OPTIMUM LINEAR ADAPTIVE DESIGN OF DOMINANT TYPE SYSTEMS WITH LARGE PARAMETER VARIATIONS*

I. Horowitz

Dept. of Electrical Engineering, University of Colorado, Boulder, Colo.

ABSTRACT

This paper presents a significant improvement over the hitherto available linear dominant system design techniques for guaranteeing system response within prescribed bounds, despite large plant parameter variations. Noteworthy features of the new technique are:

1) The mapping of the plant parameter space into the closed-loop system space is exact and permits application to a much wider and more realistic class of problems than previously possible;

2) It is shown how the loop transmission bandwidth may be made very much smaller than in the previous designs, thus considerably extending the applicability of the dominant approach, because of its drastically reduced sensitivity to internal noise.

*The research reported here was supported by the National Aeronautics and Space Administration under Research Grant NGR 06-003-083.

OPTIMUM LINEAR ADAPTIVE DESIGN OF DOMINANT TYPE SYSTEMS WITH LARGE PARAMETER VARIATIONS

I. Horowitz
Dept. of Electrical Engineering, University of Colorado, Boulder, Colo.

A. Introduction

Despite the many recent advances in control theory, one of the most fundamental problems is far from solved. This is the problem of optimum design (for a given prescribed complexity) of a system with parameter variations, so that its time response lies within specified tolerances. By optimum is here meant the very important practical problem of minimizing the effect of high-frequency sensor, amplifier etc. noise, because this is usually the dominant factor in determining the practicality of a theoretical adaptive design. It is true that many nonlinear adaptive structures have been proposed in the literature, but almost without exception there are no design procedures for tailoring their detailed design to any specific numerical problem, which is an essential step for optimization in the above sense. Thus any one of these nonlinear structures may possibly be optimum for a given numerical problem or even for a class, but neither the problem nor the class is known. Hence nonlinear adaptive theory is as yet an art, rather than a science. On the other hand, some progress has been made in developing such a science of linear adaptive theory but even here the situation is far from satisfactory.

Consider the following basic problem: (1) There is a single input-output plant with parameters which may lie (or "slowly" vary) within a given region in parameter space; (2) Specific bounds on (say) the step response are prescribed, such as acceptable range of rise-time, overshoot and settling time; (3) Linear, time invariant compensation is to be used for which the rms effect at the plant input of noise lumped at the sensor, is to be minimized. It can be categorically stated that this fundamental problem in linear adaptive theory is not as yet satisfactorily solved. The infinitesimal variations case has been treated both for statistical[1-3] and deterministic situations.[4] The deterministic case for large parameter variations has been treated by means of frequency response,[5,6] and by s-plane[6,7] techniques, with the latter confined to dominant-type formulations. The dominant approach is considered here and its shortcomings noted, which are alleviated to a significant extent by the present contribution.

The Dominant Roots Approach

In many problems it is reasonable to have the system response be determined primarily by a small number of poles and zeros, which in turn readily permits time-domain performance bounds to be translated into acceptable range of location of these few dominant poles and zeros. The inevitable additional poles are assigned "far-off". The presently available technique[6,7] may best be described around the specific example of Fig. 1, wherein ABCD is the region of variation of the complex poles of a plant transfer function, whose dominant varying part is

$$P = k/s(s^2+sS_p + P_p) \qquad (1)$$

k varies from 1 to 1000 in value; MNQR is the range of acceptable dominant pole (with parameters S_r, P_r) location of the system transfer function

$$T(s) \triangleq P_r p_f F(s)/(s^2 + S_r s + P_r)(s + p_f); [F(o) = 1]; \qquad (2)$$

p_f is the closest far-off pole and F(s), contains within it all other far-off poles and zeros; UVW is the boundary to the right of which the far-off poles may not cross. One may argue over the specific location and shape of the far-off pole boundary UVW, but the important point is that such a choice must be made if the design is to be of the dominant type. The range MNQR, $\overline{MNQR}$ is approximately that which has been suggested as acceptable in flight control.[8]

The design philosophy is to locate loop transmission [L(s)] compensation zeros (Z, $\bar{Z}$) in or near MNQR, $\overline{MNQR}$ such that, with sufficiently large gain factor, the dominant closed-loop poles are guaranteed to be in the acceptable MNQR range, despite the variations in S_p, F_p, k in (1). Suppose the far-off open-loop poles of the loop transmission L(s) are assigned say at X_1, $\bar{X}_1$, X_2 in Fig. 1. These locations must be such that over the range $k_{min} \le k \le k_{max}$, the far-off closed-loop poles remain to the left of the boundary UVW. From root-locus considerations, it is clear that the greatest danger of boundary crossing is at $k = k_{max}$. In fact, in the optimum design, there is at $k = k_{max}$ a closed-loop pole precisely on the boundary[6], say at J in Fig. 1. Write L(s) in the form $L(s) = kK'n(s)/d(s)$ where n(s), d(s) are polynomials whose leading coefficients are unity; K' is a constant. Let $Y, \bar{Y}$ in Fig. 1 mark nominal plant, and therefore loop transmission, poles. Since $1 + L(s) = 1 + kK'n(s)/d(s) = 0$ at $s = 0J$ when $k = k_{max}$, it follows that $k_{max}K' = -d(0J)/n(0J)$; i.e.,

$$k_{max} K' = - \frac{(0J)\ (YJ)\ (\bar{Y}J)}{(ZJ)\ (\bar{Z}J)} (X_1J)^2\ (\bar{X}_1J)^2\ (X_2J). \qquad (3a)$$

Let Y' represent a dominant closed-loop pole position (inside MNQR, of course) when $k = k_{min}$. Then,

$$k_{min} K' = - \frac{(0Y')\ (YY')\ (\bar{Y}Y')}{(ZY')\ (\bar{Z}Y')} (X_1Y')^2\ (\bar{X}_1Y')^2\ (X_2Y'). \qquad (3b)$$

Hence,

$$\frac{k_{max}}{k_{min}} = \left[\frac{(X_1J)^2(\bar{X}_1J)^2(X_2J)}{(X_1Y')^2(\bar{X}_1Y')^2(X_2Y')}\right] \frac{(0J)\,(YJ)\,(\bar{Y}J)/(ZJ)\,(\bar{Z}J)}{(0Y')\,(YY')\,(\bar{Y}Y')/(ZY')\,(\bar{Z}Y')} \triangleq [\lambda]\frac{(0J)}{K_1} \qquad (4)$$

with

$$K_1 \triangleq \frac{(0Y')\,(YY')\,(\bar{Y}Y')}{(ZY')\,(\bar{Z}Y')}\ \frac{(ZJ)\,(\bar{Z}J)}{(\bar{Y}J)\,(YJ)} \doteq \frac{(0Y')\,(YY')\,(\bar{Y}Y')}{(ZY')\,(\bar{Z}Y')} \qquad (5)$$

because $(ZJ)(\bar{Z}J)/(YJ)(\bar{Y}J) \doteq 1$, due to J being far-off relative to the dominant $\bar{Z}$, $\bar{Z}$, Y, $\bar{Y}$.

Suppose $k_{max}/k_{min} = 1000$. In Eq. (4), λ_{max} (which is not large) is determined by the geometrical pattern of X_1, X_2, ..., U,V,W and can be readily found in any specific problem. The ratio $0J/K_1$ in Eq. (4) must bear the brunt of satisfying the large change in the plant gain factor k_{max}/k_{min}. This leads to large 0J; for example, it will later be found that even in the optimized design $|0J| \doteq 5600$, which is at least 500 times as great as the largest magnitude dominant pole. It is next shown that large 0J means large bandwidth over which the compensation, denoted by G(s), is performing at least second-order differentiation with its noise amplification problems. One way to see this is as follows. At s = 0J, the loop transmission $G(s)P(s) \triangleq L(s) = -1$. But at s = 0J, $|P(s)| = |k/s(s^2+sS_p+P_p)| \doteq |k/s^3| = |k/(0J)^3| \triangleq \varepsilon$ is generally extremely small; therefore, the compensation magnitude $|G(s)| = 1/\varepsilon$ must be very large. Or, the above statement can be verified by the following argument. At s = 0J, L(s) = G(s)P(s) = -1. Hence, at s = 0J, $\angle G(s) + \angle P(s) = -180^\circ$, i.e., in Fig. 1 $[2\angle X_1J + 2\angle\bar{X}_1J + \angle X_2J] + [\angle 0J + \angle YJ + \angle\bar{Y}J - \angle ZJ - \angle\bar{Z}J] \triangleq \theta_f + \theta_d = 180^\circ$. But $\theta_d \triangleq [\angle 0J + \angle YJ + \angle\bar{Y}J - \angle\bar{Z}J - \angle\bar{Z}J \doteq \angle 0J = 90^\circ + \delta$, δ small. Hence, $\theta_f \triangleq 2\angle X_1J + 2\angle\bar{X}_1J + \angle X_2J \doteq 90^\circ - \delta < 90^\circ$. A pole at -a has an angle of 45° at s = ja. Since θ_f represents the totality of angles of the vectors from the poles of G(s) to 0J and $\theta_f < 90^\circ$, it follows that their effective average corner frequency must be larger than 0J. In the Bode plot, |G(jw)| therefore has a positive slope (whose asymptotic value is 40 db. per decade) commencing at ω = 0Z and continuing so beyond 0J. Thus G(s) performs second-order differentiation over a very large frequency range. If, in practice, the plant has its own additional far-off poles then the compensation network must have corresponding additional cancelling zeros and performs even higher order differentation. From Eq. (4), $(0J)(\lambda/K_1) = k_{max}/k_{min} = 1000$ and since λ_{max} is not large, it follows that minimization of K_1 defined by (5), is extremely important, in order to reduce the bandwidth of the compensation.

The previously available design techniques are deficient in the following important aspects: (1) The mapping of ABCD into a region which lies in MNQR is approximate. Viz., if A, B, ..., map into the points A;, B', ... (inside MNQR of course) it is assumed in the calculations that AA', BB', $\overline{A}A'$, $\overline{B}B'$, $\overline{Z}A'$, $\overline{Z}B'$ ... may be approximated by AR, BR, ... where R is a fixed centrally chosen point inside MNQR. Hence the design technique is satisfactory if and only if (a) the acceptable region MNQR is both relatively small in area, and (b) well removed from the plant pole range of variation ABCD and (c) from the real axis. These are all significant shortcomings for in many, if not most, control problems there is a <u>fairly large</u> acceptable closed loop dominant pole range which often overlaps the range of plant pole variation. An additional significant shortcoming is that (2) over a very large frequency range, the slope of $|L(j\omega)|$ is only -6 decibels per octave so that the required gain margin of at least 20 log $|k_{max}/k_{min}|$ requires many octaves (10 in the present example, as 6 x 10 = 20 log 1000), which could be significantly reduced by using a larger magnitude average slope, say -9 or -10 db per octave. (It is important to note that the largest noise contribution is in the last few octaves. A reduction in $L(j\omega)$ bandwidth by x octaves reduces the rms noise by a factor whose order of magnitude is 2^x.) However, this would require a more complicated L(s) far-off pole-zero pattern, which is difficult to include in a systematic s-plane design approach, but is much more readily achieved by using frequency-response techniques in this relatively "far-off" region. This paper presents procedures for eliminating the above shortcomings.

B. Design in the Dominant Range - Case Plant Gain Factor Alone Varies

If the plant pole and zero variations are small, although the gain variations are large, it is possible to achieve a considerably more economical (smaller bandwidth) L(s) than by the previous methods. Ignoring the far-off poles except for the nearest, p_f (in the notation of Eqs. 1,2) let the dominant part of L(s) be

$$L_d(s) = Kk(s^2 + S_o s + P_o)/s(s^2 + S_\ell s + P_\ell) \triangleq Kk\ n_d(s)/d_d(s) \tag{6a}$$

and the corresponding dominant part of T(s) be

$$T_d(s) = P_r p_f/(s^2 + S_r s + P_r)(s + p_f) \triangleq P_r p_f/D_d(s) \tag{6b}$$

From (5,6b)

$$d_d(s) + Kk\ n_d(s) = D_d(s) \tag{7}$$

Note from (1,6a) that not all the poles of the plant P are necessarily in L_d; i.e., it may be helpful to cancel some poles of P and replace them by others in L_d, which is no problem in this case where only the plant gain factor k varies. This is also the reason for replacing the p (indicating plant) subscripts of Eq. (1) by the more general ℓ (indicating loop transmission) subscripts in Eq. (6a). Equating the zero degree coefficients in (7) gives

$$kK = P_r p_f/P_o \tag{8}$$

Let P_r^*, p_f^*, D_d^*, ... denote P_r, p_f, D_d, ... at $k = k_{min}$ with $P_r^* = (0Y')(0\overline{Y}')$. In Eq. (7), $D_d(s) = 0$ at $s = 0Y'$ when $k = k_{min}$, so $k_{min}K = |d_d(0Y')/n_d(0Y')|$, which is exactly equal to the extreme righthand side of (5), defining K_1; i.e.,

$$k_{min}\ K = P_r^* p_f^*/P_o \doteq K_1 \tag{9}$$

The importance of K_1 minimization has been previously emphasized in connection with Eq. 4. Eq. (9) indicates that there are three variables available for this purpose. The choice of p_f^* for K_1 minimization is easy. The specs. require that $-p_f$ (which is on the negative real axis) lies to the left of the boundary UVW in Fig. 1, for all $k_{min} < k < k_{max}$. According to Eq. (7), $-p_f$ lies on the root loci of $1 + Kk\ n_d(s)/d_d(s) = 0$. From a simple, rough sketch of these root loci, it is easy to see that as k increases, the real axis root moves to the left;

i.e., $P_f > p_f^*$ because p_f^* has been defined as the value of p_f at $k = k_{min}$. The minimum value of p_f^* is therefore precisely that permitted by the specs., i.e., by the intersection of UVW with the boundary (at -16.2 according to Fig. 1). The choice of P_r^* in (9) for K_1 minimization is next considered. Obviously, the best choice of P_r^* is the minimum value of P_r permitted by the MNQR specification. Since $\sqrt{P_r}$ is the magnitude of the line from the origin to any point inside or on MNQR, the minimum value of P_r^* corresponds to a point at which a circle centered at the origin, just grazes MNQR, as shown in Fig. 2. (P_r^* is used in Fig. 2 to denote a pole position but this should not cause any confusion). It is necessary to guarantee that this choice for P_r^* leads to S_r, P_r, p_f of (6b) which satisfy the specs. for all $k_{min} \leq k \leq k_{max}$. For this purpose let the system dominant characteristic equation $D_d(s)$ at $k = k_{min}$ be denoted by

$$D_d^*(s) \triangleq d_d(s) + Kk_{min}\, n_d(s) \tag{10a}$$

and subtract (10a) from (7), giving

$$D_d^*(s) + K(k - k_{min})\, n_d(s) = D_d(s) \tag{10b}$$

Hence, the zeros of $D_d(s)$, i.e. the system poles, which the specs. require to lie in MNQR in Fig. 2, are on the root loci of

$$1 + k(k - k_{min}) \frac{n_d(s)}{D_d^*(s)} = 0 \tag{10c}$$

In the root-locus pattern determined by Eq. (10c), the open-loop poles are at the points denoted by $-p_f^*$, P_r^*, $\overline{P}_r^*$ in Fig. 2. The problem is to choose Z, $\overline{Z}$, the zeros of $n_d(s)$, to guarantee that the root loci of (10c) stay in the MNQR region for $k_{max} > k > k_{min}$. To achieve this, it is certainly necessary that the direction of departure of the root locus from P_r^* be into MNQR by a comfortable margin. Consider the vector determined by the lines $H_1P_r^*$, $H_2P_r^*$ in Fig. 2. It is reasonable to require that the direction of the root locus departing from P_r^* be inside this or a somewhat similar sector, in order to ensure that the root loci remain within MNQR for $k_{max}>k>k_{min}$. Obviously, there is some cut and try involved here. In any case, with this choice the requirement is that the root locus angle of departure denoted by θ_d, is constrained by the relation $123^\circ < \theta_d < 215^\circ$ (see Fig. 2). Using the 'angle of departure' root-locus theorem (Ref. 6, p. 125), and letting $\theta_Z = \angle ZP_r^* + \angle \bar{Z}P_r^*$, there is obtained the equation (see Fig. 2) $90^\circ + 23^\circ + \theta_d^\circ - \theta_Z^\circ = 180^\circ$; giving $56^\circ < \theta_Z < 148^\circ$. The locus of Z, $\bar{Z}$ such that θ_Z is constant, is a circle through P_r^*, $\bar{P}_r^*$ and through a third point X defined by $\angle XP_r^* = 0.5\theta_Z$. The two extreme values of 56°, 148° for θ_Z thus determine corresponding two extreme circles C_1, C_2 in Fig. 2. Thus Z may be located anywhere between C_1, C_2. From (9), it is desirable to choose Z as far from the origin as possible in order to maximize P_o and so minimize K_1. But when k_{max}/k_{min} is large, the closed-loop pole at $k = k_{max}$) is very close to Z. Hence an excellent choice for Z appears to be at the corner R. However, one must check that the root loci of (10c) stay in MNQR. This is done by finding the angle of entry (θ_e) of the root locus into Z (Ref. 6, p. 125). The result is $\theta_e = 94^\circ$ if Z is at R. This is clearly unacceptable. Location of Z at $-7 + j\,0.5$ gives $\theta_e = 94^\circ$ which is satisfactory. Thus a value obviously very close to $(K_1)_{min}$ is found by means of relatively little cut and try. The procedure for further (marginal) minimization is obvious. The above choice gives $n_d(s) = s^2 + 14s + 49.25$ and from (9), $K_1 = 2.52$. To find S_ℓ, P_ℓ of Eq. (6a), Eq. (7) is solved for $d_d(s)$. The result is $d_d(s) = s(s^2 + 11.5s + 23.4)$, which will involve dominant plant pole cancellation and replacement. In order to decide whether the economy in K_1, (and thereby in L(s) bandwidth and in noise reduction), so obtained is justified, the optimum (minimum K_1) design for specified (uncancelled) $d_d(s)$ is required. This case is included in the problem next considered.

C. Design in the Dominant Range - Case Plant Poles and Gain Factor Vary

As an aid in presenting the design technique, consider the case when the plant complex pole pair may range over the region ABCD in Fig. 1. (Note the deliberate overlapping with the acceptable system pole region which could not be handled by the previous method[6,7].) Equating coefficients in Eq. (7) and in the notation of Eqs. (6a,b) (except that S_ℓ, P_ℓ are replaced by S_p, P_p, since cancellation of plant poles is not contemplated because of their large range of variation), gives

$$X \triangleq S_p + kK = S_r + (kKP_o/P_r) \triangleq x + \gamma y \text{ (with } \gamma \triangleq kKP_o,$$

$$x \triangleq S_r,\ y \triangleq 1/P_r.) \qquad \text{(11 a,b,c)}$$

$$Y \triangleq P_p + kKs_o = \gamma xy + \frac{1}{y} \qquad (12)$$

and

$$kKP_o = p_f P_r \qquad (8)$$

The relating of open-loop pole to closed-loop pole variations is easily achieved in the X,Y plane because $\Delta X = \Delta S_p$, $\Delta Y = \Delta P_p$ at fixed k. (It will be seen that the variations in k are usually more easily handled at a later stage by root-locus methods). Therefore, the next step is to use Eqs. (11), (12) to map the acceptable MNQR region of Figs. 1,2 into the X,Y plane. Since γ is not a priori known, the mapping may have to be done for several values of γ. P_o does not in practice have much of a permissible range of variation (~MNQR), so large γ means large kK; i.e., large K_1. Hence one starts with small γ and tries larger γ if the former proves unsatisfactory. This will be clarified in the later design details. [It is obvious from 11, 12, or from the older method,[6,7] that any minimum-phase problem can be solved by means of sufficiently large kK.]

A simple way of performing the mapping is by means of loci of constant S_r, P_r in the X,Y plane. Thus Eqs. (11a),

(12) may be manipulated into

$$Y = \gamma y X + (y^{-1} - \gamma^2 y^2); \; Y = x(X - x) + \gamma(X - x)^{-1} \qquad (13a,b)$$

which are readily plotted by computer and shown in Fig. 3 for $\gamma = 600$. The acceptable region M'N'Q'R' is also shown in Fig. 3.

The final step is to map the plant variation region ABCD of Fig. 1 into the X,Y plane in Fig. 3 and to see whether it can be accommodated within the M'N'Q'R' region of Fig. 3. Thus from (11a, 12) the plant pole variations $\Delta S_p = \Delta X$, $\Delta P_p = \Delta Y$. The procedure is to first map the ABCD region of Fig. 1 into an equivalent region A'B'C'D' in an S_p, P_p plane whose units are the same as those of the X,Y plane. A'B'C'D' may then be cut out with scissors and one attempts to fit it into the acceptable M'N'Q'R' region in Fig. 3. It is seen that it cannot be precisely fitted inside M'N'Q'R, at $\gamma = 600$. The mapping must be repeated at higher γ (i.e., new loci of constant S_r, P_r with a resulting larger area M'N'Q'R; A'B'C'D' is unaffected.) In this specific example, if A'B'C'D' is located as shown in Fig. 3, the resultant s-plane closed-loop region is that enclosed by the A'B'C'D boundary in Fig. 4. Let it be assumed that the indicated excursion outside MNQR is acceptable, so that $\gamma = 600$ may tentatively be used. (It is fortuitous that part of the excursion involves an overdamped range whose extremes are given by the B', for it will be shown that this permits a "far-off" L(s) pole to be inserted sooner than would ordinarily be possible.)

The tentative qualification is used in the above because one must check whether acceptable p_f results and whether the variations in k lead to satisfactory closed-loop pole variation. To check these matters, the parameters are evaluated at $\gamma = 600$. In Fig. 3 any suitable point is chosen, say $X = 32.2$, $Y = 220$; $S_r = 8$, $P_r = 25$; at which (by reading from the scales on the portable A'B'C'D' graph) $S_p = -3$, $P_p = 80$. Since (Eq. 11a), $X = S_p + kK = 32.2$, $kK = 35.2$ and $P_o = \gamma/kK = 600/35.2 = 17$. Also, Eq. (12), $80 + kKS_o = Y = 220$, so $kKS_o = 140$ and $S_o = 140/35.2 = 4$, and Eq. (8), $p_f = kKP_o/P_r = \gamma/P_r$, so $(p_f)_{min} = 600/36$, which is satisfactory (since it is on the left of UVW boundary in Fig. 1). (Note that K_1 is 35.2 here as compared to 2.52 in Section B where there was no open-loop pole variation. This is a difference of 29db which it will be seen results in a difference of about 3 octaves in L(s). This is discussed in detail in Section E.) To check the effect of the variations in k, the equivalent of Eq. (10c) is used

$$D_1(s) + K(k - k_{min})\ n_d(s) = 0 \qquad (14)$$

The roots of (14) for $k_{min} < k < k_{max}$ give the two dominant and one far-off closed-loop poles as a function of k. $D_1(s)$, replacing $D_d^*(s)$ in (10c) represents D(s) at $k = k_{min}$. Hence the zeros of $D_1(s)$ may lie anywhere in A'B'C'D' in Fig. 4, previously obtained at the fixed $k = k_{min}$. One must therefore consider all possible root-loci for the infinitude of zeros of $D_1(s)$. It suffices to check the boundary of A'B'C'D', by calculating the angles of departure. It is found that over the entire boundary of A'B'C'D' the angles of departure are all such as to lead to loci directed into the interior of A'B'C'D'. If, in practice, it should not be so, then one can be certain that sufficiently large γ will give a satisfactory design. The objective is to get by with as small a γ as possible.

D. Far-off L(s) Pole and Zero Locations

It has been noted that in the old design method, $|L(j\omega)|$ decreases at the rate of -6 db/octave over a large frequency range, so that the required minimum gain margin of $[20 \log (k_{max}/k_{min})]$ requires many octaves. In order to significantly decrease the $L(j\omega)$ bandwidth, a larger slope is required, involving a staggering of poles and zeros in the higher frequency range. This is difficult to do in the s-plane, and much easier to design on a Bode plot. However, it is then necessary to translate the UVW boundary constraint of Fig. 1 into an equivalent constraint in the frequency domain. There is a complex pole pair which threatens to cross the UVW, $\overline{UVW}$

boundary. Let this pole pair be the zeros of $s^2 + 2\zeta_f \omega_f s + \omega_f^2$. The parameters ζ_f, ω_f are closely related (Ref. 6, p. 197) to the Bode plot type frequency parameters G_m (gain margin in nepers), θ_m (phase margin in radians), ω_c (defined by $|L(j\omega_c)| = 1$), ω_π defined by Arg $L(j\omega_\pi) = -180^o$). The approximate relations are:

$$\frac{\omega_f^2}{\omega_c^2} \approx \frac{G_m^2 + \theta_m^2(\omega_\pi^2/\omega_c^2)}{G_m^2 + \theta_m^2}$$

(15a,b)

$$\zeta_f \approx \frac{\theta_m G_m}{\theta_m^2 + G_m^2}\left(\frac{\omega_\pi - \omega_c}{\omega_f}\right)$$

They have been found to be fairly accurate--see for example, Ref. 6, p. 279 and later in the present section. As k increases from k_{min} ω_c, ω_π increase and θ_m, G_m change in value. Hence, the UVW boundary constraint on the far-off system poles, may be restated as constraints on the above Bode-type frequency parameters, and used as such in shaping $L(j\omega)$ on the Bode plot in the far-off range. This may be done if one is keenly interested in extreme optimization; i.e., an average $|L(j\omega)|$ slope close to $-10 \sim -11$ db per octave. The designer may dispense with the above if he is content with an average slope of -9 db/octave because the UVW constraints will then obviously be satisfied (the phase margin is then $\sim 45^o$ over most of the range in between ω_{c_1} and ω_{c_2}, the latter denoting the crossover frequencies at k_{min} and k_{max} respectively.) However, these relations may very usefully be used at ω_{c2} itself in order that the last far-off lag corner frequencies which must be inserted at $\omega > \omega_{c2}$, may be introduced at as low a frequency as possible (see Fig. 5).

The detailed procedure in shaping $L(j\omega)$ for the "far-off" region is straightforward. From the viewpoint of the far-off region, the poles and zeros in the dominant region are equivalent to a single pole at $s = 0$. Hence one begins the

Bode plot with a Bode sketch of $L = Kk_{max}/s$. The first decision which must be made is at how low a frequency one may insert a lag-corner-frequency (denoted by lacf); i.e., at which point may the first far-off pole be placed, without significantly affecting the positions of the dominant system poles assumed to lie in A'B'C'D' in Fig. 4. (Recall that in obtaining A'B'C'D' in Figs. 3, 4, the far-off poles and zeros were neglected.) Suppose a lacf at $\omega = 30$ (i.e., a pole at -30) is used. The maximum phase effect of a pole at -30, on points in A'B'C'D', is 8^{o}. The magnitude effect is given by $|PX'/PO|$ where X' is any point in A'B'C'D' and P is at -30. There will consequently be negligible shifts in the assumed root positions in the A'C'D' region but nonignorable effects on the points near the B' region; for example, a pole at -30 is not "far-away" with respect to the root at -8.9. Is the effect desirable or undesirable? A little thought indicates it to be an extremely desirable effect; in fact the B' points are thereby forced into the more desirable MNQR region. A bit of work with the spirule predicts that the extreme B' points roots move to $-4.3 \pm j\ 1.75$, (B" in Fig. 4) which is well inside MNQR in Fig. 4. Thus there will now be only the small region near Z (Fig. 4) outside MNQR. With regard to the UVW boundary, (Fig. 1) at $k = k_{min}$, $\omega_c \triangleq \omega_{c1} \approx 28$ (see Fig. 5) and since the approximate relation of Eq. 15a always gives $\omega_f/\omega_c \geq 1$, the UVW boundary specifications of Fig. 1 are easily satisfied at $k \geq k_{min}$ so long as the average slope of $|L(j\omega|$ is ~ -9 db/octave. Such an average slope is obtained by staggering poles and zeros as shown in Fig. 5 (poles at $s = -30$, -600, zeros at -125, -3000 but of course this is not unique).

The final step is the assignment of the last far-off poles. One decides upon the desired excess of poles over zeros of L(s). In this example an excess of 5 was chosen. Eqs. (15a,b) may be used, if desired, to economize to the utmost on the bandwidth. There was no attempt to do so in this case. Rather, with a little cut and try $\theta_m = 22^o$, $G_m = 3$ db (at $k = k_{max}$) was considered satisfactory. The approximate relations of Eqs. (15a,b) then give $\omega_f \approx 5600$, $\xi_f = 0.17$ compared to the computer values of $\omega_f = 5660$, $\xi_f = 0.187$.)

The resulting L(s) =

$$\left[\frac{(35.2)k(s^2+4s+17)}{s(s^2+S_p s+ P_p)}\right] \frac{(\frac{s}{150}+1)(\frac{s}{3000}+1)}{(\frac{s}{30}+1)(\frac{s}{600}+1)\left[(\frac{s}{8000})^2+\frac{(2)(0.3)}{8000}s+1\right]^2} \quad (16)$$

Computer runs give very good verification of the design. The points A', B'', C', D' are found to be at -2.7 + j 4.06, -4.44 + j 1.64, -3.0 + j 2.32, -1.83 + j 3.34 respectively, in good agreement with the design values in Fig. 4; as k increases to k_{max}, they converge towards Z inside A'B''C'D'. The far-off roots easily satisfy the UVW boundary constraints of Fig. 1.

Design Structure and Compensation Blocks

To complete the design, a specific structure must be chosen. Any two degree of freedom structure[6] may be used; for example, that shown in Fig. 4. The design has guaranteed dominant system poles in an acceptable region, but the specs. may possibly require other fixed dominant poles and zeros in the system transfer function $T(s) \triangleq C/R$. Let these denoted by $\tau(s)$. The dominant poles and zeros of F and H are thereby completely fixed, as will be seen. The designer can arbitrarily assign far-off poles and zeros to T(s) because these have negligible effect on the system response, and a judicious assignment may lower the complexity of F(s) and H(s). Let the subscripts d,f denote dominant and far-off poles (or zeros) respectively. Let D(s), sd(s) be polynomials representing the zeros of 1+L, poles of L respectively. Let $D_f'(s)$, $D_f''(s)$, $d_f'(s)$, $d_f''(s)$ represent portions of the corresponding polynomials, with $D(s) = D_d(s)\, D_f(s) = D_d(s)\, D_f'(s)\, D_f''(s)$, $d_f(s) = d_f'(s)\, d_f''(s)$. As usual, the leading coefficients in the d(s), D(s) polynomials is unity. Then, in the structure of Fig. 4

$$T(s) = \frac{FP}{1+L} = \frac{Fk/(s^2 + sS_p + P_p)s}{D_d(s)D_f'(s)\ D_f''(s)/(s^2 + sS_p + P_p)\ s\ d_f(s)} =$$

$$\frac{Fk\ d_f(s)}{D_d(s)D_f'(s)\ D_f''(s)} \qquad (17)$$

If T(s) is set up as $T(s) = k\tau(s)\ d_f'(s)/D_d(s)\ D_f'(s)$, then equating this with (17) gives

$$F(s) = \frac{\tau(s)\ D_f''(s)}{d_f''(s)} \qquad (18)$$

Also, $L(s) = FPH = kK(s^2 + sS_o + P_o)/s\ (s^2 + sS_p + P_p)\ d_f(s)$. Combining the latter with (18) gives

$$H(s) = \frac{K\ (s^2 + sS_o + P_o)}{d_f'(s)\ \tau(s)\ D_f''(s)} \qquad (19)$$

From (18,19) it is seen that both F and H are simplified by letting $D_f''(s) = 1$, if sufficient far-off poles have been assigned to L(s) to ensure proper high-frequency behavior of F(s) and H(s). These far-off poles may be appropriately divided between F(s) and H(s) for this purpose.

E. Feasibility of Pole Cancellation When Plant Poles Vary

Section B considered the sensitivity problem for plant gain variations with no plant pole variations, while Section C considered the same spec. but with both pole and gain variations. The difference in K_1 was found to be 29 db, which means that the final rapid decrease of $|L(j\omega)|$ must in the second case be about three octaves further off (with noise effects $\sim 2^3$ worse than before). Can this be avoided? Consider, for the moment, the case where the complex plant pole pair varies a "little." It may then be feasible to cancel the poles and locate a fixed pole pair at the optimum point found in Sec. B, and thereby use the smaller K_1. However, due to the small plant pole variation, there will be a dipole in that neighborhood. Hence one must add another specification, say, the maximum tolerable residue in the pole (of the dipole) of the system step response. The latter can be related to the maximum dipole separation, as follows

Let the loop transmission be written in the form

$$L(s) = L_1(s)\,\frac{(s - s_z)}{(s - s_p)} \tag{20}$$

Suppose the closed-loop pole associated with the dipole is at s_d in Fig. 6; i.e., $1 + L(s_d) = L_1(s_d)\,\vec{A}/\vec{B} = 0$. The last complex equation is equivalent to the two real equations:
$B/A = |L_1(s_d)|$, $\theta_z + \theta_p = -\theta_{L_1}(s_d)$ with θ_z, θ_p defined in Fig. 6. Let u, v be a set of axes as shown in Fig. 6. The last two equations then become

$$u^2 + \left[v - \frac{a(1 + m^2)}{(1 - m^2)}\right]^2 = \left[\frac{2am}{1 - m^2}\right]^2, \tag{21a,b}$$

$$\left(u - \frac{a}{N}\right)^2 + v^2 = a^2\left(1 + \frac{1}{N^2}\right)$$

with $$m \triangleq |L_1(s_d)|, \quad N \triangleq -\tan\theta_{L_1}(s_d) \tag{22a,b}$$

The (u,v) values which satisfy (21a,b) are the coordinates of s_d in the u,v plane. Equations (21a,b) generate two orthogonal families of circles, which are plotted in Fig. 7 with m, N as parameters. To use them, a value of s_d is assumed near the dipole, giving m and N. A reason-able first try is to assume $L_1(s_d) = L_1(s_x)$ with s_x at the origin of the u,v axis. This determines a point in Fig. 7 (e.g., if at estimated s_d, $L_1 = 0.5\angle{-160^\circ}$ then point M in Fig. 7 results). The point M is used as the new trial value of s_d, etc. Assuming the point M is thus found, the value of A (of Fig. 6) is that of $|s_z M|$ in Fig. 7 (Note $|\vec{s_x s_p}| = 2a$ in Fig. 7). This enables one to find the value of the maximum residue in the pole at s_d (in Fig. 6), of the system step response. This residue

$$\mathcal{R} \doteq T(s_d)\ \vec{A}/s_d \tag{23}$$

(if $|\bar{s}_d s_d| \doteq |\bar{s}_z s_d|$). Since s_d, for the postulated problem, is in the dominant region, the range of $|T(s_d)$ is well known, so the range of R may be determined. If it is satisfactorily small, then plant pole cancellation is feasible and the more economical design of Section B may be used.

If $\mathcal{R}$ is too large, then a design intermediate between the two extremes of Sections B, C may be used, as follows. Let $K_{1,min}$ $K_{1,max}$ be the two values of K_1 obtained by the methods of Sections B,C respectively. If $\mathcal{R}_1$ (associated with $K_{1,min}$) is too large, it can be reduced by increasing K_1 because, it is clear from Fig. 7, the larger the value of m (which is directly proportional to K_1), the smaller the value of A in Fig. 6 and Eq. (23.) However, there is no point, of course, in going so far as to take $K_1 > K_{1,max}$ for with $K_{1,max}$, by the method of Sec. C, there is no pole cancellation and the attendant dipole and need for consideration of the residue. Thus, when there is plant pole variation as well as gain variation, the two methods of Sections B, C may be considered as the two extremes and the required K_1 will be somewhere between $K_{1,max}$ and $K_{1,min}$. When the plant pole variation is extremely large, as in the example of Sec. C, then there is no doubt that $K_{1,max}$ of Sec. C must be used.

F. Generality of the Design Philosophy

The methods given here are, of course, restricted to dominant-type systems, thereby permitting dominant s-plane design. The resulting loop transmission bandwidth is larger than that required in non-dominant designs for which presently there exist only frequency response methods.[6,9] The former is, however, better in its correlation with transient response. The treatment in Section D for the far-off poles is applicable to <u>all</u> problems of the dominant type. The detailed design technique in Section C is, however, restricted to plants with varying dominant plant poles and gain factor. It is not directly applicable to plants with varying dominant zeros. Nevertheless, the design philosophy of Section C is also applicable to this case. It is only necessary to formulate the new equations, obtain the analogs of Eqs. (11,12) and proceed in the same manner.

G. Conclusions

This paper has presented techniques for designing dominant-type systems subject to large parameter variations and with specified acceptable range of dominant system pole positions. These techniques result in reduced loop transmission bandwidth and internal noise sensitivity which is smaller by several orders of magnitude than that obtainable from the previous dominant-type design methods.

References

1. W. F. Mazer, "Specification of the Linear Feedback System Sensitivity Function," IRE Trans. Vol. AC-5, pp 85-93, June, 1960.

2. P. E. Fleisher, "Optimum Design of Passive-Adaptive Linear Feedback Systems with Varying Plants," IRE Trans., Vol. AC-7, pp 117-128, March, 1962.

3. R. A. Volz, "The Determination of Optimum Realizable Compensating Elements for Feedback System Parameter Variation," Proc. NEC, pp 578-584, 1964.

4. B. T. Rung, G. J. Thaler, "Feedback Control Systems: Design with Regard to Sensitivity," Proc. NEC, pp 531-536, 1964.

5. K. Chen, "Analysis and Design of Feedback Systems with Gain and Time Constant Variations," IRE Trans., Vol. AC-6, pp 73-79, February, 1961.

6. I. Horowitz, Synthesis of Feedback Systems, Academic Press, 1963.

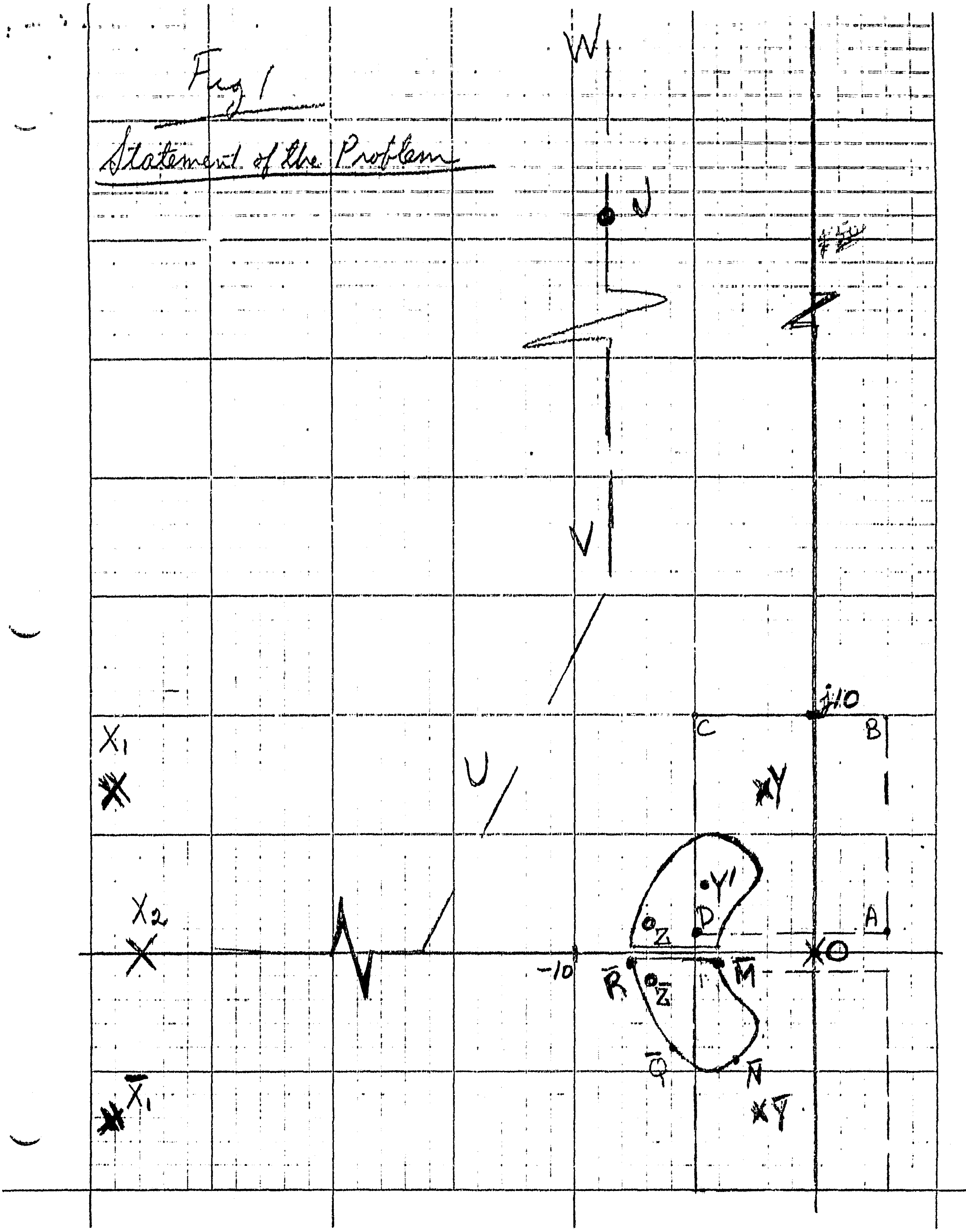
7. R. H. La Bounty, C. H. Houpis, "Root Locus Analysis of a High Gain Linear System with Variable Coefficients; Application of Horowitz's Method," IEEE Trans., Vol. AC-11, pp 255-263, 1966.

8. C. R. Chalk, "Additional Flight Evaluations of Various Longitudinal Qualities in a Variable-Stability Jet Fighter," WADC TR 57-719, Part 2, July, 1958. Figs. 6,9.

9. I. Horowitz, "Linear Adaptive Flight Control Design for Reentry Vehicles," IEEE Trans. on Automatic Control, IEEE Trans., Vol. AC-9, pp 90-97.

Fig 1

Statement of the Problem



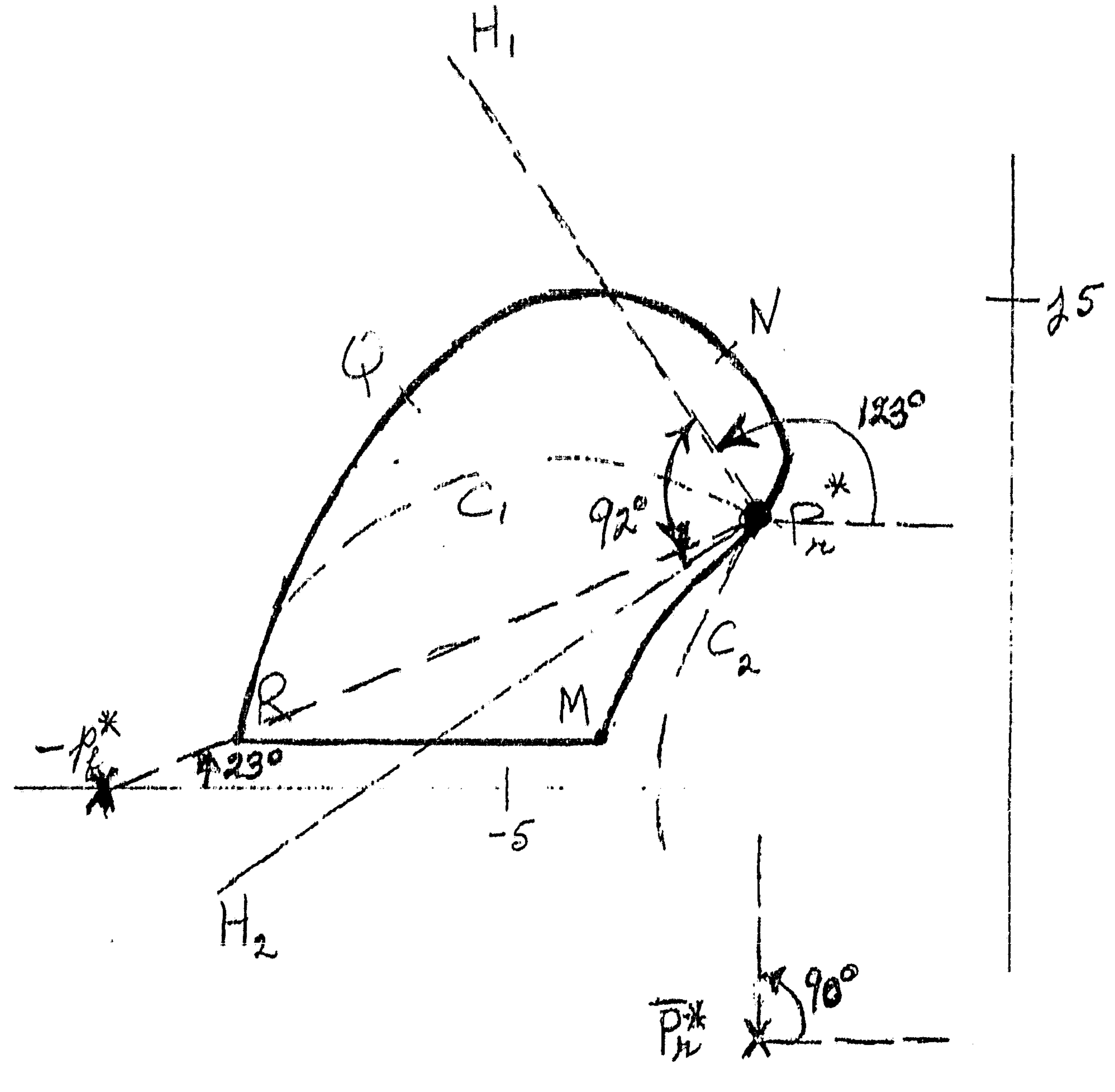
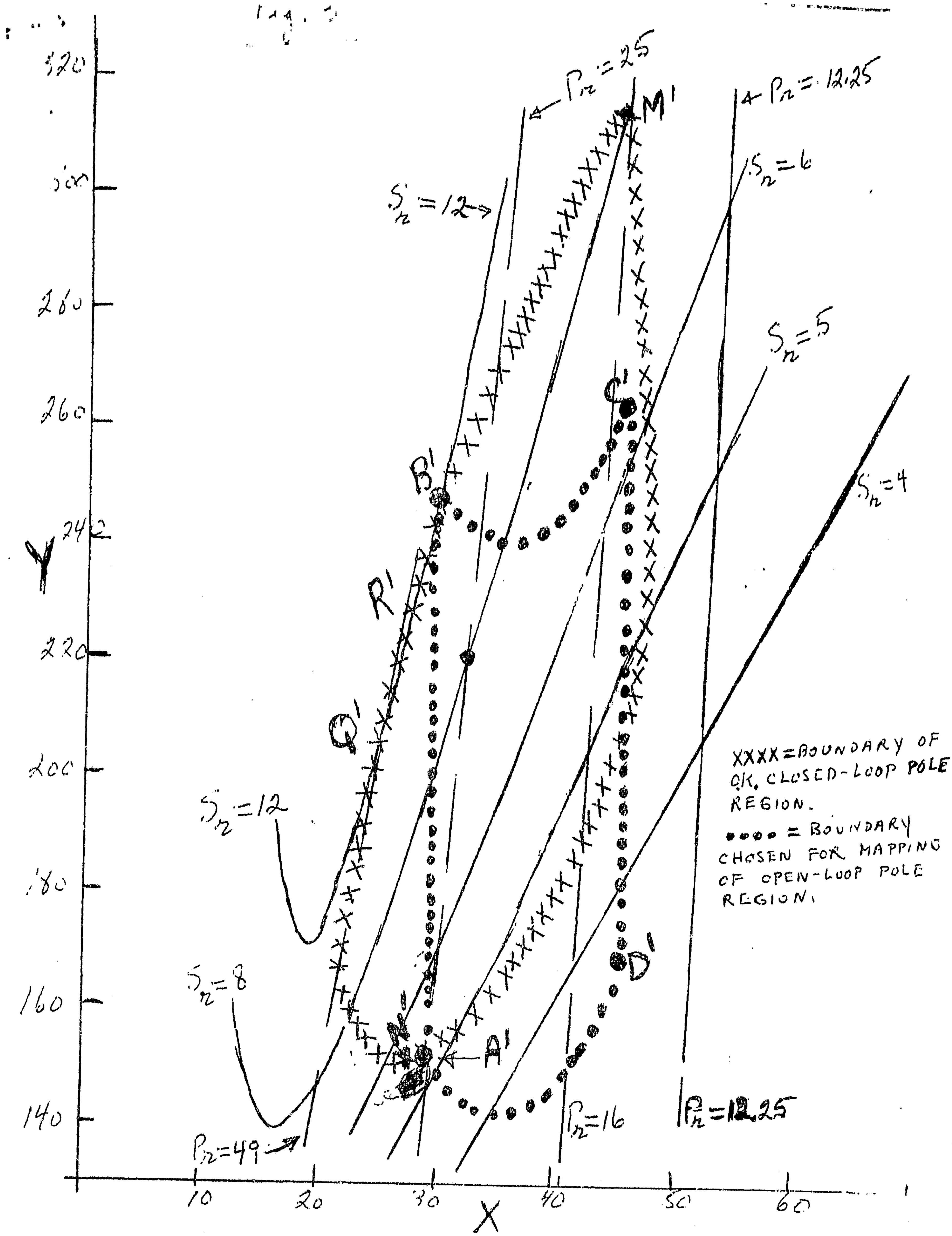
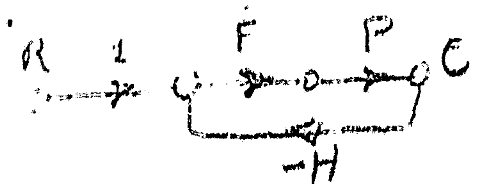


Fig 2 Design Details - Case : gain factor variations only.



Y
X
320
300
280
260
240
220
200
180
160
140
10
20
30
40
50
60
$P_n = 25$
$P_n = 12.25$
$S_n = 12$
$S_n = 6$
$S_n = 5$
$S_n = 4$
$S_n = 8$
$P_n = 49$
$P_n = 16$
$P_n = 12.25$
M'
C'
B'
A
R'
Q'
D'
A'
XXXX = BOUNDARY OF O.K. CLOSED-LOOP POLE REGION.
•••• = BOUNDARY CHOSEN FOR MAPPING OF OPEN-LOOP POLE REGION.



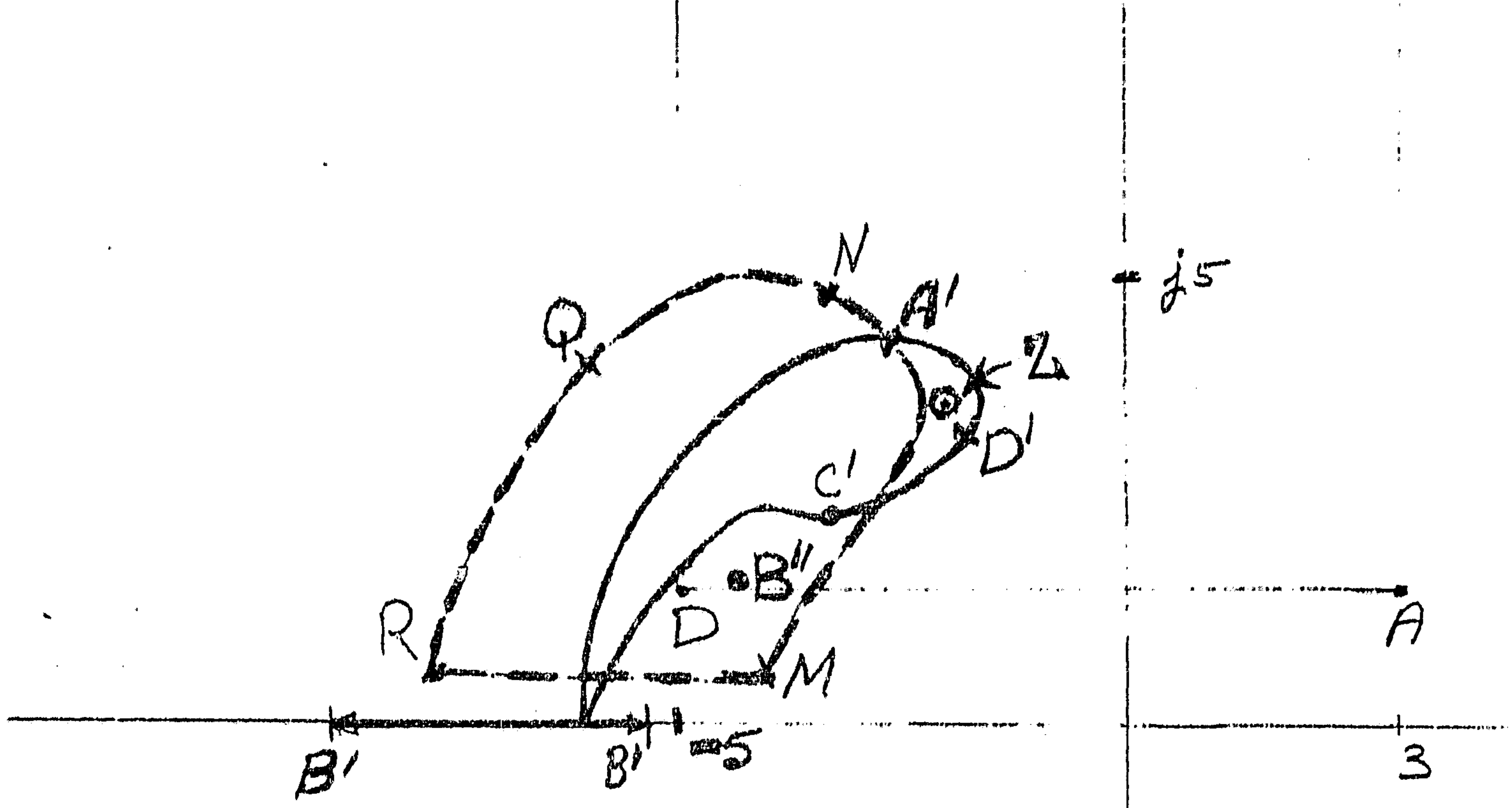
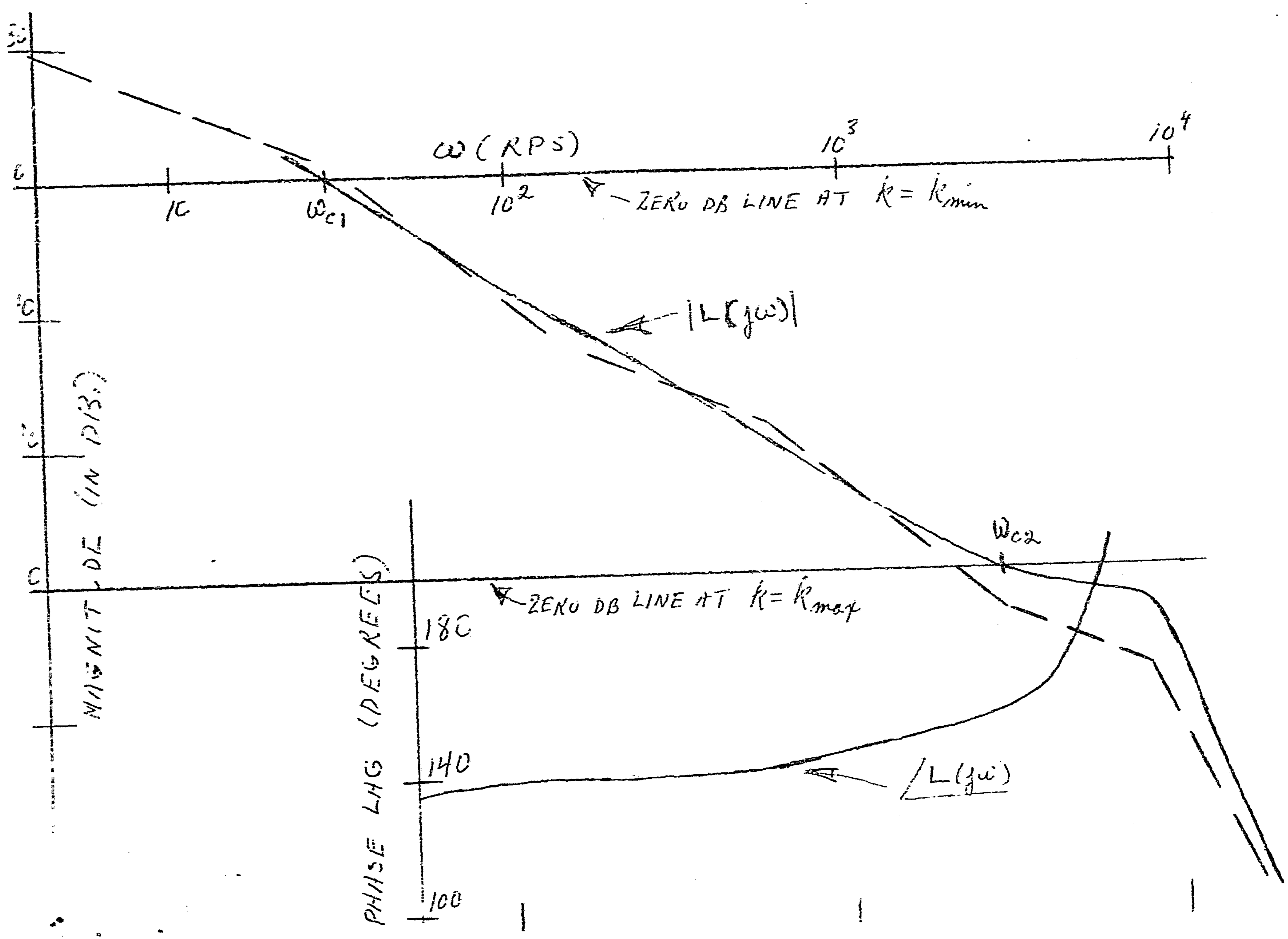


Fig 4 ~~...~~ Closed-loop dominant poles map into A'B'C'D'.

Fig 5 Design in far-off region



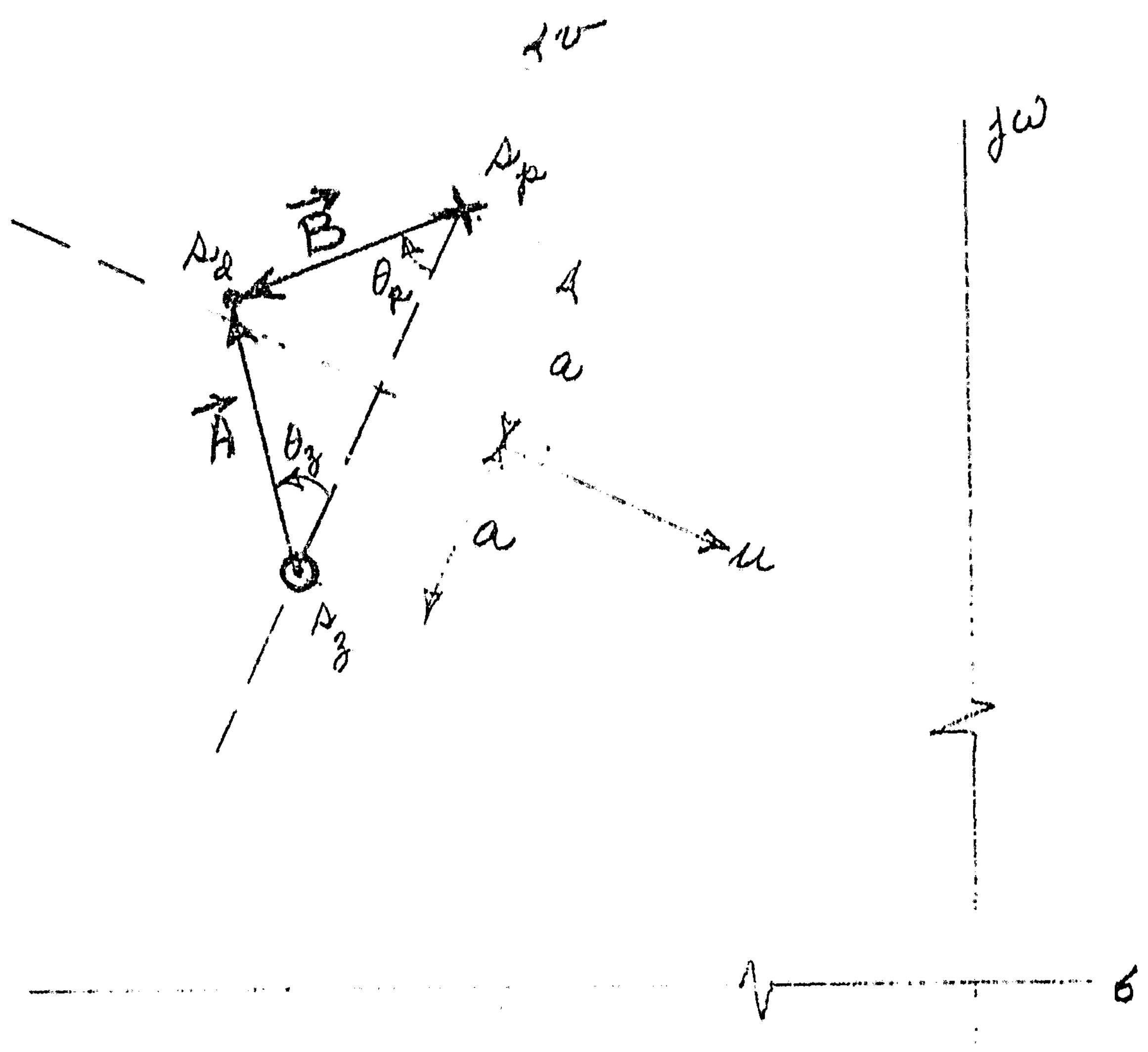


Fig 6 Feasibility of pole cancellation.

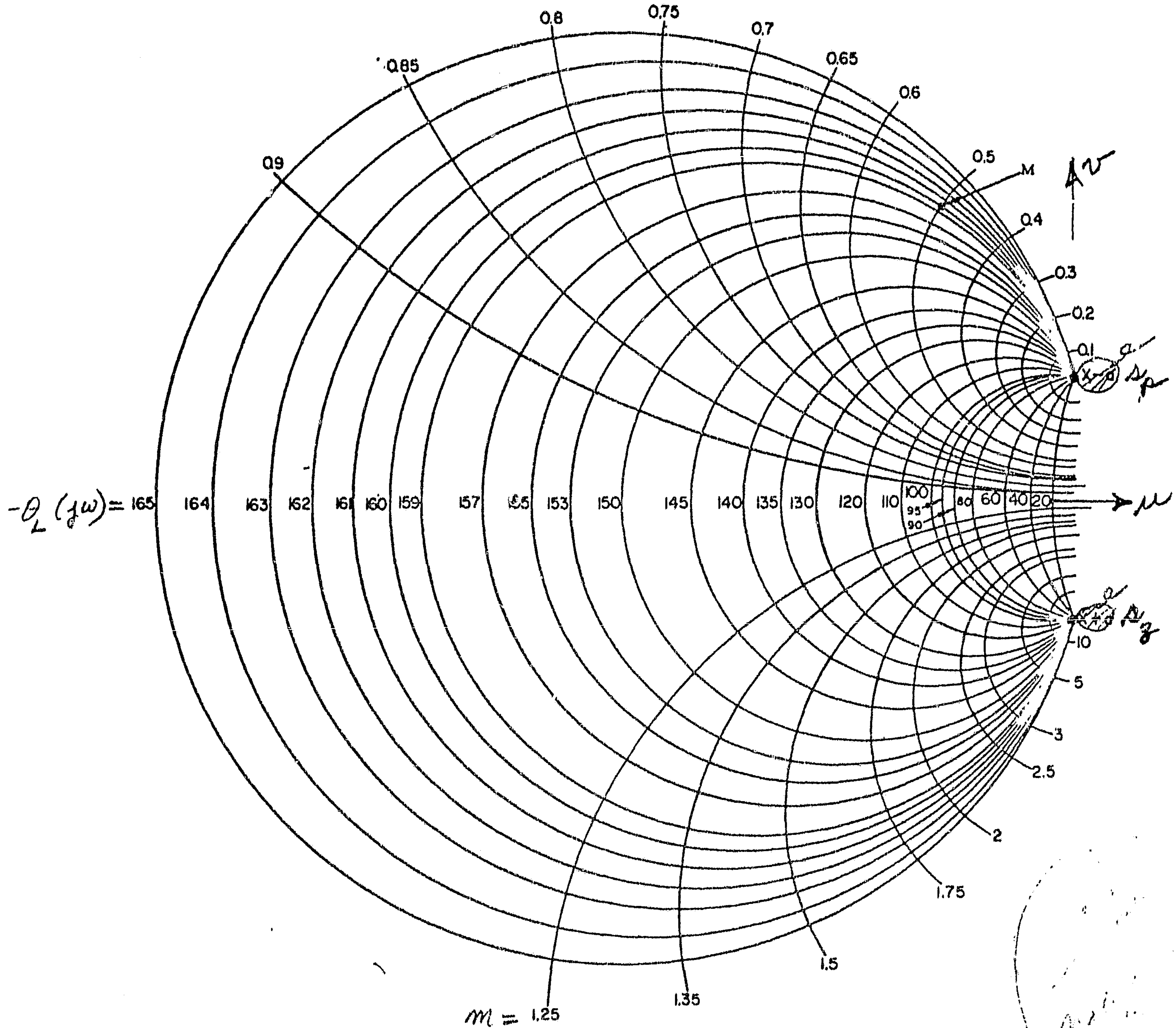


Figure ~~18.~~ 7 ~~Elastic mode loci which satisfy Eqs. 5a and 5b.~~

Loci for calculating dipole strength.