Phase-Root Locus and Relative Stability

Thomas J. Cavicchi

In this article, a new graphical tool called the phase-root locus is introduced. It is the dual of the conventional root locus, and indicates the motion of closed-loop poles in the s-plane as phase is added to the open-loop transfer function. The root locus/phase-root locus plots will be shown to facilitate destabilization diagnosis, which may help determine what part of an unstable physical system requires modification. For example, destabilization may be caused by one closed-loop pole due to a phase-shift at one value of system gain, and by a different pole due to gain variations at another value of system gain. The phase-root locus also allows relative stability information, including phase margin, to be accessed from the s-plane. Thus the s-plane can now be used for robustness analysis/design as well as transient analysis/design. The phase-root locus shows promise as a tool in compensator design as well as in the teaching of classical control theory.

Introduction

In classical single-input, single-output linear time-invariant control analysis and design, two primary measures of relative stability are in common use: gain margin (GM) and phase margin (PM). It has often been noted that neither margin alone is sufficient to characterize relative stability [1]; one margin can be large (indicating a robust system), but the other extremely small (and therefore in fact the system is not robust). A new plot, the dual of the conventional root locus (RL) named the phase-root locus (PRL), is presented in this article to help visualize relative stability from a root-locus viewpoint. The conventional RL may be referred to as "gain-root locus" because it shows the locus of closed-loop (CL) poles as the gain is varied; similarly, the phase-root locus shows the locus of CL poles as the phase is varied.

While PM often seems to students a somewhat abstract concept, GM is quite clear. Gain margin discloses the factor by which the DC forward amplifier gain K can be increased beyond the design value K_m ("m" for "mth design") and the CL system remain at least marginally stable. The increase in gain may result from either intentional adjustment or unintentional parameter variations, the latter particularly bringing in the issue of robustness. Conceptually, the GM is easily determined from the RL diagram by (i) finding the $j\omega$-axis crossing, denoted s_A, of the branch which for the lowest gain crosses over into the right-half plane (RHP); (ii) applying the CL pole magnitude condition to find the gain $K = K_0$ for which s_A is a CL pole: $K_0 = 1/|GH(s_A)|$, where $KGH(s) = KG(s)H(s)$ is the open-loop transfer function; (iii) by definition, the gain margin is thus $GM(K_m) = K_0/K_m$. As will be seen, this procedure can be automated in practice.

Phase-Root Locus (PRL)

In textbooks [1-3], stability margins are generally put off until frequency-response methods are covered. This is because although GM can be readily determined from a RL plot, there is apparently no way to determine PM from a RL plot. To help visualize PM and the relation between it and CL poles when phase is added to the open-loop transfer function, the phase-root locus (PRL) is introduced. The idea of a PRL was hinted at in [4], but was immediately dismissed because "sketching rules are not available and there is limited, if any, useful information for the designer." However, the absence of sketching rules is no longer an obstacle, given the computing power of today's personal computers. In particular, its computation becomes straightforward when it is realized exactly what the proposed locus represents. Furthermore, this article suggests ways that it can be used in a design situation. More on the relation between PRL and existing graphical tools is said later.

Recall that the conventional Evans (or "gain-") root locus plot depicts the locus of CL poles traced out in the s-plane when one adds to (or subtracts from) the log-magnitude gain *while the phase-angle added to the given design transfer function is held at zero.* Similarly, the PRL plot may be defined as one depicting the locus of CL poles traced out in the s-plane when one adds to (or subtracts from) the phase-angle *while the log-magnitude gain added to the given design transfer function is held at zero* (this of course does not imply that the gain K_m is itself unity, but rather $|K_mGH(s)| = 1$ along the PRL).

The conventional gain-root locus is equivalently the locus of values of s, $s_{G\text{-}RL}$, satisfying the angle condition

$$\angle GH(s_{G\text{-}RL}) = \pi\ell, \qquad (1)$$

(conventional ("gain-") root locus angle condition)

where ℓ is an odd integer and where zero phase-angle is added to the original open-loop transfer function K_mGH. (More generally, K_mGH could be replaced by K_mG_cGH where $G_c(s)$ is a compensator.) Notice the absence of K_m (> 0) in (1).

Analogously, the PRL is the locus of values of s, $s_{P\text{-}RL}$, satisfying the magnitude condition

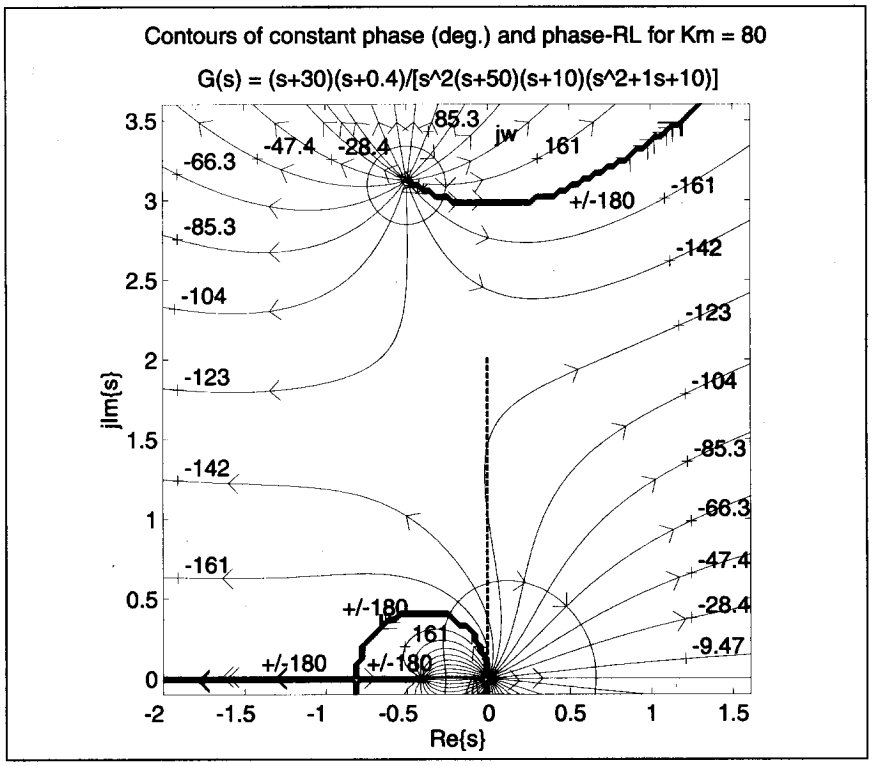


Fig 1.

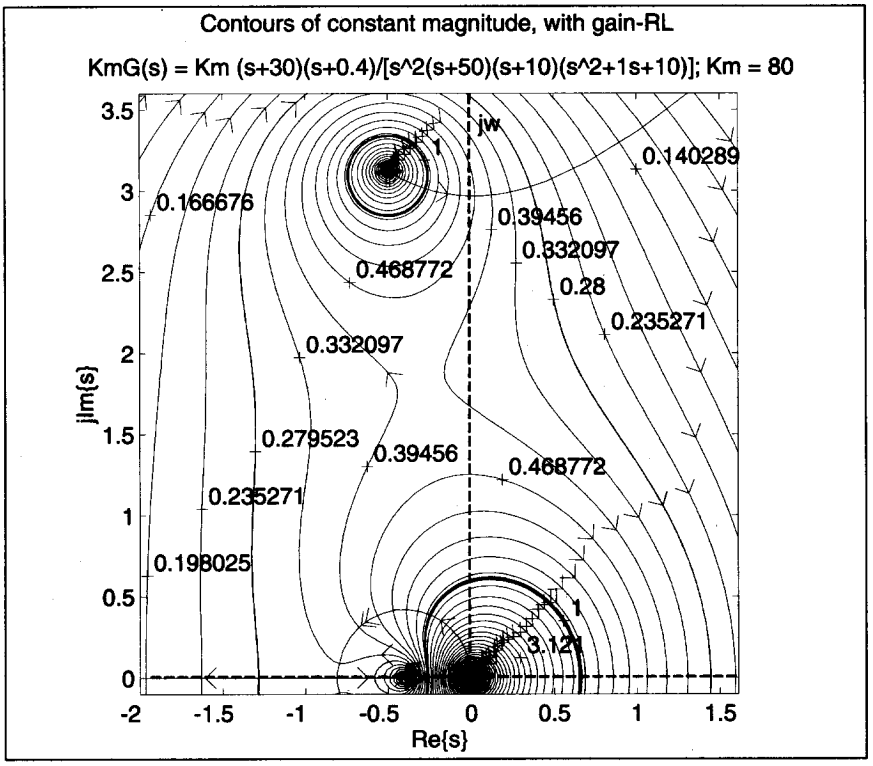


Fig. 2.

$$|K_m GH(s_{P\text{-}RL})| = 1, \qquad (2)$$

(phase-root locus magnitude condition)

where K_m is the current design value of gain K, and where zero log-magnitude gain is added to the original K_mGH open-loop transfer function.

One may obtain conventional ("gain-") root locus plots for a transfer function, albeit with reduced computational efficiency and over a chosen region of finite extent, by plotting the isocontours of ∠GH(s) with the single contour ∠GH(s) = π selected. Fig. 1 shows several constant-phase contours for the transfer function with design gain K_m and H = 1 (unity-feedback)

$$K_m G(s) = \frac{(s+0.4)(s+30)}{s^2(s+10)(s+50)\left(s^2+s+10\right)} \qquad (3)$$

The RL of (3) may be extracted from Fig. 1 as the bold curves labeled ±180°. The direction of the automatically-generated arrows indicates the usual RL CL pole movement as increasing gain is added to G(s); |G(s)| decreases from ∞ at an OL pole to zero at an OL zero. (All references to automation and code in this article are to the author's original Matlab code.) In addition, discussed later, the PRL for K_m = 80 is superimposed on Fig. 1. In all figures, only the interesting areas near the jω-axis crossings and/or the origin are shown, not entire loci.

The rather complicated open-loop system function in (3) was chosen because it has some interesting features that can be observed using PRL, as described below. G(s) could represent a positioning system with several linked mechanical elements. Specifically, a mass/damper parallel combination $\{m_1, B_1\}$ yields a simple real pole at $-B_1/m_1$ in the mechanical impedance of the combination, $Z_1(s) = (1/m_1)/(s+B_1/m_1)$ (impedance here defined as the Laplace transform of velocity/Laplace transform of applied force). Two of these mass/damper systems in series give an overall transfer function $Z_{12}(s)$ having simple poles at $-B_1/m_1$ and $-B_2/m_2$, and a real zero at $-(B_1 + B_2) / (m_1 + m_2)$ (a motor also introduces real poles into an overall transfer function). The parallel combination of a mass, damper, and spring $\{m_3, B_3, k_3\}$ has an impedance $Z_3(s) = (s/m_3)/\{s^2 + [B_3/m_3]s + K_3/m_3\}$, which has complex-conjugate-pair poles if $B_3 < 2\{m_3K_3\}^{1/2}$. If this latter system is connected in series with the above parallel combination $Z_{12}(s)$, a form of transfer function similar to that in (3) (including numerator order) results. The double pole at s = 0 in (3) results if the integral of position is considered the output variable, or if a PI controller has been applied.

There is a discontinuity in phase from -180° to +180° on either side of the RL, which is a manifestation of the phase discontinuity on the negative real axis in the G(s) plane. This discontinuity is detrimental for contour plotting on a coarse grid; notice the jagged 180°-contour in Fig. 1. Therefore, modified versions of "rlocus" in Matlab are used to generate all subsequent RL plots. To help interpret Fig. 1, note for example that the contours labeled -123° would be RL branches for the same system to which a uniform phase of -57° were added, because then the -123°-contour would become the -123-57 = -180°-contour (the RL).

In a similar way to that above, one may obtain the PRL for a transfer function, albeit without efficient plotting rules in existence (at least so far) and in a selected region of finite extent, by plotting the isocontours of $|K_mGH(s)|$ with the single contour $|K_mGH(s)| = 1$ selected. (A few isocontours are presented in [2], but no reference is made to the unit-magnitude isocontour, nor is any use made of or significance drawn from those contours, such as (2) vs. (1).) Fig. 2 shows several constant-magnitude contours for the transfer function in (3) with $K_m = 80$. The PRL is the two bold contours labeled "1" ($|K_mG(s)| = 1$). This is also the PRL superimposed on Fig. 1, for comparison with the constant-phase loci. Similarly, for reference the RL is superimposed on Fig. 2. Fortunately, there is no discontinuity problem for the PRL; the magnitude function is generally continuous near unity.

Notice in Fig. 2 that if the calibration numbers are denoted "X," then each labeled contour is the PRL for $K_m = 80/X$; e.g., the contour 0.332 is the PRL for $K_m = 80/0.332 \approx 241$. Thus, in Fig. 2 we have several PRL plots for different values of K_m, all simultaneously shown—just as Fig. 1 shows several RL plots for different added angles. When respectively relabeled with K_m or angle values, these multiple plots can be handy in system design for examining several possible alternative systems simultaneously. (We can even superimpose the multiple RL/multiple PRL plots.) While the appearance of the uncalibrated conventional RL plot $\angle K_mG(s) = \pm\pi\ell$ is identical for all positive values of K_m (and thus K_m is absent from (1)), clearly the PRL plot depends directly on K_m.

The direction of the automatically generated arrows on a given PRL contour is the direction of motion of the associated CL pole as negative phase is added. This selection of arrow directions is analogous to that used for RL: the direction of CL pole motion that for minimum-phase systems tends to be more destabilizing. Let the numbers of poles and zeros encircled by the PRL contour be, respectively, N_p and N_z. Then by Cauchy's principle of the argument [5], the phase increase over one complete clockwise (CW) revolution around the PRL is $360° \cdot [N_p - N_z]$. Thus, adding negative phase causes the -180°-point (CL pole) to move CW and thus the arrows point CW for $N_p > N_z$. If $N_p < N_z$, the arrows point counter-clockwise (CCW). An equal number of poles and zeros is never encircled.

It may be asked how phase-angle is added to an open-loop transfer function. It is agreed that all added phase is conjugate-symmetric [6]; thus, the constant-magnitude contour (and PRL) plots are always conjugate-symmetric. (Though attention is focused on the upper half-plane, a small region below the real axis is included in all plots for verification.) Details concerning a transfer function that would have such a uniform conjugate-symmetric added phase are investigated in [6]. If the concept of PM is accepted as valid and relevant, the study of addition of uniform phase to a transfer function is equally relevant because that is precisely what PM quantifies. And a highly convenient tool for study of the addition of uniform, conjugate-symmetric phase is the PRL. One might entertain the possibility that PM and added phase are irrelevant until, for example, the appearance of a system with large GM but small PM in the presence of parameter variations/model inaccuracies.

Comparison of Figs. 1 and 2 is reminiscent of the relation between electric and magnetic field lines: Electric field and RL isocontours have beginnings and ends (in the electric case because of the existence of electric charge, and in the controls case because of the presence of poles and zeros), while magnetic field and PRL isocontours close on themselves (magnetic charge does not exist). Specifically, positive charge and poles are analogous (sources of electric field lines/RL branches) while negative charge and zeros are analogous (sinks of field lines/RL branches). Moreover, the RL and PRL are orthogonal at all of their intersections, just as are $\vec{E}$ and $\vec{H}$. The analogy is striking, and a mathematical study of this might prove fruitful.

In Fig. 3, the RL and PRL are superimposed for $K_m = 150$. The intersections of the two loci, designated "P," are the locations of the actual design CL poles for $K_m = 150$; both (1) and (2) must be satisfied for any value of s to be a CL pole of $K_mGH(s)$. We see that the system is CL stable for $K_m = 150$ (and, e.g., from contour 0.332 in Fig. 2, unstable for $K_m = 241$). Relative stability and the stability margins shown on Fig. 3 are discussed below.

In teaching controls, the author has found the PRL to succeed when all else fails in bringing a student to understand what a *conventional* RL is—by means of the contrasting duality. For example, a common error made by the average student, when asked the condition for $s = s_{G\text{-}RL}$ to be on the RL, is to reply that $K_mGH(s_{G\text{-}RL}) = -1$ when in fact the answer is (1). When shown that the complete set of values of s, denoted $s_{P\text{-}RL}$, for which $|K_mGH(s_{P\text{-}RL})| = 1$ constitutes a different set of contours from the RL (namely, the PRL) drawn for a *particular* value of DC gain K_m, they see the duality and the meaning of both root loci. Another way students express the same misunderstanding is: when asked the significance of any value of s on the RL, $s_{G\text{-}RL}$, they reply that $s_{G\text{-}RL}$ must be a CL pole of the given system. This is untrue, because an infinite number of values of $s_{G\text{-}RL}$ exists, while the number of CL poles of a "given system" is finite (= N, system order). Only the *intersections* of the two loci, $s_{G\text{-}RL} = s_{P\text{-}RL}$, are the CL poles s_{CL}; only they satisfy $K_mGH(s_{CL}) = -1$. After seeing PRL, students are less likely to make these mistakes. Before seeing PRL, they may subconsciously wonder why RL is defined by only one of the conditions, and "what about the other (magnitude) condition?"

Phase-Root Locus and Relative Stability

In both RL and PRL plots, the $j\omega$ axis is the stability boundary, because any CL pole in the RHP corresponds to a rising exponential in the CL impulse response. In the conventional RL plot, the stability boundary is reached for (3) for $K = K_0 = 163.03$ at $s = s_A = j\omega_{pc}$ where ω_{pc} is the "phase-crossover" frequency of Bode analysis. This crossing is labeled in Fig. 3. Note that at $s = s_A = j\omega_{pc}$, as at all points on the RL, $\angle GH(j\omega_{pc}) = \pi$ (hence, "phase crossover" from a phase larger than π (or

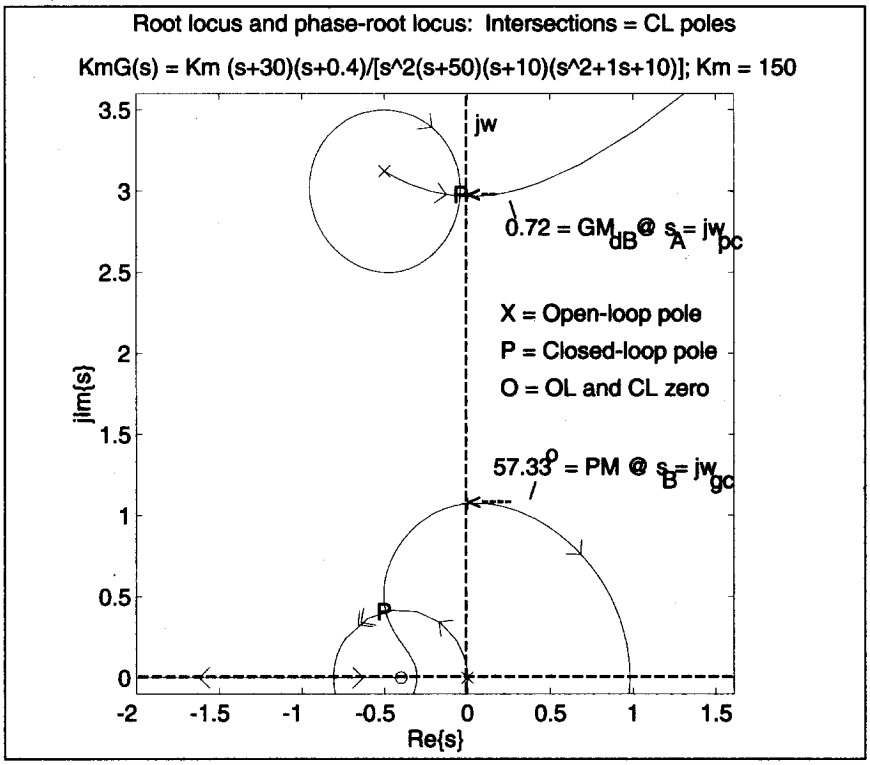


Fig. 3.

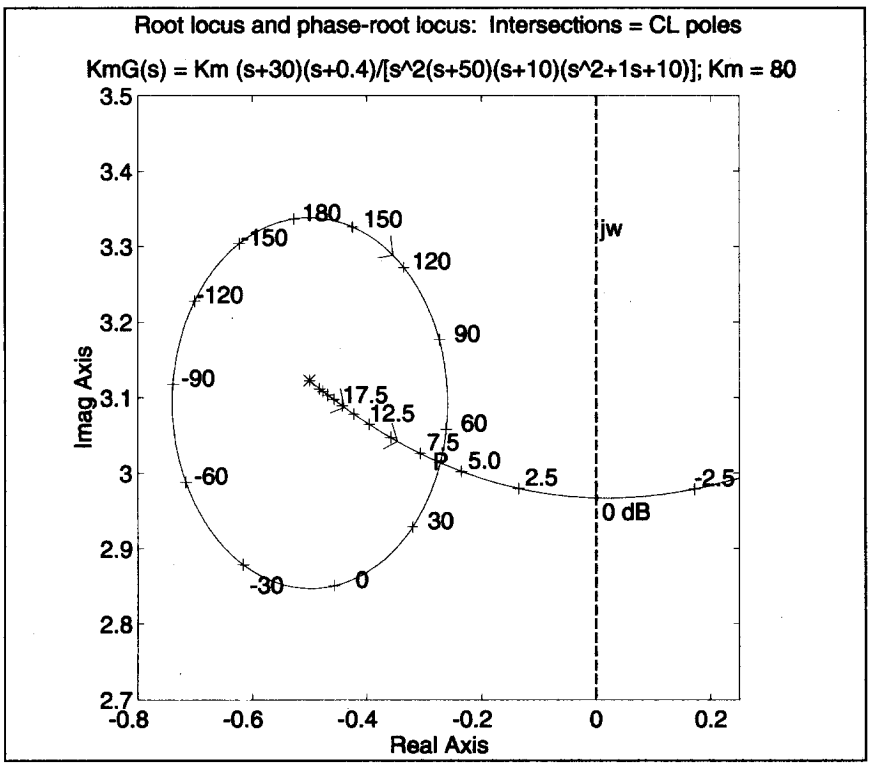


Fig. 4.

$-\pi$) to one less than π for points along the $j\omega$ axis). The GM in dB for the design value $K = K_m$ is

$$GM_{dB}(K_m) = 20\log_{10}(K_0/K_m) \quad (4)$$

The dependence of GM and PM on K_m is made explicit in this article by writing $GM(K_m)$ and $PM(K_m)$, a notation students have found helpful (there is not just one GM or PM value for any given system if the gain is adjustable). GM and PM both depend on K_m because K_m affects the location of the CL "design system" pole(s) from which the margins are determined.

To eliminate K_0 from (4), we substitute $K_0 = 1/|GH(j\omega_{pc})|$ into (4), giving

$$GM_{dB}(K_m) = 20\log_{10}(K_0) - 20\log_{10}(K_m) \quad (5a)$$

$$= -20\log_{10}(|GH(j\omega_{pc})|K_m). \quad (5b)$$

We may conventionally say that if we add M dB of gain to K_mGH uniformly for all values of s, then as long as $M < GM_{dB}(K_m)$, the resulting CL system will be stable (assuming $GM_{dB}(K_m) > 0$). The calibrated RL then shows how "P", the CL poles, move as K_m is altered by adding gain. If K_m is increased, "P" move along the RL branches in the directions of the RL arrows because the PRL contours change. If $N_p > N_z$, a PRL contour expands away from its OL pole(s) as K_m is increased (the expansion is *not* "linear," as is true for a Nyquist plot; see Fig. 2); it shrinks toward its OL zero(s) for $N_p < N_z$, as K_m is increased.

One may determine the GM of (3) for $K_m = 150$ from the RL, as previously discussed. Semi-manually, we find from Fig. 3 that $s_A = j\omega_{pc} = j2.968$ so that $1/K_0 = |G(jw_{pc})| = 0.006134$, and thus from (5b), $GM_{dB}(150) = -20\log_{10}(0.006134 \cdot 150) = 0.72$ dB, which is accurate to the resolution of the Matlab command "ginput" (approximately the value given by "margin") or, less accurately, by a manual axis-reading. This calculation is fully automated in Fig. 3 and later figures. For $K_m = 150$, we conclude that the GM is unacceptably low.

Fig. 4 shows the dual, automatically calibrated RL/{$K_m = 80$ PRL} plot in the region of the $j\omega$-axis crossing of the RL. We first focus on the RL branch plot. (The contour in Fig. 4 enclosing the OL pole, the $K_m = 80$ PRL, is calibrated with PM values, as will be described below in the discussion of Fig. 5.) From our previous analysis, we have $GM_{dB}(80) = -20\log_{10}(0.006134 \cdot 80) = 6.18$ dB. The calibrations on the RL branch are uniformly incremented "nice" values of GM in dB for the system with gain K_m adjusted so that the gain-critical CL pole is at the given calibration location. "Gain-critical CL pole" means the CL pole to cross over

to the RHP for the smallest increase in K_m; this is the pole "P" whose RL branch is depicted in Fig. 4. For example, the GM_{dB} for K_m adjusted to 163.03 so that this CL pole is located at the jω-axis crossing is $GM_{dB}(163.03) = 0$ dB. Similarly, the calibrated points nearest "P" (CL pole for $K_m = 80$) are 7.5 dB and 5.0 dB, in between which is "P", where $GM_{dB}(80) = 6.18$ dB.

Clearly, for a different value of K_m, the PRL would expand or contract, but the GM calibrations markings on the RL branch in Fig. 4 *would not change.* Displaying the calibration in this way can be a useful design tool—showing in the s-plane what various alternative gain values imply for the resulting GM, and even approximately for dynamic compensation in the vicinity of the depicted RL. Also made but omitted for brevity are automatically calibrated RL plots showing, for uniformly or logarithmically spaced sets of K_m values, the K_m value required to place a CL pole at the given position. In conjunction with Fig. 4, such plots could be useful for design. (K_m-calibrated RL plots are occasionally seen [7], but may not be automated, nor show the RL information in Figs. 4 and 5.) In fact, if one does not mind seeing a lot of numbers, each calibrated point could show the GM_{dB} *and* the corresponding K_m value required to achieve that GM_{dB} (e.g., "$GM_{dB}(80)$ = 6.18 dB" or "80/6.18 dB"), with the added benefit of also showing the locations of all the resulting CL poles in the viewing window. Furthermore, after dynamic compensation, the loci may be replotted (by re-execution of the same Matlab code) for visualizing refinements on GM and pole placement.

In Fig. 3, the stability boundary is reached along the PRL at $s = s_B = j\omega_{gc}$ where ω_{gc} is the "gain-crossover" frequency of Bode analysis. Just as $j\omega_{pc}$ is known to be an imaginary-axis crossing of the RL, $j\omega_{gc}$ is an imaginary-axis crossing of the PRL. At $s = s_B = j\omega_{gc}$, as at all points on the PRL, $|K_mGH(s_B)| = |K_mGH(j\omega_{gc})| = 1$ (hence "gain crossover" on the jω axis for $\omega > \omega_{gc}$ vs. $\omega < \omega_{gc}$). K_m determines the locations of the PRL contours and thus s_B, *and consequently the value of* $\phi_0 \equiv \angle GH(s_B)$.

Designate by s_1 the (stable) CL pole, for the design value $K = K_m$, that is on the PRL contour passing through s_B. We may call s_1 the phase-critical pole for $K = K_m$—the CL pole becoming metastable for smallest amount of added negative phase (e.g., the lower "P" in Fig. 3). At s_1, the angle condition $\angle GH(s_1) = \pm\pi$ is satisfied. The angle $\phi_0 = \angle GH(s_B)$, however, is *not* equal to π (unless the design system is already marginally stable; i.e., unless $K_m = K_0$). At both s_1 and s_B, the magnitude condition (2) is satisfied. As a negative phase-angle is added to K_mGH (e.g., from a system parameter variation) and K_m is held fixed, the point s_1 will move along the PRL contour toward s_B, just as it would move along the RL branch toward s_A if s_1 were the gain-critical pole and the gain were increased with zero added phase. Noting the analogy between K_0 and ϕ_0, (compare with (5a))

$$PM(K_m) = \phi_0 - (-\pi) \qquad (6a)$$

$$= \phi_0 + \pi \qquad (6b)$$

$$= \angle GH(j\omega_{gc}) + \pi. \qquad (6c)$$

Any integer multiple of 2π can be added to ϕ_0 for convenience. The PM is easily obtained from a PRL plot, by (i) graphically finding the metastable pole $s = s_B = j\omega_{gc}$ on the PRL (e.g., by using the Matlab command "ginput" or just reading the axis crossing visually), (ii) evaluating the phase at the metastable point, $\phi_0 = \angle GH(j\omega_{gc})$, and (iii) using (6b) to compute the PM. Again, the entire procedure has been automated, so none of this work need be done by the user.

For the example (3) with $K_m = 150$, one finds semi-manually from Fig. 3 that $s_B = j\omega_{gc} = j1.0744$ rad/s so that $\phi_0 = \angle G(s_B) = \angle G(j\omega_{gc}) = -122.66°$, and thus $PM(150) = 57.3°$, which is almost exactly what "margin" gives. Thus for $K_m = 150$, PM indicates a robust system, but we saw above that GM indicates the opposite.

If θ is a *negative* phase-angle added to the transfer function uniformly (and conjugate-symmetrically) for all values of s, then as long as $-\theta < PM(K_m)$ the resulting CL system will be stable (assuming $PM(K_m) > 0$). Recall that for a point on the PRL to be an actual CL pole, it must also simultaneously be on the RL. This condition is achieved at $s = s_B$ by adding $\theta_0 = -PM(K_m) = -(\phi_0 + \pi)$ to $K_mG(s)$ so that the new phase-angle at $s = s_B$ is $\pm\pi$ (the magnitude is already 1 because s_B is by definition on the PRL). The calibrated PRL then shows how the CL poles "P" move as negative phase θ is added. "P" moves along the PRL in the direction of the arrow because the RL swings that way (carefully review Fig. 1).

Now consider again $K_m = 80$. Using the above procedures (see the bold PRL in Fig. 2 or Fig. 5, discussed next), we may obtain $j\omega_{gc} = j0.597$, giving $\phi_0 = -130.3°$ and thus $PM(80) = 49.7°$. Thus, recalling $GM_{dB}(80) = 6.18$ dB, $K_m = 80$ gives an acceptably robust system. Figure 5 provides the same information as Fig. 4 ($K_m = 80$) for the phase-critical CL pole, near the origin. The contour crossing the jω axis is the PRL. In analogy with how the RL was calibrated in Fig. 4, the PRL in Fig. 5 is calibrated to show the new PM for the system with the phase adjusted so that the relevant CL pole is at the indicated location. It is seen that the new PM becomes increasingly more sensitive to CL pole location movement as the PRL is traversed CCW ($\theta > 0$). Equivalently, the movement of poles is *less* sensitive to added phase changes for the more CCW portions of the PRL, because it takes a larger angle variation to move a given distance along the PRL. Additional calibrations can be made to show the added phase required to move the CL pole to a given location on the PRL. Tentative dynamically-compensated systems can be analyzed in this manner for further refinements.

The new-PM calibration points on the PRL in Fig. 4 match those in Fig. 5: the "30°" markings in Figs. 4 and 5 locate two CL poles of 80G(s) with sufficient negative phase θ added such that the resulting system PM is 30°. Further, we see in Fig. 4 that the PRL encircling the OL pole -0.5+j3.123 never crosses into the RHP. Thus, when $\theta = \theta_0 = -PM(80) = -49.7°$ is added to 80G(s) to cause the CL pole on the lower PRL contour (Fig. 5) to reach the jω axis ($PM_{new} = 0°$), the CL pole in Fig. 4 merely moves to -0.46 + j2.85 (the $PM_{new} = 0°$ marking). Analogously, when sufficient gain $K_0/80$ is "added" to 80G(s) to cause the CL pole on the upper RL branch (Fig. 4) to reach the jω axis ($GM_{new} = 0$ dB), the CL pole in Fig. 5 (lower branch) merely moves to -0.54+ j0.39 (the $GM_{new} = 0$ dB marking).

In Fig. 6 are shown the RL and the $K_m = 5$ PRL. In this case, $GM_{dB}(5) = 30.27$ dB (falsely indicating robustness) while $PM(5) = 14.9°$ (system is in fact not very robust). The situation can be made arbitrarily more extreme by further reducing

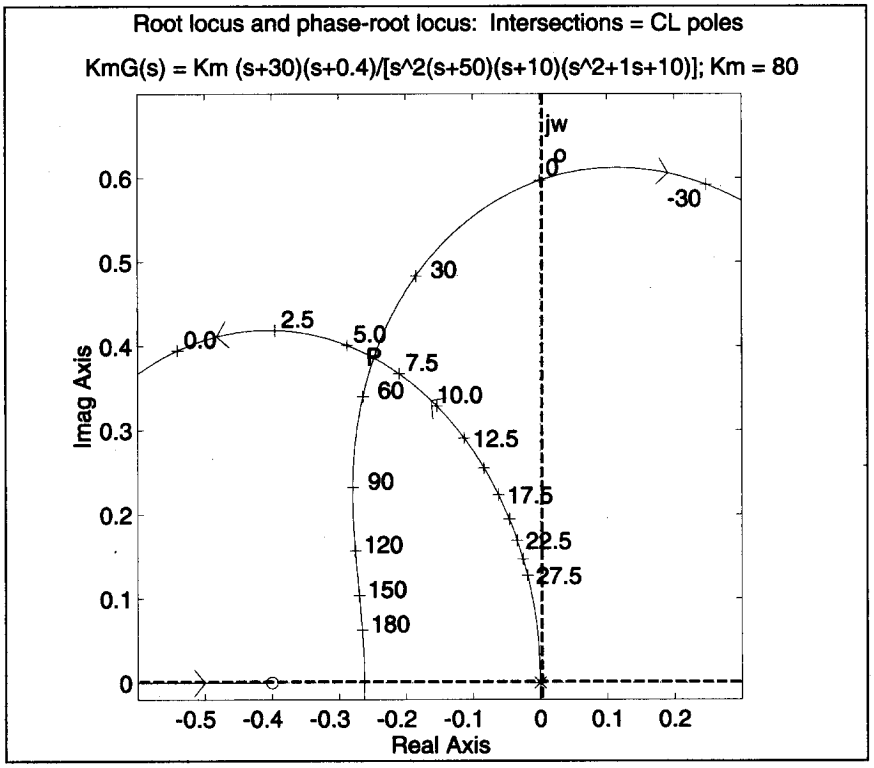


Fig. 5.

K_m. The potentially deceiving situation of great GM and poor PM is clearly illustrated when PRL is used in conjunction with gain-RL, as is the opposite situation in Fig. 3. Another interesting fact from Fig. 6 is that the most destabilizing pole for $K_m = 5$ (the "P" near the origin) is never destabilizing for added gain, but is very destabilizing for added negative phase. That is because while the RL branch for that pole never crosses the $j\omega$ axis, the corresponding PRL contour does. The opposite comments apply to Fig. 3: The gain-critical pole (on the upper branch) has a PRL contour that never crosses the $j\omega$ axis. Thus, no amount of added phase will cause that pole to become unstable, while adding a very small amount of gain will. Simultaneous RL/PRL plots reveal a cause of one margin being very large and the other very small: For low gain one pole may become unstable by adding phase, while at a high gain a *different* pole may become unstable due to added gain.

Fig. 7 shows a metastable CL system: $K_m = K_0 = 163.03$. The upper PRL contour is tangent to the $j\omega$ axis exactly where the RL intersects the imaginary axis: $s_A = s_B$. Thus, $GM_{dB}(163.03) \approx PM(163.03) \approx 0$. Also, there is more than one imaginary-axis-crossing frequency, the upper one of which (barely) coincides with $j\omega_{pc}$: For marginal stability, $\omega_{gc} = \omega_{pc}$. Looking at the lower-than-$\{X_0 = 80/163.03 = 0.491\}$ contours in Fig. 2, we see that for $K_m > K_0$ there are three imaginary-axis-crossing frequencies (e.g., see 0.395 and 0.469), until K_m is sufficiently large that the now-merged PRL contours no longer buckle into the left-half plane (the first such contour in Fig. 2 is 0.332). The PRL gives us a window into the s-plane, showing that multiple "ω_{gc}" values are caused by PRL curves passing in and out of the RHP.

An example of a CL unstable system is shown in Fig. 8. Here $K_m = 210$, so that $GM_{dB}(210) = -2.2$ dB while $PM(210) = -43°$. Now both margins are negative (both margins always have the same sign), and the PRL contours have merged into one large contour. In Fig. 8, we find that for $K_m = 210$, the upper-section of the PRL contour, not the lower is the "culprit" (the one with the destabilizing pole). Contrarily, in Fig. 6 ($K_m = 5$) the lower PRL contour is the more "dangerous" one and therefore defined the PM. In terms of the mass/damper/spring application, the upper branch/contour is dominantly controlled by the mass/damper/spring series combination, which would be responsible for deviations into the RHP for that branch/contour. Thus, if instability is caused by the upper branch/contour, this part of the system needs to be modified. The lower branch/contour is determined by the mass/damper real poles and the poles at s = 0. If one of the 1/s factors is an electronic integrator, deviation of the lower CL pole into the RHP could be caused by distortions in the integrator, for it determines the near-origin behavior of that contour. Thus, by looking at the RL/PRL plot, we may quickly diagnose the physical source of destabilization.

Further insight is obtained by comparing the upper CL poles for $K_m = 150$ (Fig. 3) with those for $K_m = 210$ (Fig. 8). For $K_m = 150$, no amount of added phase will cause that pole to go into the RHP, while for $K_m = 210$, a sufficient positive uniform angle added (above -PM = 43°) could *stabilize* the pole (and the system) with the same *"destabilizing"* value of K_m (i.e., 210)! Perhaps an approximate uniform positive phase-shifter [6] could stabilize this system.

Note that if *no* PRL contours extend into the RHP, $\omega_{gc} = \infty$. This is the s-plane visualization for the infinite ω_{gc} commonly encountered in Bode analysis. It is the analog of the well-known fact that $\omega_{pc} = \infty$ corresponds to no RL branches extending into the RHP. Incidentally, adding (conjugate-symmetric) uniform phase affects the Bode plot by shifting the phase curve up/down; the PRL shows where the CL poles move as this is done. Similarly, adding uniform gain moves the Bode log-magnitude curve up/down, and the RL shows how the CL poles correspondingly move.

Now consider a CL system that for a certain range of θ is unstable, but for additional added negative phase becomes stable. This situation is analogous to the "conditionally stable system" in RL where only for K_m within a finite range is the CL system stable. It could be diagnosed by a PRL such as the upper contour 0.469 ($K_m = 170.7$) in Fig. 2, which has a small excursion into the RHP. By examining the RL/$\{K_m = 170.7\}$ PRL upper-branch intersection in Fig. 2, we see that this is an initially unstable system ($K_m > K_0 = 163.03$). A careful study of the PRL shows that if a phase $\theta > 20°$ or $< -20°$ is added, the upper CL pole becomes stable, and the lower pole is also still in the left-half plane. The system is thus "conditionally unstable," for $|\theta| < 20°$. One might also be concerned that contours having shape similar to that of the left 0.395 con-

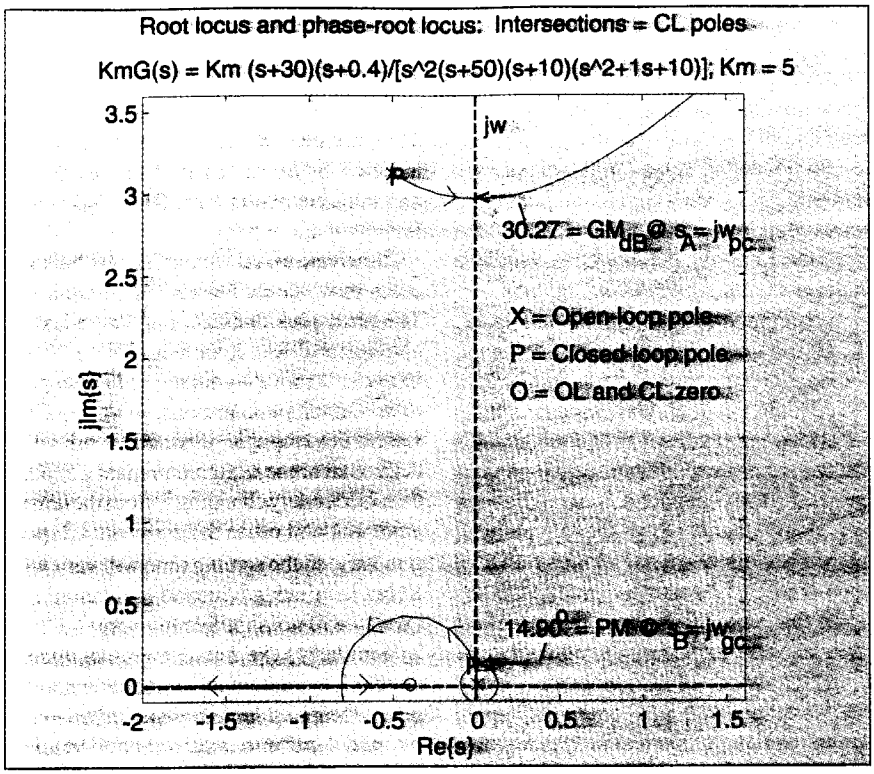


Fig. 6.

tour in Fig. 2 might for some added phase cross the jω axis, causing instability.

The PRL also facilitates understanding the relations between ω_{gc} and ω_{pc}, and thus provides a further link between Bode and s-plane analysis. The following statements are based on the assumption, commonly true away from resonances, that $|K_mGH(j\omega)|$ is a decreasing function of ω. (The frequency response is just KGH(s) for s = jω.) Let K_{m1} be a design value of K (such as 80, above) for which the CL system is stable, and K_{m2} be another (such as 210) for which it is unstable. Let the respective gain-crossover frequencies be ω_{gc1} and ω_{gc2}. Noting that (i) $|K_{m1}GH(j\omega_{gc1})| = 1$, (ii) $|K_0GH(j\omega_{pc})| = 1$, and (iii) $K_{m1} < K_0$, it follows that $|GH(j\omega_{gc1})| > |GH(j\omega_{pc})|$. Therefore, from our original assumption that $|GH(j\omega)|$ decreases with ω, we may conclude that $\omega_{gc1} < \omega_{pc}$. This is easily seen in Figs. 3 and 6, because ω_{gc} and ω_{pc} are both meaningfully shown on the same axis (unlike Bode). By identical reasoning, we can show that $\omega_{gc2} > \omega_{pc}$, which is verified in Fig. 8.

Seeing the intersections of the two root loci and simultaneously viewing the paths along constant-magnitude and constant-phase contours clarify the stability situation. The above discussion is just one example of how use of the PRL in conjunction with the conventional "gain-" root locus can help the student and controls engineer better understand exactly "what is going wrong" when a system goes unstable, and how the poles might be most easily moved to meet the temporal performance *and* robustness specifications. This is true whether or not the system is second-order dominant. Once modified by, e.g., a lead or lag compensator, the replotted RL/PRL will indicate progress made on robustness, and further optimal refinements that may be made on the compensator to meet the specifications. Such techniques are particularly helpful for the "die-hard root locus person" who prefers RL to Bode analysis and design, which previously had a monopoly on relative stability analysis.

Relation to Other Graphical Tools

The RL/PRL plot provides full stability information in the s-plane complementary to the Nyquist plot in the $K_mG(s)$-plane. Either plot easily provides GM and PM information. The calibrated RL/PRL plots in addition show, without replotting, what happens to all the CL poles and the GM or PM as variable gain or phase is added "on either side of" a given (fixed) design system $K_mG(s)$. Contrarily, the Nyquist diagram must be replotted (expanded/contracted or rotated about the axes) for each variation in gain or phase, and even then gives no direct information on the new CL pole locations. Advantages of the Nyquist plot over RL/PRL are a set of plotting rules for hand-calculation, as well as frequency response information that in the Bode form is useful in design. Also, the PRL as well as the Nyquist plot must be replotted when K_m is changed. Both the RL/PRL and the Nyquist plot give unambiguous stability information—both absolute and relative—for both minimum-phase and non-minimum-phase systems.

In [4] and [8], the information contained in RL and PRL plots is provided in an alternative graphical form. Four separate plots are required in [4], of CL poles vs. gain or added phase: |s| vs. gain, |s| vs. added angle, ∠s vs. gain, and ∠s vs. added angle. For the various plots, one must keep track of which magnitude-of-pole curve is associated with which angle-of-pole curve (all magnitude-of-poles curves are on one pair of axes, and all angle-of-poles curves are on another single pair of axes). The dominant pole is also not clearly evident. The RL/PRL is drawn on one set of axes: the familiar σ and jω. The approach in [4] may be more advantageous than RL/PRL when precise values and trends of these pole-magnitude and pole-angle parameters are individually sought. RL/PRL and [4, 8] may be used together, combining the strengths of each method.

The plots in [4] are so far unknown to be the basis of algorithms for dynamic (beyond proportional-only) compensator design. Even the Nyquist diagram is partially eclipsed by Bode techniques in design algorithms, at least for minimum-phase systems. This does not at all render these plots useless. They help the designer visualize in a variety of ways, as stability quantifiers, indirect temporal response indicators, and illustrators of non-dynamic compensation effects. Similarly, the PRL has not yet at this initial stage been articulated as the basis of a popular design method. Of course, it is true that DC gain (by electronic amplification) is more easily intentionally varied than is uniform phase. However, initial studies in [6] suggest ways that movement along the PRL can approximately be implemented using realizable compensators, for limited-magnitude shifts. Thus, the

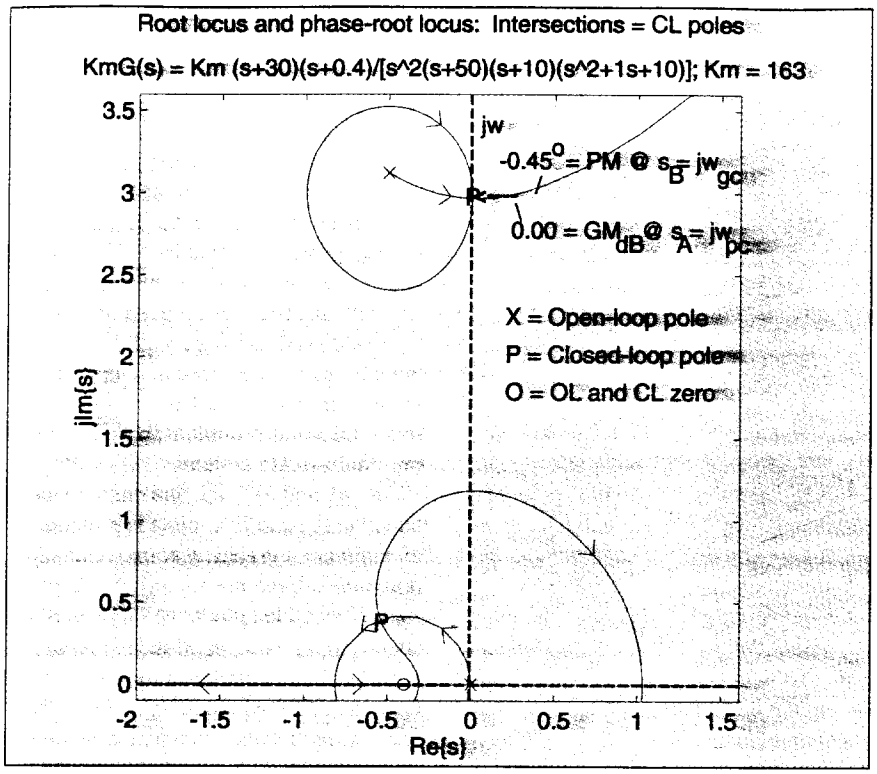


Fig. 7.

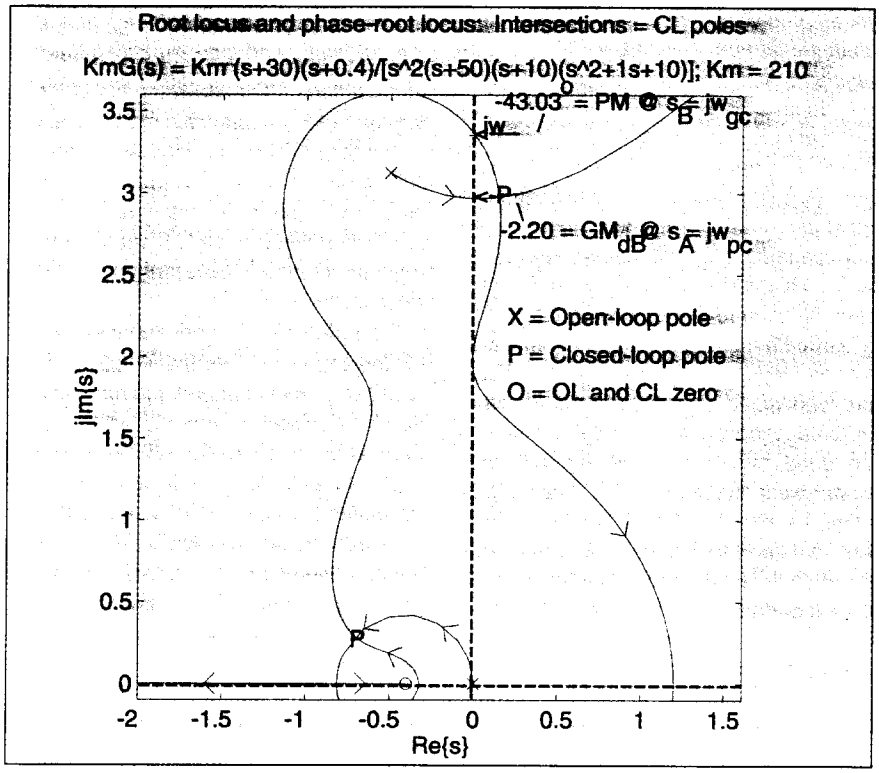


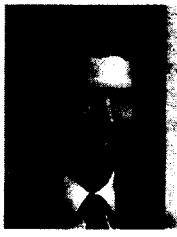
Fig. 8.

PRL may be used in practical design algorithms.

For analysis and instruction, the PRL has already proved worthwhile. Because the PRL shows CL poles on the s-plane, anyone who understands RL can easily learn to interpret PRL, which is shown on the same (s-) plane. For example, it is easy to remember that the *phase*-root locus indicates *phase* margin and the *gain*-root locus indicates *gain* margin. The PRL facilitates investigation of the temporal characteristics trends of CL poles as phase is varied, by the same method as one evaluates transient trends using RL as K_m is varied.

It may be asked what it matters how the poles move around when phase is added. The analogous question may be asked of conventional RL. Root locus shows where CL poles "move"—directly what happens to the CL temporal response—when gain is varied. For example, if the RL is primarily horizontal at the design dominant CL pole, then moderately changing K from the design value K_m will cause the dominant CL pole σ to vary, so the settling time will vary with K, but the ringing (damped oscillation) frequency will not significantly vary.

Similarly, the PRL shows explicitly what happens to the temporal response when the phase-angle is perturbed (e.g., by model parameter error/variation). For example, if the PRL is primarily vertical at the design CL pole, then the settling time will not be affected by phase variations θ, while the ringing frequency will. Bode plots do not directly give this information; this instant visualization of the temporal response is one reason among others for the preference of the RL approach by some designers. Unlike Bode plots, the RL/PRL plot provides *both* relative stability and temporal characteristics information without dependence on a second-order assumption relating them (because the effects of/on *all* CL poles are individually available).



Thomas J. Cavicchi received the B.S. degree from the Massachusetts Institute of Technology in 1982, and the M.S. and Ph.D. degrees from the University of Illinois in 1984 and 1988, respectively. Since 1993 he has been an associate professor of electrical engineering at Grove City College, where he teaches analog and digital control theory/design, and EE for non-electrical engineering majors. He wrote *Fundamentals of Electrical Engineering: Principles and Applications* (Prentice Hall, 1993). Prior to joining GCC, he was an assistant professor of EE at the University of Akron, where he published papers in the areas of DSP and biomedical ultrasonic diffraction scattering.

Finally, the computational time of RL/PRL is modest. All plots were done using Windows Matlab 6.2c.1 on a Pentium PC. The isocontour arrays are 100 x 100 pixels^2, but results are about as good with only 50 x 50 or less. With no efforts yet made on code optimization, the times for the most computationally-intensive plots in this article (Figs. 4-5) are as follows: The 100 x 100 arrays take 30 s (50 x 50 takes only 7 s) to generate and the multiple-curve plotting and full calibration of all contours, including arrows, takes another 28 s (50 x 50 takes 22 s). Noting that an uncalibrated, unlabeled "rlocus" call alone takes 5 s (20 s on a laptop), the wait is quite reasonable and may be further reduced by code optimization.

Conclusion

The PRL plot is a means of easily determining phase margin (and gain-crossover frequency) using a root locus approach, which previously could not be done. More important, it offers a new dimension of potential design information: Now the s-plane can be used for robustness as well as transient design, and it can help diagnose the faulty elements in an unstable CL control system. A design technique based on approximate PRL and uniform phase modifications in the resonance region of the s-plane has already been developed [6]. By adding uniform phase, the RL branches swing and deform (see Fig. 1), causing the intersections with the PRL and thus the CL poles to move in a predictable manner. The exact trajectories of CL poles toward and away from instability for added gain and phase are provided by the RL/PRL plot. As more is learned about the causes of phase variation (e.g., via parameter error/variation), the PRL may provide the resulting temporal effects and be increasingly practical as a design tool. Conventional RL-based design methods may be enhanced by the complementary information in the same format provided by the PRL.

The control system designer with experience learns patterns that occur in conventional RL plots when zeros and poles are added or subtracted, or plant or controller parameter values are altered. However, only one of two "dimensions" of root-locus behavior is then examined. Also important, particularly for robustness analysis, may be how the CL poles vary with (often unintentionally) added phase. Again there will be patterns—heretofore unexamined—in the effects that pole/zero addition/relocation have on the contours that CL poles take when phase is added. Thus new insights and a sense of completeness not attainable with only conventional RL are now available with its dual, the phase-root locus.

References

[1] B.C. Kuo, *Automatic Control Systems,* Englewood Cliffs, NJ: Prentice Hall, 1995 (7th ed.).

[2] K. Ogata, *Modern Control Engineering,* Englewood Cliffs, NJ: Prentice Hall, 1990 (2nd ed.).

[3] N. Nise, *Control Systems Engineering.* Reading, MA: Benjamin/Cummings, 1995 (2nd ed.).

[4] M.L. Nagurka and T.R. Kurfess, "Gain and Phase Margins of SISO Systems from Modified Root Locus Plots," *IEEE Control Systems Mag.,* June 1992, vol. 12, pp. 123-127.

[5] N. Levinson and R.M. Redheffer, *Complex Variables,* San Francisco, CA: Holden-Day, 1970, pp. 216-218.

[6] T.J. Cavicchi, "Minimum Return Difference, Phase Margin, and Compensator Design," submitted to IEE Proceedings—Control Theory and Applications, March 1996.

[7] C.-T. Chen, *Analog and Digital Control System Design.* New York: Saunders College Publishing, 1993. p. 258.

[8] T.R. Kurfess, M.L. Nagurka, "Understanding the Root Locus Using Gain Plots," *IEEE Control Systems Mag.,* August 1991, vol. 11, pp. 37-40.