Journal of The Franklin Institute

Devoted to Science and the Mechanic Arts

Vol. 218 SEPTEMBER, 1934 No. 3

THEORY OF SERVO-MECHANISMS.

BY

H. L. HAZEN, Sc.D.,

Massachusetts Institute of Technology.

INTRODUCTION.

In this age characterized by huge resources of mechanical and electrical power, these agencies have in many fields almost completely replaced human muscular power. In a similar way the functions of human operators are being taken over by mechanisms that automatically control the performance of machines and processes. Automatic control is often more reliable and accurate, as well as cheaper, than human control. Consequently the study of the performance of automatic control devices is of particular interest at the present time.

Automatic control devices are of two principal kinds. In the first kind the control is actuated by some quantity such as time which is more or less independent of the result of the control operation. Time-operated traffic-signal control is an example of this first kind. With this type of control the flow of traffic resulting from the action of the signal lights in no way affects the cycle of operation of the lights. In the second kind, the control is actuated by a quantity that is affected by the control operation, and for this reason it may be called a "closed-cycle" control. An example of this kind is the thermostat in which the temperature-sensitive element controls the amount of heat supplied to some object such as a process furnace or a house, while the temperature of the object actuates the thermostat. The first kind of control may be superimposed upon the second kind as in the case of a house thermostat whose setting is changed periodically by a clock. The second kind is also illustrated by certain traffic light controls in which the cycle of light operation is automatically adjusted by the flow of traffic. Another example of the second kind of control is automatic ship steering. Here the rudder operation is actuated by the angle between compass and hull, while this angle is in turn affected by the rudder operation.

The distinction between these two kinds of control is made because of important differences in the nature of the control mechanism required in the two cases. In the first kind the source of the control usually has sufficient power to operate the control mechanism directly. Thus synchronous motors are available which can cheaply and accurately effect time control for practically any device. In the second kind, however, the source of control is some form of measuring instrument which very seldom can deliver any appreciable force or power and give usefully accurate indications. The operation of the actual control element nearly always requires much greater power than that available from the measuring instrument. Consequently some intermediate device for amplifying the power of the measuring instrument, while preserving its indications, is essential. When the output element of such a device is so actuated as to make the difference between the output and input indications tend to zero the device is called a "follow up" mechanism or servo-mechanism. In what follows a "servo-mechanism" is frequently called merely a "servo."

In the thermostat or controlling pyrometer,[1, 2, 3] the measuring instrument takes various forms. One of the simplest is the bimetal element or other differential expansion device in which the indication is a mechanical displacement. Other forms are the thermocouple and the temperature-sensitive resistance in which the indication appears as an electric potential difference or current. In nearly all cases, the energy associated with these indications is very small while the actual control is effected by a valve, switch, rheostat, damper, or other device requiring considerable energy for its operation. A servo-mechanism or other amplifier bridges this gap in energy level.

---

[1] For numbered references see bibliography.

Considering the ship-steering example, this disparity exists between the energy magnitude associated with the measuring instrument, a compass, and that associated with driving the rudder. In small craft with a direct wheel-to-rudder drive, the helmsman serves as a human servo-mechanism. In larger craft, the steering engine is a servo-mechanism between wheel and rudder, while the helmsman is still a human servo-mechanism between compass and wheel. With automatic steering the helmsman is replaced by a servo-mechanism which in turn controls the steering engine. Alternatively, the entire compass-to-rudder drive may be considered as a single servo-mechanism.[4, 5, 6, 7, 8]

Thus it is seen that a servo-mechanism is very likely to form an essential part of any closed-cycle control system. Servo-mechanisms are also used merely to produce the indications of measuring instruments at a relatively high power level. Many recording instruments contain illustrations of this use.[1, 2, 3, 22, 23]

Applications of servo-mechanisms, in addition to those already mentioned, include the speed control of steam turbines and water wheels by governors;[9] the control of these governors by master clocks or by power-indicating instruments;[9, 10, 11, 12] the stabilization of ships by gyroscopes;[13, 14, 15, 16, 17] the operation of gyro-compass repeaters;[6, 7] the automatic stabilization and guiding of aircraft;[18, 19, 20, 21] and in fact the automatic recording or control of almost any measurable or measurable and controlable physical quantity.[1, 2, 3, 22, 23, 24, 25, 26, 27, 28]

It is apparent that servo-mechanisms form a vital link in the application of automatic control and that a study of their performance is of importance. The purpose of this paper is to present an analysis of the operating characteristics of certain important types of servo-mechanisms.

Although the subject of servo-mechanisms has been treated in a qualitative or semi-quantitative way in some of the references given above, to the writer's knowledge no systematic quantitative treatment of even the simple common types has previously been given.* It seems worth while, therefore, in

* Minorsky's first paper, ref. 4, gives an excellent analysis of the rudder-hull dynamic system in the ship-steering problem which bears some resemblance to the problem of the continuous-control type of servo here treated. Specific references to his results are given later in the paper.

view of the rapidly expanding field of application of servo-mechanisms, to present the beginnings of such a treatment, outlining quantitatively at least a few of the important properties of the familiar types. Preceding the quantitative treatment, a general discussion of servos and of properties of three important types will be given.*

GENERAL CHARACTERISTICS OF SERVO-MECHANISMS.

Before going into further detail regarding the performance of servos it will be well to define more explicitly what is meant by the term servo-mechanism. As stated before, a servo-mechanism is a power-amplifying device. However, its action differs in one essential particular from that of a simple vacuum-tube or mechanical power amplifier.† Such an amplifier preserves approximately a given functional relation, usually linear, between input and output quantities, due to the properties of the amplifier itself. Thus in a good vacuum-tube amplifier, the current output is very closely proportional to the voltage applied to the grid of the first tube. This linearity of response is due to the constancy of the parameters within the amplifier. Any departure from constancy of these parameters affects the relation between input and output directly.

The servo-mechanism differs from the simple amplifier in that the responsibility for the functional relation is not placed directly on the amplifying element of the servo. Here it will be necessary to distinguish between the input to the servo-mechanism, which is the indication of the measuring instrument, and the input to the amplifier element in the servo which is something different. The output of the servo amplifier element can be considered as the output of the servo, however. In a servo-mechanism, the input to the servo amplifier element is connected to the *difference* between the servo input and output. When this difference is finite the servo output is driven in such a manner as to tend to make this difference zero. Thus the only function of the servo am-

---

* In a companion paper, the design and test of a high performance servo-mechanism, designed on the basis of the analysis in the present paper, is given.

† For a description of an ingenious and successful mechanical torque (and hence power) amplifier developed by Mr. Neiman of the Bethlehem Steel Corporation see reference 29, bibliography.

plifier element is to apply sufficient force to the servo output to bring it rapidly to correspondence with the servo input. Such an amplifier element can be a relatively crude affair. In fact it may consist merely of a suitable relay or switch controlling an electric motor. An illustration of this action is furnished by a servo-mechanism used on an early model of a machine for solving differential equations.[30] The servo input was the angle of a rotating shaft, the output was the angle of another rotating shaft. When these two angles differed, a contact started the electric servo-motor which drove the output shaft in the direction to restore coincidence of the angles.

A servo-mechanism *may thus be defined as a power-amplifying device in which the amplifier element driving the output is actuated by the difference between the input to the servo and its output.* An ideal servo-mechanism is one in which the input and output indications (expressed in common units) are equal at all instants of time. Although the ideal servo is never realized in practice, its operation furnishes the standard by which the operation of actual servos is judged.

When compared to the ideal, an actual servo-mechanism is subject to two principal defects, oscillation and lag. Oscillation is a periodic deviation of output from input. Lag is an average or unidirectional deviation. Either or both of these forms of deviation are always present in some degree and it is the purpose of design to reduce these deviations to a magnitude which is negligible in any particular application.

A servo-mechanism is by nature the type of device in which oscillation would be expected to occur if definite preventative means were not employed. To simplify the discussion, assume that the input indication is a constant. The output is acted upon by a force tending to return it to coincidence with the input. In the absence of damping forces such a system would oscillate indefinitely at any amplitude at which it was started. Suitable damping reduces these oscillations to a small amplitude or prevents them altogether, depending on the type of servo. In certain applications, small oscillations may be useful in minimizing the effects of friction by making the dynamic rather than the larger static coefficient effective. At best this scheme reduces one evil at the cost of introducing another. The friction evil might better be treated by increasing the ratio of driving torque to friction torque. Continuous oscillation necessarily introduces relatively large wear and tear in moving parts. Oscillation, then, is to be tolerated only as an evil justified by cost or by being the lesser of two optional evils.

Lag occurs in some degree in practically all servo-mechanisms when the input has motion. That is, the output indication at a given instant corresponds to that of the input a short time before. In some applications the effect of lag may be negligible while in others it may be very serious.[30] It does no harm, for example, in a recording-instrument servo provided the magnitude of the lag is less than the permissible error, and provided this source of error is taken into account in the design of the instrument. In a closed-cycle control system, however, lag in the servo-mechanism frequently introduces negative damping into the system as a whole, which must be overbalanced by positive damping from some other source if the system is to be stable.* Consequently it is an effect that must be carefully considered.

The quantitative study of oscillation and lag requires a separate analysis for each type of servo. Some servos are inherently oscillatory while others can be made non-oscillatory by suitable design. Three main types in common use are treated in this paper and will be described briefly before proceeding to the mathematical analysis.

The first type of servo-mechanism and one that is widely used because of its simplicity, may be called a relay servo, because of the essentially "off" and "on" nature of the forces acting on the output element. Several forms of gyro-compass repeaters, automatic pilots, and gyro-stabilizers for ships use this type of servo.[6, 13, 14, 15, 18, 20, 21, 27] In this type the restoring force applied to the output element is usually substantially constant in magnitude, while operative. Ideally, this force would be brought into operation by an infinitesimal deviation of the output from the input but practically it is usually necessary to have a small range of deviation over which the restoring force is inactive. If the inactive deviation range were infinitesimal, and if there were no time lag in the applica-

* For example see ref. 4 (1922), pp. 305–8, and ref. 30.

tion of the restoring force, this servo could be made to operate with an infinitesimal amplitude of oscillation, an infinite frequency of oscillation, and an infinitesimal lag error. It is quite evident that such operation represents a limiting case which at best can only be roughly approached with actual physical apparatus. Practically the amplitude and frequency of oscillation are finite, and in most cases lag error is present. The limiting case is of interest however from the point of view of analysis and from its significance as an ideal.

This type of servo has certain disadvantages. The rather sudden application of the entire available driving torque, first in one direction and then in the other is conducive to large wear and tear of the entire mechanism. In practice the restoring force is usually initiated by electric contacts which may be somewhat troublesome especially when only very small forces are available for their operation. This type of servo has the asset of simplicity, however, and may be useful where static friction in the mechanism is troublesome, and a relatively crude type of control suffices.

The second type of servo-mechanism is one in which the correction of the output is made in finite steps at definite time intervals. This type is extensively used in recording instruments and controllers for quantities that vary relatively slowly.[1, 2, 3, 9, 10] It is well illustrated by a temperature recorder in which the position of the pen is periodically tested to ascertain whether or not it corresponds to the indicated temperature. If in error, the pen position is given a small finite change in the correct direction and a short time later its position is again tested, and so on. In this scheme, the magnitude of the step correction is usually made approximately proportional to the deviation between output and input. When properly designed and adjusted this type of servo is non-oscillatory. The necessary conditions for non-oscillation are: first, that any correction and the indication of it shall be substantially completed before a new test is made; second, that the correction applied shall not be greater than that required to reduce the deviation approximately to zero; and third, that the inactive range over which the output can vary without causing a correction to be made shall be at least as large as the smallest possible correction step.

A modification of this type is used in an automatic steering gear suitable for small craft.[8] The rudder, which is the output of the servo, is periodically displaced from mid position an amount approximately proportional to the error in the ship's heading, and then allowed to trail back to mid position before the heading is again tested and another temporary displacement is given to the rudder.

Evidently the finite-step or impulse type of servo-mechanism can be non-oscillatory. However it does have a persistent lag error when the input is so varying that corrections must be applied continually in one direction. It is also inherently somewhat slower than the other types of servos, since in order to be non-oscillatory, the driving force can be effective during only a fraction of the cycle of operation.

The third type of servo considered here is one in which the restoring force, acting continuously on the output element, is approximately proportional to the deviation of the output.

This type is rapidly coming into use and is now employed in a number of devices such as certain gyro-compass followers,[6, 7] and various automatic recorders.[22, 23, 24, 25, 26] By the use of suitable damping, its action can be rendered aperiodic or oscillatory to any desired degree. Where a high speed of response, high sensitivity, and freedom from hunting are desired this type undoubtedly has the greatest possibilities of the three. Small deviations call into play only small restoring forces, hence its operation lacks the somewhat violent nature characteristic of relay servos. Moreover, these small forces can be called into play promptly and large deviations are thereby avoided. This type can be built to have very rapid response.

A servo-mechanism of this third type, described in a companion paper, capable of operating on the output of a photo-electric cell and of delivering about one-tenth of a horse power has been built which substantially completes the correction of a small deviation in about one-twentieth of a second. If desired this speed could undoubtedly be improved by a factor of five using the same basic design. This servomechanism as normally operated is aperiodic in response.*

---

* See footnote on p. 282.

Having discussed the general characteristics of servo-mechanisms, and the method of operation of three of the most important types, attention is turned to the analysis. In what follows only a very simple dynamic system is treated in detail, one in which friction, inertia, and a restoring force are associated with the output element. This representation is very close to the truth for many servos. For example, in the case of the continuous-control servo referred to above, test and calculation agreed within a small experimental error. In other cases, the idealizations made in order to make a usefully simple analytical treatment possible may depart somewhat more from the facts. Nevertheless, the analysis of an idealized case gives a real insight into the characteristics of a given servo. In most specific cases in which a general analytical solution is too cumbersome to be useful, a restricted analytical or a numerical solution can be readily made, taking into account departures from more idealized conditions. The methods of treating these more involved cases will be outlined briefly at suitable points in the analysis. The first type to be considered is the relay-servo.

RELAY-TYPE SERVO-MECHANISMS.

The relay type of servo-mechanism is characterized primarily by a constant-magnitude restoring force brought into play when the output deviates by some predetermined amount from the input. This characteristic alone does not fix the performance however as the effects of a number of other factors are important. Among these others are included: first, the type of dynamic system to which this force is applied, which includes one or more inertia and friction parameters, and possibly elastance parameters; second, the nature of the friction parameters, whether the frictional force is independent of, or a function of, the output velocity; third, the inactive range, or the limits of the output deviation within which the restoring force is not in action; fourth, the time delay between the indication that the restoring force is to be applied or removed, and the actual application or removal of this force; fifth, the nature of variation of the input, whether this variation is relatively slow so that the servo maintains the output at a substantially constant value, or is rapid and unidirectional so that the servo must move the output more or less continuously in one direction. Each of these factors must be considered in the analysis of any particular case. A comprehensive treatment of all cases is evidently beyond the scope of this paper, but a few significant cases will be analyzed and from the results obtained deductions can be made covering other cases.

In this section an analysis is made of the cases indicated in the following outline:

Relay Servo-Mechanisms.

A. No inactive zone.
  1. Friction force proportional to velocity (viscous friction).
  2. Friction force independent of velocity (Coulomb friction).

B. Finite inactive zone.
  1. Friction force proportional to velocity.
  2. Friction force independent of velocity.

C. Cases *A* and *B* with time delay.

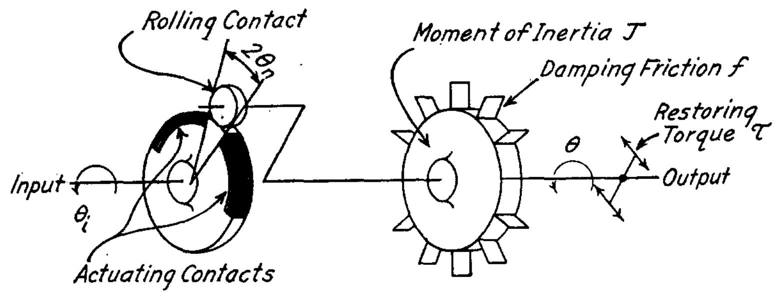
The properties of particular interest are the amplitude of steady-state oscillation, and the magnitude of the average lag error as a function of the input speed. These quantities will be determined in what follows.

Because the output of many servos is in the form of mechanical rotation the analysis is given in these terms. However, the equations can be used as given for certain other systems, by merely redefining the symbols in terms of the analogous parameters of other types of motion such as mechanical translation or electric current. A more specific discussion of this subject is given further on in the paper.

Case A1. Relay-type servo having no inactive zone and output-element characterized by inertia and viscous friction. Three different conditions of operation of this servo will be considered: first, stationary input and zero time lag in application of restoring force; second, stationary input, and finite time lag; third, uniformly varying input, zero time lag. For each of these conditions the amplitude of steady-state oscillation and the lag error will be determined.

For the first condition consider the action of the physical system shown schematically in Fig. 1.

FIG. 1.

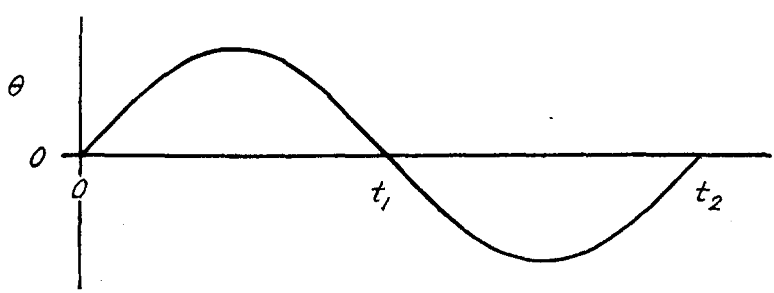


Schematic diagram of the significant dynamic elements of the relay-type servo-mechanism.

Let θ_i = angle of servo input (radians),
θ = angle of servo output (radians),
$2\theta_n$ = inactive range (radians) (= 0 for case A),
J = moment of inertia of servo referred to output (gm. cm.^{2}),
f = friction or damping torque per unit angular velocity of output (dyne cm. per radian per sec.),
$\pm \tau$ = restoring torque acting on output (dyne cm.),
$p = \dfrac{d}{dt}$, the time derivative operator,
$\omega = p\theta$ = angular velocity (radians per sec.).

The restoring torque τ is normally of sign opposite to that of the quantity $\theta_i - \theta$. A sketch of θ for one cycle of steady-state operation is shown in Fig. 2. To find the amplitude of

FIG. 2.



steady-state oscillation it is necessary merely to solve for the motion of the output over the range $0 < t < t_1$ in terms of

ω_0, the value of $p\theta$ at $t = 0$, and to make $p\theta$ at t_1 equal to $-\omega_0$. The resulting equation in ω_0 and the system parameters determines ω_0 for steady-state operation.

For the interval $0 < t < t_1$ the differential equation of motion is

$$-\tau - fp\theta = Jp^2\theta. \tag{1}$$

The known terminal conditions are

$$\left.\begin{matrix} t = 0 \\ \theta = 0 \\ \omega = \omega_0 \end{matrix}\right\}, \quad (2) \qquad \left.\begin{matrix} t = t_1 \\ \theta = 0 \\ \omega = -\omega_0 \end{matrix}\right\} \quad (3)$$

Integrating (1) and inserting conditions (2) there results

$$\omega = -\omega_s + (\omega_s + \omega_0)\epsilon^{-t/T}, \tag{4}$$

$$\theta = -\omega_s t + T(\omega_s + \omega_0)(1 - \epsilon^{-t/T}), \tag{5}$$

in which

$$\omega_s = \frac{\tau}{f} = \text{runaway velocity of output,}$$

$$T = \frac{J}{f}.$$

Solving (4) for t_1 with $\omega = -\omega_0$ and using this value of t_1 in (5) for which $\theta = 0$, there results

$$\ln\frac{1 - \dfrac{\omega_0}{\omega_s}}{1 + \dfrac{\omega_0}{\omega_s}} + 2\frac{\omega_0}{\omega_s} = 0. \tag{6}$$

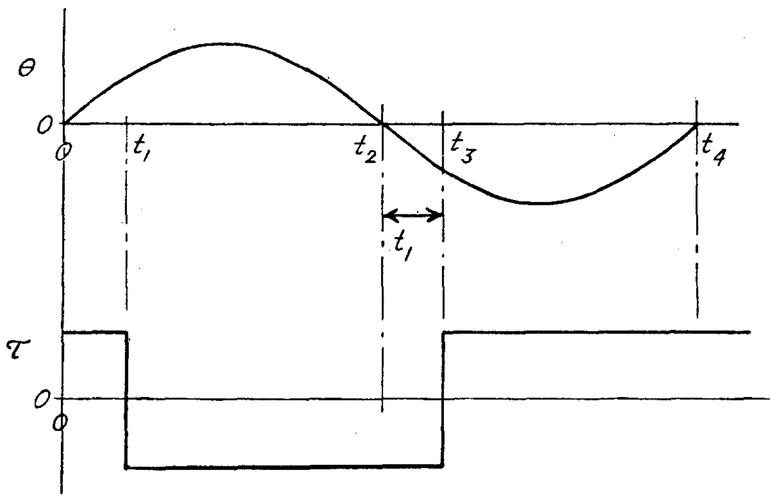
Equation (6) is satisfied only for vanishingly small values of ω_0/ω_s. Since ω_s is finite, ω_0 must be vanishingly small. From this result it is seen that the servo output in this case oscillates, but at an infinitesimal amplitude and an infinite frequency. Physically this result is absurd for any actual servo, hence the necessary conclusion is that this case is too greatly idealized to represent the facts. It is of interest, however, as a limiting case.

Because (6) is very nearly satisfied for small finite values of ω_0/ω_s, it is to be inferred that only a slight departure from the assumed conditions could result in a finite amplitude and frequency of oscillation.

By taking time lag in the change of the restoring force into account, a practical case results. As the second condition then the results of time lag will be considered.

This condition differs from the one just treated only in that the restoring torque is changed in sign, not when $\theta = 0$ but t_1 seconds later. A cycle of output angle and restoring torque is shown in Fig. 3a. The plan of solution is similar to that just used.

FIG. 3.



Restoring torque and output angle for case A1 with time lag present. Input stationary.

For the interval, $0 < t < t_1$, the differential equation of motion is

$$\tau - fp\theta = Jp^2\theta. \tag{7}$$

For the interval $t_1 < t < t_3$ this changes to (1). The known terminal conditions are:

$$\left.\begin{array}{l} t = 0 \\ \omega = \omega_0 \\ \theta = 0 \end{array}\right\} \quad (8) \qquad \left.\begin{array}{l} t = t_2 \\ \omega = \omega_2 = -\omega_0 \text{ for steady state} \\ \theta = 0 \end{array}\right\} \quad (9)$$

and that at $t = t_1$, ω and θ are continuous. Integrating (7) and (1), inserting the terminal conditions (8) and (9), and equating the two resulting expressions for ω at $t = t_1$, there results

$$\epsilon^{-t_1/T} + \epsilon^{(t_2-t_1)/T} = \frac{2\omega_s}{\omega_s - \omega_0}. \tag{10}$$

Equating the two expressions for θ at $t = t_1$,

$$\omega_s(2t_1 - t_2) + T(\omega_0 - \omega_s)(2 - \epsilon^{-t_1/T} - \epsilon^{(t_2-t_1)/T}) = 0. \tag{11}$$

Substituting (10) in (11), solving the result for t_2 and putting this in (10), gives, after some algebraic reduction, the following equation in t_1 and ω_0:

$$a^2(1 - \Omega_0)\epsilon^{2\Omega_0} - 2a + (1 - \Omega_0) = 0, \tag{12}$$

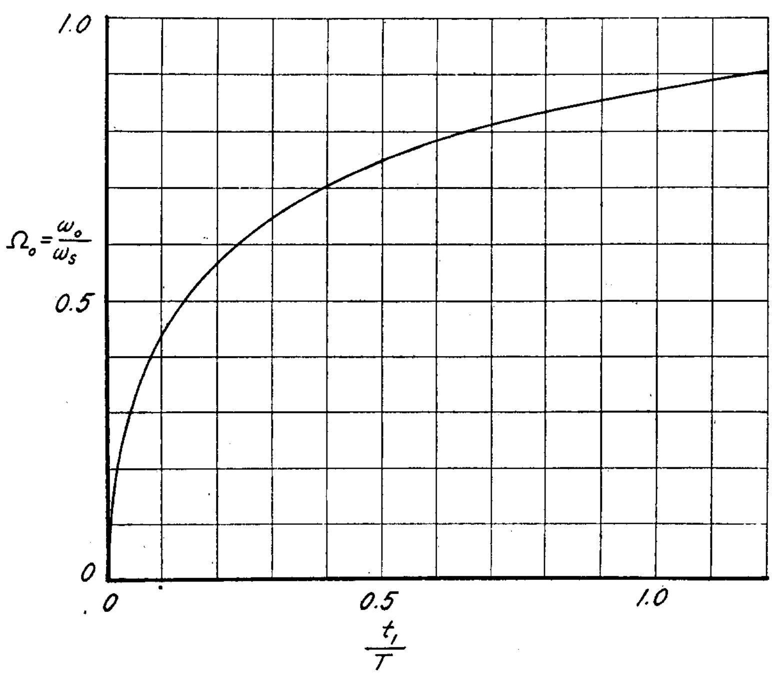
in which

$$a = \epsilon^{t_1/T},$$

$$\Omega_0 = \frac{\omega_0}{\omega_s}.$$

This relation between ω_0 and t_1 plotted in dimensionless form as Ω_0 vs. t_1/T is shown in Fig. 4. The conclusion drawn by

FIG. 4.



Amplitude of steady-state oscillation as a function of time lag for case A1. Input stationary.

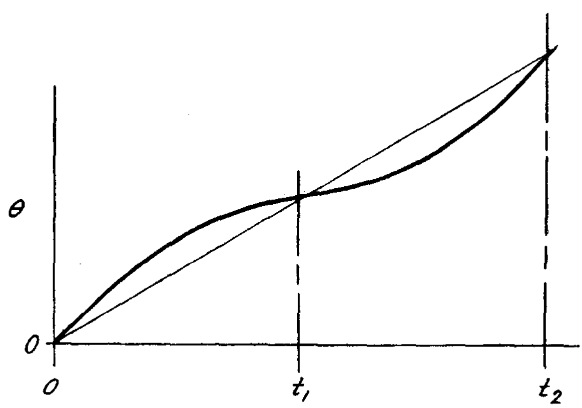
inference from (6) above that ω_0 might acquire a relatively large finite amplitude for even a small departure from the assumed ideal conditions is well borne out by Fig. 4. Evidently if a small amplitude is desired from this type of servo, it is very important to make the time lag in the application of the restoring-torque change very small. This lag is the effective lag due to all causes such as the time required to actuate relays and that required to establish a torque after the necessary contacts, or valves, etc., have been operated.

The third condition will next be considered. This condition is similar to the first, i.e., time lag is not considered, but the input instead of being stationary is moving at a constant speed,

$$\left.\begin{aligned} p\theta_i &= \omega_a \\ \theta_i &= \omega_a t \end{aligned}\right\}, \tag{13}$$

where ω_a is constant. A curve of the input angle θ_i and the output angle θ for a sample cycle beginning at $t = 0$ is shown in Fig. 5. The differential equations of motion are

FIG. 5.



$$-\tau - fp\theta = Jp^2\theta \quad \text{for} \quad 0 < t < t_1 \tag{14}$$

and

$$\tau - fp\theta = Jp^2\theta \quad \text{for} \quad t_1 < t < t_2. \tag{15}$$

For steady-state operation the known terminal conditions are

$$\left.\begin{aligned} t &= 0 \\ \omega &= \omega_0 \\ \theta &= 0 \end{aligned}\right\}, \quad (16) \qquad \left.\begin{aligned} t &= t_1 \\ \omega &= \omega_1 \\ \theta &= \omega_a t_1 \end{aligned}\right\}, \quad (17) \qquad \left.\begin{aligned} t &= t_2 \\ \omega &= \omega_0 \\ \theta &= \omega_a t_2 \end{aligned}\right\}. \quad (18)$$

Intergrating (14) and using the terminal conditions (16), there results for $0 < t < t_1$:

$$\omega = -\omega_s + (\omega_s + \omega_0)\epsilon^{-t/T}, \tag{19}$$

$$\theta = -\omega_s t + T(\omega_s + \omega_0)(1 - \epsilon^{-t/T}). \tag{20}$$

For the interval $t_1 < t < t_2$ the corresponding expressions as obtained from (15) and (17) are

$$\omega = \omega_s - (\omega_s - \omega_1)\epsilon^{-(t-t_1)/T}, \tag{21}$$

$$\theta = \omega_s(t - t_1) + \omega_a t_1 - T(\omega_s - \omega_1)(1 - \epsilon^{-(t-t_1)/T}). \tag{22}$$

From the relations (17) to (22) the steady-state amplitude of oscillation can be obtained in terms of the relation of ω_0 and ω_1 to ω_a, that is in terms of the relation of the maximum and minimum angular velocities to the average angular velocity of the output. Let

$$\left.\begin{aligned} \frac{\omega_a}{\omega_s} &= \Omega_a \\ \frac{\omega_0}{\omega_s} &= \Omega_0 \\ \frac{\omega_1}{\omega_s} &= \Omega_1 \end{aligned}\right\}. \tag{23}$$

Using (17) in (19) and expressing the result in terms of (23),

$$\epsilon^{-t_1/T} = \frac{1 + \Omega_1}{1 + \Omega_0}. \tag{24}$$

(17) in (20) gives

$$\frac{t_1}{T} = \frac{1 + \Omega_0}{1 + \Omega_a}(1 - \epsilon^{-t_1/T}). \tag{25}$$

Using (24) in (25) and solving the result for t_1/T, and substituting this in (24) gives one equation relating Ω_a, Ω_0 and Ω_1 thus:

$$\frac{1 + \Omega_1}{1 + \Omega_0} = \epsilon^{-(\Omega_0 - \Omega_1)/(1+\Omega_a)}. \tag{26}$$

Another expression relating these same quantities can be obtained from (21) and (22) with (18). Using (18) in (21),

$$\epsilon^{-(t_2 - t_1)/T} = \frac{1 - \Omega_0}{1 - \Omega_1}. \tag{27}$$

(18) in (22) gives

$$\frac{t_2 - t_1}{T} = \frac{1 - \Omega_1}{1 - \Omega_a}(1 - \epsilon^{-(t_2 - t_1)/T}). \tag{28}$$

Using (27) in (28) and solving the result for $(t_2 - t_1)/T$ and substituting this in (27) gives the second equation relating Ω_0 and Ω_1 to Ω_a. This equation is

$$\frac{1 - \Omega_0}{1 - \Omega_1} = \epsilon^{-(\Omega_0 - \Omega_1)/(1-\Omega_a)}. \tag{29}$$

Equations (26) and (29) taken simultaneously suffice to determine Ω_0 and Ω_1 in terms of Ω_a. No general analytical solution can be obtained, at least by elementary means. However from the result obtained for the previous case, i.e., for $\Omega_a = 0$ it is worth trying to see if a small amplitude of oscillation satisfies (26) and (29). To do this let

$$\begin{aligned} \Omega_0 &= (1 + \rho_0)\Omega_a, \\ \Omega_1 &= (1 - \rho_1)\Omega_a, \end{aligned} \tag{30}$$

where ρ_0 and ρ_1 are small positive quantities. Furthermore if ρ_0 and ρ_1 are small, they are presumably approximately equal. If a solution to (26) and (29) can be obtained on the assumption that they are equal, this assumption will be justified. Assume tentatively then that

$$\rho_0 = \rho_1 = \rho. \tag{31}$$

Using (30) and (31) in (26) and (29), the following expressions are obtained:

$$\frac{1 - m\rho}{1 + m\rho} = \epsilon^{-2m\rho}, \tag{32}$$

$$\frac{1 - n\rho}{1 + n\rho} = \epsilon^{-2n\rho}, \tag{33}$$

in which

$$m = \frac{\Omega_a}{1 + \Omega_a}, \qquad n = \frac{\Omega_a}{1 - \Omega_a}.$$

Expanding the left side of (32) by division and the right side by the exponential series,

$$1 - 2m\rho + 2m^2\rho^2 - 2m^3\rho^3 + \cdots$$
$$= 1 - 2m\rho + \frac{4m^2\rho^2}{2} - \frac{8m^3\rho^3}{6} + \cdots. \tag{34}$$

Doing the same with (33),

$$1 - 2n\rho + 2n^2\rho^2 - 2n^3\rho^3 + \cdots$$
$$= 1 - 2n\rho + \frac{4n^2\rho^2}{2} - \frac{8n^3\rho^3}{6} + \cdots. \tag{35}$$

(34) and (35) are satisfied for vanishingly small values of $m\rho$ and $n\rho$ and are approximately satisfied for values of these quantities for which their cubes are negligible in comparison with unity. Barring the limiting case in which $\Omega_a = 1$, i.e., $\omega_a = \omega_s$, (34) and (35) are satisfied by vanishingly small values of ρ. Since ρ is the maximum fractional deviation of the instantaneous output angular velocity from the average or input velocity, it is seen that the oscillations are of vanishingly small amplitude and of infinite frequency.

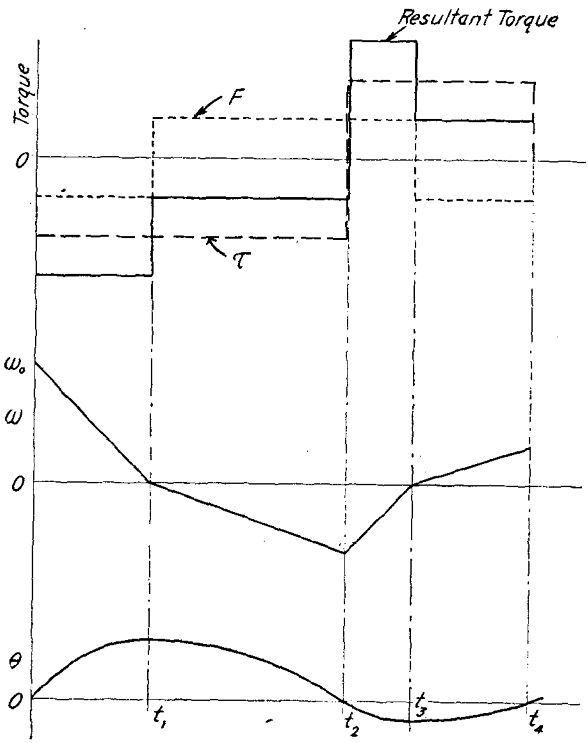
Although as in the first condition, this is an idealized case that cannot be realized physically, the result has interest and significance. It shows that an ideal relay-type servo is capable of following a constant input velocity with precision, i.e., without finite deviation in the nature of lag or oscillation. That this condition could be approached even under ideal conditions is interesting.

The inevitable presence of time lag in the change of the restoring torque causes a finite amplitude and frequency in any actual servo. This fact could be demonstrated by explicit formulation but the analysis will not be given here. Instead the effect of such time lag in this and other cases in which the analysis is unwieldy is discussed further on in the paper.

Case A2. In the case just treated, the friction or damping force was assumed to be proportional to the velocity. In the present case, the frictional force is assumed to follow the Coulomb law, that is, to be independent of velocity. Otherwise the two cases are identical. As was done above, the response of the servo output for the condition of a stationary input will be treated first. As a second condition the effect of time lag is considered with the input stationary. A third condition with constant input velocity and zero time lag is considered last.

In Fig. 6 the variations of the torques, the output angle

FIG. 6.



Torques and angular velocity and angle of output for case A2. No time lag. Input stationary.

and the output velocity during a cycle of operation for the first condition are sketched. As before, the amplitude of steady-state oscillation and the output lag are the quantities of interest.

Let

$$\pm F = \text{friction force, sign opposite to velocity.}$$

The other quantities are as previously defined. The differential equations of motion are

$$-\tau - F = Jp^2\theta = Jp\omega \quad \text{for} \quad 0 < t < t_1 \tag{36}$$

and

$$-\tau + F = Jp^2\theta = Jp\omega \quad \text{for} \quad t_1 < t < t_2. \tag{36a}$$

The known terminal conditions are

$$\left.\begin{aligned} t &= 0 \\ \omega &= \omega_0 \\ \theta &= 0 \end{aligned}\right\}, \quad (37) \qquad \left.\begin{aligned} t &= t_1 \\ \omega &= 0 \\ \theta &= \theta_1 \end{aligned}\right\}, \quad (38) \qquad \left.\begin{aligned} t &= t_2 \\ \omega &= \omega_2 \\ \theta &= 0 \end{aligned}\right\}. \quad (39)$$

Integrating (36) and using (37) to determine the constants of integration give, for the interval $0 < t < t_1$,

$$\omega = -\alpha_1 t + \omega_0 \tag{40}$$

and

$$\theta = -\frac{\alpha_1}{2} t^2 + \omega_0 t, \tag{41}$$

where

$$\alpha_1 = \frac{\tau + F}{J}. \tag{42}$$

Using (38) in (40) to obtain t_1, and putting this value of t_1 in (41) gives

$$\left.\begin{aligned} t_1 &= \frac{\omega_0}{\alpha_1}, \\ \theta_1 &= \frac{\omega_0^2}{2\alpha_1}. \end{aligned}\right. \tag{43}$$

Integrating (36a) and evaluating the constants of integration by (38) and (43) give

$$\omega = -\alpha_2(t - t_1), \tag{44}$$

$$\theta = -\frac{\alpha_2}{2} t^2 + \frac{\alpha_2}{\alpha_1}\omega_0 t + \frac{\omega_0^2}{2\alpha_1} - \frac{\alpha_2\omega_0^2}{2\alpha_1}, \tag{45}$$

where

$$\alpha_2 = \frac{\tau - F}{J}.$$

(39) in (45) gives

$$t_2 = t_1\left(1 \pm \sqrt{\frac{\alpha_1}{\alpha_2}}\right), \tag{46}$$

in which the plus sign alone is of interest since $\alpha_1 > \alpha_2$. Using (46) in (40) it is found that

$$\frac{\omega_2}{\omega_0} = -\sqrt{\frac{\alpha_2}{\alpha_1}}. \tag{47}$$

Since the second half of the cycle is identical in form with the first half, the ratio of initial to final velocity will be the same in both cases. The value ω_4 of ω at $t = t_4$ is then given by

$$\frac{\omega_4}{\omega_2} = \frac{\omega_2}{\omega_0} = -\sqrt{\frac{\alpha_2}{\alpha_1}}$$

and

$$\frac{\omega_4}{\omega_0} = \frac{\alpha_2}{\alpha_1}. \tag{48}$$

Equation (48) shows that if the servo starts with a finite amplitude of oscillation, this amplitude is reduced by a constant factor each cycle, and thus the amplitude approaches zero asymptotically. In the steady-state, therefore, this ideal servo, like the previous one, oscillates with zero amplitude and infinite frequency when the input has a constant value.

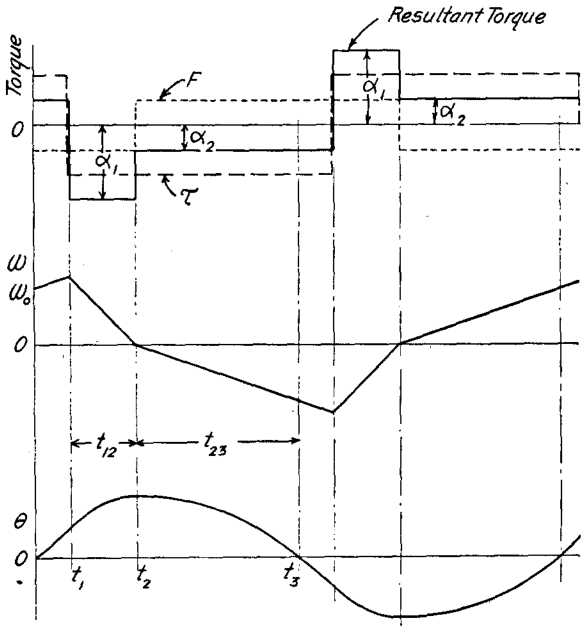
Although this condition if of interest as a limiting case, the effect of time lag is always present practically. As the second condition then, the effect of a definite time lag of t_1 seconds in the reversal of the restoring torque is considered. The nature of the variations of angle, velocity, and torque with time in this condition is sketched in Fig. 7. Remembering that during each interval when the resultant torque is constant, the motion is that of a uniformly accelerated body and using the following terminal conditions,

$$\left.\begin{aligned} t &= 0 \\ \omega &= \omega_0 \\ \theta &= 0 \end{aligned}\right\}, \ (49) \qquad \left.\begin{aligned} t &= t_2 \\ \omega &= 0 \\ \theta &= \theta_2 \end{aligned}\right\}, \ (50) \qquad \left.\begin{aligned} t &= t_3 \\ \omega &= -\omega_0 \text{ for} \\ &\quad \text{steady state} \\ \theta &= 0 \end{aligned}\right\}, \ (51)$$

the following expression is obtained for θ_2 starting from the conditions at $t = 0$:

$$\theta_2 = \omega_0 t_1 + \frac{\alpha_2}{2} t_1^2 + \frac{1}{2\alpha_1}(\omega_0 + \alpha_2 t_1)^2. \tag{52}$$

FIG. 7.



Same as Fig. 6 except that time lag is present.

Using (51) and writing an expression for θ_2 starting from $t = t_3$,

$$\theta_2 = \frac{\omega_0^2}{2\alpha_2} . \tag{53}$$

Equating (52) to (53) and solving the result for ω_0 there results for steady-state operation:

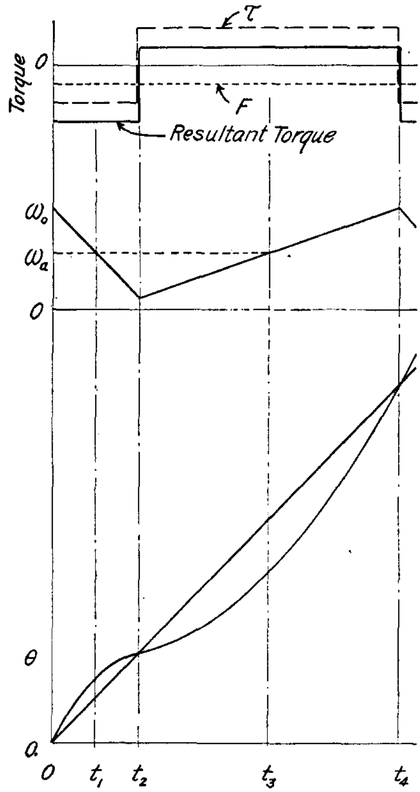
$$\omega_0 = \alpha_2 t_1 \frac{\alpha_1 + \alpha_2 \pm \sqrt{2\alpha_1(\alpha_1 + \alpha_2)}}{\alpha_1 - \alpha_2} . \tag{54}$$

Equations (54) and (53) give the interesting result that when the friction obeys the Coulomb law, the relay servo with a stationary input oscillates with an amplitude proportional to the square of the time lag in the restoring force change. In contrast to the previous case, similar except that the damping was due to viscous friction instead of Coulomb friction, the amplitude due to a small time lag is very small. From the point of view then of making the amplitude of oscillation small, a servo with Coulomb friction acting on the output is preferable to one with viscous friction. This statement applies, as will be seen below, only to the condition with a substantially stationary input.

The third condition is similar to the first for this case, except that the input, instead of being stationary, has a constant angular velocity ω_a.

In Fig. 8 are shown the general relations between the

FIG. 8.



Similar to Fig. 6 except that input has constant input velocity ω_a.

torques, velocities, and angles, and time for this condition. It will be assumed that the output velocity is always positive, i.e., that although the output velocity varies above and below the average or input velocity, it never becomes negative, and that as a consequence the frictional force always acts in the same direction. The implications of this assumption become evident in the analysis.

The differential equations of motion are:

$$-\tau - F = Jp^2\theta = Jp\omega \quad \text{for} \quad 0 < t < t_2, \tag{55}$$

$$\tau - F = Jp^2\theta = Jp\omega \quad \text{for} \quad t_2 < t < t_4. \tag{56}$$

The known terminal conditions are:

$$\left.\begin{matrix} t = 0 \\ \omega = \omega_0 \\ \theta = 0 \end{matrix}\right\}, \quad (57) \qquad \left.\begin{matrix} t = t_2 \\ \omega = \omega_2 \\ \theta = \omega_a t_2 \end{matrix}\right\}, \quad (58) \qquad \left.\begin{matrix} t = t_4 \\ \omega = \omega_4 \\ \theta = \omega_a t \end{matrix}\right\}. \quad (59)$$

From (55) and the terminal conditions (57) the expressions for output velocity and angle for the interval $0 < t < t_2$ are found to be

$$\omega = -\alpha_1 t + \omega_0, \tag{60}$$

$$\theta = -\frac{\alpha_1}{2} t^2 + \omega_0 t \tag{61}$$

as before. From the value of θ at $t = t_2$ in (58) substituted in (61), the value of t_2 is found to be

$$t_2 = 0 \text{ or } 2\left(\frac{\omega_0 - \omega_a}{\alpha_1}\right), \tag{62}$$

the second value only being of interest. Putting this value in (60) gives

$$\omega_2 = 2\omega_a - \omega_0. \tag{63}$$

Integrating (56) and evaluating the constants by (58) and (63), the following expressions for velocity and angle during the interval $t_2 < t < t_4$ are obtained:

$$\omega = \alpha_2 t + 2\omega_a\left(1 + \frac{\alpha_2}{\alpha_1}\right) - \omega_0\left(1 + 2\frac{\alpha_2}{\alpha_1}\right), \tag{64}$$

$$\theta = \frac{\alpha_2}{2}(t^2 - t_2^2) - (\omega_a - \omega_0)\left(1 + 2\frac{\alpha_2}{\alpha_1}\right)t_2 + 2\omega_a\left(1 + \frac{\alpha_2}{\alpha_1}\right)t - \omega_0\left(1 + 2\frac{\alpha_2}{\alpha_1}\right)t. \quad (65)$$

The value of t_4 in terms of known quantities can be found by using the value of θ given in (59) in (65) and is

$$t_4 = t_2 \quad \text{or} \quad 2\frac{(\alpha_1 + \alpha_2)}{\alpha_1\alpha_2}(\omega_0 - \omega_a). \quad (66)$$

The second value is, of course, the only one of interest. Putting this value of t_4 in (64) gives the angular velocity at the end of the cycle in terms of the velocity at the beginning. This value is

$$\omega_4 = \omega_0. \quad (67)$$

This result is surprising at first glance for it shows that this servo, under the given conditions, continues to oscillate with any finite amplitude at which it may happen to be started. That this result is sound physically may be seen by considering the curve of resultant torque in Fig. 8. Even though friction is present, when it always acts in the same direction and is of constant magnitude, its effect is merely to increase the magnitude of the restoring force during the first portion of the cycle, and to decrease it during the second portion. Thus no damping action whatever is produced. This is the first case in which the relay-type of servo has been found to have the possibility of a finite amplitude of oscillation in steady-state operation under the ideal condition of zero time lag.

This result, of course, holds only for the postulated condition that the output velocity is unidirectional. If the amplitude of velocity oscillation is sufficiently large in comparison with the input or average output velocity so that reversal of the output velocity occurs, the action is quite different. Due to reversal of the output velocity and therefore of the direction of the friction force, a damping action is produced and the amplitude of oscillation is reduced. This reduction continues until the output velocity just fails to reverse, when the damping action ceases. Thereafter the output oscillations continue indefinitely at such amplitude that reversal of the output velocity just fails to occur. Time lag aggravates this condition in a way that is discussed further on.

Evidently this type of servo-mechanism with constant-magnitude restoring and friction torques is unsuited for use with other than a fixed or very slowly varying input, because of the absence of useful damping when the input has appreciable velocity.

In the cases treated thus far the restoring torque has been acting at all instants either in one direction or the other. Actually there is always some inactive period which may be so small, however, in many instances that its effect is insignificant. In other instances, a finite inactive range, over which the deviation of input from output may vary without bringing the restoring torque into action, is purposely provided. The analysis of the performance of servos with such an inactive range is therefore important and will be considered next.

Case B1. In this case is considered a servo-mechanism having a constant-magnitude restoring torque and a viscous-friction torque acting on the output element.

It differs from the servo with these properties treated under A1 in that the restoring torque is not brought into action until the output differs in magnitude from the input by some predetermined angle θ_n. The analysis of this servo, while straight-forward and simple when carried out numerically for a particular case, is rather clumsy analytically because of the large number of time intervals into which a cycle of operation must be subdivided. For each of these subdivisions a different differential equation applies. Each integration to obtain an expression for angle involves two constants determined by conditions at the beginning of the time interval. Since at least five such time intervals appear in one cycle the analytical expressions become rather cumbersome by the end of the cycle.

Certain conclusions of value can be drawn without explicit formulation, however. If this servo has a stationary or slowly moving input, the inactive zone allows the servo-mechanism to remain inoperative much of the time, since the restoring torque would be called into action only when the deviation exceeded the inactive or tolerance range θ_n. It is readily seen that under these conditions the output angle is uncertain within the range $\pm\ \theta_n$ from the input.

This servo will not oscillate because all of the damping forces exist that were present with no inactive zone. Consequently any initial oscillation will damp down to an amplitude that lies within the inactive zone, and then cease entirely. The effect of time lag in the application or removal of the restoring torque is discussed further on.

When the input of such a servo has an appreciable velocity, conclusions cannot be so readily drawn without a mathematical formulation, which as seen above is cumbersome. Numerical computation is perhaps as direct a method as any for obtaining an explicit result in a particular case. For a general study covering all cases, mechanical methods are the most attractive. Thus on a device such as the differential analyzer [32] many particular solutions could be run off with relative ease. These solutions could take the form of curves relating the dimensionless variables θ/θ_n and t/T, with ω_0/ω_a and ω_s/ω_a as dimensionless parameters whose values would be different for each particular solution. All possible solutions could be represented on a series of curve sheets, each sheet containing a nest of curves of θ/θ_n vs. t/T with ω_0/ω_a as a parameter, all for a given value of ω_s/ω_a. Each other sheet would be similar except that it would be for a different value of ω_s/ω_a.

The case under discussion has not sufficient interest to warrant the presentation of such a set of solutions in this paper. However, the operation of this servo-mechanism will be illustrated by a numerical example using the following values:

$$\omega_s = 2,$$
$$\omega_a = 0.5,$$
$$\omega_0 = 1.5,$$
$$\theta_n = 0.1,$$
$$T = 1.$$

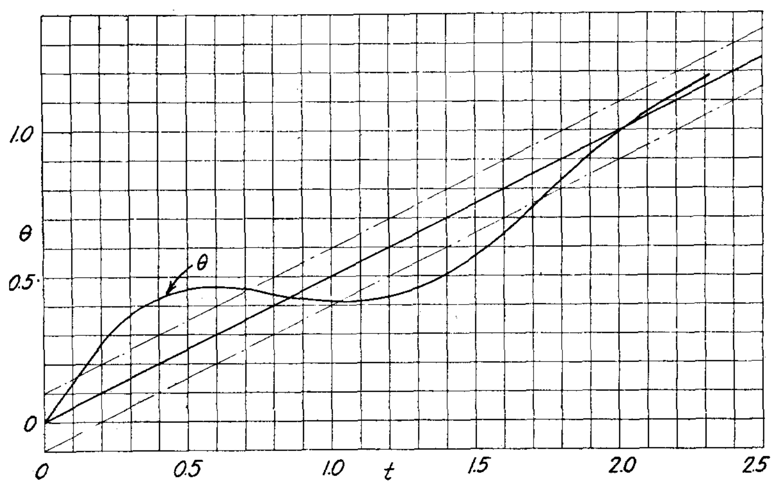
Any consistent set of units can, of course, be used. The results of this calculation are plotted in Fig. 9 which shows a rather rapid damping out of oscillations. In the steady state the output oscillates with an infinitesimal amplitude about the lagging limit, i.e., the limit at which the positive-direction restoring torque comes into action, or at an output angle

$$\theta = \theta_i - \theta_n.$$

This servo then has a definite lag angle θ_n which is half the inactive range, assuming that this range is centered on the output position for exact correspondence between input and output. From these results it appears that an inactive range has little usefulness from the theoretical point of view when the input is moving unidirectionally. Practically it may be necessary because of electrical contact difficulties or the like.

That the conclusions drawn from Fig. 9 for a constant

FIG. 9.



Curve showing output angle vs. time for a sample calculation of case B1 with constant input velocity ω_a, and no time lag.

input velocity are of rather general application can be seen from the following considerations. Suppose the initial amplitude of oscillation is very large in comparison with the width of the inactive zone. Then the analysis given under case A1 will apply approximately, leading to a smaller amplitude, when conditions will become similar to those shown in Fig. 9. If the initial amplitude is very small, the steady-state will be reached in less time than that required for the initial conditions shown in Fig. 9.

Case B2. In the case just considered the damping was due to viscous friction. In the present case Coulomb friction is assumed. The inactive zone is the same as assumed above, i.e., the output has a permissible deviation of $\pm\, \theta_n$ from the input before the restoring torques are effective. Except for the presence of the inactive range this case is identical with case A2.

Considering first that the input is moving slowly or not at all, a consideration of the motion of the output and the forces acting upon it will show that oscillations damp out relatively rapidly, and that the servo may remain inactive much of the time as in the foregoing case. The same uncertainty as to output position also exists.

When a constant velocity input is considered, the prediction of behavior is somewhat more involved. First assume that the initial amplitude of output oscillation is not so large as to cause reversal of the output velocity. A cycle of operation can then be represented as shown in Fig. 10. It can be shown readily that steady-state operation of this servomechanism can occur with any amplitude of oscillation for which the output velocity is never negative as assumed above. This is done as follows: From the differential equations and terminal conditions, or by inspection, remembering that only constant accelerations are encountered, the following relations are obtained:

$$\omega_0 - \omega_a = \omega_a - \omega_1, \tag{68}$$

$$\omega_a - \omega_2 = \omega_3 - \omega_a, \tag{69}$$

$$\omega_2 = \omega_1 - \alpha_0 t_{12}, \tag{70}$$

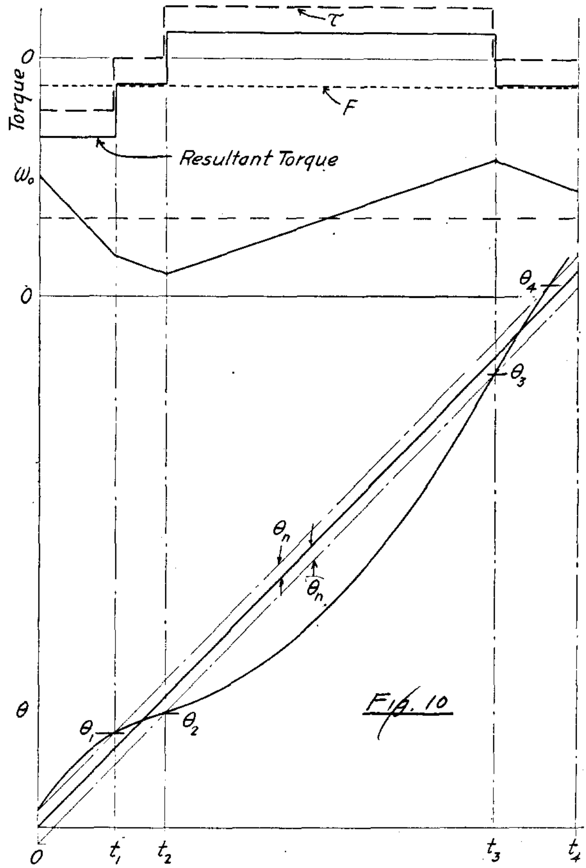
$$\theta_2 = \theta_1 + \omega_1 t_{12} - \frac{\alpha_0}{2} t_{34}{}^2 = \theta_1 + \omega_a t_{12} - 2\theta_n, \tag{71}$$

$$\omega_4 = \omega_3 - \alpha_0 t_{34}, \tag{72}$$

$$\theta_4 = \theta_2 + \omega_3 t_{34} - \frac{\alpha_0}{2} t_{34}{}^2 = \theta_3 + \omega_a t_{34} + 2\theta_n. \tag{73}$$

The significance of the subscripts is indicated in Fig. 10. Equations (68) and (69) express the fact that, in uniformly

FIG. 10.



Torques, and angular velocity and angle of output for case B2 with constant input velocity ω_a and no time lag.

accelerated motion, the average velocity is equal to the average of the initial and final velocities. From these six equations it can be shown that

$$\omega_4 = \omega_0 \tag{74}$$

which will prove the statement made above. To demonstrate (74) substitute ω_1 from (68) in (70), substitute the resulting expression for ω_2 in (69) and solve for ω_4. This result substituted in (72) gives the following relation between ω_4 and ω_0:

$$\omega_4 = \omega_0 + \alpha_0(t_{12} - t_{34}).$$

To prove (74) it is therefore sufficient to show that

$$t_{12} = t_{34}.$$

This can be done by solving (71) and (73) for t_{12} and t_{34} respectively, and expressing ω_1 and ω_3 in the results in terms of ω_0.

It follows then that the introduction of an inactive range in the constant-friction relay servo does not damp out oscillations when the input has a constant velocity and the output velocity is unidirectional. These conditions are representative of normal operation, hence this servo is inherently unsatisfactory. To damp out oscillations under these conditions, the frictional force must have a component that varies in magnitude as some positive power of the velocity.

This completes the analysis of the various cases of relay-type servo-mechanisms considered in this paper. As this analysis has involved a considerable amount of detail, it is well to summarize the important results. This is done in Table I, which is self-explanatory except for certain of the effects of time lag discussed below. These results apply to a servo-mechanism in which the output element has inertia, and is acted upon by constant magnitude restoring forces, and a frictional force of some type.

C. Effect of Time Lag in Change of Restoring Torque.—Although in two cases, the effect of time lag in the application and removal of restoring torques was treated mathematically, a treatment of its effect in other cases was postponed for a general discussion which will be given here.

Time lag, as stated before, is the delay between (*a*) the time at which the restoring torque would be applied or removed were the relay to act and the resulting torque to build up instantaneously when the need for a change is indicated by a predetermined discrepancy, between input and output and (*b*) the time when the torque actually becomes effective. This time lag may be different for the application and the removal of the torque when there is an inactive zone, but it suffices for this discussion to consider the combined effect of these two time lags. Actually there is always some delay or time lag. In some cases it may be so small as to have an insignificant

TABLE I.

	Stationary Input.			Constant Velocity Input.		
	Steady-state Oscillation.		Average Lag Error.	Steady-state Oscillation.		Average Lag Error. No Time Lag.
	No Time Lag.	Time Lag.		No Time Lag.	Time Lag.	
A. No inactive zone 1. Viscous friction	Zero amplitude infinite frequency	Finite and function of time lag	Zero	Zero amplitude infinite frequency	Finite amplitude	Zero
2. Coulomb friction	Zero amplitude infinite frequency	Finite and proportional to square of time lag	Zero	Indeterminate	Amplitude increases until postulated conditions fail to hold	Indeterminate
B. Finite inactive zone 1. Viscous friction	None	May or may not oscillate depending on parameters	± half of inactive range with no time lag	Zero amplitude infinite frequency	Finite amplitude	± half of inactive range
2. Coulomb friction	None			Indeterminate	Amplitude increases until postulated conditions fail to hold	Indeterminate

effect upon the action of the servo. In many cases, however, its effect is significant. This is demonstrated by the great sensitiveness of the viscous-friction damped relay servo to time lag, where a small time lag results in a relatively large amplitude of steady-state oscillation. On the other hand, a relay servo having Coulomb friction is relatively insensitive to small values of time lag. These facts have been demonstrated for a servo with an approximately constant input and no inactive range.

When the input has appreciable velocity, or a finite inactive zone, the formal analysis, while straightforward, is somewhat cumbersome because of the number of points during a cycle at which the differential equation of motion changes and constants of integration must be reëvaluated. In any particular case this process can be carried through numerically with relative ease.

A qualitative analysis of the effect of time lag in these cases is readily made however. In general the result of a time lag in changing the restoring forces is to continue to increase the output velocity in a given direction after the instant at which this velocity should be decreased. Thus in an ideal relay servo with no inactive zone, if the output is ahead of the input, but is moving toward the input, the restoring torque should reverse in sign at the instant when the input and output coincide. From this instant the output is accelerated toward the input. When time lag is present, the restoring torque acting prior to this instant continues to act for the time-lag interval after this instant. During this interval the restoring torque is tending to drive the output away from the input rather than toward it. As a result, the overshoot is greater than when no time lag is present, and the relative velocity between input and output is greater when they again are in coincidence than it would have been in the absence of time lag. At the time of this second coincidence the time lag in changing the sign of the restoring torque again causes an increment of output velocity in the wrong direction. The net result, as is seen, is to tend to increase the amplitude of oscillation. Looked at in a slightly different way, the impressed torque, when time lag is present, may be resolved into two components, one with no time lag, the effect of which has been studied in each case, and a second which gives an impulse to the output in the direction of its motion, at a time when this motion should be retarded. This impulse evidently adds energy to the oscillation, a fact demonstrated by the results in the two cases analyzed with time lag present.

The point of real interest is the effect of this increment of energy on the steady-state amplitude of oscillation. The cases previously considered are divided into two groups with respect to this effect. The first group includes the cases where positive damping exists as indicated by a zero amplitude of steady-state oscillation. The second includes the cases where no damping exists as indicated by the indefinite persistence of any existing finite amplitude.

The effect of time lag on the first group is to increase the amplitude of oscillation until the friction or damping work per cycle due to increased amplitude is equal to the work per cycle put into the output as a result of the time lag. This amplitude is finite for the cases considered in the first group.

In the second group time lag results in an amplitude of oscillation that increases until the postulated conditions of operation cease to hold. It has been seen that in the Coulomb-friction relay servo there is no damping so long as the output moves unidirectionally. When the amplitude is so large as to cause reversal of the output during some portion of the cycle, however, net work is done against friction, and positive damping is thereby introduced. Due to time lag, then, the amplitude of oscillation will increase until the work per cycle resulting from time lag is equal to the net friction work per cycle resulting from reversal of the output unless something breaks before this limit is reached. Evidently, unless the input velocity is so small that sufficient positive damping to overcome the negative damping effect of any time lag present can be introduced by a reasonable amplitude of oscillation of the output, the second group of servos is quite useless for practical purposes.

The analysis of the operation of relay servos given in the foregoing pages has assumed that the dynamic elements of the system under consideration can be reduced to an output member that has inertia, and that only frictional and constant-magnitude restoring forces act upon this member.

This is the simplest type of relay servo-mechanism but one which includes many practical cases.

This concludes the discussion of relay-type servo-mechanisms given here, which could, of course, be almost indefinitely extended. Attention will now be turned to the second, or definite-correction type.

DEFINITE-CORRECTION SERVO-MECHANISM.

Because of the relative simplicity of the operating characteristics of this type of servo-mechanism no extensive mathematical analysis will be given. The criteria for nonoscillatory response were given under the discussion of general characteristics of servo-mechanisms. Assuming that these criteria are satisfied, the other important operating characteristics are the maximum input speed which this servo can follow and its lag error. The maximum input speed is easily found.

In the definite-correction servo-mechanism, the value of the output is periodically measured to determine whether or not it differs by more than half the inactive range from the value of the input. The time interval between successive measurements is a definite quantity, independent of the manner in which the input may be varying. As a result of a measurement, a definite correction is applied to the output. If the time interval is Δt and the maximum correction that can be applied to the output in one interval is $\Delta\theta_m$, the maximum input speed ω_{im} that this servo can follow is evidently

$$\omega_{im} = \frac{\Delta\theta_m}{\Delta t}.$$

If this servo is to follow a high-speed input, either $\Delta\theta_m$ must be large or Δt small, or both. Δt is fixed by the time required for the correction to be substantially completed and for the measuring device to take up an indication substantially in accordance with the corrected value of the output. The lower limit for Δt is then established by design considerations relating to the time required for the correction to be effected, and to the time constant of the measuring device.

If only one size correction $\Delta\theta$ can be made, then $\Delta\theta$ is fixed by the permissible error, since, to prevent oscillation, $\Delta\theta$ can be no larger than twice the range of deviation of the output from the input over which no correction is applied. This limitation can be mitigated by grading the size of correction, making the smallest size satisfy the error condition, and the largest satisfy the maximum input-speed condition. Of course the correction should be proportional to the measured deviation, and just sufficient to restore the deviation to zero. Nearly all definite-correction servos have this proportional correction feature.

One limitation to which this type of servo is subject should be noted. It can follow satisfactorily only those components of input velocity that are sufficiently continuous in a mathematical sense to make the input velocity approximately constant over at least two successive time intervals Δt. That is, the output varies according to a sort of step function in which the size of a given step is determined largely by the average input velocity during the preceding interval Δt. If the input velocity is sufficiently continuous, this output step function will be a good approximation to the input function. In this type of servo, however, the output always lags behind the input, because a deviation is corrected only after it has occurred. The input is always one jump ahead.

The definite-correction servo has a number of advangages for certain classes of work. When properly adjusted it is non-oscillatory. When the input is constant or varies only slowly, the output drive may be inactive for considerable periods, thus saving wear and tear. A malfunctioning of the servo is more likely to result in inaction than in a racing away of the output toward infinity or the limit stop. The relay servo is always hunting a balance position, whereas the definite-correction servo moves only when balance is disturbed. In a loose way it may be said that the relay servo keeps the output in equilibrium only by continuous juggling, while the output of the definite-correction servo is in static equilibrium. This fact makes for long life of the latter. The definite-correction servo is, however, somewhat slower than other types because of the relatively small fraction of the time during which the restoring force can be effective. This slowness imposes a restriction upon the field of application rather than upon its effectiveness in any application for which it is suitable. The definite-correction servo is widely used, especially for recording and control in industrial processes where it is normally highly successful.[1, 2, 3, 9, 10, 11, 12]

CONTINUOUS-CONTROL SERVO-MECHANISM

The third type of servo-mechanism, in which the indicating element continuously controls the restoring force acting on the output element in both magnitude and direction, is perhaps the most interesting of the three. General considerations indicate that this smooth torque control should be superior to the somewhat crude "off-on" control of the relay servo. Also the continuous use of the input indications should be superior to the occasional use of these indications characteristic of the definite-correction servo. Thus the continuous-control type appears on casual inspection to have inherent advantages over the other two types, especially where accurate, rapid following is required.

The continuous-control type is interesting for another reason. Because of their nature, the forces acting in this servo are easily expressed mathematical functions of displacement and velocity. Consequently a thorough-going analysis of the performance of this servo is more readily made than of the two previous types.

The analysis given here applies to a servo in which an output member, which has inertia, is acted upon by restoring and damping torques, the restoring torque being a function of the deviation between input and output and its derivatives, and the damping torque proportional to the output velocity. Most continuous-control servos fall into this category. Especially in high-speed servos the forces obeying the Coulomb friction law are likely to be quite negligible when compared with the forces having the effect of viscous friction.

Two cases are considered in what follows. In the first, the restoring torque is proportional to the input-output deviation and the necessary damping is secured by viscous-friction-law torques. In the second, the restoring force is a linear function of the first and second time derivatives of the deviation as well as of the deviation itself. A third case, not considered in detail here because of certain practical limitations involved in its utilization in servo-mechanisms, but of interest in other similar dynamic systems, is that in which the first time derivative of the restoring torque is a linear function of the deviation and its first two time derivatives. Minorsky [4] treats these last two cases as applied to the rudder-hull dynamic system of the ship-steering problem. The reason that the third case will not be generally useful in servo-mechanisms is that this type of control of the restoring torque relies largely on the derivatives of the deviation and these derivatives in the usual system utilizing a servo-mechanism are not sufficiently continuous to provide smooth operation of the servo. In case a servo is used in connection with a large vessel or other large inertia system, such a control of servo restoring torques might be attractive. However, the second-case type of control has not yet been exploited in servos and until this has been done that of the third case can wait. Analysis of the third case is not difficult and has been carried out to the point of determining the conditions for stability and the nature of the steady-state deviations by Minorsky for the rudder-hull dynamic system.[4] A general solution involves the solution of a cubic or higher degree equation making the results cumbersome at best. Again, however, numerical solutions for particular cases are quite straightforward.

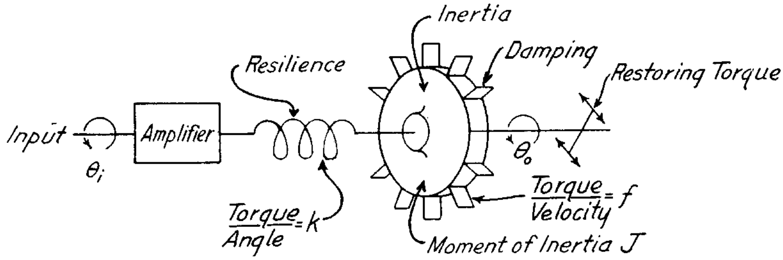
With this preliminary discussion, attention is now turned to a mathematical consideration of the first case, in which the servo restoring-force control is a linear function of the input-output deviation. Expressions for the steady-state and transient behavior are derived and a criterion of merit for such a servo is developed.

(a) Restoring Torque Proportional to Deviation.

For this analysis consider the physical system shown in Fig. 11. This system correctly represents the dynamic elements of a continuous-control servo on the condition that the effect of all inertia and damping torques can be assumed to be concentrated at the output shaft. For many continuous-control servos these assumptions correctly represent the facts.

The amplifier shown in Fig. 8 has the effect of reducing the reaction on the input of the resilience torque acting on the output by the amplification factor. Such amplification in some form is nearly always present in this type of servo. It may take electrical, mechanical or other form but the net

FIG. 11.



Schematic diagram of the significant dynamic elements of the continuous-control type servo-mechanism.

effect is the same in any case. It should be noted that the amplification factor of the amplifier affects the resilience constant, and to that extent the amplifier and resilience element are not independent as shown in Fig. 11.

In the analysis which follows a number of physical quantities not previously used appear, hence the notation for this analysis will be entirely redefined. Let

θ_i = input angle (radians),
θ_0 = output angle (radians),
$\theta = \theta_i - \theta_0$ = lag of output with respect to input (radians),
k = resilience constant (dyne-cm. per radian),
J = moment of inertia of servo referred to output shaft (gm.-cm.^{2}),
f = damping constant (dyne-cm. per radian per sec.),
t = time (seconds),
1 = Heaviside's unit function,
$p = \frac{d}{dt}$,
$\omega + p\theta$.

c.g.s. units are indicated in parentheses but of course any other consistent set is equally suitable since no empirical factors are used.

The equations governing the motion of the elements shown in Fig. 11 are:

$$k(\theta_i - \theta_0) - fp\theta_0 = Jp^2\theta_0, \tag{75}$$

$$\theta_i = \Phi(t). \tag{76}$$

$\Phi(t)$ is some arbitrary function of time representing the input motion. These equations may be solved in terms of any of the angle variables. The deviation or error angle θ is of primary interest, and from it and the given θ_i, the output angle is readily obtained. Therefore, the solution will be made in terms of θ.

Substituting

$$\theta_0 = \theta_i - \theta$$

in (75) gives

$$(k + fp + Jp^2)\theta - (fp + Jp^2)\theta_i = 0$$

or

$$\theta = \frac{f + Jp}{k + fp + Jp^2} p\theta_i. \tag{77}$$

To obtain the error angle θ as an explicit function of time some $\theta_i = \Phi(t)$ must be chosen for which the error is to be determined. An interesting case and one that imposes a severe test on a servo consists in suddenly applying a constant velocity to the input, previously at rest. Furthermore, from the response to this particular time function, the response to any $\Phi(t)$ can be determined by the use of the superposition integral.[31] If the magnitude of the suddenly-applied velocity is ω_1,

$$p\theta_i = \omega_1 1. \tag{78}$$

This expression in (77) gives

$$\theta = \frac{p + \frac{f}{J}}{p^2 + \frac{f}{J}p + \frac{k}{J}} \omega_1 1. \tag{79}$$

Equation (79) can be evaluated directly from a table of operational formulas.[31] To do this the denominator of (79) should be written in the form,

$$(p + \alpha - \beta)(p + \alpha + \beta),$$

in which

$$\alpha \pm \beta = \frac{f}{2J} \pm \sqrt{\frac{f^2}{4J^2} - \frac{k}{J}}. \tag{80}$$

Three cases of solution of (79) exist corresponding to positive, zero, and negative values of the quantity under the radical of (80). Two of these cases are of particular interest in this problem, the first or critically-damped case for which

$$\beta = 0 \tag{81}$$

and the second or oscillatory case for which

$$\begin{aligned}\beta &= j\sqrt{\frac{k}{J} - \frac{f^2}{4J^2}} \\ &= j\phi,\end{aligned} \tag{82}$$

ϕ being a real number. The third case for which β is real, corresponds physically to overdamping. This case is of relatively minor interest from the design point of view because the steady-state lag error is greater and the speed of response less than in the critically-damped case, and there are no compensating advantages in the use of over-damping.

Although (81) is a limiting case of (82), the solution of (79) using (81) is somewhat simpler than that for the oscillatory case and is therefore given first. Putting (80) and (81) in (79) there results

$$\theta = \left[\frac{p}{(p+\alpha)^2} + \frac{2\alpha}{(p+\alpha)^2}\right]\omega_1 1. \tag{83}$$

From a table of operational formulas,* the time function corresponding to (83) is found to be

$$\theta = \omega_1\left[\frac{2}{\alpha} - \left(t + \frac{2}{\alpha}\right)\epsilon^{-\alpha t}\right].$$

Substituting the time constant T for $1/\alpha$ and dividing the resulting expression by $\omega_1 T$ to obtain a dimensionless equation,

$$\frac{\theta}{T\omega_1} = 2 - \left(\frac{t}{T} + 2\right)\epsilon^{-(t/T)}. \tag{84}$$

This equation is plotted as curve 1 of Fig. 12 in terms of the dimensionless variables $\theta/T\omega_1$ and t/T. These results are discussed below in connection with those which will now be obtained.

* Ref. 31 of Bibl., Formulas (7) and (8) of Appendix C.

When ϕ is finite and real, the response is oscillatory. A solution for this case is important. Here (79) can be written in the form,

$$\theta = \frac{p + 2\alpha}{(p + \alpha - j\phi)(p + \alpha + j\phi)} \omega_1 1. \tag{85}$$

Equation (85) integrates into the form,

$$\theta = \psi_1(t) + 2\alpha\psi_2(t). \tag{86}$$

By formula * the first and second terms of (86) are found to be respectively:

$$\psi_1(t) = \frac{\omega_1}{\phi} \epsilon^{-\alpha t} \sin \phi t \tag{87}$$

and

$$\psi_2(t) = \frac{\omega_1}{\alpha^2 + \phi^2}\left[1 - \epsilon^{-\alpha t}\left\{\frac{\alpha}{\phi}\sin \phi t + \cos \phi t\right\}\right]. \tag{88}$$

The latter is obtained after some algebraic manipulation. Substituting (87) and (88) in (86), replacing $1/\alpha$ by the time constant T, and dividing by $\omega_1 T$ to make the expression dimensionless, there results

$$\frac{\theta}{\omega_1 T} = \frac{2}{1 + T^2\phi^2} \times \left[1 - \epsilon^{-(t/T)}\left\{\cos \phi t + \left(\frac{1 - T^2\phi^2}{2T\phi}\right)\sin \phi t\right\}\right] \tag{89}$$

or in polar form

$$\frac{\theta}{\omega_1 T} = \frac{2}{1 + T^2\phi^2} - \frac{1}{T\phi}\epsilon^{-(t/T)} \sin (\phi t + \xi), \tag{90}$$

where

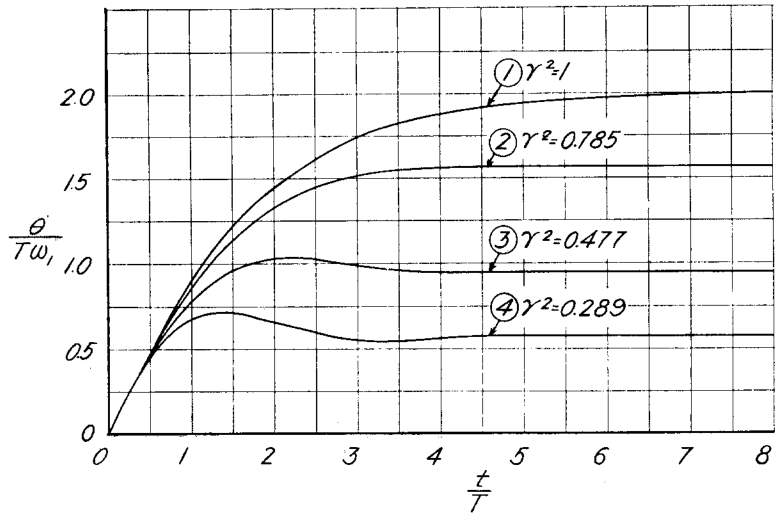
$$\xi = \arctan \frac{2T\phi}{1 - T^2\phi^2}. \tag{91}$$

In Fig. 12 (89) is plotted for various values of $T\phi$ as indicated on the curves 2, 3, and 4.

---

* Bibl. 31, formulas 17 and 11 of Appendix C.

FIG. 12.



Response of continuous-control type servo to a suddenly-applied constant input velocity. Curves of input-output deviation as a function of time with the relative damping γ^2 as a parameter. Dimensionless variables are used making these curves applicable to a servo with any constants.

Before discussing the curves of Fig. 12, it is of interest to have a single quantity which characterizes the degree of oscillation in the servo response. One such quantity, which may be called the relative damping factor, is defined by

$$\gamma^2 = \frac{f^2}{4Jk}. \tag{92}$$

From the values of T and ϕ in terms of f, k, and J it is easily shown that

$$\gamma^2 = \frac{1}{1 + T^2\phi^2}. \tag{93}$$

From the condition for critical damping, that $\phi = 0$, it will be seen that $\gamma^2 = 1$ characterizes this case, while for less than critical damping

$$\gamma^2 < 1.$$

For large values of t/T, that is when the steady-state has been reached, the substitution of (93) in (89) or (90) yields the following very simple relation between the steady state lag angle θ_s and the impressed velocity ω_1:

$$\theta_s = 2\gamma^2 T\omega_1. \tag{94}$$

The implications of equations (84), (89) and (94) will now be considered with the aid of Fig. 12. Several interesting facts can be deduced.

In the first place, the steady-state lag error is shown to be proportional to the time constant T, to the angular velocity ω_1 of the input, and to the relative damping factor γ^2. This shows that a servo with a high speed of response, i.e., a small time constant T, is inherently more accurate than a slower servo. This is true not only for the particular input function investigated here but for any form of $\theta_i = \Phi(t)$ because the factor T is carried through the evaluation of the superposition integral used to obtain the error for any $\Phi(t)$.

The significance of the relative damping factor γ^2 can perhaps best be seen by reference to Fig. 12. γ^2 characterizes the form of the deviation or error curve, that is, the amount of overshoot and persistence of oscillation. A small γ^2 correponds to a large overshoot and subsequent oscillation. A small γ^2 also results in a small steady-state error. In designing this type of servo for small error, it is evident that some compromise must be made between the tolerated error and the amount of oscillation permitted. Fortunately the steady-state error can be very materially reduced by using a small γ^2 without introducing a very large overshoot. In certain applications, however, where the input may contain periodic components of a period comparable to that of the natural period of the servo, an aperiodic adjustment, i.e., $\gamma^2 = 1$, may be necessary.

From the design point of view perhaps the most important deduction from these equations follows from the establishment of a figure of merit for servos. Suppose this figure of merit is arbitrarily taken as the product of the numerical measure of two desirable properties of a servo-mechanism. One of these factors is taken as the maximum attainable speed ω_m. The other is taken as the smallness $1/\theta_s$ of the steady-state error θ expressed as a fraction of the total angle ω_1 turned per unit time, or ω_1/θ_s. Let the product of these two factors be the figure of merit M thus:

$$M = \omega_m \frac{\omega_1}{\theta_s}.$$

It has been seen that θ is proportional to ω in the steady-state, hence values corresponding to the maximum speed may be used and

$$M = \frac{\omega_m^2}{\theta_m}, \tag{95}$$

where θ_m is the steady-state error at the speed ω_m. The expression for M can be written in terms of the design constants k, f, and J as follows: From (94) using $\omega_1 = \omega_m$,

$$\frac{\omega_m}{\theta_m} = \frac{1}{2\gamma^2 T},$$

$$\frac{1}{T} = \alpha = \frac{f}{2J} = \gamma\sqrt{\frac{k}{J}}$$

and

$$M = \frac{\omega_m}{2\gamma}\sqrt{\frac{k}{J}} = \frac{k\theta_m}{4\gamma^2 J} = \frac{\tau_m}{4\gamma^2 J}, \tag{96}$$

where τ_m is the maximum torque that can be applied to the output element. Taking (95) as a measure of the desired performance of a servo-mechanism, equation (96) shows that in design, the ratio of the maximum torque τ_m to the moment of inertia J should be as large as practicable. It also shows that the relative damping should be as small as can be permitted. These are important results as they give a very definite guide for design.

In the companion paper above referred to this criterion is used as a basis for design. Tests of the servo thus designed show very fast response and an excellent agreement between test and calculation.

(b) Restoring Torque Proportional to Deviation and its First Two Time Derivatives.

The second case of the continuous-control servo will now be considered, the case in which the restoring torque is a linear function not only of the deviation θ but of its first two time derivatives. Let the restoring torque τ be given by the following relation:

$$\tau = (k + lp + mp^2)(\theta_i - \theta_0), \tag{97}$$

where l and m are constants of proportionality.

Consider the response to a suddenly-applied input velocity ω_1 just as in the first case. Using (97) in the differential equation of motion, the operational expression for the deviation or error angle θ is

$$\theta = \frac{\frac{J}{J_1}p + \frac{f}{J_1}}{p^2 + \frac{f_1}{J_1}p + \frac{k}{J_1}}\omega_1 1, \tag{98}$$

where

$$\begin{aligned} f_1 &= l + f, \\ J_1 &= m + J. \end{aligned} \tag{99}$$

Redefining α and ϕ in terms of the parameters of the denominator of (98),

$$\alpha \pm j\phi = \frac{f_1}{2J_1} \pm \sqrt{\frac{k}{J_1} - \frac{f_1^2}{4J_1^2}} \tag{100}$$

and evaluating (98) as an explicit function of time in the same way that (85) was evaluated gives the following expression for the deviation or error angle θ:

$$\theta = \frac{\omega_1 f}{J_1\alpha^2}\left[1 - \epsilon^{-\alpha t}\left\{1 + \alpha t - \frac{\alpha^2 J t}{f}\right\}\right] \tag{101}$$

for the critically-damped case, i.e., for $\phi = 0$, or

$$\theta = \frac{\omega_1 f}{J_1(\alpha^2 + \phi^2)} \times\left[1 - \epsilon^{-\alpha t}\left\{\cos\phi t - \frac{J(\alpha^2 + \phi^2) - \alpha f}{f\phi}\sin\phi t\right\}\right] \tag{102}$$

for the oscillatory case, i.e., for ϕ finite and real.

These are very interesting results for they show that the steady-state error of this servo can be made zero by making the factor f zero. f is the coefficient of damping for the viscous-friction torque acting on the output. Although this damping is made zero, the servo operation can be made aperiodic or oscillatory in any desired degree by the damping effect introduced by the component of restoring torque depending on the first derivative of θ. By this method the damping effect is due not to the velocity $p\theta_0$ of the output but to the relative velocity $p\theta$ of the output with respect to the input. The magnitude of this damping is proportional to the coefficient l in (97). Physically the torque corresponding to the term $lp\theta$ in (97) is readily introduced, when a d-c. vacuum-tube amplifier is used, by an inductance component of interstage coupling.

The effect of introducing a component of restoring torque proportional to the relative acceleration of output and input, i.e., of introducing the effect represented by the term $mp^2\theta$ of (97) is only to alter the equivalent output inertia from the actual inertia J to an equivalent value $J_1 = J + m$. By making m negative numerically, the equivalent inertia can be given any value positive or negative. Negative values evidently are barred practically because the coefficient of the exponent in the transient-term exponential would be positive and the servo would be unstable. So long as $m > -J$, however, the system will be stable assuming of course that

$$f_1 = f + l > 0.$$

Subject to the above conditions for dynamic stability, a given servo with adjustable coefficients l and m can be made arbitrarily accurate, and fast in the sense that transient effects disappear rapidly, by giving suitable values to l and m. This is a very important result.

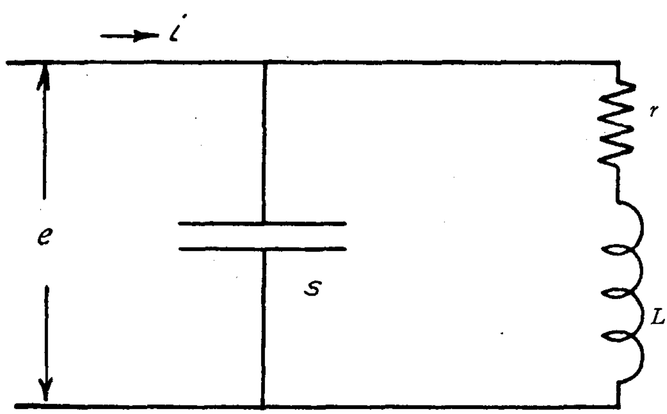
Physically the effect of a coefficient m can be introduced in a d-c. vacuum-tube amplifier by the use of two successive inductive interstage couplings. A negative m is secured by reversing the sense of one of these couplings. Mechanically these coefficients l and m could presumably be introduced by suitable viscous friction and inertia couplings respectively. Gyroscopic methods could also be applied to this end.[4]

The results of the analysis of this last case open up a great opportunity for improvement in the performance of servo-mechanisms of the continuous-control type. At sufficiently high speeds of response, however, the dynamic system of a simple servo-mechanism fails to reduce to the system postulated in this analysis and the analysis no longer applies. This fact sets one limitation upon indefinite improvement. Another important practical factor is that of stability of operation which may fix a limit below which the effective inertia cannot be reduced. Nevertheless, very real gains in the performance of even present high speed response servos using deviation control should be possible by the addition of deviation derivative control.

SIMILAR DYNAMIC SYSTEMS.

Earlier in the paper it was stated that although the analyses were made in terms of rotation, they could be applied without change to a dynamically similar system involving translation, electric currents, etc. As an example of this parallelism an electric circuit which is equivalent to the continuous-control servo analyzed is shown in Fig. 13. Using

FIG. 13.



Electric circuit equivalent to continuous-control servo-mechanism.

the parallelism shown in Table II it is seen that the charge on the susceptance S is analogous to the servo error.
Thus, writing the operational expression for q_s, the charge on the condenser, there results.[31]

$$q_s = \frac{e}{S} = \frac{iZ(p)}{S}$$

$$= \frac{1}{S} \frac{(r + pL)\dfrac{S}{p}}{r + pL + \dfrac{S}{p}} i$$

TABLE II.

Quantity.	Rotation.	Translation.	Electric Circuit.
Displacement	Angle θ	Distance x	Charge q
Velocity	$p\theta$	px	Current i
Acceleration	$p^2\theta$	p^2x	pi
Force	Torque τ	Force F	e.m.f. e
Inertia	Moment of inertia J	Mass M_x	Inductance L
Resilience	Spring constant K	Spring constant K_x	Elastance $S = \frac{1}{\text{capacitance}}$
Loss constant	Viscous friction f	Viscous friction f_x	Resistance r
Time	t	t	t

$$= \frac{r + pL}{S + rp + Lp^2} i. \tag{103}$$

The resemblance of (103) to (77) is evident, remembering that $pq = i$. If the current i is taken as

$$i = I1, \tag{104}$$

where I is a constant, (103) can be written

$$q_s = \frac{p + \frac{r}{L}}{p^2 + \frac{r}{L}p + \frac{S}{L}} I1. \tag{105}$$

Equation (105) has solutions of exactly the same form as (84) and (89) but in terms of the electrical quantities corresponding to rotational quantities as shown in Table II.

The translational problem is strictly analogous to the rotational problem so it is unnecessary to go into details of the analogy here.

As mentioned before, the rudder-hull dynamic system involved in the automatic steering of ships is similar to the servo-mechanisms here analyzed. These examples suffice to illustrate the correspondence between analogous systems.

ANALYSIS OF OTHER SERVO-MECHANISMS.

In the foregoing analysis of servo-mechanisms, a dynamically simple system was assumed both because most servo-mechanisms are adequately represented by this system and because this system is readily analyzed. This system consists of an output element having inertia, and acted upon by a frictional force of some type and a restoring torque which is some function of the input-output deviation.

For servo-mechanisms involving additional dynamically-significant elements, the analysis will be somewhat more complex. The differential equations of motion will, in general, be of higher order than the second and hence somewhat more lengthy algebra will be involved in their solution. The method of formulation is, however, essentially the same as that used here.

In the case of the continuous-control type of servo; the procedure is easily stated. By writing the differential equations of motion, a relation between the deviation angle θ (using rotational terms as illustrative), the input angle θ_i and the time derivative operator p of the form,

$$\theta = F(\theta_i, p),$$

can be obtained. Equations (77) and (98) are of this form. Operational methods of evaluating such expressions as explicit time functions such as (89) and (102), due to Heaviside and later workers,* are sufficiently well developed so that the process is one of routine algebra for equations resulting from lumped-parameter systems. Practically all servo-mechanisms will be included in this class. Numerous distributed parameter systems can also be treated but may involve more than algebraic work. For lumped-parameter systems the greatest difficulty is likely to be associated with the determination of the roots of higher degree algebraic equations in one variable, a cumbersome but straightforward process numerically. Such an analysis presupposes constant parameters, i.e., a linear system, an assumption widely justified particularly for the small variations which are usually of primary importance.

It should also be mentioned that entire closed-cycle control systems are dynamically similar to servo-mechanisms and their operation is investigated by the same methods. Often a

* See ref. 31 for these methods and references to other works on the subject.

closed-cycle control contains a servo-mechanism. If the time constants of the servo and of the system being controlled are widely different, it is often possible to simplify the problem considerably by analyzing the two parts separately. Otherwise the equations of the entire system must be solved as a unit.

CONCLUSIONS.

In this paper the performance of three important types of servo-mechanisms has been analyzed. These types include the relay servo, the definite-correction servo, and the continuous-control servo.

The first or relay type is always oscillatory in response to a varying input but under certain conditions the amplitude of oscillation and lag error can be made small. If unidirectional motion occurs, the presence of Coulomb friction alone will not damp out an initial amplitude of oscillation. Time lag in the application of the restoring forces tends to increase the amplitude of oscillation.

The second or definite-correction type is aperiodic in operation when properly adjusted and is quite suitable for use with slowly varying quantities. There is a slight lag error.

The third or continuous-control type is probably the best type where high-speed response and smoothness of control are required. This type can be made to have a response which is aperiodic or oscillatory with any given decrement, by suitable design and adjustment. By using the first and second derivatives of the deviation of the output from the input, as well as the deviation itself, to control the restoring force applied to the output, a very high rate of response and a very small steady-state deviation should be attainable. This type of servo has the advantage of being susceptible to rather easy and complete analysis. Tests on a high-speed of response servo of this type with deviation control alone show an excellent check with the theory. The design and test of this unusually fast servo are given in a companion paper.

ACKNOWLEDGMENTS.

The writer is indebted to Dr. V. Bush who suggested the nature and scope of this study, and who critically reviewed the manuscript. He also wishes to thank those of the Elec-

trical Engineering staff at the Massachusetts Institute of Technology who have contributed helpful suggestions.

BIBLIOGRAPHY.

1. STEIN, I. M.: "Precision Industrial Recorders and Controllers," *J. F. I.*, Vol. 209, No. 2, pp. 201–228, Feb. 1930.
2. HARRISON, T. R.: "The New Brown Potentiometer Recorder," *Rev. of Sci. Inst.*, Vol. 2, No. 10, pp. 618–625, Oct. 1931.
3. "Kent Recording and Controlling Apparatus," *Engineering* (London), Vol. 132, No. 3428, pp. 407–408, Sept. 25, 1931.
4. MINORSKY, N.: "Directional Stability of Automatically Steered Bodies," *Jour. Am. Soc. of Naval Engrs.*, Vol. 34, No. 2, pp. 280–309, May 1922. "Automatic Steering Tests," *Jour. Am. Soc. of Naval Engrs.*, Vol. 42, No. 2, pp. 285–310, May 1930.
5. SAUVAIRE-JOURDAN: "Comment On Peut Faíre Gouverner Automatiquement un Navire," *La Nature* (Paris), No. 2848, pp. 355–358, April 15, 1931.
6. FERRY, E. S.: "Applied Gyro-Dynamics," John Wiley and Sons, Inc., 1932.
7. "New Compass and Path Indicator," *Shipbldg. and Shipping Rec.*, p. 684, Dec. 5, 1929.
8. "Automatic Yacht Steerer," *Sci. Amer.*, p. 238, Apr. 1933.
9. "Hydraulic Turbine Governors and Frequency Control," *Proc. N. E. L. A.*, Vol. 88, pp. 542–572, 1931. Also as Pub. No. 13. Report Hydraulic Power Committee. (Contains extensive Bibliography.)
10. MCCREA, H. A.: "Automatic Control of Frequency and Load," *G. E. Rev.*, Vol. 32, No. 6, pp. 309–313, June 1929.
11. JONES, D. M.: "Controlling Load, Maintaining Frequency," *El. Wld.*, Vol. 95, pp. 1072–1076, May 31, 1930.
12. SPORN AND MARQUIS: "Frequency, Time and Load Control on Interconnected Systems," *El. Wld.*, Vol. 99, pp. 618–624, April 2, 1932.
13. CHALMERS, T. W.: "The Automatic Stabilization of Ships," Chapman and Hall, Ltd., London, 1931. This material also appeared in *The Engineer*, London, Apr. 4 to June 27, 1930.
14. "The Gyroscopic Stabilizing Equipment of the Lloyd Sabado Liner Conte de Savoia," *The Engineer* (London), Vol. 153, I, pp. 32–35, 45, Jan. 1, 1932; II, pp. 62–65, Jan. 8, 1932.
15. FRISCH, E.: "Controlling a Gyro Stabilizer," *El. Jl.*, Vol. 29, No. 7, pp. 408–410, July 1931.
16. SCHILOVSKY, P. P.: "The Gyroscopic Stabilization of Ships," *Engineering* (London), Vol. 134, No. 3491, pp. 689–690, Dec. 9, 1932; Discussion, Vol. 135, pp. 53–54, Jan. 13, 1933; p. 107, Jan. 27, 1933.
17. VITERBO, F.: "Girostabilizzatori antirullanti (Anti-Rolling Gyro Stabilizer)," *Ingegnere*, Vol. 5, No. 3, pp. 140–152, Mar. 1931.
18. HAUS, F.: "Stabilité automatique des Avions," *L'Aéronautique*, Vol. 14, Nos. 156–159, May, June, July and August 1932. English translation, "Automatic Stability of Airplanes," N.A.C.A. Tech. Memo. No. 695, Dec. 1932.
19. GREEN AND BECKER: "Radio Aids to Air Navigation," *Elect. Engrg.*, Vol. 52, No. 5, pp. 307–312, May 1933.

20. SPERRY, E. A., JR.: "Description of the Sperry Automatic Pilot," *Aviation Eng.*, pp. 16–18, Jan. 1932.
21. HUGGINS, M.: "Gyropilot Goes Cross-Country," *Aero Digest*, Vol. 17, No. 1, pp. 51–52, July 1930.
22. LA PIERRE, C. W.: "An Improved Photoelectric Recorder," *G. E. Rev.*, Vol. 36, No. 6, pp. 271–274, June 1933.
23. BERNARDE AND LUNAS: "A Recorder for Minute Quantities," *El. Jl.*, Vol. 30, No. 3, pp. 108–109, March 1933.
"A new Electronic Recorder," *Elect. Engrg.*, Vol. 52, No. 3, pp. 168–170, March 1933.
24. GULLIKSEN, F. H.: "Recent Developments in Electronic Devices for Industrial Control," A. I. E. E., Paper No. 33–24, 1933.
25. ALFRIEND, J. V.: "Light Sensitive Process Control," A. I. E. E. Paper No. 33–45, 1933.
26. HARDY, A. C.: "A Recording Photoelectric Color Analyzer," *Jour. Opt. Soc. Am.*, Vol. 18, No. 2, pp. 96–117, Feb. 1929.
27. BAYLE, M. A.: "Un Héliostat Sans Mouvement d'Horlogerie," *Rev. d'Optique*, Vol. 10, No. 12, pp. 495–511, Dec. 1931.
28. OSBOURNE, A.: "Teaching the Robot to Think Ahead," *Chem. and Met. Engr.*, Vol. 38, No. 4, pp. 245–247, April 1931.
29. NEIMAN, *American Machinist*, Vol. 66, pp. 895–897, May 26, 1927.
30. BUSH AND HAZEN: "Integraph Solution of Differential Equations," *J. F. I.*, Vol. 204, No. 5, pp. 575–615, Nov. 1927; also as M. I. T. E. E. Dept. Bul. No. 59.
31. BUSH, V.: "Operational Circuit Analysis," John Wiley and Sons, Inc., 1929.
32. BUSH, V.: "The Differential Analyzer," *J. F. I.*, Vol. 212, No. 4, pp. 447–488, Oct. 1931.