CONTROL SYSTEM DYNAMICS BY DIGITAL COMPUTER

J. B. X. Devotta

1 Introduction

Analysis of Control System Dynamics is usually done either in the time domain or in the frequency domain. In the former, the system dynamics appear in the form of differential equations, and in the latter, by Laplace transformation, they appear as ordinary algebraic equations. It may seem that the task of solving algebraic equations is much easier than solving differential equations. However this is not true. There is no straightforward method of solving algebraic equations when the degree of the equation exceeds 2. Lin's[3] method of Trial Divisors is often used in such cases, but the labour involved is quite excessive. In dealing with Control Systems in particular, a knowledge of the roots of the characteristic equation is essential, because the stability of the system and its transient response is entirely governed by the roots of the characteristic equation. Furthermore the designer often wishes to know how these roots are affected by any intended or unintended changes in one or more system parameters. The frequency response method of Nyquist[4] and the Root Loci method of Evans[5] are some of the well established techniques in this field.

In the Nyquist method, knowledge of the open loop transfer function (either explicitly or deduced experimentally), enables deduction of useful information regarding the stability of the closed loop system. The concept of gain margin, phase margin and crossover frequencies, as obtained from this method provide some information regarding the degree of stability. However only in those cases where it can be assumed that the system response is dominated by one pair of complex poles is it possible to relate directly the concept of phase margin and gain margin and crossover frequency with some of the time domain specifications such as damping ratio, overshoot, settling time etc. The question is to know beforehand whether such an assumption is valid, and this cannot be known until all the closed loop poles are located.

In the Root Loci method, the migration of the closed loop poles of the system, as the loop gain constant is varied, is shown in the form of curves. These plots yield valuable information to the system designer. The main difficulty encountered is the accurate construction of these loci themselves, as for example,

(i) Several search points have to be located to obtain an accurate plot.

(ii) The location of breakaway and break in points involve the solution of reciprocal equations (trial and error method).

(iii) The calibration of the curve with respect to the gain constant parameter is tedious.

The use of the spirule, can reduce the labour involved, to a certain extent. In addition, this method permits the variation of only one particular parameter, namely the system loop gain. Modifications have been devised[6] for the study of the variation of other parameters but these are complicated. Another important feature of this method is that, even a slight modification of the system causes radical changes in the nature and shape of the Loci plots. For example, the introduction of error derivative or integral of error type of compensation, which is equivalent to adding an extra zero or pole in the open loop transfer function, results in an entirely new shape of the Root Loci plot.

When viewed in this light, the method of Mitrovic[1,2] has several distinct advantages such as (i) possibility of simultaneous adjstments in two parameters (ii) work is entirely in the real variable domain (iii) all system poles are located simultaneously (iv) only one set of curves is needed in most cases and (v) application to a wide variety of cases. Basically this method can be regarded as a modification of the Contour Integration method of Nyquist, mapped onto the $(a_0 - a_1)$ parametric plane with the important exception that the function considered is the characteristic polynomial of the closed loop system, instead of the open loop transfer function. In this method, a family of curves are generated, each one corresponding to a particular value of damping ratio g, satisfying the characteristic equation. A set of simple algorithms is used, repetitively, to generate a table of coefficients from which the curves are generated. However, without the help of a calculating machine the numerical work involved can be quite formidable, and perhaps this is the reason, why this method has not become popular so far. In this era of computers, this is no longer a problem. The programming part is quite simple (section 5) and even a small computer or a desk calculator with a small memory unit is sufficient for the purpose.

Dr. Devotta, BSc (Eng), Phd, graduated from the University of Sri Lanka in 1965 and joined the staff of the University in the same year. In September 1969, he joined the Power Systems group of the Imperial College of Science and Technology as a research student. His work was on system stability with special reference to the D.W.R. Machine and included both experimental and theoretical work on the Micromachine installations and digital computer respectively. He returned to Sri Lanka in January 1973 and rejoined the University Staff.

2 Formulation

Consider the feedback control system, having a forward transfer function G(s) and feedback transfer function H(s) (Fig. 1). Assuming N_1, N_2, D_1 and D_2 to be polynomials in s, the characteristic equation of the closed loop system can be written as,

$$f(s) \equiv N_1(s).N_2(s) + D_1(s) = \sum_{k=0}^{n} a_k s^k = 0 \quad \text{....(1)}$$

A necessary (but not sufficient) condition for stability is that all the coefficients a_k be positive. Rearranging (1) by substituting the generalised variables x and y in place of a_1 and a_0,

$$f(s) = \sum_{k=2}^{n} a_k s^k + xs + y = 0 \quad \text{.....(2)}$$

Substituting for the complex frequency variable S, in terms of w_0 the natural frequency and g the damping ratio,

$$s = w_0(-g + j\sqrt{1-g^2}) = w_0 \,\angle\, \tfrac{\pi}{2} + \theta \quad \text{.....(3)}$$

where $0 < \theta < \frac{\pi}{2}$

and if g is constant, and w_0 is varied, we obtain the locus of a line such as BOA of Fig. 2; It can be shown from (1), (2) and (3) that

$$x = \sum_{k=2}^{n} a_k w_0^{k-1} \phi_k(g) \quad \text{.....(4)} \quad \text{and}$$

$$y = -w_0^2 \left(\sum_{k=2}^{n} a_k w_0^{k-2} \phi_{k-1}(g)\right) \quad \text{.....(5)}$$

where $\phi_k(g)$ is given by the recursive relation,

$$\phi_k(g) = -\left[2g\phi_{k-1}(g) + \phi_{k-2}(g)\right] \quad , \text{.....(6)} \quad \text{with}$$

$$\phi_0(g) = 0 \text{ and } \phi_1(g) = -1$$

Equations (4) and (5) define the mapping of the line BOA into the O_{xy} plane, as mapped by the function f(s). The functions ϕ_k are independent of the degree of f(s), and of the coefficients a_k of the characteristic polynomial. ϕ_k can be calculated once and for all, for any desired value of g.

Now consider the contour S_1 in the s-plane (Fig. 2), Suppose we wish to determine if all roots of $f(s) = 0$. have a damping ratio greater than a given value g. For the assumption to be valid, the contour S_1 should enclose all these roots. Since all the coefficients a_k are real, it follows that, contour S also encloses all these roots.

Assuming that f(s) does not vanish on the contour S, we can deduce from Cauchy's theorem that the increase in the argument of f(s), when the contour S is traversed once, should be equal to $2m\pi$, where m is the number of roots of f(s) lying within S. If m equals n, (the order of the system polynomial), then all the roots of f(s) have a damping ratio greater than the value corresponding to S. Now, we can assume without loss of generality that $a_n > 0$, and since f(s) tends to $a_n s^n$ as w_0 tends to infinity, the contribution to the argument by the semicircular path ADB tends to $n\pi$. Therefore the contribution by the path BOA should also tend to $n\pi$.

Instead of considering the mapping of S by f(s) explicitly we consider the mapping of S into the O_{xy} plane as defined by equations (4) and (5), and find the corresponding relationship between this curve in the O_{xy} plane and the mapping as defined by f(s) directly. From equations (2), (4) and (5), we obtain the relationship,

$$f(s) = (a_0 - y) + s(a_1 - x) \quad \text{......... (7).}$$

Knowing the curve in the O_{xy} plane, the shape of the curve in the s-plane can be easily deduced using equation (7). An accurate plot of the curve in the s-plane is not necessary as we are interested only in the increase in argument of the curve. Furthermore, as will be evident from the subsequent discussion, we need not draw the curve in the s–plane at all. We can simplify the procedure still further by considering the mapping of the portion OA only, instead of BOA, as given below.

Draw the mapping of the line OA in the O_{xy} plane as defined by (4) and (5). By allowing w_0 to vary from zero to infinity, corresponding to the particular value of g. Let this contour be named C_g. The necessary and sufficient conditions for all the roots of $f(s) = 0$, to have a damping ratio greater than the value corresponding to the contour S, have been deduced by Mitrovic[1,2] and are given below:

(i) The point M ($x = a_1$, $y = a_0$) should lie in the first quadrant.

(ii) The lines $y = a_0$ and $x = a_1$ should cut the Cg curve alternately, starting with the line $y = a_0$.

(iii) The number of such intersections described in (ii) is m where m is the largest positive integer satisfying the inequality

$$m\frac{\pi}{2} + \frac{(1-(-1)^m)}{2} < n\left(\frac{\pi}{2} + \theta\right) \quad \ldots\ldots(8)$$

where $\theta = \sin^{-1}(g)$

Thus, the problem of regional location of roots has been simplified to the drawing of the curve Cg, locating the operating point M (a_1, a_0) and inspecting the number of intersections of the lines through M, parallel to the x-axis and y-axis, with the curve Cg. A table of values of n, g and m corresponding to the inequality (8) is given below for ease of reference.

n	Range of g	m
2	0 $0 < g \leq 1$	1 2
3	0 $0.0 < g \leq 0.50$ $0.50 < g \leq 1$	2 3 4
4	0 $0.0 < g \leq 0.50$ $0.50 < g \leq 0.707$ $0.707 < g \leq 1.0$	3 4 5 6
5	0 $0.0 < g \leq 0.309$ $0.309 < g \leq 0.707$ $0.707 < g \leq 0.809$ $0.809 < g \leq 1.0$	4 5 6 7 8
6	0 $0.0 < g \leq 0.309$ $0.309 < g \leq 0.588$ $0.588 < g \leq 0.809$ $0.809 < g \leq 0.866$ $0.866 < g \leq 1.0$	5 6 7 8 9 10

TABLE I

For example, in the specific case of $n = 3$, $g = 0.45$, m should be 3. From Fig. 3, it is seen that the operating point M_1 satisfies all the conditions whereas the operating point M_2 does not.

3 Computation

The application of the Cg curves just explained, cover a very wide field. Therefore the actual computation of these curves will be described in detail. We can rewrite equations (4) and (5) in the form,

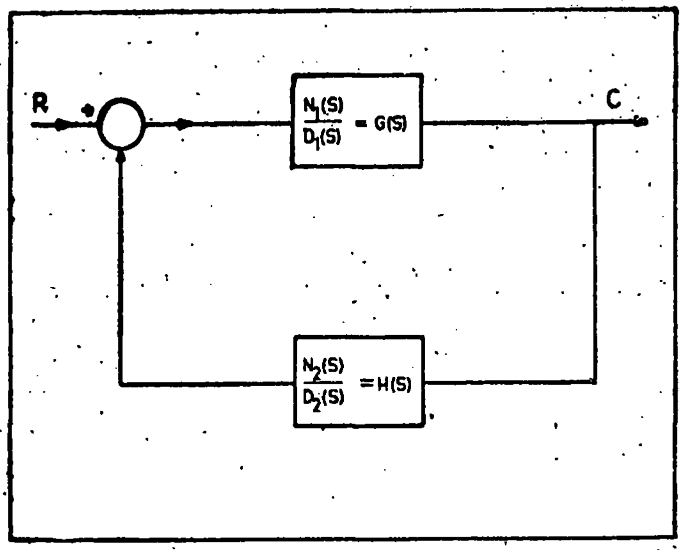

Fig. 1 Block diagram of a Feedback Control System

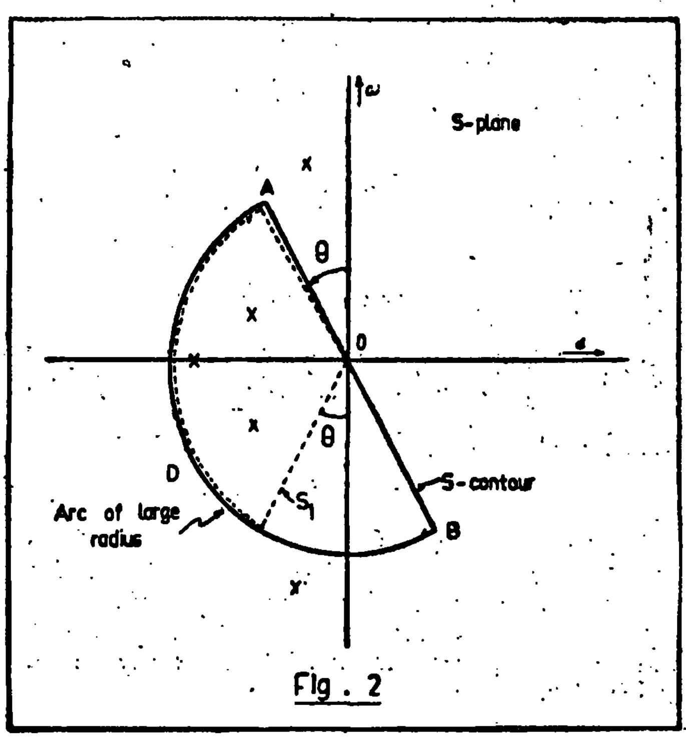

Fig. 2 Mitrovic Contours in the s-domain

$$x = \sum_{2}^{n} a_k F_k \qquad \ldots\ldots (9)$$

$$y = -w_o^2 \sum_{2}^{n} a_k F_{k-1} \qquad \ldots\ldots (10)$$

where $F_k \equiv \phi_k(g) \cdot w_o^{k-1}$

F_k, which is a function of g, w_o and k will hereafter be referred to as the Mitrovic Coefficient. Note that F_k is independent of the system order n, and the coefficients a_k. These Metrovic coefficients can be calculated once and for all, and stored for as many values of g and w_o as we choose. The computation of x and y then simplifies to the simple summation algorithm defined by equations (9) and (10).

When these curves are actually computed, it is found that the range of values of x and y vary very widely, since they depend on the values of the coefficients a_k, (k> 2) and also on the range of w_o selected. Theoretically the range of w_o should be from zero to infinity. It is desirable to know beforehand how large a value of w_o should be

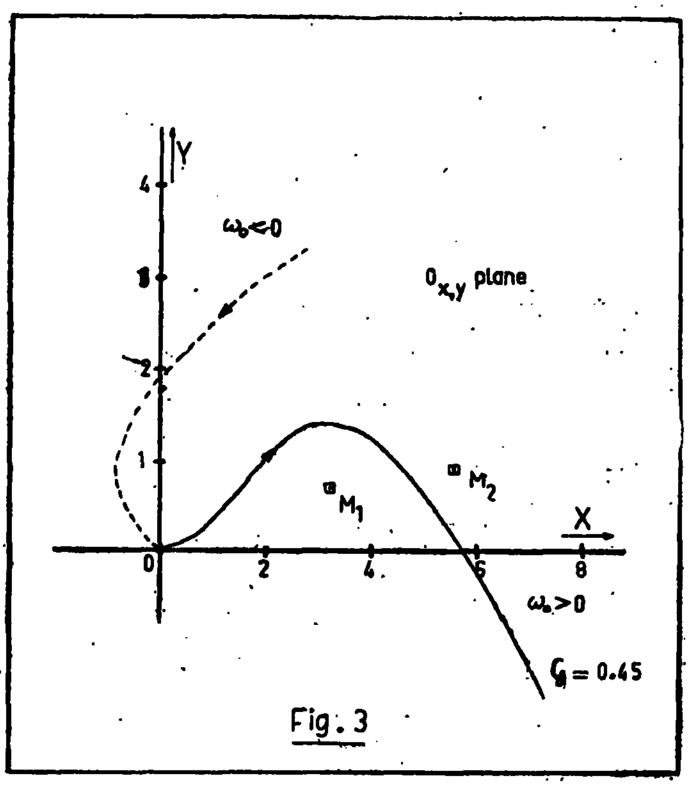

Fig. 3 Mitrovic Curve in the parametric plane

used in a practical computation. The process of numerical computation can be generalised for all problems by a simple linear transformation of the variable s. By transforming the characteristic equation to a form,

$$p^n + p^{n-1} + a'_{n-2} p^{n-2} + a'_{n-3} p^{n-3} + \ldots\ldots a'_o = 0 \qquad \ldots\ldots (11)$$

where $a'_k = \dfrac{a_k}{a_n \cdot C^{n-k}}$

the natural frequency w_o, and the co-ordinates of the operating point M become,

$$w'_o = \frac{w_o}{C}, \quad a'_1 = \frac{a_1}{a_n C^{n-1}} \text{ and } a'_o = \frac{a_o}{a_n C^n} \qquad \ldots\ldots (12)$$

For problems concerning root location, stability and design, a range of w_o from zero to 1 is sufficient (see examples I, II and III) and for most applications, variation of w_o from zero to a value of 2 or 3 is all that is needed. (See Appendix A).

4 Applications

1. *Stability of Closed Loop System.* The system is stable, if the damping ratios of all roots of the characteristic equation are greater than zero. Thus the curve C_0 should be used for determining stability and inequality (8) reduces to the simple form m = n-1, by substituting g = o = θ. That is, the operating point M = (a_0, a_1) should lie in the first quadrant and the lines $y = a_o$ and $x = a_1$ should intersect $C_{o.o}$ curve alternately (n - 1) times, starting with the line $y = a_o$ (see example Ib).

2. *Determination of Real Roots.* Consider the case of $C_{1.0}$ curve. This is the curve corresponding to the existence of double negative real roots, that is θ of Fig. 2 tends to $\frac{\pi}{2}$. For this condition the inequality (8) reduces to the form m = 2(n - 1). The interpretation of the statement "damping ratios of all roots greater than 1.0" is that all roots of f(s) = o are real and negative. Therefore the necessary and sufficient conditions for the existence of a totally overdamped system response is that the operating point M = (a_o, a_1) should lie in the first quadrant, and the lines $y = a_o$, $x = a_1$ should intersect the curve $C_{1.0}$ 2(n - 1) times alternately, commencing with the line $y = a_o$.

The curve $C_{1\cdot 0}$ also yields information regarding the actual magnitudes of the real roots of $f(s) = 0$. However, some of the properties of the $C_{1\cdot 0}$ curve should be known first, as they provide much assistance in drawing this curve. These are given below.

For $g = 1$, $s = -r$ and hence by (6), we get

$\phi_k = (-1)^k . k$ .....(13) and $F_k = (-1)^k . k . r^{k-1}$ .....(14)

Also,

(i) $C_{1\cdot 0}$ curve has only one extremum and that occurs at the origin.

(ii) $C_{1\cdot 0}$ curve has no points of inflexion.

(iii) $dy/dx = -$ at r all points along the $C_{1\cdot 0}$ curve.

(iv) The $C_{1\cdot 0}$ curve may have cusps at certain singular points. The number of cusps is either (n-2), (n-4) ... etc., where n is the order of the equation. These cusps occur at values of 6 which are the roots of $f''(-r) = dx/dr = 0$. For there to be cusps, these roots must be real and simple.

(v) The real roots of $f(s) = 0$ are given by the negative of the slope of the tangents from the operating point $M = (a_0, a_1)$ to the curve $C_{1\cdot 0}$. (For a proof of these results see Appendix B).

For illustration of the property regarding the cusps, consider a third degree equation which will have either one cusp or none. (A fourth order equation will have either two or none, a fifth order equation, either 3 or 1 or none etc.). An explicit solution of the equation $f''(-r) = dx/dr = 0$ is not necessary; it is only necessary to keep track of the sign of dx/dr, as r, the explicit parameter of the $C_{1\cdot 0}$ curve is varied. When dx/dr changes sign, a cusp is indicated and the $C_{1\cdot 0}$ curve passes from one side of the tangent at the cusp, to the other (see Fig. 4a). A separate computation of dx/dr is not necessary. It can be shown (see Appendix B) that,

$$\frac{dx}{dr} = -\sum_{1}^{n} k . a_k . F_{k-1}$$

Hence dx/dr can be calculated at the same time x, y are being calculated.

The determination of the real roots by drawing tangents could become rather inaccurate if attempted by hand, especially when the real roots are large. However these tangents can be drawn accurately even though the point of tangency is inaccessible or cannot be known exactly. The slope of the line is known to be (- r) and the point at which it cuts the x axis is given by

$$x_0 = \sum_{1}^{n} \frac{a_k F_k}{k} \quad \text{(see Appendix B)}$$

which also can be calculated. Since the slope and one point of the line are known precisely, it is possible to draw accurately the family of tangents for as many values of r, as may be needed (Fig. 4b).

It should be clear from what has been discussed so far, that this method is ideally suited for programming on a calculating machine. With the help of a small digital computer, capable of storing the Mitrovic coefficients, the computation of the Cg curves, and the tangents to the $C_{1\cdot 0}$ curves is quite a simple task. The program devised for this purpose is explained in the next section.

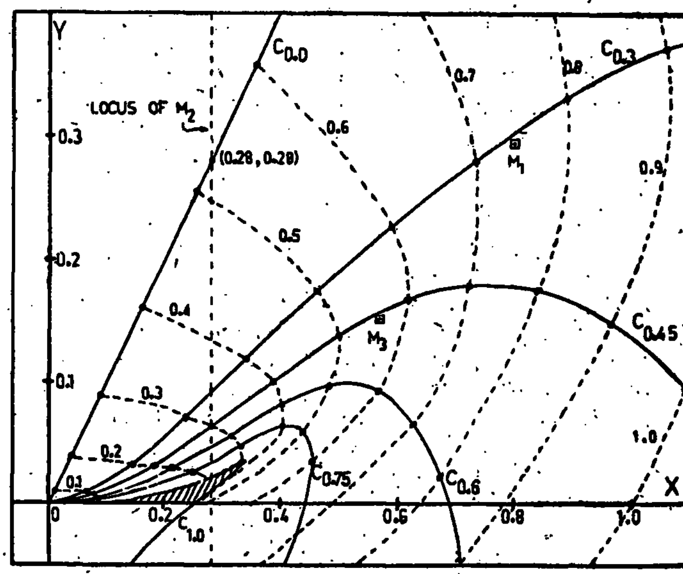

Fig. 4a Mitrovic curves for the third order system (Example 1 of Section 6).

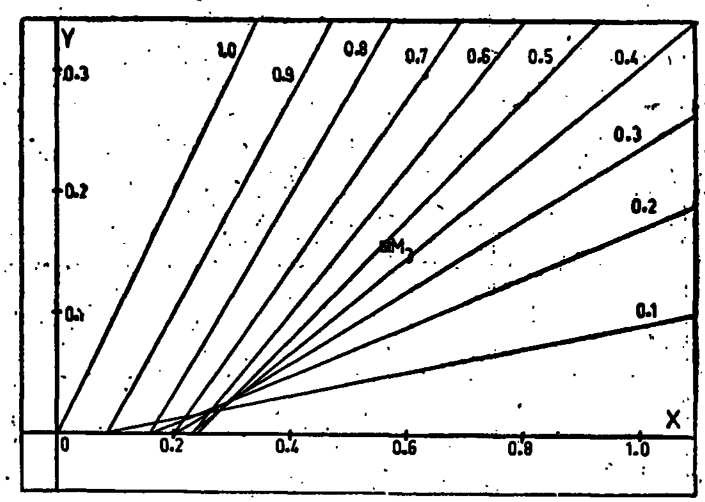

Fig. 4b Family of Tangents to the Mitrovic Curves $g = 1$. (Example 1 of Section 6).

5 Computer Program

The computer program for the application of this method was developed in two stages, and was run on the IBM 1130 computer at the Faculty of Engineering, Peradeniya. A Plotter is available as a peripheral. In the first stage a program was written for the evaluation of the functions ϕ_k for different values of g and k. These were then used to generate the matrix of Mitrovic coefficients F_k, which was then stored on disc. The system order considered was limited to a maximum value of 7. The damping ratios selected were from zero to unity in steps of 0.05. The normalised frequency range was from zero to 2.0 also in steps of 0.05. This subdivision permits sufficient accuracy for most practical applications. By plotting the Mitrovic curves on a sufficiently large scale, the values of g and w_o can be estimated to an accuracy of + 0.01 p.u. which was considered adequate. It should be noted that enhanced accuracy is not limited by the method adopted and this can be obtained at the expense of computer storage and execution time. While calculating the points on any particular Mitrovic curve Cg, g does not change. Hence the 5820 Mitrovic coefficients were stored as ordered arrays in such a way as to facilitate their recall.

The second stage was the development of the program MITRO. The user has to supply the values of n, the system order m, the number of Cg curves desired, a_k, the array of coefficients of the characteristic polynomial, and g, the array of values of damping ratio desired. The coefficients a_k are first normalised and then using the corresponding F_k coefficients, the points on Mitrovic curve are computed. The output is given in the form of a table of the normalised frequency w_o, and the co-ordinates of the points on the Cg curve. Facilities are provided for a graphical display of some or all of these curves on the plotter. An optional facility is the enlargement of the region around the origin in the case of the $C_{1.0}$ curve, to show more clearly the location of real roots. Whenever unity damping ratio is specified in the data by the user, the program also calculates dx/dr, and data pertaining to the family of tangents of the $C_{1.0}$ curve. These results are supplied in the printed output. The family of tangents can also be plotted at the same time (see Fig. 4b). Note that any problem concerning third order systems can be solved by just one set of curves, which need be computed only once. This chart, called the Mitrovic Chart is extremely useful in the solving of many types of problems concerning Control Systems which can be approximated to the third order.

(To be continued)

References

1. Mitrovic, D., '*Graphical Analysis and Synthesis of Feedback Control Systems* Part I'. Trans. A.I.E.E., Vol. 77, Part II, (1959), pp. 476 - 487.
2. Mitrovic, D., '*Graphical Analysis and Synthesis of Feedback Control Systems*, Part II', ibid., pp. 487 - 496.
3. Lin, S. N., '*A Method of Finding the Roots of Algebraic Equations*', Journal of Mathematics and Physics, Cambridge (Mass). Vol. 22. June 1943. pp. 60 - 77.
4. Nyquist, '*Regeneration Theory*', Bell Systems Technical Journal, Vol. II, 1932. pp. 126 - 147.
5. Evans, W. R., '*A Graphical Analysis of Control Systems*', Trans. A.I.E.E., Vol. 67, (1948). pp. 547-551.
6. D'Azzo and Houpis, '*Feedback Control System Analysis and Synthesis*' (Book). McGraw-Hill, 2nd Edition.