ADAPTIVE CONTROL SYSTEMS PHILOSOPHY

Dr. Yao Tzu Li
Associate Professor of Aeronautical Engineering
Massachusetts Institute of Technology

INTRODUCTION

Literally, an adaptive control system or self-adapting system is a control system which can produce a desirable performance under adverse changes in the operating conditions. To be sure, all control systems are designed to have some degree of adaptive ability. The glorified title of this new interest in the instrumentation field is, therefore, to emphasize the need for more adaptive ability in a system which did not have enough.

To accomplish this object, within the capability of the primary components, one may have to use several known schemes instead of only one, as in the case of simple control systems. When combination schemes are used in a system, a knowledge is required of the exact role and proportion for each scheme so that the overall performance may be improved while the physical components of the system may actually be simplified. The object of the present paper, therefore, is to establish from a basic philosophical point of view most of the possible schemes and arrange them in a general pattern. Some of the schemes are devised in existing works. By inquiring into basic philosophy several new schemes are thereby generated. With the classification of the design philosophy the true merits and possible limitations of each scheme can also be compared.

OBJECTS AND METHODS OF ADAPTIVE CONTROL SYSTEMS

Control systems may be classified according to the objective for which they will be used. For instance, a control system may be desinged to yield an output following an input with a minimum steady state error as in the case of an ordinary positional servo. It may also be designed to make some other physical operating system yield an optimum efficiency index as in the case of the cruise control of an aircraft. All these systems are designed to accomplish specified objectives under changing operating conditions and are therefore adaptive systems or self-adapting systems. But the current practice is to associate the name " adaptive control" only with the dynamic response of a control system. This is because most of the simple feedback systems we now have can provide enough self-adaptive capability for static performance, but not enough for dynamic response. At the same time, with the advance of modern aircraft and missiles, the system must go through an environmental change much larger than in conventional situations so that tighter specifications regarding the dynamic response of any individual system involved are needed. Application to aircraft and missiles, in fact, has stimulated the growth of current interest in adaptive systems. People in this field recognize that an adaptive system is identified with the dynamic response of a control system. But to people in many other fields some confusion does exist as to the exact meaning of an adaptive system. For this reason, Table I is prepared to show the classification of adaptive control systems based upon their objectives and the possible methods that may be used to accomplish these objectives.

The objectives as listed in Table I are three, but these are neither exhaustive nor mutually exclusive. The first objective is to obtain the static performance and this is usually accomplished by feedback. Comparable results may be obtained by programming the interferences and the input. The second objective is to make some other physical operating system yield an optimum performance as measured by such factors as efficiency or economic index. For this type of control, the optimum performance is often related to the control inputs in a nonlinear manner and exhibits an optimum condition which changes with environmental effects. Programming and optimalizing systems are two possible methods to accomplish this objective. The third objective is to control the dynamic response of a system. Table I shows a few possible schemes that can be used to improve the self-adaptive ability as applied to dynamic response. A more detailed treatment is discussed after a brief review of the characteristics of the reference system.

REFERENCE SYSTEMS

Once the objective is fixed in an adaptive system, the next thing to be determined is the reference standard and the associated specifications regarding the deviation of the actual performance from the reference standard. For static performance, the reference is usually the input. For economic index, the reference is the best possible performance. For dynamic response, the reference may be described in the form of a reference system either analytically or represented by an analogue model. In addition to the reference system, a typical input test function and a method to measure the deviation of the actual system output from the reference system output may be needed. For example, by the use of a random signal as a test input and the use of correlation methods, one can establish the dynamic characteristics of the actual system or the reference system. For most analytical works the performance of the reference system is often taken as unity. But for the operation of a dynamic response adaptive system, the reference system should represent the most practical and feasible performance consistent with the consideration of the types of the disturbance input signal. Generally speaking, there are two types of disturbance input. The first type of disturbance input produces an output which may be considered as noise. The second type of disturbance may change the response of the system to the command input, but produces no output directly. When the disturbance is an active input signal, then the reference system should be one which yields the minimum combined error due to the dynamic effect of the input signal plus the noise output. If the dynamic characteristics of the input and the disturbance are known in a convenient form such as a frequency spectrum then the desirable characteristics of the reference system are theoretically known. For instance when signal and noise are random in natural order, then the reference system would be the ideal filter formulated by Wiener.

When the disturbance is not a direct signal but a modifying input which changes the parameters of the actuating system, then two desirable conditions may exist. In one case, the system is forced to produce the best possible dynamic response for each environmental condition. In the other, the reference system is chosen according to the capability of the system and the characteristics of the output function under the worst environmental conditions.

ADAPTIVE CONTROL WITH LINEAR FEEDBACK SYSTEMS

Feedback is the most powerful scheme for increasing the adaptive capability of the static performance. The same principle can certainly be applied to systems for improving the dynamic response adaptability with somewhat more difficulty. For static performance, adaptive ability is achieved when the forward loop sensitivity is very high so that the overall performance is dominated by the characteristics of the feedback component. To extend this principle to dynamic response one should have a forward loop with a high sensitivity and low phase shift over the desirable operating frequency range. In the feedback branch, a system equivalent to the reciprocal of the reference system is installed as shown in Figure 1-a. Thus the overall system performance is forced to be equal to the reference system despite the possible change in characteristic of the variant parameters in the forward loop. As a variation of this scheme one may have a reference system in cascade with a unit feedback system as shown in Figure 1-b. In this latter arrangement, the feedback system should yield a performance function of unity over the entire range of operation in order to make the complete system behave like an adaptive system. This can only be accomplished when the forward loop continues to give high gain and low phase shift under the entire range of environmental conditions. One possible method to enforce this property is to use a type of compensation which introduces signals as a function of the rate or acceleration of the input to balance the forward loop lag. In doing this, one must bear in mind that if the lag is due to a certain parameter which is affected by environmental conditions, then only the part which is not affected can be balanced. For instance, if the mass of an airplane constitutes a certain forward loop lag, and a large part of the mass remains unchanged, then the lag due to this unchanged part of mass can be compensated for. When the lag due to invariant parameters is all properly compensated for, then a much faster system with a much higher forward loop gain can be established. With this modification a practical adaptive system utilizing linear feedback may be realized.

ADAPTIVE SYSTEM WITH RELAY SYSTEMS IN THE FORWARD LOOP AND A FEEDBACK TO OVERPOWER THE VARIANT PARAMETERS

As described in the last section, the system shown in Figure (1) would operate as a dynamic response adaptive system only when the forward loop gain is high. This may be accomplished by compensation. Another method is through the use of relay control. A relay control can provide large amplification of power at relatively high speed. One basic drawback of this type of controller is the inherent exaggerating of oscillation when the power drive system has second or higher order performance characteristics. The oscillation may be illustrated by Figure (2-a) and Figure (2-b). In Figure (2-a), a first order drive system is assumed. The output signal for this system is a saw tooth function bounded by the switching zone. This type of oscillation is, in general, considered satisfactory because the amplitude is no larger than the switching zone. In Figure (2-b) the output drive is assumed to be a second order system. Now the amplitude of the oscillation is larger than the switching zone. Since an airplane is usually a higher order system, it would show a rather exaggerated oscillation when controlled by simple two-way relay system. For this reason when a relay system is used some scheme must be incorporated to limit the oscillation; and the oscillation as shown in Figure (2-a), with the switching zone as a limit, may be considered as the ideal condition. When this is fulfilled, assuming the switching zone can be squeezed down smaller than a given tolerance, we have an adaptive system for all environmental conditions in which the output drive system has enough power to maintain the small oscillation straddling the ideal output.

The above discussion shows that for higher order drive systems, such as in airplanes, the use of relays as a means to get fast power amplification must incorporate schemes to minimize the oscillation. The ideal result is to make the relays approach a fast first order system. By emphasizing fast speed, it insures the ability for the output to straddle the ideal output when the latter is making a fast change.

To illustrate the possible schemes for minimizing the oscillation of a relay control used with higher order drive systems, a concept block diagram of this control system is shown in Figure (3). In this diagram, the command and the output is fed into a computer which in the simplest form would be a comparator. The output of this unit controls the timing of the relay. In the simplest form, a two position relay is usually used. The power is supplied from a source through a magnitude adjusting device. The output from the drive system is split to show an average output per cycle and the associated oscillatory component.

The concept block diagram of Figure (3) illustrates the basic characteristics and ingredients involved in the operation of a relay servo. One primary requirement is to keep the average output within the switching zone.

The other is to minimize the oscillation. Oscillation is caused by the dynamic characteristics of the drive system and is affected by the magnitude of the drive power and the relay switching timings. To achieve the design objective of a closely controlled average output and minimized oscillation, the following schemes may be used.

1. Programming

a. Compensate for the invariant parameters of the drive system to make it behave like a fast first order system. Compensation may be done by feedback from the output signal to modify the drive power magnitude as shown by the dotted line of Figure (4). Equivalent drive power magnitude modification may also be achieved with constant drive power but gated to give an adjustable duration.

b. Use multiple relay having switch timing of fixed relationships with error functions. As a typical example, the relay action may take place when the error is equal to zero and when the error rate is equal to zero. For each relay engagement the drive power may have fixed magnitudes at various levels. The purpose of the multiple control can be shown by Figure (5). In this figure one typical cycle of the oscillation is shown. The desired characteristics are to make section " a" of the output to be driven back toward the ideal output with highest power, and section " b" to coast over with sufficient speed to minimize the time and yet without too high an approaching speed at the crossing point. A similar situation is repeated at section " c" and " d" . The overall criterion is to minimize the integrated error with some constraint such as the maximum acceleration of certain components.

c. Fixed multiple relay timing with magnitude of the drive power is programmed according to input function as shown in Figure (6). The sensitivity used in the programming is based upon the drive system invariant parameters. Figure (7) illustrates the situation for three hypothetical cases. Case " a" - the input is moving slowly and a normal output oscillation results. Case " b" shows the output lagging behind when the input experiences a fast acceleration and when only normal drive power is used. Case " c" shows the desirable output condition for similar inputs as shown in case " b" when the drive power level is properly adjusted.

d. Use fixed multiple relay timing with the drive power programmed according to the output, as shown in Figure (8).

Schemes " c" and " d" aim at the same problem, namely to make the system follow a certain command signal faster. In doing so, the power drive for the various sections of output function, as shown in Figure (5), would be asymmetrical with respect to the average output. When scheme " d" is used, the output signal used to actuate the programming should be the average output instead of the instantaneous output. This is one reason why scheme " d" is not as easy to execute as scheme " c" .

The logic of the relay servo developed by Flugge-Lotz and Taylor, involving the switching of the feedback sensitivity, seems to be based upon an intuition to achieve the goal as outlined in scheme " b" . In their system, however, a certain part of scheme " d" is also involved, and yet is not blended in the most desirable proportion.

2. Non-Linear System with Feedback Regulation to Limit the Oscillation Amplitude

The undesirable oscillation of a relay system depends upon the magnitude of the drive power and the timing. It is possible to arrange a feedback system to regulate the oscillation amplitude to a tolerable limit by means of the following two schemes.

a. Oscillation amplitude limited by controlling the magnitude of the drive power, as shown in Figure (9).

b. Oscillation amplitude limited by controlling the relay engaging timing with respect to the error signal as shown in Figure (10). Figure 11 shows that the relay engaging timing may be shifted ahead by the following three methods.

(1) By a controlled bias added to the zero error.

(2) By a controlled bias added to the zero error rate.

(3) By a controlled delay time applied to the zero error rate.

All the above feedback schemes must derive the feedback signal from the oscillatory output amplitude. Some suitable scheme would be needed to yield a signal proportional to the amplitude. Other constraining factors such as acceleration limit of certain control members may also be added when desired.

Like all feedback systems there is a dynamic stability problem that requires limiting of the loop gain and therefore the response speeds; whereas the programmed types, as outlined in Scheme 1, are not subject to this limit. Since programmed schemes can only be applied to the nonvarient parameters, a combination of the two schemes may prove to be most advantageous.

In all the systems thus far presented, the adaptive ability of a system is achieved by a strong forward loop paired with a feedback to render the effects of the variant parameters in the drive system relatively unimportant.

Programming of the nonvariant parameters is recommended to strengthen the forward loop whenever necessary. The schemes to be described in the following sections approach the problem in a different manner. In these systems the effects of the variant parameters upon the dynamic response of the system is counterbalanced by the adjustment of some compensating parameters in the controller. The first necessary condition is the realization of such an adjustable parameter or a set of parameters which can offset the effects of the variant parameter. The next problem is to find a suitable scheme to execute the adjustment. The easiest one to be compensated for is the sensitivity of the system. This is because the overall sensitivity of the system is equal to the product of the sensitivities of various components; and a change of the sensitivity of one component can be compensated for by the adjustment of the sensitivity of some other component. But the physical parameters of a component usually effect all the dynamic characteristic parameters such as natural frequencies and damping ratios. In other words, a change of one physical parameter may shift the position of all the poles and zeros of a physical system. To compensate exactly for these effects the controller should create a reciprocal dynamic characteristic function of the system to be compensated for, so that the shifting of the corresponding parameter of the controller may counterbalance exactly the effect of the variant parameter in the controlled system. In practice this exact compensation is too complicated to be useful and some sort of approximation is always necessary. For instance, the transient response of a second order system with normal damping ratio may be approximated by the transient response of a low damped system cascaded with a first order system. Generally speaking, removing the effects of a variant parameter in a higher order system by parameters of a lower order system can usually be accomplished under the following two conditions: (1) when some of the poles of the higher order system are not effected significantly by the variant parameter, and (2) when the total effects of the shifting of the poles of the higher order system yield approximately the same effects, as the shifting of a set of poles in the controller of lower order. The cascaded first and second order system, described before, illustrates this latter condition.

ADAPTIVE SYSTEM BY PROGRAMMING OF THE COMPENSATING PARAMETER

The first possible scheme to make the necessary adjustment of the compensating parameter described in the previous section involves the use of a programmed control system which makes the adjustment based upon the actual measurement of the variant parameter or the environment conditions which affect the variant parameter. The programming control is designed based upon knowledge of the relationship between the variant parameter and the compensation parameter. Generally speaking, a precise programming control is difficult to realize; but a moderately accurate system is quite practical.

ADAPTIVE CONTROL SYSTEM BY ADJUSTING THE COMPENSATING PARAMETER THROUGH FEEDBACK

The adjustment of the compensating parameter would change the system dynamic response. If a device is available to interpret the dynamic response in terms of a performance index, and if a particular value of the performance index is known to represent the desired dynamic response, then it is simple to design a feedback system to adjust the compensating parameter to force the system performance index equal to the desired performance index. The key to the realization of this scheme depends upon the performance index generating system. Generally speaking, it is difficult to express the dynamic response in terms of a performance index on an absolute scale upon which the reference system corresponds to a fixed value. On the other hand, a reference system can be represented by an analogue model. The difference of the output between the actual system and that of the model, when both are subjected to a suitable test input, may be minimized through the adjustment of the compensating parameter. Thus an actual model is used to represent the reference system instead of a performance index. For practical reasons, the model can be made corresponding to the dominating poles of the reference system. This simplification is subjected to the same qualification as discussed before for the adjusting of the parameters. Despite this simplification, a model can represent the reference system much better than a performance index. The use of a model requires an optimalizing controller instead of a simple feedback controller.

ADAPTIVE CONTROL SYSTEM BY ADJUSTING THE COMPENSATING PARAMETER THROUGH OPTIMALIZING CONTROLLER

An optimalizing controller is one type of feedback control which makes adjustment of certain parameters of a controlled system, to force a certain output of the controlled system to reach the optimum level. The difference between an optimalizing control system and ordinary feedback lies in the fact that the optimum output level does not represent a definite output level. Without a definite desired output level, it is impossible to generate an error signal to be used in ordinary feedback control system. An optimalizing system incorporates some suitable form of test adjustment and is followed by a waiting period to observe its effect upon the controlled system and to make further adjustment after the correct direction of proper adjustment is established. For this reason, an optimalizing controller may be regarded as possessing one degree higher of intelligence than an ordinary feedback system. For dynamic response control system, an optimalizing controller may be used to adjust the compensating parameter under one of the three following conditions.

(1) When the desirable dynamic response is the best possible dynamic response a system can produce under a given environment. This

means the reference system dynamic characteristic function is unity. If a performance index signal generating system is available to convert the dynamic response of the controlled system into the performance index, then the optimalizing system may receive the performance index as the input signal to control the compensating parameter and thus close the loop. A typical functional block diagram is shown in Figure 12.

(2) When the reference system represents a compromised performance and the purpose is to make the controlled system behave exactly the same as the reference system, but, because of practical reasons the compensating parameter is not a perfect one, the performance index generating method can only give an approximate value. For these reasons, the relationship between the performance index and the compensating parameter may become quite nonlinear so that the optimum value of the performance index represents the desirable operating condition. An optimalizing controller should therefore be used for this situation instead of the ordinary feedback as described earlier.

(3) When the reference system is represented by a model and the controlled system is matched with the reference system by comparing the difference of the output function subject to some suitable inputs. Several methods can be used to compare the two output functions. All involve some sort of integration of the function of the absolute value of the difference. As a result, the best match is obtained when the difference between the functions of the two outputs is minimized. For this reason an optimalizing controller is needed as shown in Figure 13. The adaptive system designed by M.I.T. is a typical example of this type. This system seems to be the most promising type among all systems involving compensating parameter adjustment.

CONCLUSION

The above analysis shows that a large number of possible schemes may be used to accomplish dynamic response adaptability. Since the characteristics of an actual controlled system and its environment may vary very much from one to the other, it is rather difficult to draw a conclusion to say which system is definitely better than the other. A well-engineered feedback adaptive system with relay control in the forward loop, as developed by a Minneapolis-Honeywell group, and an equally recommended system utilizing an optimalizing controller, as developed by Massachusetts Institute of Technology group, have both produced satisfactory results. It is true that when changes in environmental conditions are moderately severe and the desirable output is moderately fast, many schemes can yield satisfactory results. But when the output demand is high and the environmental conditions are severe, then a real test of the capability of a given scheme is on hand.

A suitable figure of merit may possibly be established to assess the achievement of a design with different types of input functions, various response speeds of the reference system, specified tolerance of the deviation of the dynamic response, and variation of the operating conditions. A good design is one which can achieve a high figure of merit and yet not be too complicated in construction through the use of a well balanced scheme of the various principles.

TABLE I

CLASSIFICATION OF ADAPTIVE CONTROL SYSTEMS

I. According to objectives

  A. A constant static performance between command and output.

  B. Optimum performance in the form of an economic index.

  C. Matching the dynamic response of an operating system against that of a reference system, with the reference system dynamic response established as following:

    1. A fixed reference system representing the best possible performance for the entire range of operating conditions.

    2. A fixed reference system representing the best possible performance for the worst operating conditions.

    3. A reference system with performance programmed according to the maneuver requirements of the command signal and the capability of the drive system.

    4. A reference system with performance programmed according to the types of command signal and active interference signal.

II. Possible schemes for the matching of the dynamic response of an operating system against that of a reference system.

  A. Use feedback to overcome the effects of the variant parameters.

    1. Linear system with compensation for invariant parameters.

    2. Use relay control to manipulate the drive power.

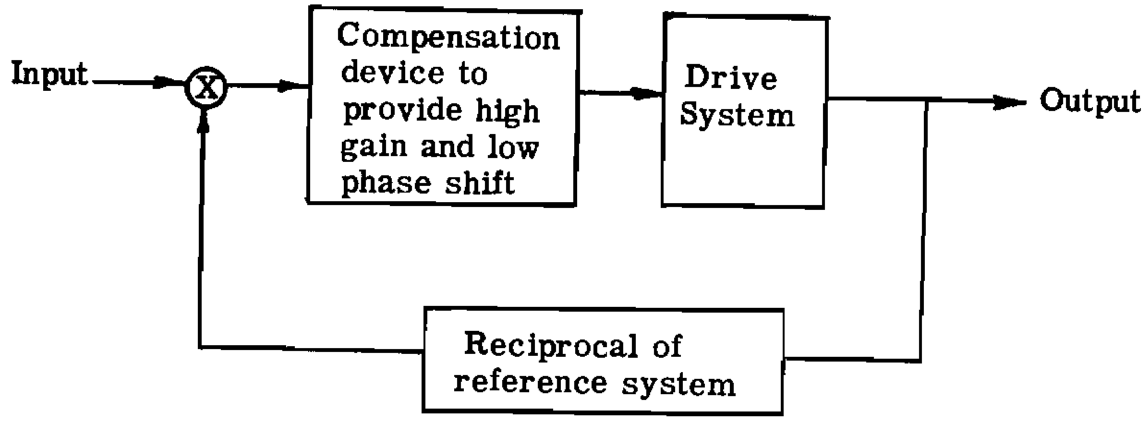
      a. Use various programming methods to minimize the inherent oscillations.

      b. Use various schemes of feedback to minimize the amplitude of oscillation of the limit cycles.

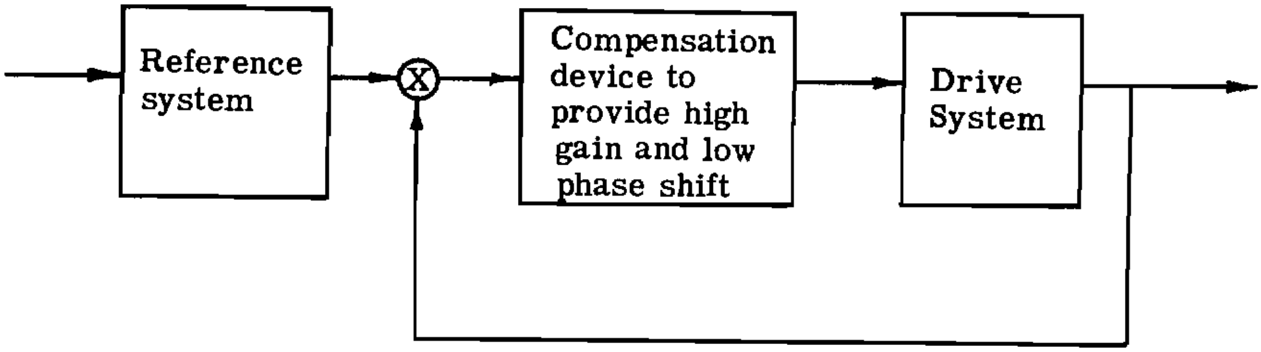
  B. Use adjustable parameters to compensate for the effect of the variant parameters.

TABLE I (CONT'D)

1. By programming with suitable computor.
2. By linear feedback with suitable dynamic performance index indicator.
3. By optimalizing system, with either a suitable dynamic performance index indicator or a reference system model.

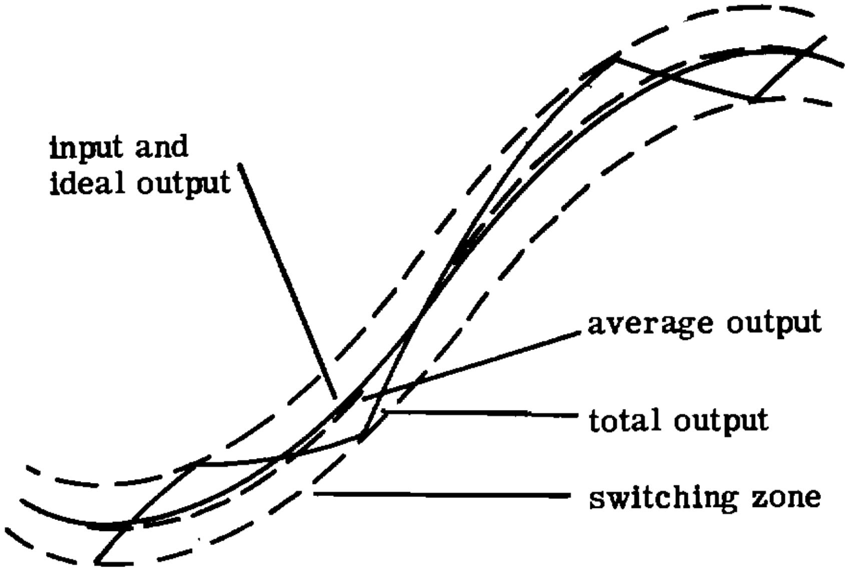


a. Use reference system in feedback loop.

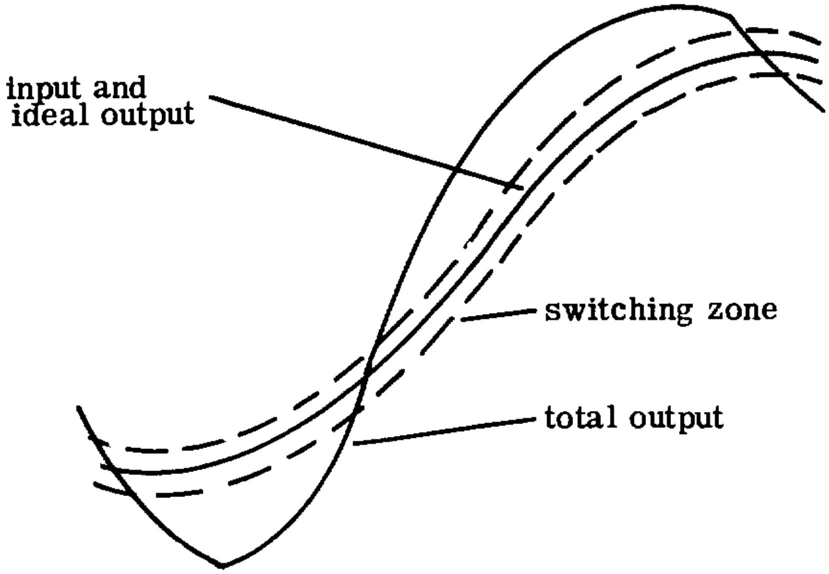


b. Use reference system in cascade with a unit feedback system.

Fig 1. Basic schemes using feedback to overcome effects of the variant parameters in the drive system.



a. Relay control with first order drive system



b. Relay control with high order drive system

Fig. 2 Input-output relationship with relay control system for different types of drive systems.

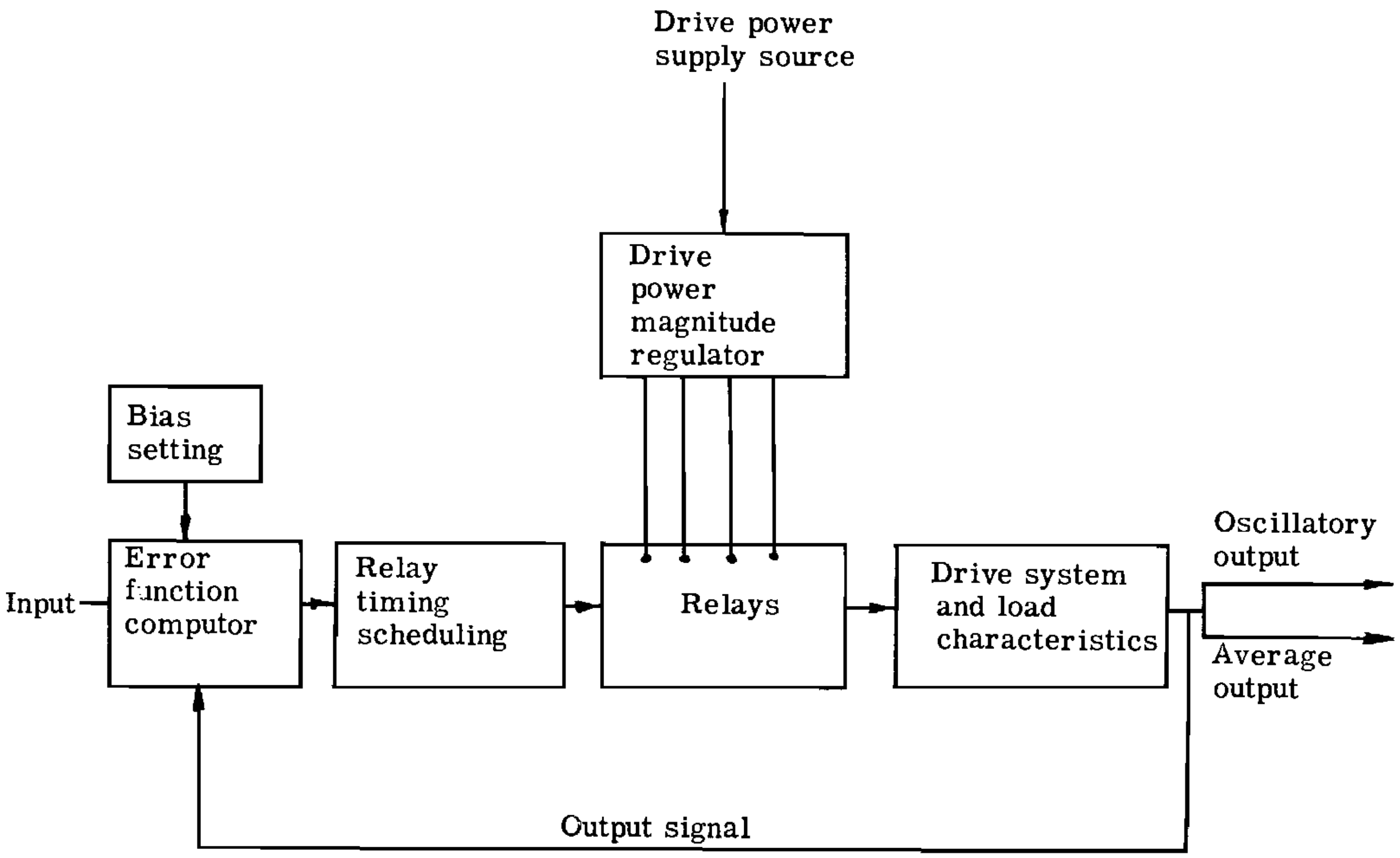


Fig. 3 Basic concept block diagram of relay control system

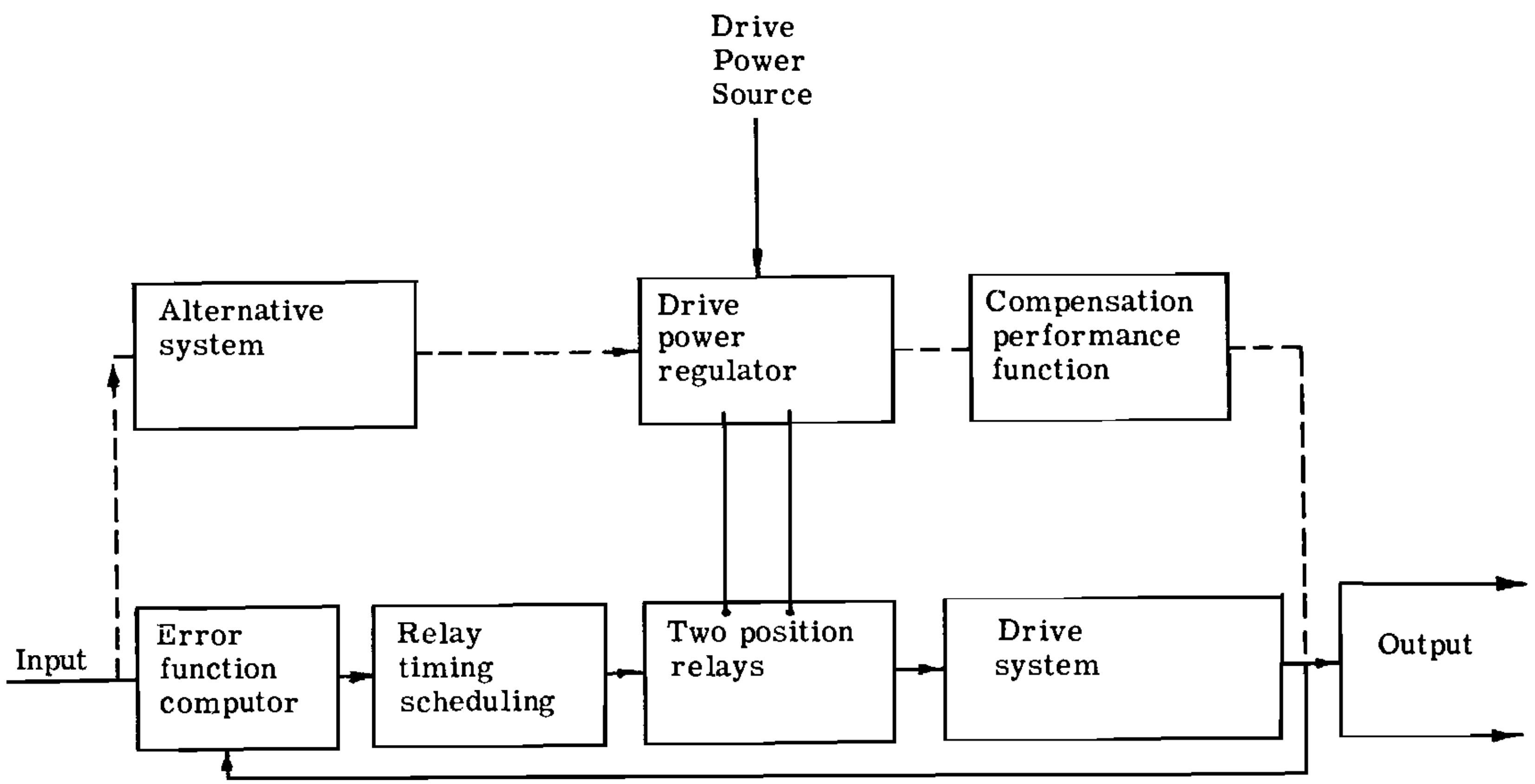
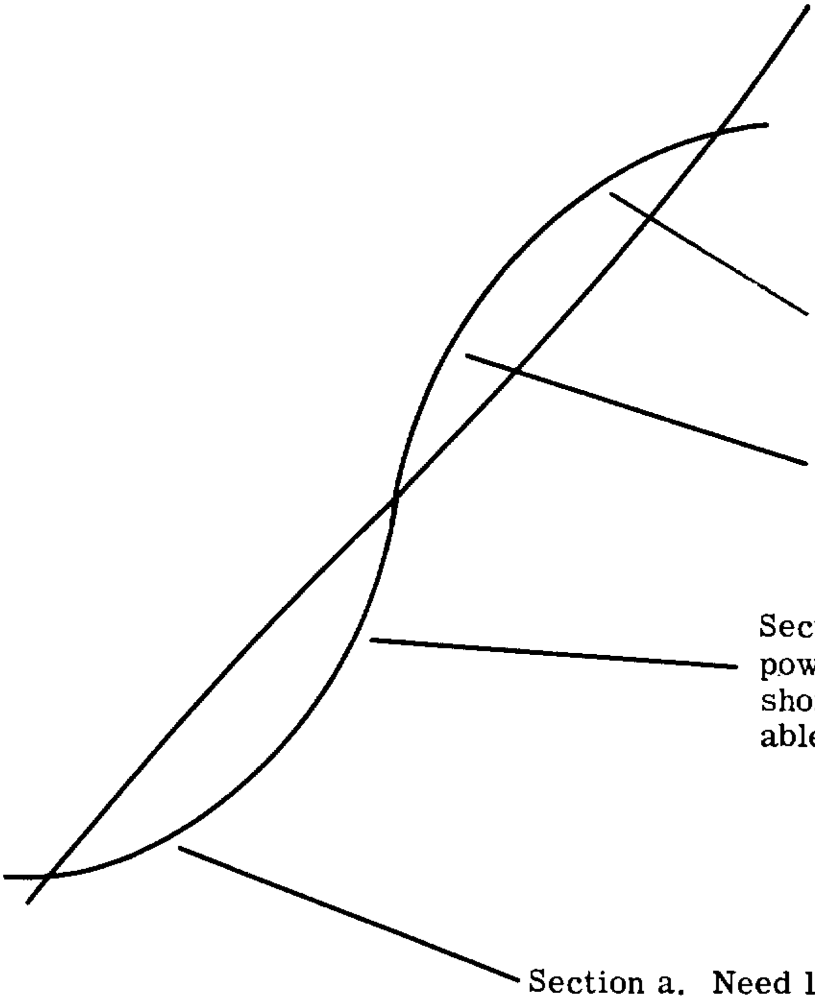


Fig. 4 Block diagram to show the use of feedback from output or input to compensate for invariant parameter and thereby make the system performance function approach that of first order system between switching



Section d same as b

Section c same as a

Section b. Need small drive power to balance the desirable short coasting time and undesirable approaching speed.

Section a. Need large drive power to speed recovery from diverging course.

Fig. 5 Power requirement at different section when four way relay is used.

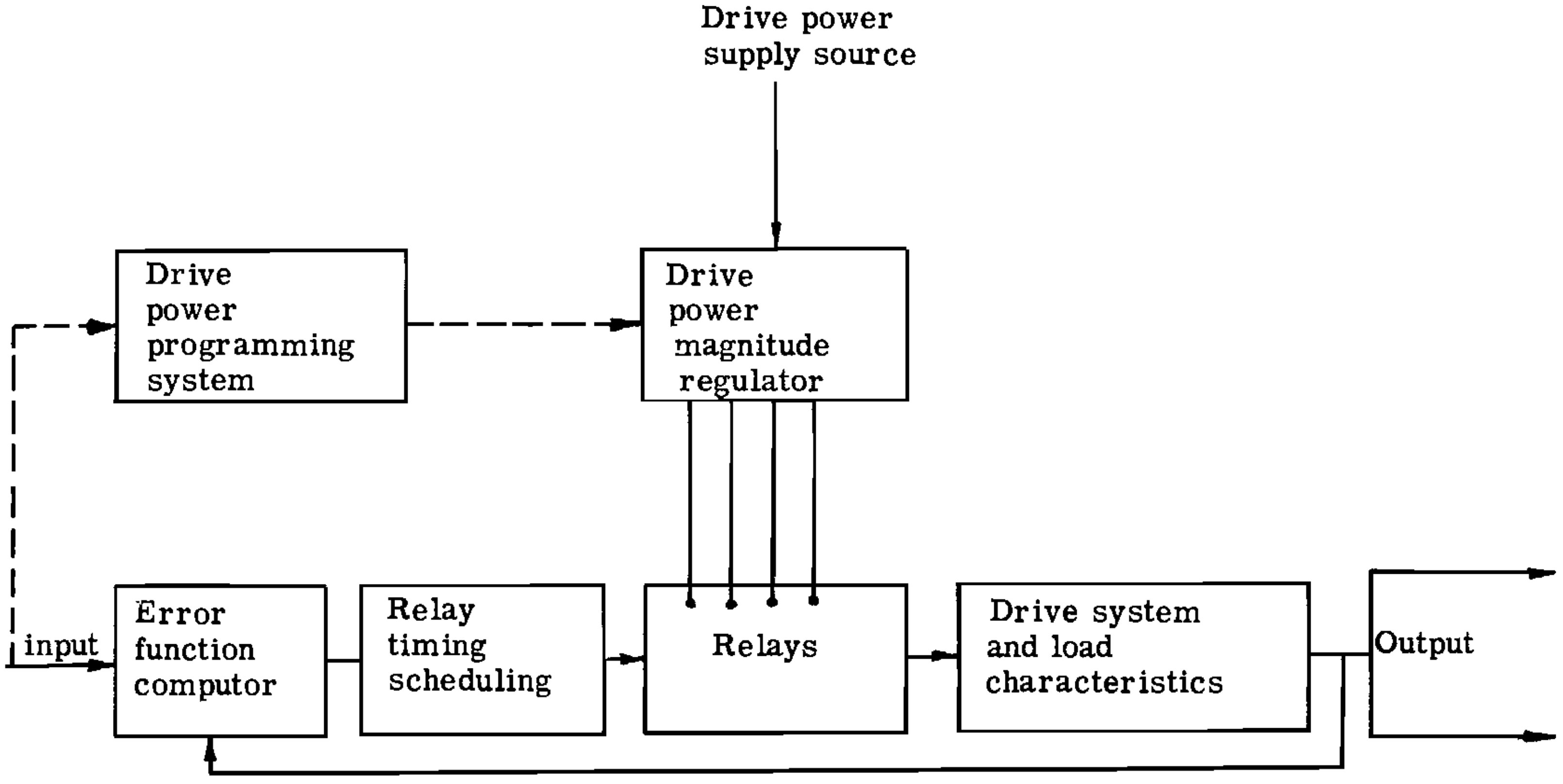
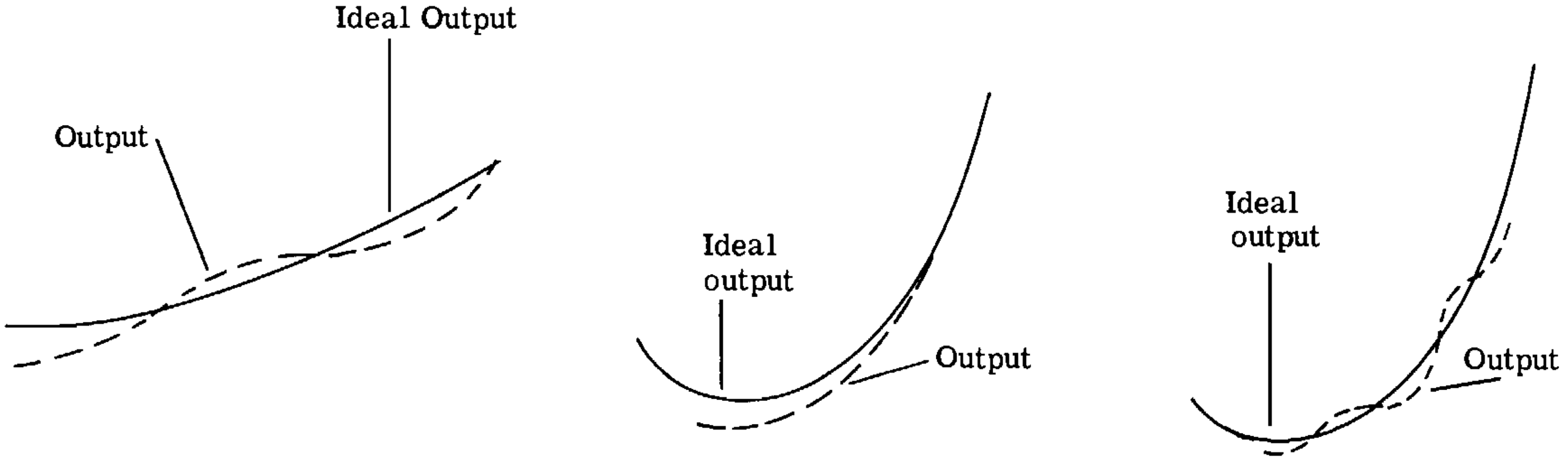


Fig. 6 Multiple relay with drive power magnitude programmed according to the input function.



a. Low load demand and low drive power b. High load demand and low drive power c. High load demand with programmed drive power.

Fig. 7 Output oscillation with drive power programmed according to input function.

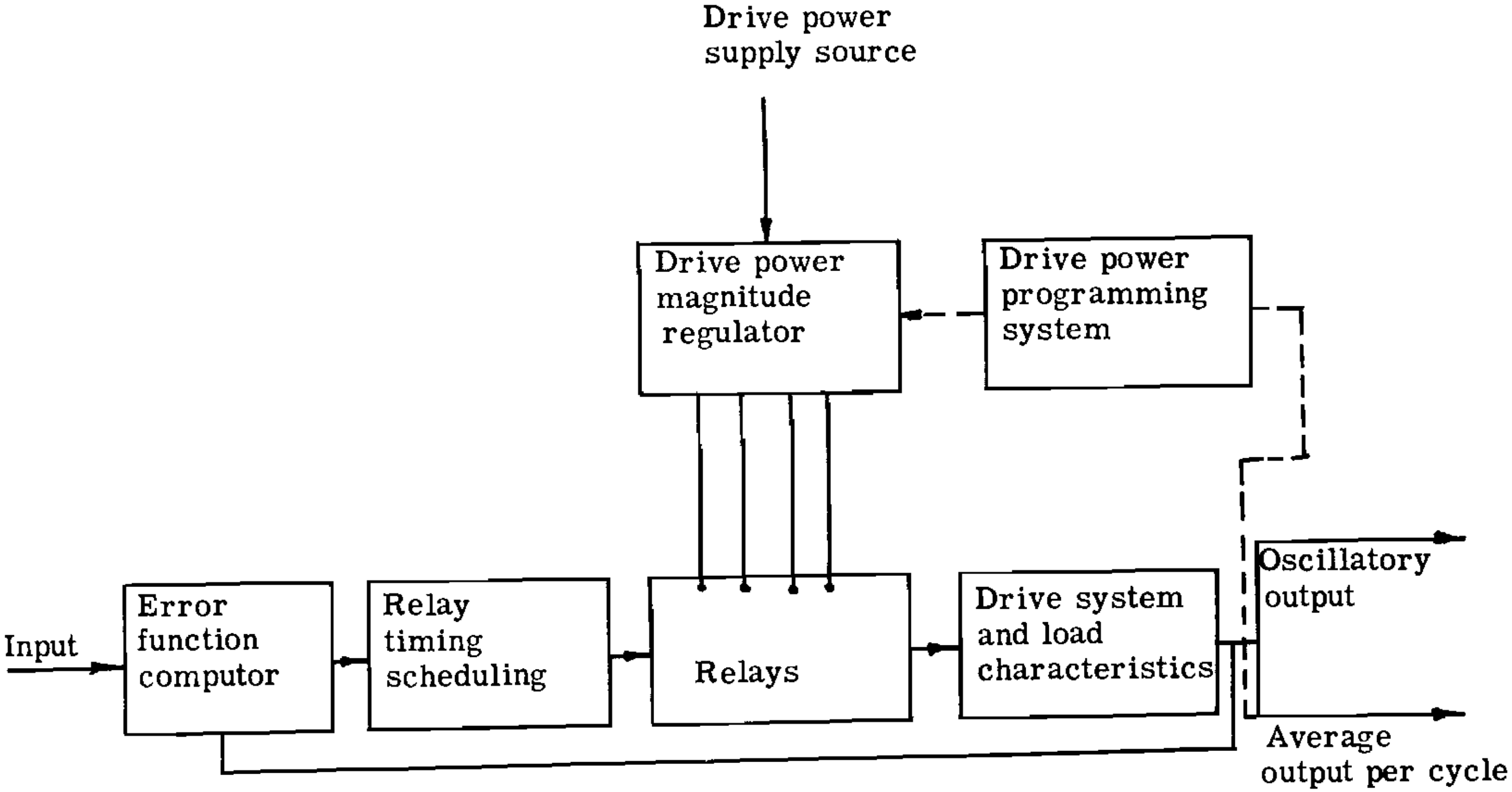


Fig. 8 Multiple relay with drive power programmed according to the average power per cycle.

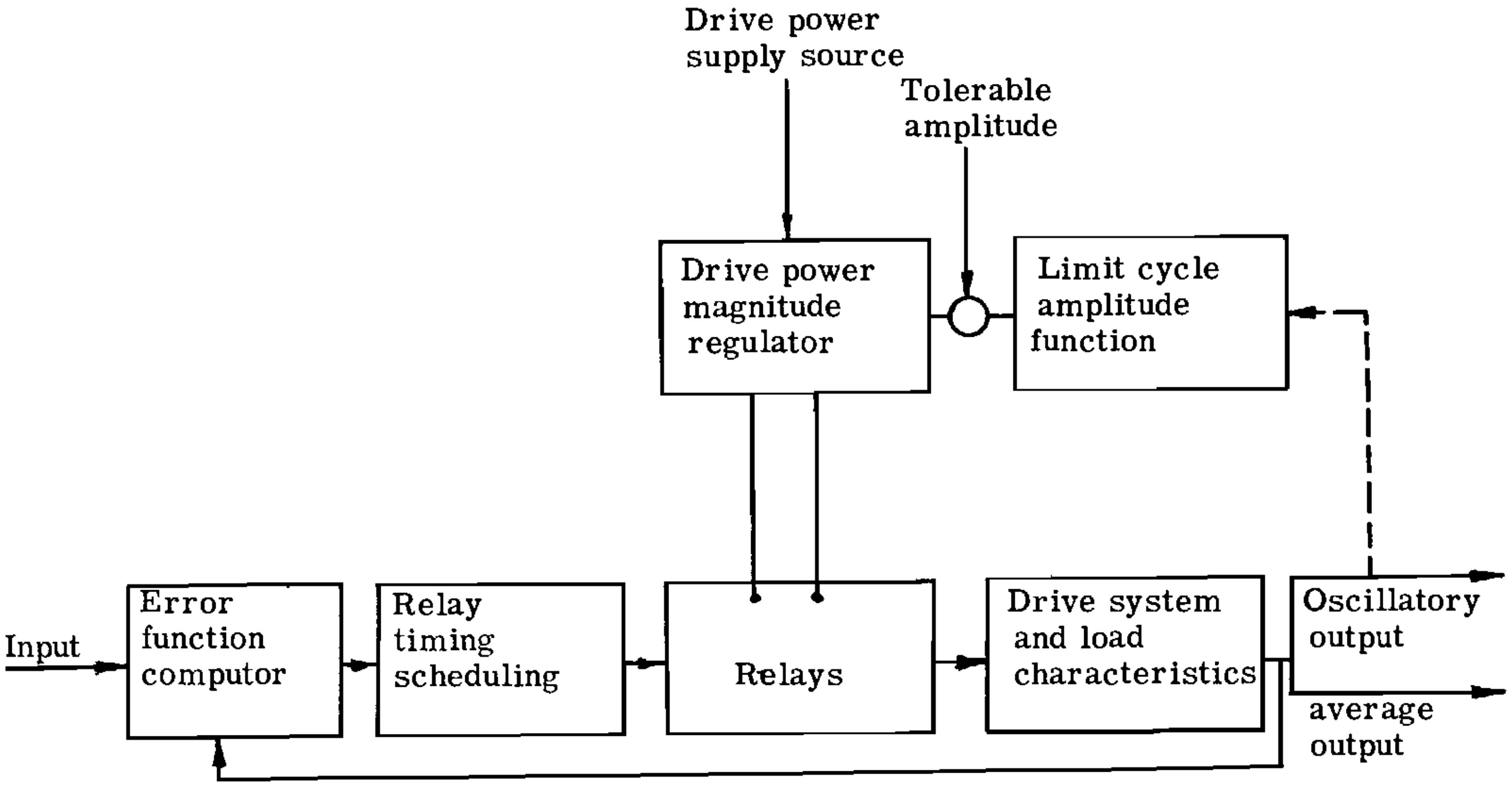


Fig. 9 Drive power magnitude adjustment through feedback of amplitude of oscillatory output.

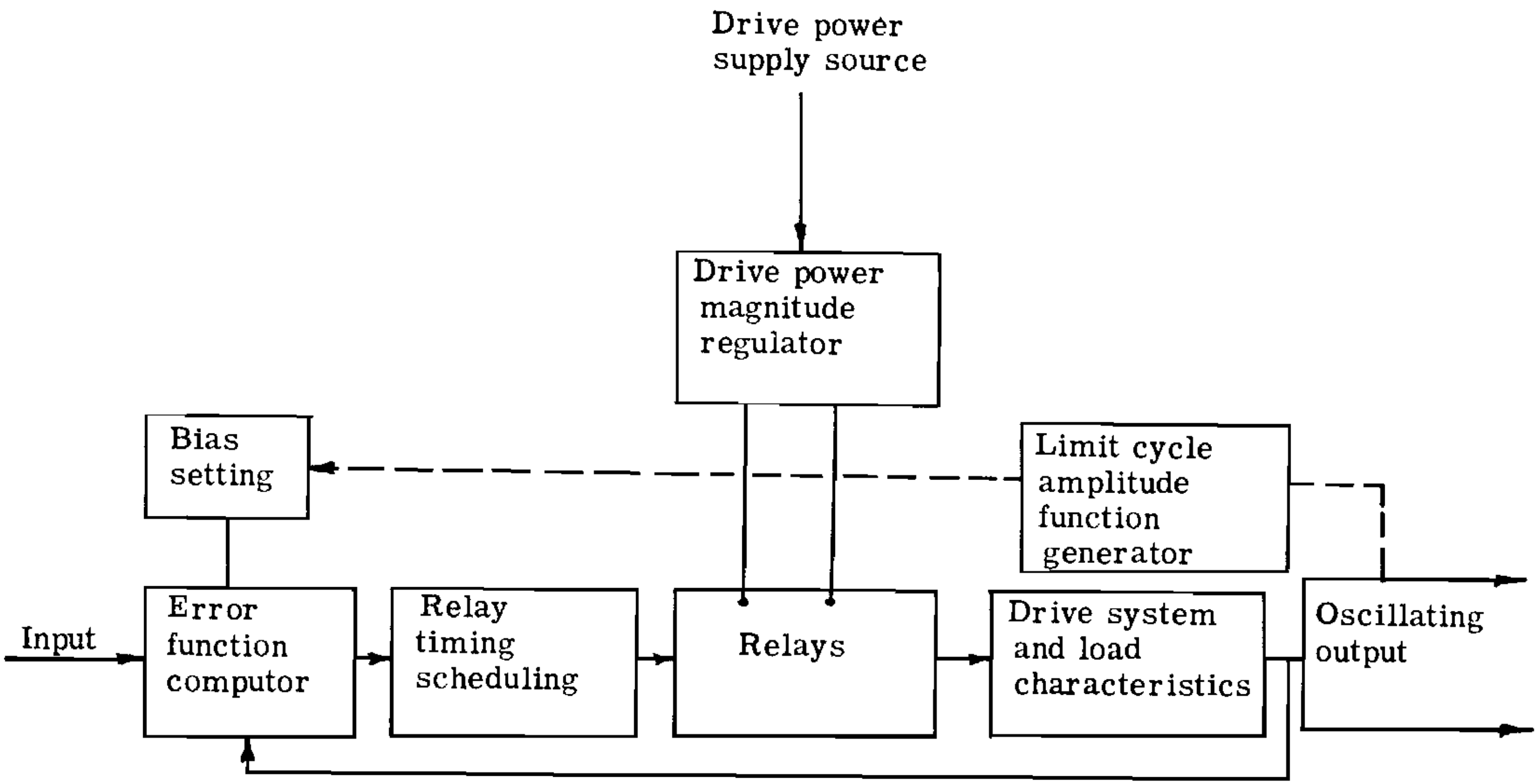
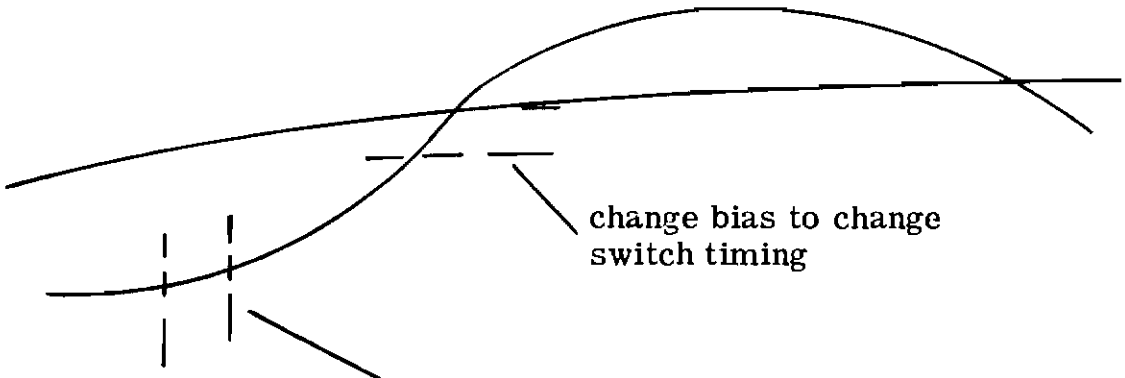


Fig. 10 Drive power switch timing adjustment through feedback of amplitude of oscillatory output.



change bias to change
switch timing

add time delay to change
switch timing

Fig. 11 Methods to change switch timing.

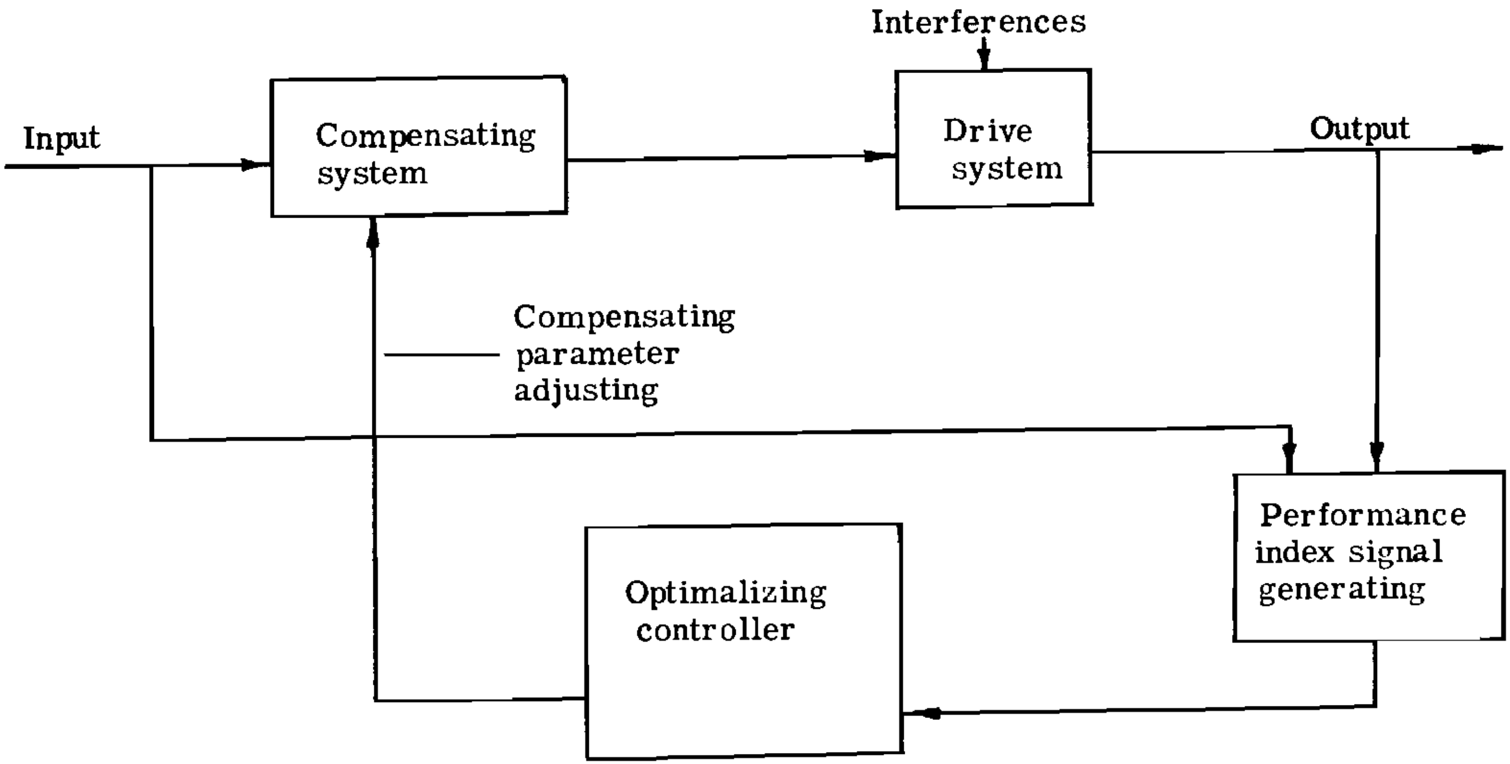


Fig. 12 Typical dynamic response optimalizing system with the use of a performance index signal generating system.

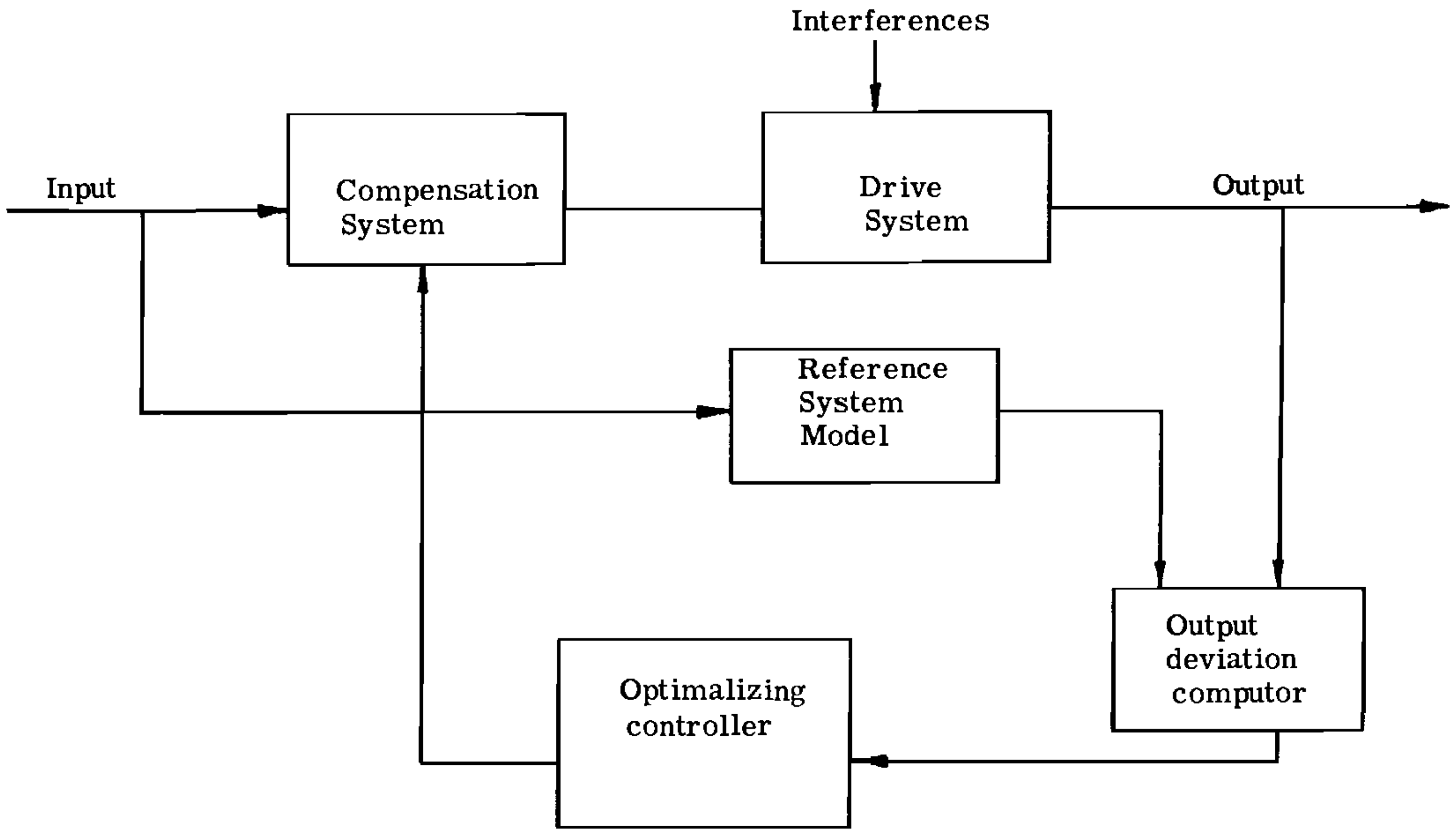


Fig. 13 Typical dynamic response optimalizing system with the use of a reference system model and an output deviation computor.