought of as the result of applying negative feedback of the output quantity to the input quantity. Recognition of the similarity between servo-systems and negative feedback amplifiers in electronic circuits has enabled many of the methods already existing for the analysis of such amplifiers to be judiciously applied to analysing servo-systems. This principle, together with the feature

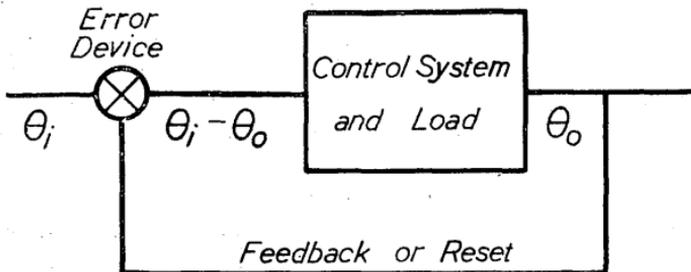


Fig. 1.—Block diagram of servo-system.

in practice, that electrical servo-mechanisms are more flexible in adjustment than other types, has led to much of the theory being developed in electrical terms.

GENERAL FEATURES OF SERVO-SYSTEM PERFORMANCE

Self-oscillation. The most important feature of a servo-system, in common with any system having feedback over an amplifying-means, is the possibility of sustained self-oscillation without the application of an external input quantity. The tendency to go into an uncontrolled state, in which the output and all other variables in the control sequence periodically increase and decrease to large amplitudes in either direction, depends upon the degree of corrective action put into the control, or more briefly, on the sensitivity of the control. If the corrective action is slight, the output will be very sluggish in responding

to a sudden fixed deviation of the input and an unnecessarily long time will elapse before coincidence with the input occurs. At the other extreme, by the overshooting of an inertia load, the corrective action may be over-violent and may produce a greater error than the original. Such a system will increase its errors alternately in either direction and is therefore totally unstable. The actual frequency of the oscillation depends on the size of the system, being higher for small fractional-horsepower systems than for larger power-type servos, for example, a 5-h.p. system which might have a natural frequency of oscillation of the order of 1 to 2 cycles per sec. Clearly, the control action must be intermediate between the above states, producing a sufficiently fast response and one in which any overshooting is restricted in magnitude and rapidly caused to die away.

Input Signals. The type of input quantity variation to which a servo-mechanism will be subjected in practice will depend on the particular application. For an anti-aircraft gun following a target moving at constant velocity in a horizontal plane, the following speed of the gun (considering the training motion only) will increase as the target approaches the gun and will decrease as the target recedes, the maximum angular velocity depending on how near the plane approaches and on its speed. This type of input quantity will, therefore, be relatively gradual in application, showing first an increasing rate of change followed by a decreasing rate of change. A converse state of affairs exists when any control system is suddenly made operative with, in general, an initial difference between input and output. This results in a sudden "step-function," as it is called, being applied to the system, which will recover after the general manner shown in Fig. 2a.

This step-function response, or transient response, on account of the approximation to it occurring in practice and also on account of the ease of testing an experimental system, is a valuable method of specifying the performance of any particular system in respect of its speed of response and degree of damping. In general terms, one overshoot of 15 per cent. of the input magnitude with a small subsequent undershoot might be acceptable, provided that the final decay of the error was not unduly long.

Other specifications for a following system are the errors

allowable with certain maximum rates of change of input or certain maximum accelerations of the input. In a servo following an input moving with constant velocity, Fig. 2b, it frequently happens that the output, although running at the same speed, lags a fixed distance behind the input, this distance being known as the steady-state velocity error. Certain applications require this error to be zero; such systems are designated *zero-velocity-error systems*. Analogously one has *zero-displacement-error systems* for those applications in which no static error results in the step-function response. Curve C in Fig. 2a shows a system having a steady-state error in displacement, that is, static error.

Another form of input that gives much insight into the performance of any system is one which has a harmonic variation

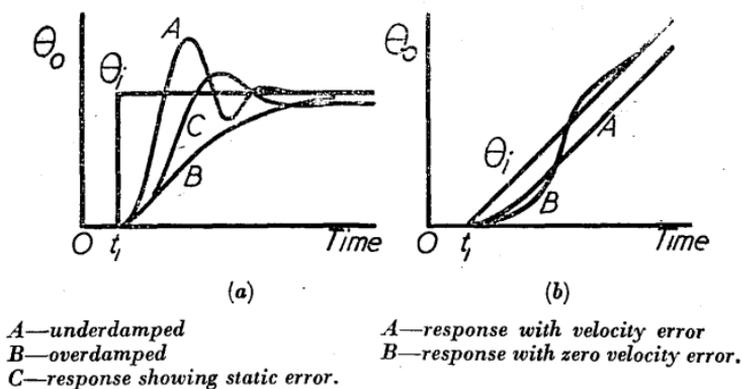


Fig. 2.—Servo-mechanism input signals and responses.

with time. Although the input itself seldom has such a variation, an equivalent effect is obtained when the input is fixed and a periodically-varying torque is externally applied at the output shaft of the system. Such a situation may arise due to the unstabilizing effect of a ship's roll on an unbalanced gun mounting. The application of such an input motion and the resulting output of a servo-mechanism is shown in Fig. 3a, from which the main effects are first, a change of the magnitude of the variation and second, a lag of the output quantity behind the input quantity. Whether the output magnitude is greater or less than that of the input depends on frequency of the imposed harmonic variation; at a high frequency it is fairly evident that the input motion cannot be reproduced by a massive load, which will

follow quite well at very low frequencies. A convenient method of displaying the relationship of Fig. 3a is the vector diagram of Fig. 3b.

Regulators and Stabilizers. These are essentially servo-control systems having a fixed value of input quantity and subject to extraneous influences. Under these conditions the error produced by an external torque or load being applied, can only be such as will call into existence an equal opposing torque from the control action. A particularly good example occurs in ship stabilization by the use of activated fins. The input quantity is fixed and, for this reason, let it be called the datum quantity. In this application it is the vertical direction provided by a vertical-keeping gyroscope. Roll of the ship causes a signal to

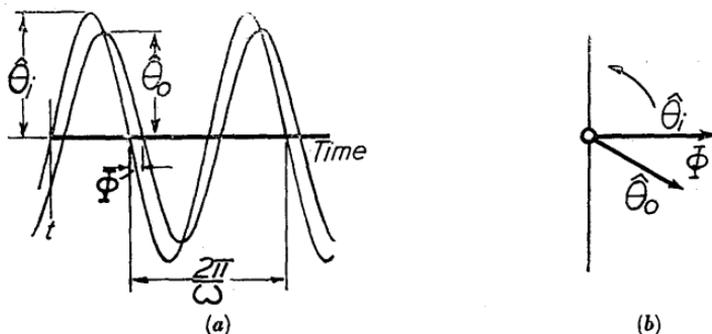


Fig. 3.—Sinusoidal input signal and steady-state response.

be supplied by the gyro and this, after modification and several stages of amplification by hydraulic means, eventually causes the fins to tilt and through the ship's forward motion provide a stabilizing couple in the correct direction.

In connection with regulators, it frequently happens that the datum quantity is not of the same physical nature as the regulated quantity. In a steam-engine governor the datum quantity is a certain setting of the force provided by the governor controlling spring. The speed of the engine is expressed by an axial force exerted on the governor sleeve, this conversion being effected by the centrifugal action of the rotating balls. Another example of an implicitly-contained datum quantity is the balance-point of a non-linear electric circuit, used as a voltage-sensitive device.

The general problem with regulators is in the main a steady-state one. While it is obviously advantageous for a regulating system to have as quick a response as possible commensurate with stability, the primary concern is the steady-state effect upon the controlled quantity of some constantly applied torque or load demand. The analysis of regulators, therefore, differs slightly from that of servo-mechanisms in that firstly, the output response to an output disturbance is required (sometimes to a supply system disturbance in addition) rather than to a variation in the datum quantity and that secondly, the steady-state error or droop is of more importance than the transient effects. These differences are points of operation only and do not alter the fact that regulators and servo-mechanisms are basically the same class of control, the former with a fixed pre-determined input and the latter with a random, time-varying input. The generic name of closed-cycle, closed-loop or closed-sequence control includes all such systems.

BASIC COMPONENTS

The purpose of this section is to give a very brief descriptive review of commonly occurring servo-components. Though there are numerous methods of securing a particular result in practice, attention will be drawn only to those elements that may be termed typical and which tend to recur in association with other typical members. This will be done for electric and hydraulic servo-mechanisms only. Pneumatic servos of a few watts output, such as occur in aircraft applications, are not considered. The description "basic" refers to the essential features only of the control, that is, to error-determination, power-amplification and load movement. Additional connections of apparatus to this basic structure for obtaining improved servo-responses are taken up in a later section.

Electrical Servo-mechanisms. Fig. 4 shows the simplified diagram of the metadyne system for the remote-position-control (r.p.c.) of a gun mounting. This refers to one motion only of the gun, two separate servos being required to control it completely in both training and elevation. The gun-driving motor is a D.C. motor operating with a fixed shunt field current and fed with variable armature current in either direction by a metadyne-

generator, which is essentially a rotating-machine amplifier. The small input power required by the metadyne generator for its differential or push-pull excitation is supplied in turn by a high-gain thermionic amplifier whose input is the error signal. This is obtained in the form of a small alternating voltage from a pair of magslips, one driven by the output shaft and the other by the input shaft. The sequence then is error-determination (magslips), power-amplification (thermionic amplifier and metadyne generator), power drive (D.C. servo-motor) and load (training moment-of-inertia). Such elements represent a selection

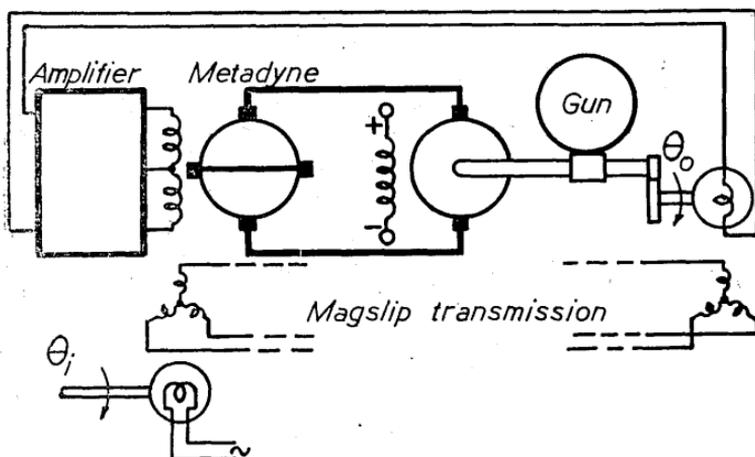


Fig. 4.—Metadyne system for remote-position-control of gun mounting (diagrammatic).

of the practical components that frequently occur in electrical servo-circuits. The features of each are described below, together with brief mention of alternative arrangements.

Error Devices. The magslip family of data-transmission systems was developed by the Admiralty Research Laboratory, and of the many forms now in use the coincidence transmission system and the transmitter-hunter-resetter system will be described.

Coincidence transmission produces the angular error between two remotely-situated shafts as an alternating voltage. Referring to Fig. 5a, it will be seen that the 3-phase stators are connected

and that the rotor of the transmitter has its single-phase winding connected to an alternating supply and is rotated by the servo input shaft. The rotor of the second magclip is driven or reset by the servo output shaft and the magnetic axis of its winding is so arranged that in the position of zero error no voltage is induced by the stator currents. The A.C. voltage obtained if there is either a fixed positive or negative error is shown in Fig. 5*b*. The magnitude and phase of this voltage with a variable

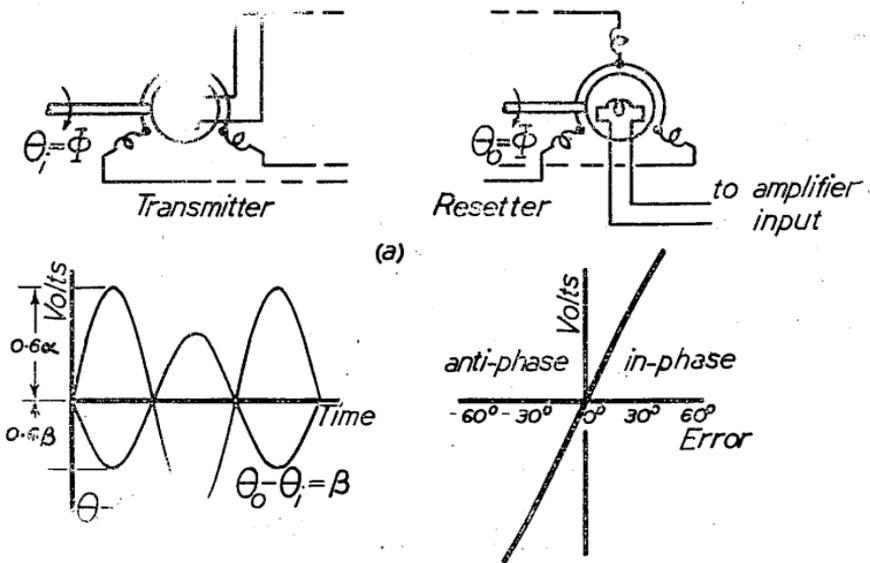


Fig. 5.—Magclip coincidence transmission.

error is shown in Fig. 5*c*. The sensitivity of these units is 0.6-volt per 1° misalignment, their accuracy $\pm \frac{1}{4}^\circ$, and normal operation is from a 50-volt, 50-cycles per sec. supply or from a 20-volt, 1,100-cycles per sec. supply.

The transmitter-hunter-resetter chain produces the angular error between two remotely-situated shafts as an angular displacement. Fig 6 shows an additional unit, the hunter, carrying 3-phase windings on both rotor and stator which are connected to the transmitter and resetter stators respectively. The

transmitter and resetter rotors are connected to an A.C. supply and the torque acting upon the hunter rotor is, for small errors, proportional to the misalignment between the input and output shafts. The movement of the hunter rotors is constrained by springs to a few degrees either way and, by the linkage shown, the rotor moves the sensitive pilot-valve of an oil unit. The torque produced in a 3-in. hunter is about 0.16 oz.-in. per 1° misalignment.

For applications in which a local control of power is sufficient a mechanical differential gear is a simple form of error-measuring device. It is not, however, without reaction on the input shaft, which must be capable of resisting the torque.

Power Amplification. A thermionic amplifier is normally required to give a D.C. output of the correct magnitude and

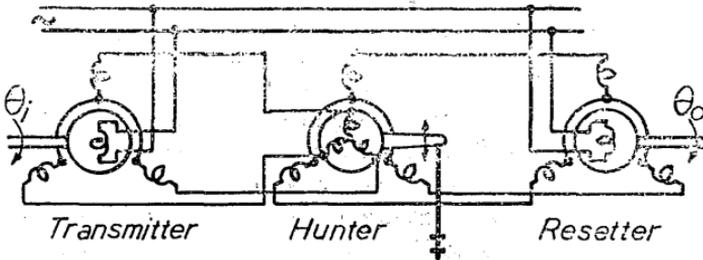


Fig. 6.—Three-element transmitter-hunter-resetter chain.

direction in response to an A.C. input of variable magnitude and which is either in-phase or anti-phase with a fixed datum. This type of amplifier, therefore, has a preliminary stage of phase-sensitive rectification followed by two or more stages of balanced, push-pull D.C. amplification. In addition the amplifier is the point where further signals or interstage networks may be added for improved servo response. It may also function in such a way to reduce the power supply to the servo as the extremities of the load range are reached and any limit switches are operated. The power output of thermionic amplifiers, however, is reached at about 50 W. and for power purposes a rotating-machine or dynamo-electric amplifier follows.

Dynamo-electric Amplifiers. Developed originally for electric traction, the metadyne generator obtains its high power amplification and rapid output current response to the control-field

current by using armature-reaction fluxes. On one axis of the machine strong armature-reaction flux is produced by a short-circuited pair of brushes and on a quadrature axis to this, negative feedback of the output current from a secondary set of brushes is applied to the control-field current. This feature tends to make the output current rather than the output voltage of the generator proportional to the control-field current, and in effect constitutes a torque control on the D.C. servo-motor, especially at low speeds. The differential control-field currents are of the order of 40 ± 40 mA. Power amplifications range from 100 (for about 500 W. output) to 10,000 (for about 50 kW. output).

Similar to the metadyne generator but for the omission of negative feedback of output current, the amplidyne generator lacks the rapid current response of the metadyne. Positive feedback of output current (over-compensated amplidyne) may be adopted to increase the output voltage.

In the interests of power amplification the straightforward D.C. generator (Ward-Leonard control) requires to be preceded by an exciter system. Since, neglecting the effect of time lags, the generator voltage is proportional to the amplifier input, this type of control gives a D.C. servo-motor speed proportional to error, but excessive currents and torques arise when the error changes suddenly. Negative feedback of the generator output current provides this torque limitation, and tends also to result in the control of the servo-motor torque rather than speed.

Servo-motors. For power-type servos, the D.C. motor with its armature input controlled and having a fixed shunt field current, is normally used and results in the developed torque being proportional to the motor armature current. In fractional-horsepower servos, the converse is arranged, with the field current variable and armature current constant. This method of connection enables the amplifier to supply the field power of a small motor directly and, approximately, makes the servo-motor torque proportional to the error.

The 2-phase A.C. motor is frequently used in low-power applications thereby providing a wholly A.C. operated servo. Difficulties of designing stabilizing circuits in this case restrict its use to systems having a certain amount of inherent damping.

Hydraulic Servo-mechanisms. Fig. 7 illustrates a simple valve-controlled hydraulic servo. The error is determined by

the transmitter-hunter-resetter chain, the resulting small pilot-valve movement being amplified by the hydraulic relay, the output of which moves the valve controlling the oil motor. This drives the load, shown as a simple inertia, and also resets the local magslip. Apart from the A.C. supply to the magslips, the system uses no electrical power and constitutes an all-hydraulic servo. In large hydraulically-powered servos it frequently happens that coincidence transmission magslips are used, followed by thermionic amplification of the voltage and conversion to a mechanical displacement before passing it to the hydraulic relay circuits. The final power drive, apart from simple rams used to obtain straight-line motions is by a pump- or valve-

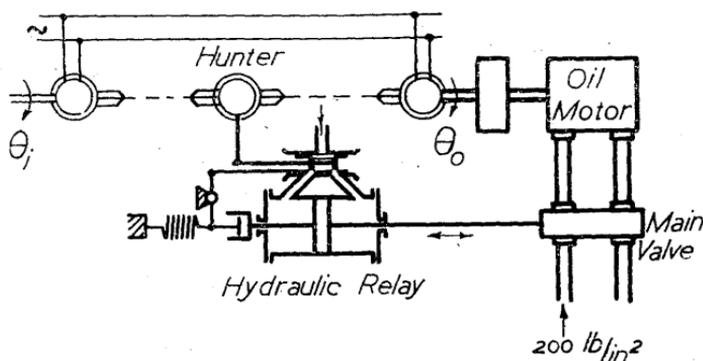


Fig. 7.—Valve-controlled hydraulic servo-mechanism (diagrammatic).

controlled oil servo-motor, both examples requiring some form of variable-delivery pump. In a pump-controlled power drive the delivery is variable in either direction and determined by the pump stroke which is dependent upon the control action. The pump supplying a valve-controlled power drive is pressure-regulated and gives up to full delivery in one direction only. Descriptive details of some typical methods used in hydraulic or electro-hydraulic systems are as follows.

Power Amplification. The simple double-acting ram or hydraulic relay is an amplifier giving very high force magnifications, and is operated generally with a source of oil at a constant pressure lower than that used for the final power drive. Hydraulic amplifiers are capable of practically unlimited power output,

but owing to friction and difficulties of mechanical construction there is a minimum input force below which the sensitive valve cannot be made to work.

In the basic hydraulic relay, consisting of a double-acting ram, the velocity of the ram or output can be taken as proportional to the effective port opening of the valve, provided that the pressure drop across the ram does not exceed about 50 per cent. of the supply pressure. Assuming that the port opening is proportional to the valve displacement from the mid position, the basic relay therefore produces an output displacement proportional to the integral of the input displacement.

The integrating relay with negative displacement feedback, more commonly called a "proportional," "corresponding," or "follow-up" relay, has the valve moved according to the relative displacement of the ram and the input acting at opposite ends of a floating or follow-up lever, pivoted at the end of the valve rod. An alternative method is to have a movable liner, operated by the ram, in the pilot-valve cylinder. The addition of this negative displacement feedback produces an amplifier which, in the steady state, has an output displacement proportional to the input displacement.

The normal form of the electro-hydraulic relay has as its input quantity the D.C. voltage supplied by the previous stage of thermionic amplification. By means of a coil and cylindrical-magnet system (similar to that of a moving-coil loudspeaker), the input is converted to a mechanical force which operates the hydraulic relay.

Pumps and Control Valves. Variable delivery pumps are normally of the radial-piston type, the action of the stroke control causing the piston-assembly to be driven eccentrically and produce pumping action, the amount and direction of which depends on the amount and direction of the eccentricity imparted by the stroke-control lever.

For requirements of appreciable power an approximately constant pressure source is obtained by regulating the stroke and hence delivery of a variable-delivery pump, according to the pressure changes brought about by the load. In this instance it is arranged that the pump delivery will be zero at some maximum pressure rating P_1 and will be increased to the full capacity in one direction at some lower pressure, about $0.8P_1$ thus giving a

total pressure change of 20 per cent. from zero to full capacity of the pump. In practice this can be done by the movement of a spring-controlled piston in a cylinder supplied with the output pressure of the pump, the resulting displacement operating the stroke control lever.

The reciprocating-piston control valve is much used in servo technique in a wide range of sizes. In large valves working at high pressures, the forces due to hydraulic reaction exceed the frictional forces present, but with sensitive pilot-valves the converse holds. The main feature that limits the operation of small sensitive valves is in fact the friction and stiction effect. The latter is overcome to a certain extent by superimposing on the operative force a relatively high frequency "dither" force, provided by a linkage to an eccentrically driven member.

Servo-motors. The hydraulic motor used in servo-control is in effect a reversed radial-piston or swash-plate type pump with a fixed stroke. Neglecting leakage, the velocity of such a motor is directly proportional to the volume of liquid passing through it per unit time. In practice, however, the velocity is not quite independent of the torque but decreases as the torque, pressure difference and hence leakage across the pump, increase. While valve-control of an oil motor is basically wasteful due to the energy loss in the valve itself, its use enables a number of motors to be operated simultaneously from one constant-pressure source. Pump-control on the other hand demands a separate pump and driving motor for each installation but against this it is likely that the wear on the pump, which is not generating at constant pressure, will be less. To a first approximation, the velocity of a pump-controlled motor is proportional to the amount of stroke given to the pump. A more complete analysis includes the effect of compressibility of the oil and in effect inserts a certain resilience into the system.

SERVO-MECHANISM RESPONSE CHARACTERISTICS

In this section, details are given of the problems that arise in predicting the response of a servo under certain conditions of input motion and load characteristics. The problems are discussed as far as possible without formal mathematical treatment.

Stability. As any practical control system must not only be

stable, but must possess a certain stability margin, the root causes of instability are the first points to be investigated. In this respect three aspects of the problem emerge, each of which is capable of causing unstable operation. These are :

- (a) The form of the basic control characteristic, assuming all time lags in the operation of the sequence are absent.
- (b) The presence of time lags in the control sequence.
- (c) Imperfections of the mechanical drive, for example backlash, and variable friction characteristics of the load or other mechanical members.

Effects (a) and (b) lend themselves to analytical treatment in linear systems, that is, whenever the system "constants" or parameters do not vary with the magnitude of their related varying quantity. Frequencies of oscillation resulting from these effects are termed natural frequencies. Item (c) is essentially a non-linear property. Any resulting oscillations may be either continuous or interrupted, and not of the same frequencies or amplitudes as in the linear systems.

(a) *Basic Control.* The relationship between the output quantity and a constantly applied error quantity determines the basic nature of the control action. For example, by removing the A.C. supply to the magslips in Fig. 4 and instead supplying a constant A.C. voltage, v , to the amplifier (this in effect represents an error $v/0.6^\circ$), it is found that the motor eventually rotates at constant velocity. It is emphasized that this steady-state relationship is only established after the lapse of a certain period, the duration of which is dependent on the magnitude of the time delays occurring in the system, and that, in this steady state, time-lags have ceased to have any effect. Any linear system in which a steady-state output velocity is produced by the application of a constant error is represented by a basic control equation, $d\theta_o/dt = K_v \epsilon$, where K_v is the gain factor of the system. Such a *velocity-controlled* system is inherently stable.

The type of control existing in Fig. 7, however, results in a steady-state output acceleration proportional to error* and

* Owing to the finite travel of the relay piston, an experimental trial of this would require the steady state to be attained before the relay piston moved this much.

hence has a basic control equation, $d^2\theta_o/dt^2 = K_a\epsilon$. This type, namely *acceleration-control*, is inherently unstable and must always have additional stabilizing means. It may be compared to an undamped pendulum where the restoring force is proportional to the swing ϵ from the vertical. Thus, since

$$\begin{aligned} -Mld^2\epsilon/dt^2 &= Mg\epsilon, \\ -d^2\epsilon/dt^2 &= (g/l)\epsilon, \end{aligned}$$

while, for the above control type,

$$\begin{aligned} d^2\theta_o/dt^2 &= d^2(\theta_i - \epsilon)/dt^2 = K_a\epsilon \\ \text{i.e.} \quad -d^2\epsilon/dt^2 &= K_a\epsilon \end{aligned}$$

for a stationary input. Hence with no motion whatever given to the input, continuous oscillatory errors are formed as the output hunts about the input position. This type of instability is therefore inherent in the control characteristic and without some stabilizing means this form, which is shown later to be highly desirable for certain applications, cannot be used. The remaining type, which does not occur in position control, however, is that in which a steady-state error produces a steady-state output displacement. This form is confined to certain types of regulating systems only.

(b) *Time Lags*. The two forms of time lag considered here are finite time lag and simple exponential time lag. Finite time lag is also known as "dead-time." It is the time-interval between the occurrence of a signal and the response, due to an intermittent action in the system. Finite time lag does not appear very often in servo-circuits but is of common occurrence in process-control systems, where due to the finite speed of propagation of the signal change, a lag results. The input-output relationship for an element having finite time lag and subjected to a sudden change in its input is shown in Fig. 8a and the corresponding relationship for the more common exponential time lag in Fig. 8b.

Exponential time lag occurs, for example, in the response of the piston of a proportional-type hydraulic relay to a change in position of the input member of the relay. It arises, also, in the growth of the current in an inductive circuit such as the field winding of a generator or exciter. The greatest time lag in an

electrical servo, however, is due to the output inertia and to the viscous damping effect on the load introduced by the servomotor operation. The falling torque/speed characteristic of D.C. servo-motors causes a considerable interval to elapse before an inertia load reaches the steady-state speed corresponding to a change of the control field current of the generator supplying the motor.

Time lags, whether present in the forward control sequence, or incurred in the process of output measurements, accentuate any tendency to instability by virtue of their supplying to the component immediately following them information concerning the past state of the output rather than its actual state. In exponen-

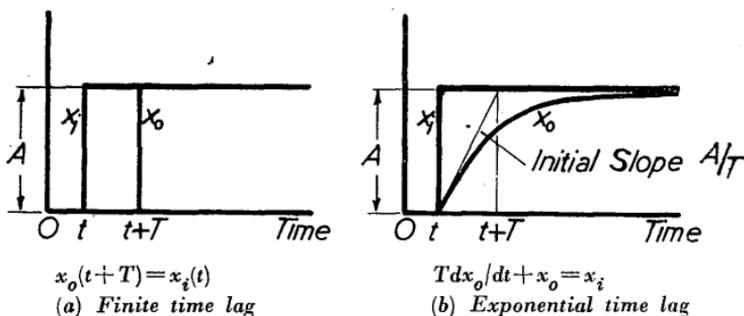


Fig. 8.—Responses of time-lag elements to sudden change in input.

tial time lags, for given values of the lags and the other constants, for example, sensitivity, of any system, the existence of stability or not, can be checked by Routh's¹ or Hurwitz's² Criterion. This is not, however, particularly helpful for assessing the damping or relative stability. Much design work with this second aim in view is carried out in terms of the behaviour of the system components when they are carrying continuous sinusoidally-varying quantities as in Fig. 3. This arises from the fact that a system which is in a state of persistent self-oscillation operates with such signals at all points of the control sequence. Thus, for continuous self-oscillation without the application of an input, the error quantity must be formed wholly from the output. That is, if

$$\theta_o = \theta_{om} \sin \omega t,$$

then

$$\varepsilon = -\theta_o = \theta_{om} \sin (\omega t - 180^\circ).$$

¹ See bibliography, p. 296.

The conditions for continuous self-oscillation are therefore that the ratio of the amplitudes of the output and error should be unity, that is, $\theta_{om}/\varepsilon_m = 1$ while the phase shift is 180° .

Considering now the effect of a simple exponential time-lag element having a sinusoidally-varying input, Fig. 9a shows the resulting output reduced in amplitude and lagging in phase. The amplitude ratio x_{om}/x_{im} is $1/\sqrt{1+\omega^2T^2}$ while the phase shift is $\tan^{-1}\omega T$ (Appendix 1). For a given time lag, the phase shift increases with the frequency of the input while the amplitude ratio x_{om}/x_{im} decreases, Fig. 9b. For a given frequency the same result is obtained as the time lag is increased.

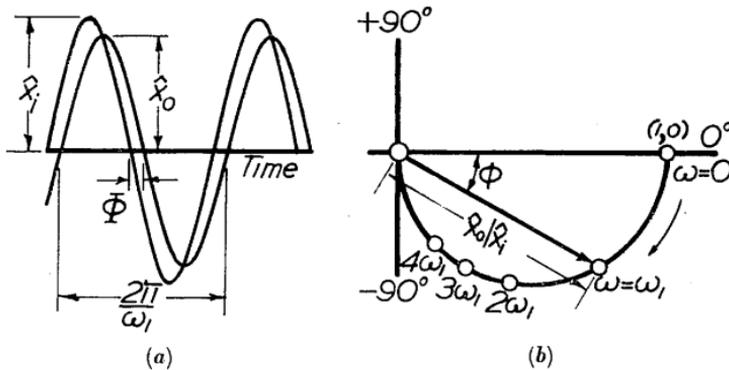


Fig. 9.—Steady-state frequency response of exponential time lag.

The implication of this result as far as stability is concerned is depicted in Fig. 10. Curve A is the polar plot of the phase shift and amplitude reduction or attenuation, introduced by a simple exponential time lag. It is essentially a repeat of Fig. 9b. Curve B, which coincides with the -90° radius vector, is the polar plot of the basic control characteristic of a velocity-controlled system.* Curve C is the resultant curve giving the amplitude ratio and phase shift for θ_o/ε in a velocity-controlled system with one time lag. It is derived by "multiplication" of the two curves A and B at corresponding frequencies. That is, for the frequency $\omega_1/2\pi$,

$$OR = OP \times OQ \text{ and } \theta = \Phi + 90^\circ.$$

* Since $d\theta_o/dt = K\varepsilon = K\varepsilon_m \sin \omega t$
 $\theta_o = -(K\varepsilon_m/\omega) \cos \omega t = (K\varepsilon_m/\omega) \sin(\omega t - 90^\circ) = \theta_{om} \sin(\omega t - 90^\circ)$
 Therefore the amplitude ratio, OQ in Fig. 10, = $K/\omega i$. The phase shift is equal to -90° for all frequencies.

The result of the time delay has been to increase the phase lag of the output with respect to the error quantity. A further step now is to add another time lag to the system, as in Fig. 11.

Curve A is the polar plot of the resultant of two time lags acting one after the other. Curve C is the resultant θ_o/ε polar plot for a velocity-controlled system with the two time lags. In this instance, the phase shift introduced by the time lags is sufficient to turn the resultant C curve through angles in excess of 180° at high frequencies. With the sensitivity or gain-factor of the system as shown, the amplitude ratio at 180° phase shift, namely,

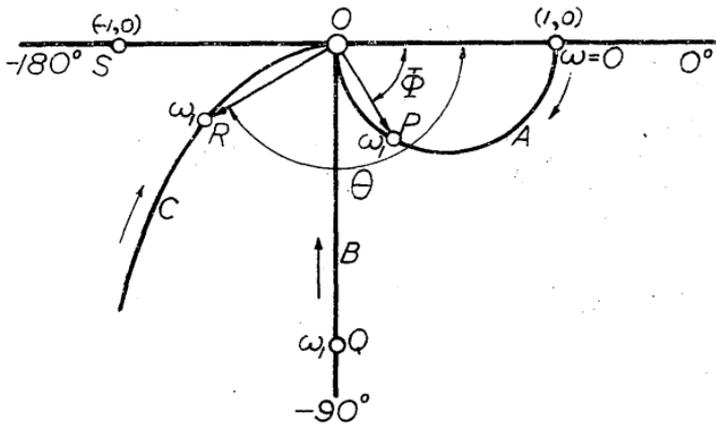


Fig. 10.—Nyquist diagram for velocity-controlled servo-mechanism with one exponential time lag. (Arrow with each locus, indicates direction of increasing frequency.)

the length ON , is less than unity. Since, however, the control sensitivity K is a variable quantity, the amplitude ratio $\theta_{om}/\varepsilon_m$ at "cross-over" will depend wholly on the value given to K . If it is increased from that shown in Fig. 11, a stage is eventually reached when ON is equal to unity, that is, N coincides with S . Under these adjustments the amplitude ratio $\theta_{om}/\varepsilon_m$ is unity while the phase shift is 180° . This, as stated previously, is the critical stability limit at which the servo-system is capable of supplying its own error solely from the output quantity. Any attempt to increase the sensitivity further merely causes continuous oscillations to occur. This particular technique for investigating stability results in the test is generally known as the Nyquist³

gear ring of large diameter and as the drive may be required to operate between certain extremes of temperature, a certain freedom of operation must be allowed. This, together with the normal machining errors, limits to several minutes of arc the accuracy it is possible to achieve in a high-torque drive. Again, with such a drive, mechanical resilience occurs in the shafts themselves and in any comparatively soft materials used in mechanical couplings. The result of this backlash and resilience is that the motor and load tend to oscillate in anti-phase at either end of the power drive and if the system possesses high amplification oscillations of a relatively high frequency result within the backlash. This tendency to instability can be partially overcome by driving the resetting transmitter from the motor instead of from the load, since with resilience and backlash present the motor is effectively in advance of the load position at any time. Such a procedure necessarily places the load outside the control loop and though stabilizing may result, the accuracy of the system is decreased by the errors which still exist in the power drive. A further improvement due to Belsey⁶ consists in deriving part of the reset from the motor and part from the load. They are then combined in such proportions that resetting is effectively carried out at the "nodal-point" of the power drive, that is, at the position which would be occupied by a solidly-coupled motor-load inertia. This divided-reset as it is termed, gives greater accuracy by allowing a further increase in the gain of the system before instability occurs and by obtaining the feedback information at a point, in effect, nearer the final load.

Steady-State Following. The basic control equation referred to in a previous section is an expression of the conditions under which a servo-mechanism can sustain a constant error. Thus for an acceleration-controlled system, $d^2\theta_o/dt^2 = K_a\epsilon$. A constant following error of Φ/K_a therefore results when the output, and hence the input also, possesses the constant acceleration Φ . Such an input is shown in Fig. 12, curve C. An approximation to this may occur in practice over a very limited time. More commonly, the inputs B and A result and as neither possesses any acceleration, no permanent error can exist in an acceleration-controlled servo following either input. For this reason, acceleration-controlled servos are also termed zero-velocity-error systems.

Similarly in the velocity-controlled type, $d\theta_o/dt = K_v \varepsilon$, a constant error $\varepsilon = \Omega/K_v$ will exist for an output velocity and hence an input velocity Ω , as in B. This type requires no error, however, to maintain a constant output angle θ . It is therefore also termed a zero-displacement-error system.* From the expression for the constant-velocity following error, it is evidently desirable that K_v , that is, the gain factor or control sensitivity of the system, should be as large as possible. The displacement-controlled type is not capable of following steadily moving inputs and therefore has no application in servo-mechanism work.

The property of following a constant-velocity input with zero steady-state error is very advantageous and can be realized in

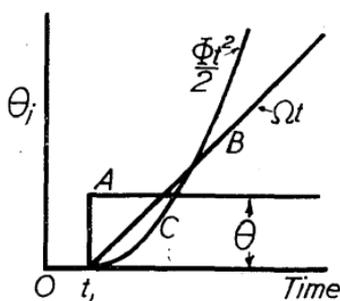


Fig. 12.—“ Step-function ” constant-velocity and constant-acceleration inputs.

both electrical and hydraulic servos, but probably more easily by the latter. Such a system will require a fair degree of stabilizing in order to counteract the inherently unstable basic characteristic and the inevitable time lags which occur. In order to attain sufficient constant-velocity following accuracy with a velocity-controlled type the gain setting requires to be very high, which again entails considerable stabilizing in the transient region. The design problem, therefore, resolves itself into one of stabilization without loss of steady-state accuracy. Methods of achieving this are taken up in a following section.

Performance at Creep Speeds. The variable friction characteristic of a normal load or other mechanical member is shown in Fig. 13.

* It must be remembered that the above results are established for linear systems only. In practice, stiction may cause a small static error to appear.

Considering the motion of output as the input shaft is slowly turned, the error must build up until the torque developed is sufficiently large to overcome the stiction of the output. Once movement occurs, the friction torque reduces suddenly and as a result, overshooting of the output generally occurs so that it comes to rest ahead of the input. A stationary period of the output then follows until the input motion results in sufficient error to move the output again. Depending upon the adjustments of the system, the output may even overshoot the input

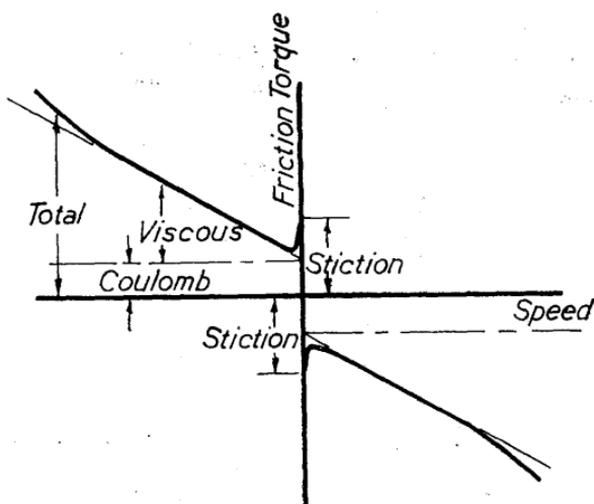


Fig. 13.—Dependence of friction on speed.

position sufficiently to cause an intermittent type of oscillation. In practice also there is usually a wide variation of the stiction value or stiction/friction ratio over the range of load movement. The net result of such irregularities is that smooth following at low creep speeds presents a formidable problem, which is only partially solved. Attempts have been made to superimpose a "dither" torque on the servo-motor in the same manner as for the sensitive pilot-valve of an oil unit.

A similar state of affairs exists in the "static" behaviour of a system in which the motor torque is proportional to the integrated error. An initial static error is then integrated until the torque which has been built up suddenly moves the output,

which then overshoots and comes to rest. The integration of error then proceeds once more and a slow cycle of load movement about the input position results.

STABILIZATION OF SERVO-MECHANISMS AND REDUCTION OF STEADY-STATE ERROR

There may be said to be two sources of error in a servo-system, namely, transient error and steady-state error. Transient error is the discrepancy between output and input while the system is in the course of responding. The natural tendency of any system to oscillate appears as overshooting and undershooting before the output is finally constrained to obey the input. In a completely unstable system this period, during which the system behaves according to its natural modes, lasts indefinitely. Stabilization consists essentially in giving large damping to the natural frequencies of oscillation of the system. From the standpoint of the Nyquist Criterion, it implies some means of putting a local "dent" into the polar plot of θ_o/ϵ , in order to place it as far as possible from the critical $(-1, 0)$ point. It is not, however, intended to go into this aspect of design here.

On the other hand, steady-state error is dependent on the system behaviour once this natural settling-period is over, and as already shown depends only on two features. These are, firstly, the basic type of control, and secondly, the amount of sensitivity or gain that can be allowed. Thus the requirements of high steady-state accuracy conflicts with that of reasonable transient performance. Since also the steady-state velocity error existing in a velocity-controlled system cannot be made zero without in effect changing over to an acceleration-controlled system, additional modifying networks or connections which produce zero-velocity error do in fact change the class of the control system.

There are numerous methods in practice for bringing about a particular system response. Since one, and only one, control equation can exist for any specified time response, which is the solution to that equation, all practical methods of improving servo performance are merely methods of altering the coefficients of that equation to a certain set of values. From the practical viewpoint they are by no means equivalent, some being more

easily or advantageously applied than others; realization of this analytical equivalence is often a design aid in attempting to replace a rather cumbersome linkage by an electric network, or conversely in knowing a stabilizing method in electric terms and translating this into mechanical methods.

As to the means employed in improving either the transient or steady-state performance, three main techniques arise, which may be described as cascade compensation, feedback compensation and load compensation or vibration damping.

Cascade compensation is the insertion into the forward control sequence, at the thermionic amplification stage, of specially designed networks chosen for the modifying effect on the error signal as it changes. Feedback compensation, employed in both electric and hydraulic servos, consists in supplying additional information to the low-power end of the forward control sequence, that is, the amplifier input, as to the rates of change of the output quantity. Load compensation, employed for small power instrument-type servo-systems, makes use of some form of vibration damper. It is particularly useful in wholly A.C. operated servo-mechanisms where the design of corrective networks presents difficulty. A further general point already mentioned is that stabilizing methods are more easily derived for electric systems than for their hydraulic counterpart. For this reason high-power hydraulic servos usually have a first stage of thermionic amplification in which the primary stabilizing means are provided. Purely mechanical methods of cascade or feedback compensation are restricted to the possible varieties of a spring dashpot combination.

Stabilization Methods. Output Velocity Feedback. All loads possess a certain amount of natural friction which damps the load movement to some extent. Such friction is of a very variable nature, however, and the amount of damping it provides is negligible in comparison with that which can be obtained electrically by means of an additional feedback of output velocity. Thus, by driving a small D.C. tachogenerator from the output shaft, a D.C. voltage proportional to the velocity of the output is obtained and this can be added to the normal error voltage at a suitable stage in the amplifier. The connections are so arranged that the torque produced by this voltage at all times opposes the velocity of the output and thus effectively introduces

viscous friction on the output shaft. An alternative to providing a tachogenerator is to derive a current approximately proportional to the output speed by connecting directly across the servo-motor armature and supplying the auxiliary field winding of a metadyne generator or a Ward-Leonard exciter. Although an effective method of stabilization, output velocity feedback is not suitable for systems required to follow constant velocity inputs, since the error must increase in order to overcome the

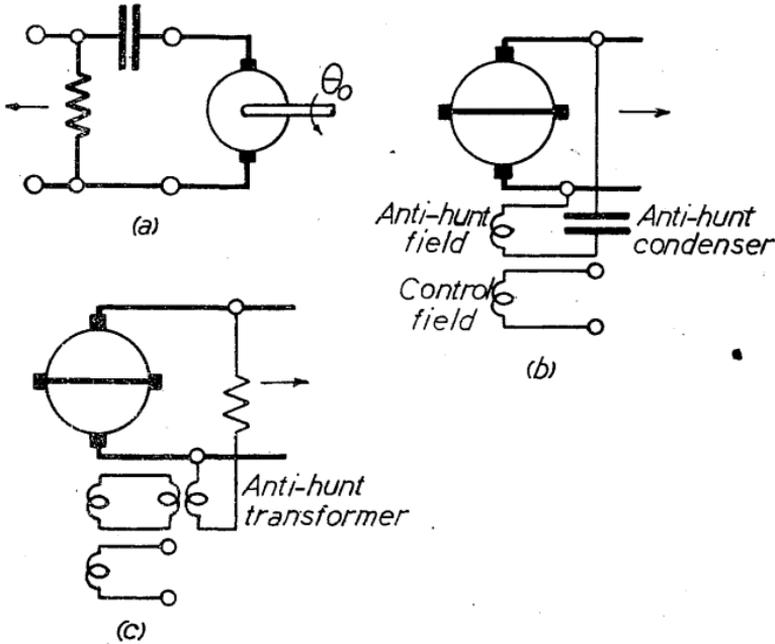


Fig. 14.—Stabilizing arrangements. (a) Transient output-velocity feedback.
 (b, c) Amplidyne anti-hunt connections.

voltage feedback due to the constant speed of rotation, and provide the necessary driving torque in addition. The above defect may be remedied by making the feedback operative only when the output changes its velocity. A simple CR network, Fig. 14a, cascaded with the output of the tachogenerator, will accomplish this transient velocity feedback.

The auxiliary or anti-hunt field winding is also used frequently in amplidyne regulating systems, and is then supplied through a condenser or anti-hunt transformer, Figs. 14b and 14c. Both

arrangements allow the feedback ampere-turns to be operative only when the output voltage changes, the current flowing being approximately proportional to the rate of change of output voltage.

Addition of Derivatives of Error. That the addition of time-derivatives of the error quantity can stabilize a servo-system may be shown by considering the response of a simple lagging element such as might occur in the forward control sequence. For an inductive field circuit having a voltage proportional to the error impressed on it, the rise of current is given by

$$Ri + Ldi/dt = K_1\varepsilon.$$

If we add a component proportional to the rate of change of the error, this becomes

$$Ri + Ldi/dt = K_1\varepsilon + K_2d\varepsilon/dt,$$

so that by making $L/R = K_2/K_1$, the response becomes instantaneous and the time lag has been cancelled. This is not necessarily the best procedure in practice but it gives approximately the amount of derivative control to be used. Similarly from the standpoint of the output motion, by feeding in a term proportional to the derivative of error, the motor has two components of torque. The action of the two components can be seen by considering the torques acting as the output overshoots from an initial point behind the input, that is, initially positive error. As zero error is being approached by the output with increasing velocity, the derivative component opposes the normal error component and provides an increasing retarding torque reaching maximum at zero error, while immediately after this point the normal and derivative components act together. This provides an effective stabilizing action without affecting the steady-state velocity error under constant-velocity following. Practical methods of obtaining the derivative of error term are shown in Fig. 15*a* and *b*. Both networks operate on D.C. and are inserted between D.C. stages of the thermionic amplifier. It should be stated that network *b* produces only an approximation to the "error+derivative of error" condition, but this in no way makes it any less effective. Theory indicates there is no particular advantage in a *pure* derivative component.

Proportional Feedback over a Time-Lag. This simple method is used partially to counteract the effects of time-lags as they occur. By opposing the input with the output or a proportion of it, the effective time lag of the modified element is reduced but at the expense of steady-state gain. (Appendix II.)

Load Compensation. Two types of vibration damper are in use for servo stabilization, resonant and non-resonant types. In the former an inertia disc with suitable damping is spring-coupled to the output shaft. The spring constant and inertia are so chosen to give a natural frequency equal to that of the oscillation requiring damping. In these conditions the disc removes energy from the hunting shaft and dissipates it in its viscous brake. In the non-resonant type an inertia is friction-coupled to the output, which therefore experiences friction

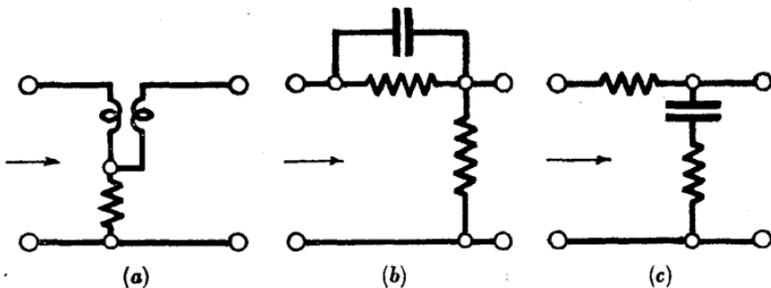


Fig. 15.—Cascade (or interstage) stabilizing networks.

torque whenever there is relative velocity between it and the coupled inertia. This occurs when output oscillates, since the inertia disc cannot follow under these conditions. The energy of the output oscillation is thus converted into heat and the output shaft stabilized. For small servos this method of stabilization is very effective.

Reduction of Steady-State Error. The primary method of reducing steady-state error is to have as large a value as possible of the control sensitivity. Given that a certain servo is stable but suffers from undue velocity-error, one solution of the problem is to incorporate in the amplifier a network which reduces the amount of signal passed for fairly rapid changes of the error but does not affect the transmission in the steady state. By this means the sensitivity of the amplifier may be increased without affecting the original stability but at the same time

providing a higher steady-state amplification than initially. The steady-state errors are thereby reduced. The circuit in Fig. 15c which operates on D.C. signals is commonly used and sometimes referred to as an "error + integral of error" network.

Since a zero velocity-error system depends basically on an acceleration-control, the connection of a simple hydraulic integrating relay in series with either a pump- or valve-controlled power drive will give the necessary characteristic, but will require stabilizing in view of its inherent instability. An arrangement much used in oil servo units is shown in Fig. 16. Here the simple integrating relay is provided with "transient dis-

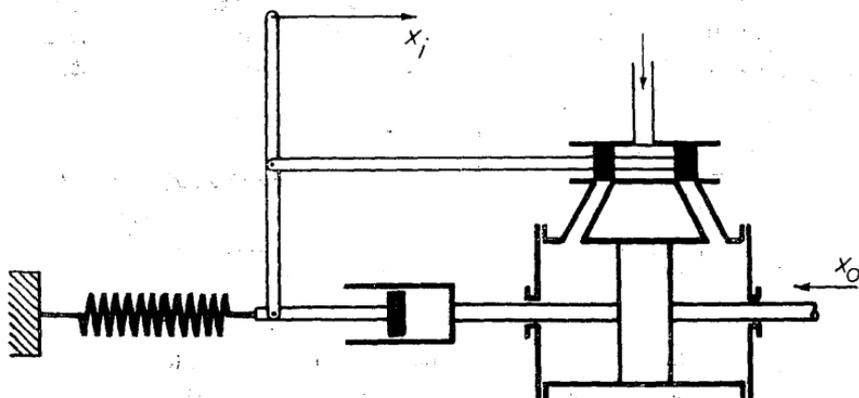


Fig. 16.—Hydraulic integrating relay with transient displacement feedback

placement feedback" by means of the spring-dashpot combination. Since in the steady state, the spring will always return the end of the lever to its zero position, the result of a constant displacement of the input end of the lever is to cause the piston to move with constant velocity. The steady-state integrating property therefore remains. Due, however, to the "solid" action of the dashpot when initial movement of the piston takes place, the valve temporarily closes before being slowly displaced to its steady-state position by the spring. The result of this transient behaviour is highly stabilizing due to the initial peak of velocity given to the piston.

As previously stated, although producing the same result certain practical methods of stabilization possess distinct advan-

tages over others. This particularly applies in the comparison of cascaded networks in contrast to additional feedback loops. In the latter, since the signal is derived from a later stage in the sequence there is always sufficient power available for supplying the feedback network. A cascaded network on the other hand must be arranged so that it does not affect the purity of the error signal and if more than one network is required these must be placed between stages of the amplifier, and not directly connected in cascade. There is the additional feature that corrective networks normally operate with D.C. signals and thus a feedback connection from rotating D.C. machinery is particularly applicable. Even if a feedback source does not exist in the sequence it can be provided, for example, by a D.C. tachogenerator. Finally, feedback stabilization usually makes the system less responsive to output disturbances than cascaded networks, due to information being more directly provided at the input end of the control sequence. The action, however, of feedback networks is less obvious than with networks cascaded in the control sequence, and the best stabilizing arrangement is very often determined by experience or by trial and error rather than wholly analytical considerations.

APPENDIX I

PHASE SHIFT AND ATTENUATION DUE TO SIMPLE EXPONENTIAL TIME LAG

In the differential equation relating the input x_i and output x_o of the delay element, whose time constant is T seconds,

$$T dx_o/dt + x_o = x_i$$

put $x_i = x_{im} \sin \omega t$.

Then if $x_o = A \sin \omega t + B \cos \omega t$ in the steady-state,

$$\omega TA \cos \omega t - \omega BT \sin \omega t + A \sin \omega t + B \cos \omega t = x_{im} \sin \omega t$$

$$\therefore A - \omega BT = x_{im}$$

$$A + B/\omega T = 0,$$

whence $A = x_{im}/(1 + \omega^2 T^2)$, $B = -x_{im} \omega T / (1 + \omega^2 T^2)$

$$\begin{aligned} \therefore x_o &= x_{im} (\sin \omega t - \omega T \cos \omega t) / (1 + \omega^2 T^2) \\ &= x_{im} \sin (\omega t - \Phi) / \sqrt{1 + \omega^2 T^2}, \quad \Phi = \tan^{-1} \omega T. \end{aligned}$$

The phase shift is $\tan^{-1} \omega T$ and the attenuation x_{om}/x_{im} is

$$1/\sqrt{1 + \omega^2 T^2}.$$

APPENDIX II

REDUCTION OF EFFECTIVE TIME-CONSTANT OF A SIMPLE EXPONENTIAL TIME LAG, BY NEGATIVE FEEDBACK.

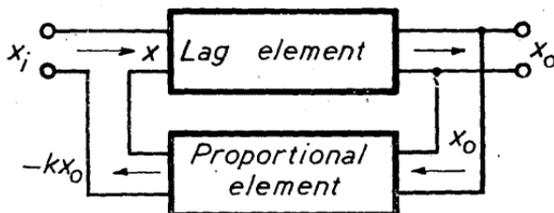


Fig. 17.

For the exponential delay element, with the fraction k of its output negatively fed back,

$$T dx_o/dt + x_o = x = x_i - kx_o$$

$$\therefore T dx_o/dt + (1+k)x_o = x_i$$

$$\text{or } T' dx_o/dt + x_o = x_i / (1+k)$$

where $T' = T/(1+k)$ is the effective time constant. Note that steady-state gain x_o/x_i has been reduced from unity to $1/(1+k)$.

BIBLIOGRAPHY

- (1) "Dynamics of a System of Rigid Bodies," by E. J. Routh. 1877. Macmillan & Co., London.
- (2) "Ueber die Bedingungen, unter welchen eine Gleichung nur Wurzeln mit negativen reellen Theilen besitzt," by A. Hurwitz. *Mathemat. Annalen*, 1895, vol. 46, p. 273.

- (3) "Regeneration Theory," by H. Nyquist. *Bell System Tech. Journ.*, 1932, vol. 11, p. 126.
- (4) "Contributions to the Theory of Automatic Controllers and Followers," by D. G. Prinz. *Journ. Sci. Inst.*, 1944, vol. 21, p. 53.
- (5) "Theory of Servo Systems, with Particular Reference to Stabilisation," by A. L. Whiteley. *Journ. Inst. Elec. Eng.*, 1946, vol. 93 II, p. 353.
- (6) "Naval Applications of Electrical Remote-Positional Controllers," by W. E. C. Lampert. Discussion by F. H. Belsey. *Journ. Inst. Elec. Eng.*, 1947, vol. 94 II A, p. 236 (Convention Paper). A great deal of information is contained in the Proceedings at the Convention on Automatic Regulators and Servo Mechanisms, *Journ. Inst. Elec. Eng.*, 1947, vol. 94 II A, Nos. 1 and 2.
- (7) "Principles of Servo-mechanisms," by G. S. Brown and D. P. Campbell. 1948. John Wiley & Sons, New York.
- (8) "Theory of Servomechanisms," by H. M. James, N. B. Nichols and R. S. Phillips, 1947. McGraw-Hill Book Co., New York.

Discussion

Prof. B. HAGUE, D.Sc., Ph.D. : The Institution is to be complimented on the inclusion of this paper in its *Transactions* ; and the Author is to be congratulated on the clarity and precision with which the principles have been stated. The paper is a clear illustration of the late Prof. Perry's dictum that engineering knowledge should not be separated into watertight compartments, a statement which is particularly true of the subject of servo-mechanisms. In this subject the old principle of servo-control, due to the mechanical engineer, has been combined with a large body of theoretical knowledge, developed by the electrical engineer for the discussion of regenerative action or feedback in electrical amplifiers. The result has been the development of the modern technique of automatic control so ably discussed in the paper.

Two applications of servo-mechanism recently came under the writer's observation. One was the adjustment of the electrode position in an electric furnace, which was carried out by servo equipment operating through an amplidyne generator. The furnace current, which was to be maintained constant, operated a control mechanism with amplification through the amplidyne ; this in turn supplied current to the motors which adjusted the

position of the electrodes so as to keep the furnace current at the desired value. The other application was the automatic control of cutting force and speed of cut in a large planer. This is often done by some form of Ward-Leonard control gear, involving the use of several auxiliary electrical machines. In this particular example the performance of the planer was controlled by a purely electronic servo device, the power supply being taken from a thermionic rectifier.

Mr. V. R. PALING, B.Sc. : The writer would support Prof. Hague in complimenting the Author on his lucid presentation of the subject. Such expositions which make the subject accessible to the layman hardly exist at present. That is unfortunate, because it is a subject which should be of interest to engineers whose main work lies in other fields. The terminology and nomenclature are particularly confusing, and one American writer* has summed up the situation by saying that the control experts have got their terminology so balled up that no one but themselves can understand it. That is a fair statement of the position.

The British Standards Institution is preparing a glossary on this subject,† and a reference to that might usefully be included. The preparation of such a glossary must be a difficult task, and is likely to end up by pleasing no one. In the first section that has been issued, in addition to a list of definitions, a list of rejected terms is given as an appendix, and the rejected terms outnumber the accepted terms.

The, by now, classical papers of Routh, Hurwitz and Nyquist are given in the bibliography, but these are not likely to prove helpful to an engineer who wants to find out what it is all about. None of those authors was writing with control systems in view, and none was writing for engineers. Nyquist indeed was writing for electronic engineers, but everybody knows that they are different from other people and have a language of their own.

On p. 272 the Author divides control systems into those with a fixed predetermined input and those with a random time-varying input. It is desirable to recognize the existence of an intermediate class which has a predetermined input which

* Ed S. Smith. Trans. Amer. Soc. Mech. Eng., 1946, vol. 68, p. 523.

† British Standard 1523. (In course of publication.)

nevertheless varies with time ; that would be the case in a furnace in which the temperature was required to go through a prescribed cycle of operations. Prof. Hayes* has suggested that a more logical classification is simply into systems with a predetermined input and those with a random input.

Mr. J. STEFENSON : The Author has described both electrical and hydraulic servo-mechanisms. How does the relationship between power input and output for the two systems compare with the ratio of say 1 to 10,000 in ordinary hand-controlled plants (for example, tram cars and small marine engines) ? For gun-mountings both systems are used by different firms and in the Denny-Brown stabilizer the first servo stage is electrical and the rest hydraulic. Are there any special reasons for this ?

Mr. D. MORRISON, B.Sc. (Associate Member) : Can the Author comment on the suitability of servo control systems for comparatively small power installations in view of the extra equipment required ? The writer has in mind the particular problem of controlling closely the speed of a D.C. motor for aircraft resonance testing. This can be accomplished by using a 5-h.p. motor with manual speed control and an ample reserve of power, or by using a 1-h.p. motor with automatic speed regulation.

It will be found, however, that the extra weight, space and expense of the control system for the second scheme more than outweighs the original saving on the motor.

With regard to the Author's comments on the capstan torque amplifier, it is the writer's recollection that this system was used for remote control of gun turrets on some German aircraft (ME 210 and 410).†

Mr. A. SILVERLEAF, B.Sc. (Associate Member) : The Author states that the problem of gun stabilization on a rolling ship had not yet been satisfactorily solved. Does this imply that it could be completely solved ? In a true servo-mechanism, is it possible to obtain complete accuracy of control ? The writer has the impression that in servo-mechanisms it is essential to have some error, for it is only through an error that control can

* Proc. Inst. Mech. Eng., 1948, vol. 159, p. 36.

† *Flight*, 1943, vol. 43, p. 144.

be exercised. In this connection, is there any difference in principle between steady-state regulators, and servo-mechanisms linked to a varying datum quantity, which affects the degree of control which it is possible to achieve?

Author's Reply

Mr. CRUICKSHANK : Prof. Hague mentions the common ground which the subject of servo-mechanisms provides for the electrical and the mechanical engineer. In this respect, the two-way nature of the development has perhaps not been fully recognized. While recent theory has predicted seemingly new means of achieving certain practical results, it has also happened that a few of these have been already in existence in the field of mechanical applications and that a closer understanding of their operation is now provided.

Of the two industrial applications cited by Prof. Hague, the first is an example of regulating to a constant value and represents the majority of industrial problems, while the second more nearly approaches true servo action. Industrial servo controls of the second type have on the whole been rather slow in forthcoming, possibly because their complexity offers a certain maintenance problem. There has always been a tendency to avoid installing complex apparatus if accuracy of control is gained at the expense of ease of maintenance. Standardization of components in future control apparatus may successfully solve this problem.

With regard to Mr. Paling's remarks on terminology, the Author agrees with him in sympathizing with engineers whose main work lies in other fields and who may be called on to take up a problem in servo control. The fluid state of nomenclature may be cleared up when the British Standards Institution publishes the glossary he refers to, but it may quite simply mean that process-control engineers and servo-control engineers will have to learn not only each other's language, but also that of the glossary. The Author also agrees with the more logical classification of control systems suggested by Prof. Hayes, which takes in those with predetermined time-varying inputs and which are not mentioned in the paper.

In reply to Mr. Stefenson, the total power amplification provided in a large, high-accuracy angular-position control servo

is very great indeed. The final power-amplifier, whether electric or hydraulic, has, of course, its output determined by the h.p. of the system and an input power in the region of 5 to 25 watts. These figures imply power-amplifications up to 10,000 for the final stage alone of a system having a 50-kW. output. Prior to this, a large power-amplification takes place in the thermionic amplifier in electric and electro-hydraulic systems. As the power-output of a metadyne generator or of a hydraulic amplifier is set only by the physical size of the unit, it is difficult to see how a comparison with the manually-controlled systems can really be made at all.

Turning to the second point made by Mr. Stefenson, namely the relative merits of electric and hydraulic servos, space precludes a detailed answer to the question. Some of the advantages of the hydraulic type, however, are the lighter construction required for a given power, the high torque/inertia ratio of the oil motor which thus enables it to accelerate a load without itself absorbing about an equal quantity of power, as in the case of an electric motor, and the ease of maintenance of hydraulic systems if once properly manufactured. Against these the electric servo can show great flexibility in the thermionic amplifying stage and additional connections for improved responses can be incorporated more easily than in the hydraulic counterpart. It is in the main composed of standard units, for example, motors and generators that can be supplied in large numbers in the normal course of industry. Frequently the best solution is a combination of both electric and hydraulic, the former for the initial stages of error-detection, amplification and primary stabilizing means, while the latter is used for the final power stage. In the Denny-Brown stabilizer the final fin-tilting stage is much the same as a hydraulic steering gear, pre-amplification is by hydraulic means and the mag-slip-resetter system for determining the ship's roll and roll-velocity constitutes the only electrical means employed in the actual control chain. It would be more correct, therefore, to term this a hydraulic system rather than electro-hydraulic.

Mr. Morrison states that a 1-h.p. automatic speed-regulating system for aircraft resonance testing demanded more weight, space and expense than the original system employing a 5-h.p. motor with ample reserve of power. While the space and expense of an extra control system may not be in its favour, it is surprising

that weight considerations of the smaller motor and its equipment should compare unadvantageously. In the scheme mentioned, which concerns mechanical resonances, it is clear that an extremely sensitive control would be required to maintain accurate exciting frequencies in the face of large energy fluctuations in the load. For the holding of such constant exciting frequencies and for the measurement of resulting amplitudes, as opposed to the measurement of merely the resonant frequencies, an electronic control system would be essential.

The Author thanks Mr. Morrison for his interesting reference to the use of the capstan torque amplifier in gun turrets on certain German aircraft.

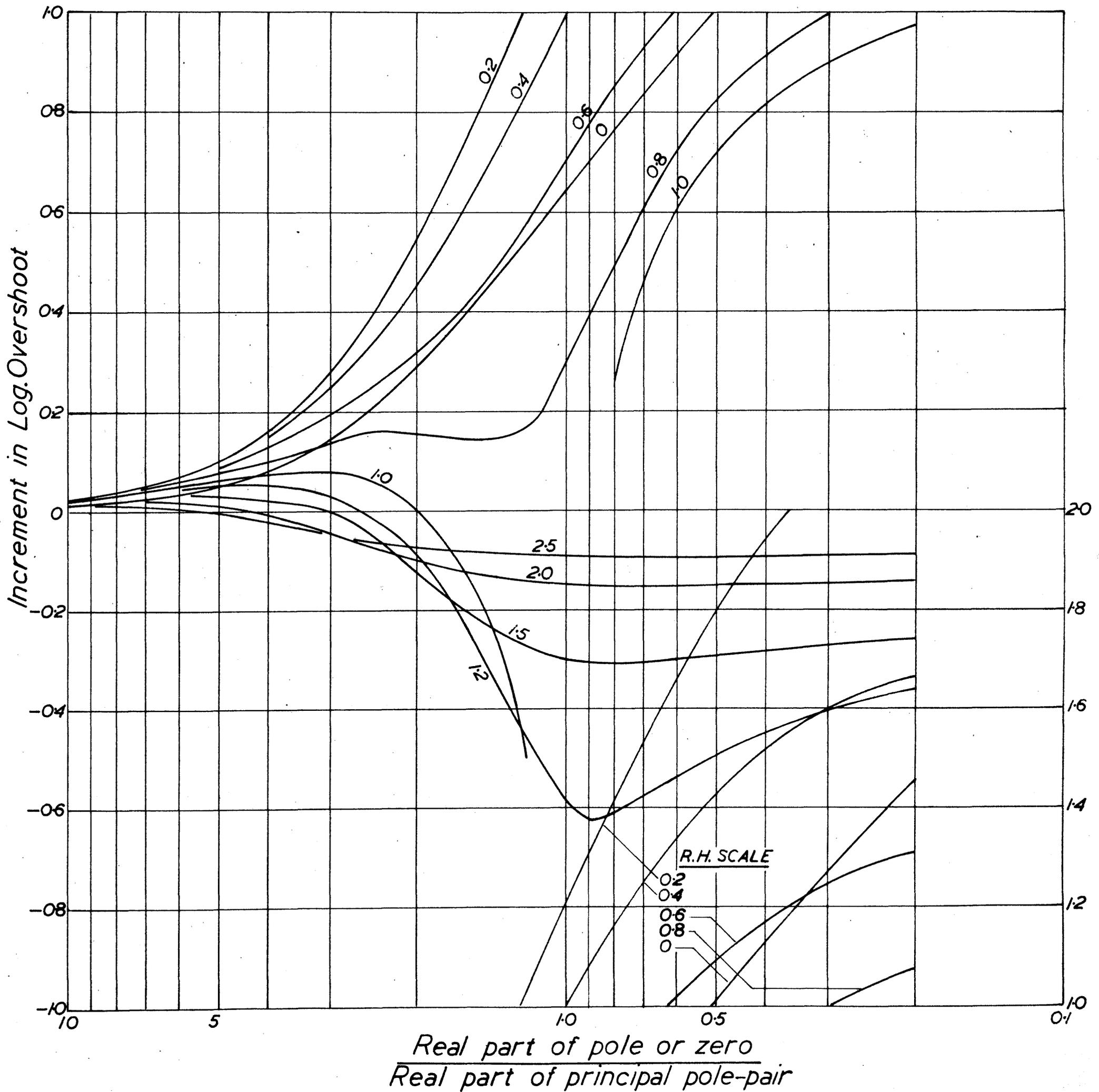
In reply to Mr. Silverleaf, if the Author has given the impression that the complete stabilization of a gun on a rolling ship is possible, he hastens to correct it. Complete accuracy of control cannot be obtained either in stabilizing or in following systems in which the corrective action depends on an error. With regard to regulators, however, which, as stated previously, present mainly steady-state problems, it is true to say that in the steady-state, the accuracy may be brought to within any desired limits however small, and according to the design of the system. There must, however, be certain transient errors before such accuracy or, in the limit, even zero steady-state error, is achieved. Analytically speaking, there is no difference between a steady-state regulator and a servo-mechanism.

A. J. O. CRICKSHANK

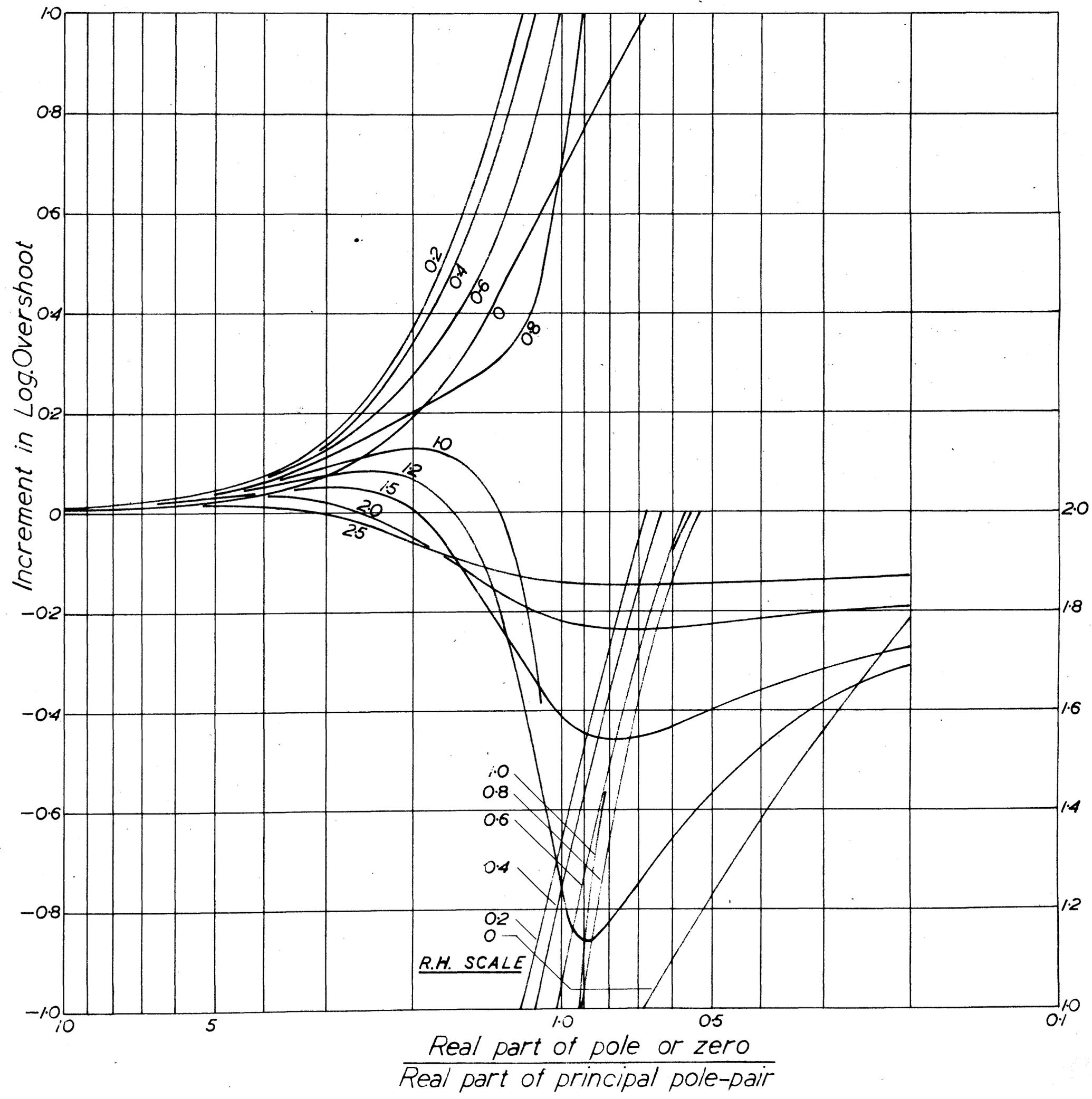
DESIGN CHARTS.

- | | | | |
|----|---|---|---------------|
| 1. | $\theta_o/\theta_i(p)$ | Overshoot Chart 1, for principal pole-pair at | $-0.5 \pm j1$ |
| 2. | $\theta_o/\theta_i(p)$ | Overshoot Chart 2, " " " " " | $-1 \pm j1$ |
| 3. | $\theta_o/\theta_i(p)$ | Overshoot Chart 3, " " " " " | $-2 \pm j1$ |
| 4. | Angle Chart 1, for principal pole-pair at | | $-0.5 \pm j1$ |
| 5. | Angle Chart 2, " " " " " | | $-1 \pm j1$ |
| 6. | Angle Chart 3, " " " " " | | $-2 \pm j1$ |
| 7. | $\epsilon/\theta_i(p)$ | Overshoot Chart 1, for principal pole-pair at | $-0.5 \pm j1$ |
| 8. | $\epsilon/\theta_i(p)$ | Overshoot Chart 2, " " " " " | $-1 \pm j1$ |
| 9. | $\epsilon/\theta_i(p)$ | Overshoot Chart 3, " " " " " | $-2 \pm j1$ |

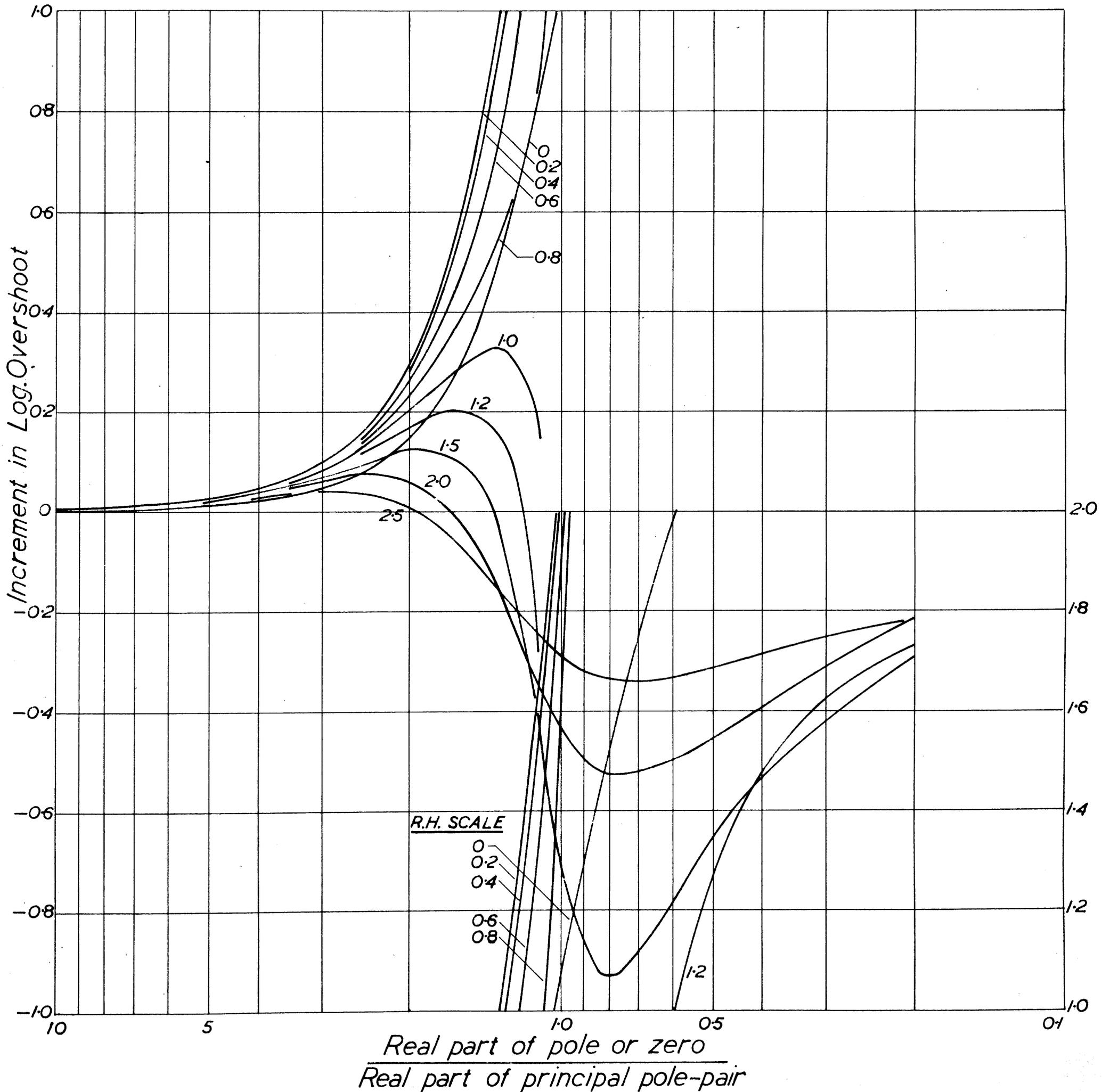
θ_0/θ_1 OVERSHOOT CHART 1
Principal pole-pair at $-0.5 \pm j1$



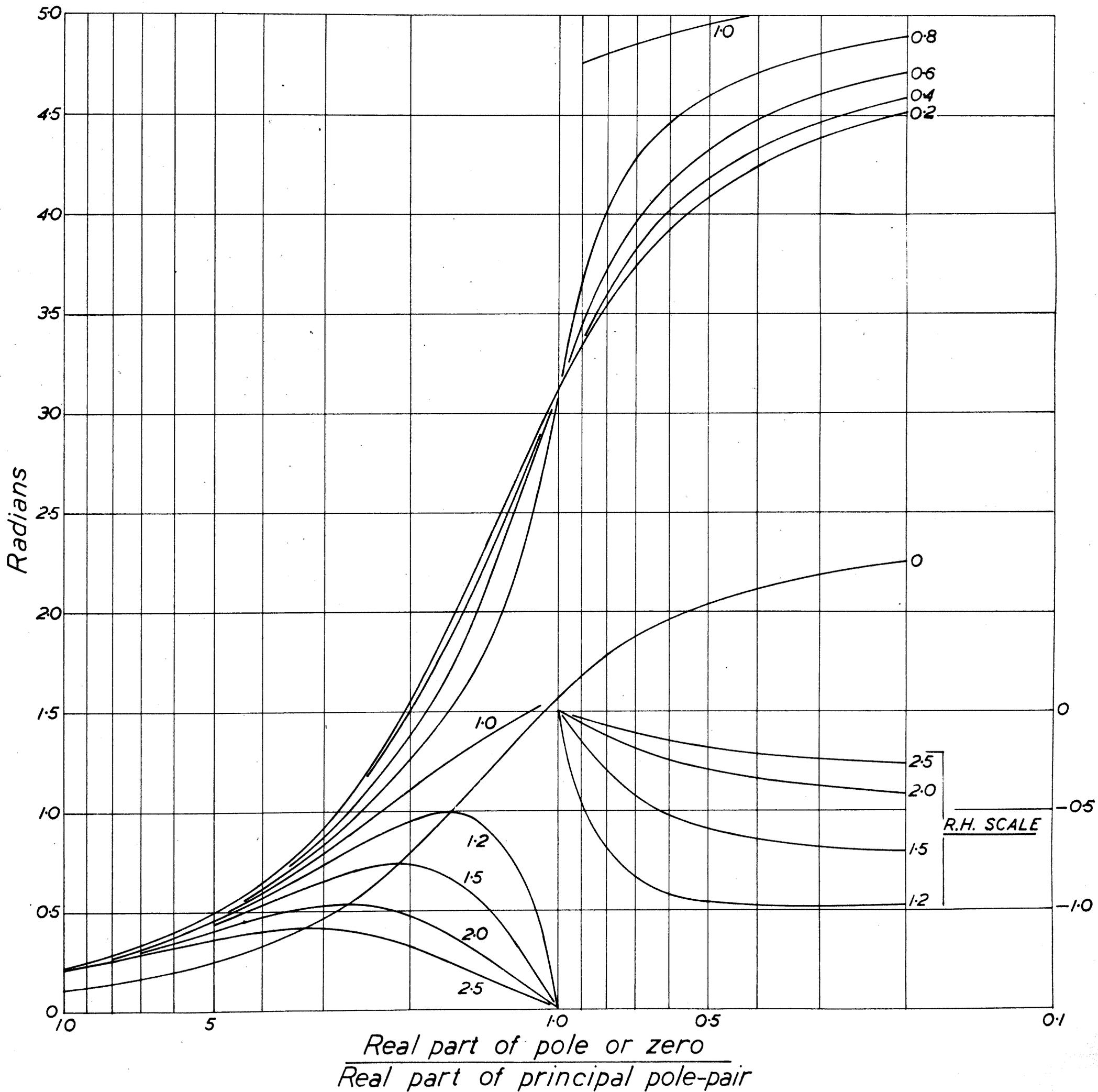
θ_2/θ_1 OVERSHOOT CHART 2
Principal pole-pair at $-1 \pm j1$



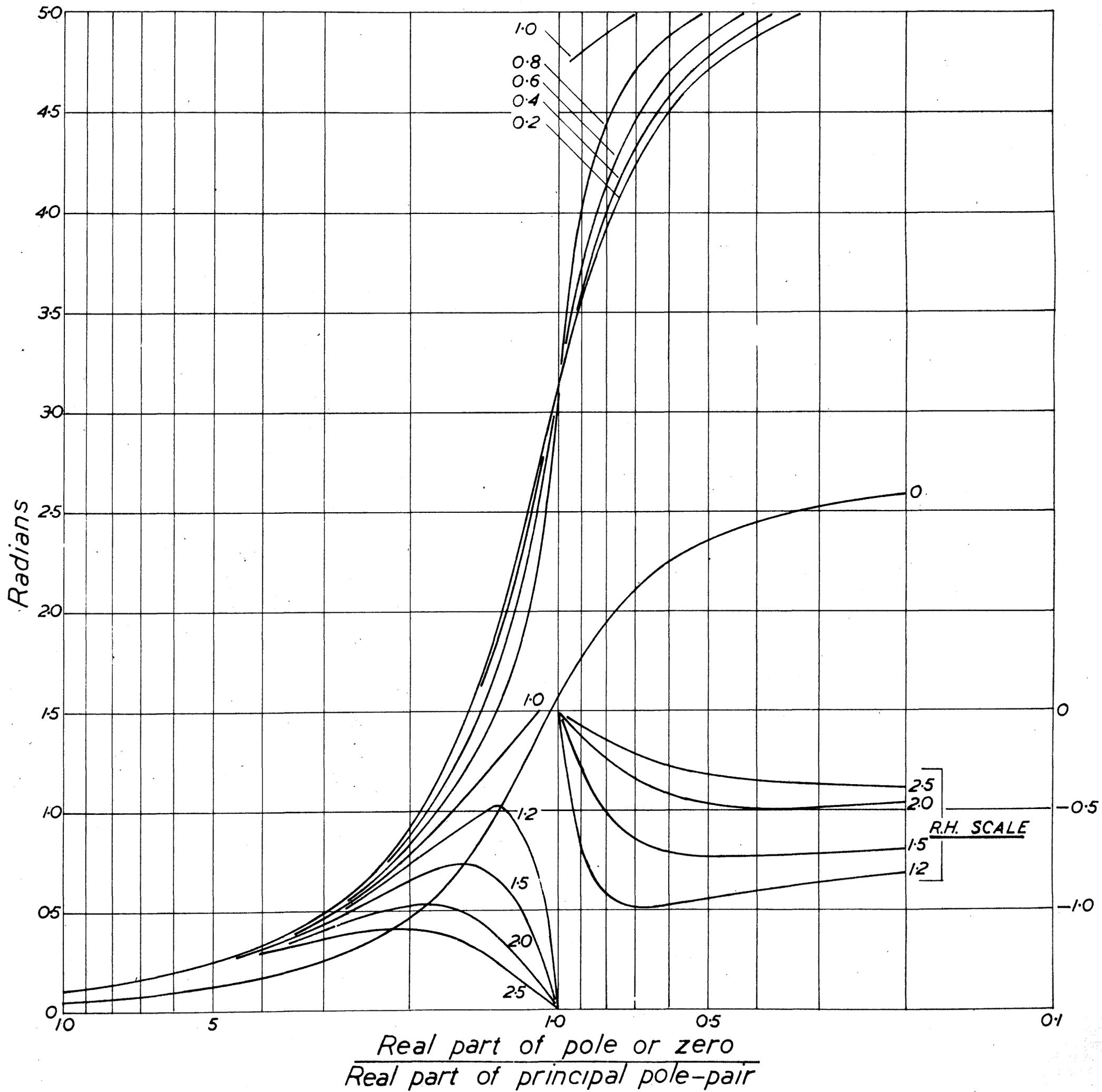
θ_0/θ_1 OVERSHOOT CHART 3
Principal pole-pair at $-2 \pm j1$



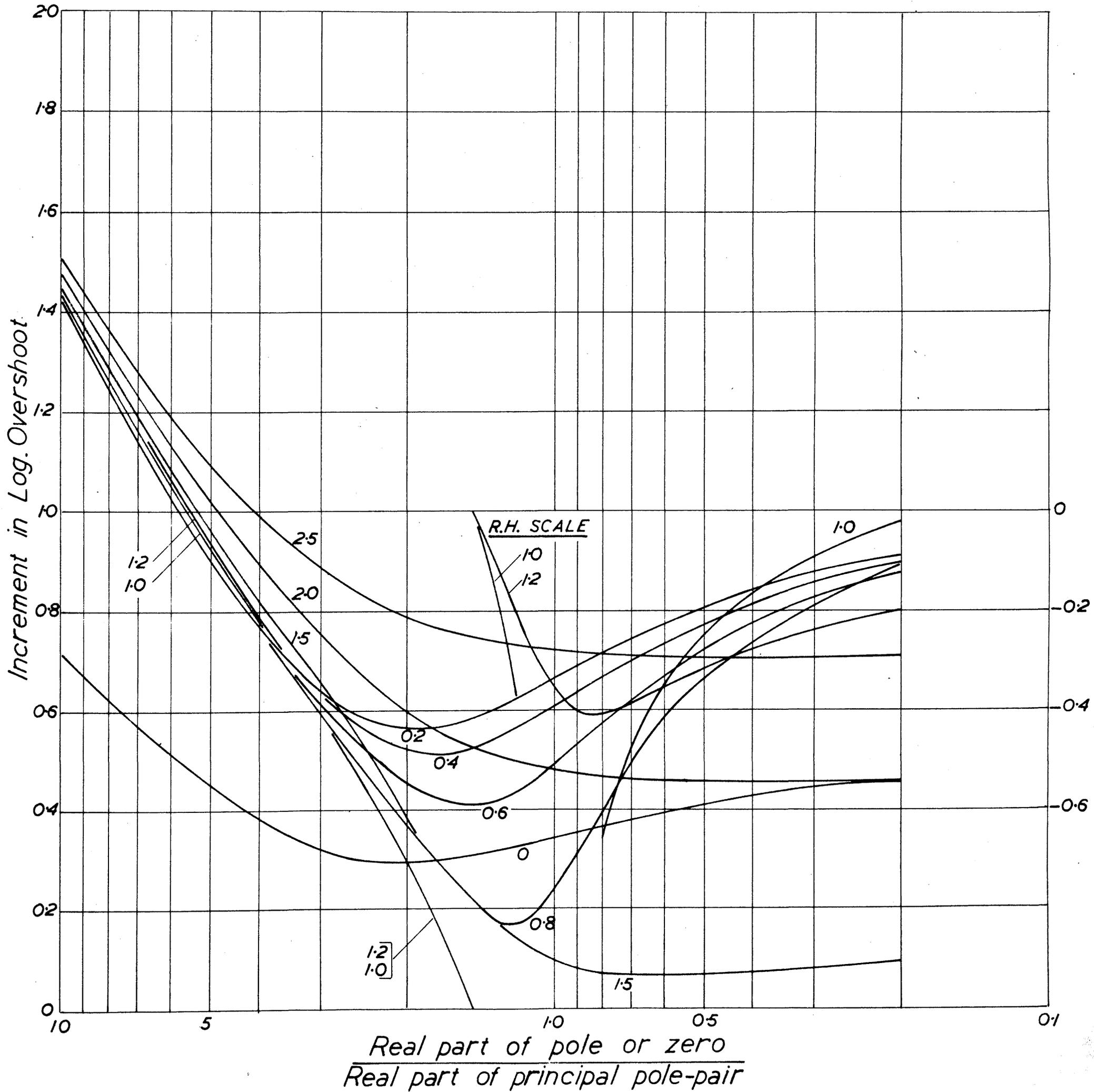
ANGLE CHART 2
Principal pole-pair at $-1 \pm j1$



ANGLE CHART 3
Principal pole-pair at $-2 \pm j1$



ϵ/θ , OVERSHOOT CHART 1
 Principal pole-pair at $-0.5 \pm j1$



ϵ/θ OVERSHOOT CHART 3
Principal pole-pair at $-2 \pm j1$

