AIAA 82-4307

# Factors for Cubics and Quartics

Dunstan Graham*
*Custom Computer Consultants, Ypsilanti, Michigan*

**Methods of solving cubic and quartic polynomial equations are reconsidered and compared. Particular reference is made to the most efficient strategy for each case irrespective of the values of the (real) coefficients. A number of useful "exact approximations" are collected or developed and displayed. These are appropriate starting estimates for iterative, numerical calculations; or, in literal terms, they serve to enhance understanding. Alternatives to the popular method originally due to Bairstow are re-examined, and it is suggested that the possibility of systematic choice of improved starting estimates makes these alternates much more attractive than they have seemed heretofore. Considerations appropriate to the cubic equation are extended to quintics.**

## Introduction

THE solution of polynomial equations with real coefficients has a surpassing importance in engineering and applied science. This may be particularly the case in connection with aeronautics. Since the discovery of the quadratic formula, cubic and quartic equations have represented the prototype problems.

More than 5000 years ago, the Sumerians invented the city, writing, and the wheel. Professor Dirk Struik[1] has pointed out that, in their schools, they taught the solution of cubic and quartic equations. Their technique, however, apparently was lost for a long time. The first generally appreciated results were published in 1545 by Girolamo Cardan. Cardan's student, Ludovico Ferrari, about the same time, found a method for the solution of quartic equations by means of a "resolvent" cubic. Since then, the "theory of equations," and, in particular, the solution of cubics and quartics with real coefficients has been studied and improved upon by some of the very greatest mathematicians including Descartes (1596-1650), Newton (1642-1727), Euler (1707-1783), Lagrange (1736-1813), and Gauss (1777-1855). However, about 1926, there developed a consensus perception that the theory was complete. "Pure" mathematicians ceased work on it, and the subject was dropped from the college curriculum.

Neumark,[2] comparatively recently, has presented a comprehensive survey of both the practice and relevant science. One is bound to agree with his statement: "It is, of course, impossible to add anything now to the basic theory of these equations."

Nevertheless, it is possible for an engineer to persist in discontent. He still often must deal with cubics and quartics which are difficult; and these are a key to understanding equations of higher degree. In spite of an enormous effort to develop useful charts,[3-8] tables,[2,9-12] approximate[7,8,13-15] and iterative solutions,[16-20] it would seem there are no methods which are equally effective for each cubic or quartic irrespective of the values of the coefficients. The several methods, perhaps fortunately separately, may involve division by zero (or the very small difference between large numbers), an unsatisfactory algorithm for forming an iterative estimate, or a very slowly convergent procedure. It is known that such difficulties may plague the routines of stored program computers. They are often fatal with hand-held calculators.

All this suggests that a retrospective comparison of the methods may shed light on the ineluctable connections, and clarify the choice of the most efficient alternative solution strategies in each case. Along the way, it will be convenient to point to a number of approximate solutions. Some perhaps are novel. When more widely appreciated, these shall improve the insight of designers if not that of transcendental mathematicians.

Received Oct. 19, 1981. 

*Chief Engineer. Associate Fellow AIAA.

## Cubic Equations

Consider the complete, standard form cubic equation with real coefficients

$$x^3+px^2+qx+r=0 \tag{1}$$

It is solved by finding the factors of the polynomial left-hand side;

$$\begin{aligned} x^3+px^2+qx+r &= [x+(1/T)]\,[x^2+2\zeta\omega x+\omega^2] \\ &= [x+(1/T)]\,[(x+\alpha)^2+\beta^2]=0 \end{aligned} \tag{2}$$

These factors, separately equated to zero, yield the roots

$$x_1=-1/T;\quad x_{2,3}=-\zeta\omega\pm j\omega\sqrt{1-\zeta^2}=-\alpha\pm j\beta \tag{3}$$

The "fundamental theorem of algebra" says that there are three roots. These must be either real or occur in complex conjugate pairs $(-1<\zeta<1)$. The sum of the roots, $x_1+x_2+x_3=-p$. The product of the roots, $x_1x_2x_3=-r$, and the sum of the inverses $1/x_1+1/x_2+1/x_3=-q/r$. In general, if $x_1$ is a root of Eq. (1), $1/x_1=z_1$ is a root of an equation formed by reversing the order of the coefficients

$$rz^3+qz^2+pz+1=z^3+\frac{q}{r}z^2+\frac{p}{r}z+\frac{1}{r}=0 \tag{4}$$

Sometimes, as in Cardan's method, the solution of Eq. (1) is facilitated by the substitution $x=y-p/3$. This yields the "reduced" equation[21] in which the second-degree term is eliminated

$$y^3+(q-p^2/3)y+(r+2p^3/27-pq/3)=0 \tag{5}$$

Most often, however, in practical work the fact that $x_1$ is a solution of Eq. (1) is demonstrated by means of "synthetic" division.[21] This appears as follows.

$$\begin{array}{ccccc} 1 & p & q & r & \lfloor x_1 \\ & x_1 & x_1(p+x_1) & x_1[q+x_1(p+x_1)] & \\ \hline 1 & p+x_1 & q+x_1(p+x_1) & r+x_1[q+x_1(p+x_1)] & \end{array} \tag{6}$$

When the remainder, $R=r+x_1[q+x_1(p+x_1)]=0$; $x_1$ is a solution. The other two roots (real or complex conjugate) then, conveniently, are found by means of the quadratic formula from the "depressed" equation[21]

$$x^2+(p+x_1)x+[q+x_1(p+x_1)]=x^2+(p+x_1)x+r/x_1=0 \tag{7}$$

There are three widely applicable methods for (numerically) iteratively improving a real root estimate. They comprise Newton's method, linear interpolation (or extrapolation), and the process called "cross divide." Each of these has particular uses. For comparison and convenient reference they are set forth concisely in the Appendix.

On the other hand, Janke and Emde[7] have presented a parameter plane nomogram for the solution of the "normalized" cubic equation. The normalization (in which both the leading and trailing coefficients appear as ones, and the influence of three coefficients is subsumed in only two parameters) was first introduced by Vischnegradsky.[22] Representation of the properties of cubic (and higher degree) solutions on a two dimensional graph has, since then, been much elaborated.[23-26] It has an outstanding usefulness with which we are not yet done.

Figure 1a shows a simplified version of the Jahnke and Emde diagram. [Here, $b=p/r^{1/3}$, $c=q/r^{2/3}$, and $x_1=-r^{1/3}/T$ is a root of Eq. (1).] In order to avoid confusing clutter, only a few of the straight lines

$$b=Tc-T^2+(1/T) \tag{8}$$

are shown. (In problems in dynamics, $T$ is the "time constant" of the exponential mode.) Likewise, only five of the contours of constant "damping ratio parameter," $\zeta$, of a possible complex pair of roots are drawn. These five, however, are particularly significant.

Note that the drawing is geometrically symmetrical about the inclined $b=c$ axis. This is a reflection of the fact that reversing the order of the coefficients in the equation, $s^3+bs^3+cs+1=0$, makes any solution of the new equation the inverse of a solution of the original. For the most part, therefore, we can confine our attention to the region below and to the right of the axis of symmetry.

### Exact Approximations

It is well known[7,8,14] that if there is a "large" real root of Eq. (1), its approximate value is $x_1=-p$. Equally, if there is a "small" real root its approximate value is $x_1=-r/q$. But along the "stability boundary" where the critical Routh test function[27,28] $pq-r=0$, or $bc=1$; $x_2+x_3=0$, and, therefore

$$x_1=-p=-r/q;\quad x_{2,3}=\pm\sqrt{-q} \tag{9}$$

This is not an approximation; it is exact. Both estimates give the same answer. The damping ratio, $\zeta=0$. It is not necessary that the root $x_1$ be large or small. Consider the equation $s^3+s^2+s+1=0$. The root $s_1=-1$ is quite evidently neither.

Estimates[7,8] of roots intermediate between large and small include

$$x_1 \doteq -q/p;\quad \text{and}\quad x_1\doteq -r^{1/3} \tag{10}$$

Represented on the parameter plane, these statements yield the same exact answer whenever $b=p/r^{1/3}=c=q/r^{2/3}$. This is the case of the "quasi-symmetric" cubic.[2] It includes the denominator zero (pole) constellations corresponding to a triple real root (binomial filter) and the Butterworth filter[29] as subcases.

In addition, we may remark that when $b+c=-2$ (or $pr^{1/3}+q=-2r^{2/3}$); $x_1=+r^{1/3}$ is a root of Eq. (1).

### Special Cases

Neumark[2] has further enumerated three special cases in which particular relations among the coefficients allow the direct inscription of the solution. These comprise what he calls solutions by square roots and cube roots, and the case of a double root.

Where $27r+2p^3-9pq=0$, a solution of Eq. (1) is $x=-p/3$. Then, $x_{2,3}=(-p\pm\sqrt{3p^2-9q})/3$. This is the case of equal real parts if $3q>p^2$. When $p^2-3q=0$, a solution is

$$x_1=(-p+\sqrt[3]{p^3-27r})/3$$

If, in addition to $p^2-3q=0$, $p^3-27q=0$; there is a triple real root $x_{1,2,3}=-p/3$. Where $p^2-3q>0$ and the determinant,

$$\begin{vmatrix} 3 & p & q \\ p & 4q-p^2 & 3r \\ q & 3r & pr \end{vmatrix}=0$$

there is a double (real) root: $x_{2,3}=(-p\pm\sqrt{p^2-3q})/3$, and $x_1=(-p\mp 2\sqrt{p^2-3q})/3$. The choice of signs here is

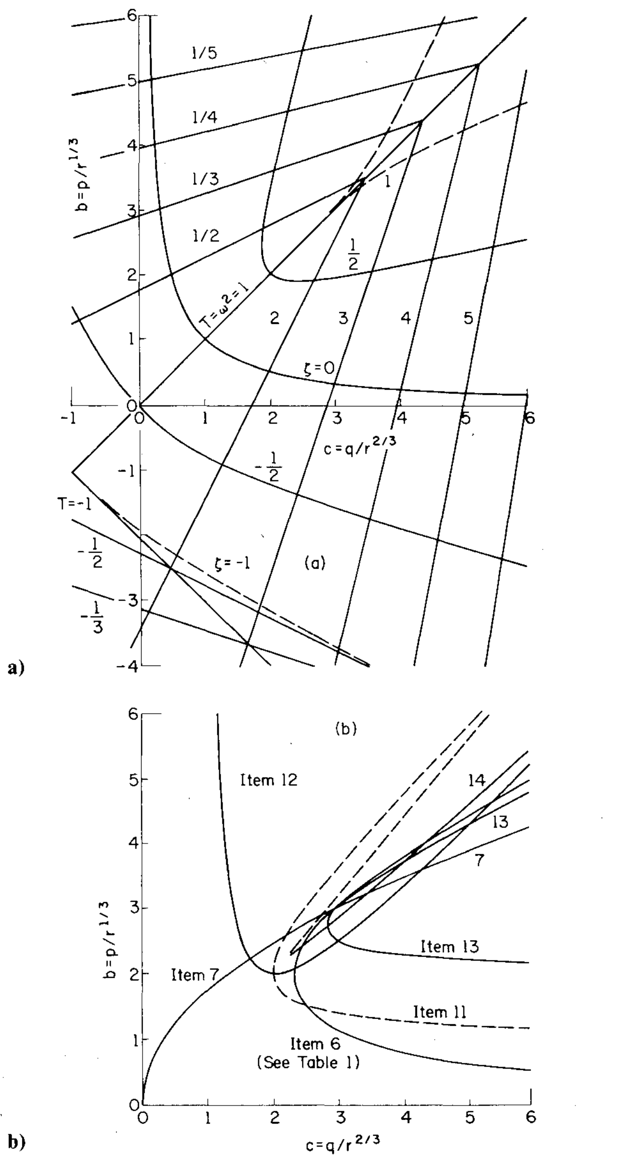


**Fig. 1 Solutions of the normalized cubic equation, $u^3+bu^2+cu+1=0$; a) Jahnke and Emde nomogram, b) conditions of validity for several exact approximate factors.**

predicted on sgn $(27r+2p^3-9pq)$. The upper one is assigned if $27r+2p^3-9pq>0$, and the lower one in the converse case. These are the explicit conditions for the damping ratio, $\zeta=\pm 1$ already illustrated in Fig. 1a.

Solutions by square roots and cube roots just now discussed are delineated on Fig. 1b. Together with the previous exact approximations, the several special cases are summarized in Table 1.

**Additional Exact Approximations**

Less often remarked[14] is the fact that the large root estimate may be "improved" by means of the expression $x_1' \doteq -p+q/p$. This can be shown to be exact whenever $b^3-b^2c^2+c^3=0$, or $p^3r-p^2q^2+q^3=0$.

Furthermore, Anderson's method[30] can be used to "improve" the estimate of a small root. This starts with the assumption that the quadratic equation $px^2+qx+r=0$ has a small root $x_1 \doteq -r/q$. If this were the case, the $x^3$ term in Eq. (1) would be comparatively negligible, and we might write $x(px+q)\doteq -r$, or $x_1 \doteq -r/q(1+px/q)$. Substituting the expression for the small root on the right-hand side

$$x_1 \doteq -\frac{r}{q}\,\frac{1}{1-pr/q^2} \tag{11}$$

**Table 1 Exact "approximate" factors for cubics**
$x^3+px^2+qx+r=0$

| Case | Factors | Conditions of validity |
|---|---|---|
| 1) Large root | $(x+p)\,[x^2+q]=0$ | $pq=r$ |
| 2) Small root | $(x+r/q)\,[x^2+q]=0$ | $pq=r$ |
| 3) Intermediate root, $-q/p$ | $\left(x+\frac{q}{p}\right)\left[x^2+\left(p-\frac{q}{p}\right)x+\frac{pr}{q}\right]=0$ | $pr^{1/3}=q$ |
| 4) Intermediate root, $-r^{1/3}$ | $(x+r^{1/3})\,[x^2+(p-r^{1/3})x+r^{2/3}]=0$ | $pr^{1/3}=q$ |
| 5) Intermediate root, $+r^{1/3}$ | $(x-r^{1/3})\,[x^2+(p+r^{1/3})x-r^{2/3}]=0$ | $pr^{1/3}+q=-2r^{1/3}$ |
| 6) Solution by square roots | $\left(x+\frac{p}{3}\right)\left[x^2+\frac{2}{3}px+\frac{3r}{p}\right]=0$ | $2p^3+27r=9pq$ |
| 7) Solution by cube roots | $(x+\nu)\left[x^2+(p-\nu)x+\frac{r}{\nu}\right]=0$ | $p^2=3q$ |
| | $\nu=(p-\sqrt{p^3-27r})/3$ | $(p,q,r>0)$ |
| 8) Double root | $(x+p-\mu)\,[p+\mu]^2=0$ | $p^2-3q>0$; and |
| | $\mu=(p\mp\sqrt{p^2-3q})/3$ | $\begin{vmatrix} 3 & p & q \\ p & 4q-p^2 & 3r \\ q & 3r & pr \end{vmatrix}=0$ |
| | here the sign is chosen as $-\operatorname{sgn}(27r+2p^3-9pq)$ | |
| 9) Improved large root | $\left(x+p-\frac{q}{p}\right)\left[x^2+\frac{q}{p}x+\frac{r}{p-q/p}\right]=0$ | $p^3r+q^3=p^2q^2$ |
| 10) Improved small root | $\left(x+\frac{1}{\tau_1}\right)\left[x^2+\left(p-\frac{1}{\tau_1}\right)x+r\tau_1\right]=0$ | $p^3r+q^3=p^2q^2$ |
| | $\tau_1=q(1-pr/q^2)/r$ | |
| 11) Alternate improved large root | $(x+p-r^{1/3})\left[x^2+r^{1/3}x+\frac{r}{p-r^{1/3}}\right]=0$ | $p^2+qr^{2/3}+2r^{4/3}$ $=2pr+pqr^{1/3}$ |
| 12) Alternate improved small root | $\left(x+\frac{1}{\tau_2}\right)\left[x^2+\left(p-\frac{1}{\tau_2}\right)x+r\tau_2\right]=0$ | $q^2+pr+2r^{4/3}$ |
| | $\lvert\tau_2=q(1-r^{2/3}/q)/r\rvert$ | $=\lvert 2r^{2/3}+pqr^{1/3}\rvert$ |
| 13) Alternate intermediate root | $(x+\rho)\left[x^2+(p-\rho)x+\frac{r}{\rho}\right]=0$ | $2p^2+2q+9r^{2/3}$ |
| | $\rho=r^{1/3}-\frac{qr^{1/3}-pr^{1/3}}{q-2pr^{1/3}-3r^{2/3}}$ | $=\frac{pq}{r^{1/3}}+8pr^{1/3}$ |
| 14) Second convergent | $\left(x+\frac{1}{\tau_3}\right)\left[x^2+\left(p-\frac{1}{\tau_3}\right)x+r\tau_3\right]=0$ | $0.85q+0.3r^{2/3}=pq^{1/3}$ (approximate) |
| | $\tau_3=\frac{q}{r}\,\frac{(q^2-2pr)}{(q^2-pr)}$ | |

(It is interesting to note that this expression is identical to the one which is the result of applying cross divide to the small root estimate.) Again, this can be shown to be the exact value of the root whenever $b^3 - b^2c^2 + c^3 = 0$ or $p^3r - p^2q^2 + q^3 = 0$. In turn, these are the explicit conditions, already illustrated, under which the damping ratio of the complex pair, $\zeta = \pm 1/2$ (Ref. 31).

Now, if $x_I = -p + q/p$ should be an improved estimate for a large root, it seems (in a way) equally likely that $x_I = -p + r^{1/3}$ should be one also. In fact, the latter alternate improved large root estimate is exact whenever $p^2 + qr^{2/3} + 2r^{4/3} = 2pr + pqr^{1/3}$. The same sort of reasoning leads to $x_I = -r/q(1 - r^{2/3}/q)$ as an alternate improved small root estimate. This one is exact whenever $q^2 + pr + 2r^{4/3} = 2r^{2/3} + pqr^{1/3} = 0$.

The several additional exact approximations involving improvements and alternate improvements have been added to Table 1; and curves representing the conditions of validity also have been drawn on the *b-c* parameter plane in Fig. 1b. There we may observe, possibly with some satisfaction, that, together with $bc = 1$ and $b = c$, they neatly section the portion of plane which is most often of the greatest interest. This implies the usefulness of these additional approximations in conducting a sectioning search via synthetic division for a root which is not one of those considered up until now.

Of course, any reasonable or unreasonable root estimate can be shown to be exact somewhere; and we might go on. As the root estimate expressions become more sophisticated, however, this procedure gets tedious; and the conditions of validity become very complex. There are diminishing returns. It may appear that we have a sufficiency for nearly every purpose except possibly in the case of nearly equal roots.

#### Nearly Equal Roots

It is evident in Fig. 1b that the conditions of validity for the expressions connected with the solutions "by square roots" and "by cube roots" (as well as the inverses of these solutions) serve to section the coefficient space in the vicinity of the conditions for three equal roots ($b = c = 3$). Since, however, the roots here are very sensitive to the coefficient values, we might wish that the sectioning curves were more dense in this vicinity.

Consideration of perturbations in the values of the parameters leads to the expression for an "alternate intermediate root" (see item 13 in Table 1). The conditions of validity are represented in Fig. 1b by an additional curve tangent to $b = c$ at $b = c = 3$. Of course, the inverse of the alternate intermediate root would be valid under circumstances in which the roles of $b = p/r^{1/3}$ and $c = q/r^{2/3}$ were interchanged.

Penultimately, in connection with factors for cubic polynomials, we may remark that the expression for an improved small root suggests that the second convergent of a continued fraction,

$$x_I \doteq \left(\frac{r}{q}\right) \frac{1}{1 - \dfrac{pr}{q^2} \dfrac{1}{1 - pr/q^2}} \tag{12}$$

might also be a useful expression.[31] Indeed it is, although the conditions of validity cannot be simply stated. It is, of course, necessary that $pr/q^2 < 1/2$. Then a graphical solution of the implied relations yields the approximate condition of validity: $0.85q + 0.3r^{2/3} - pr^{1/3} = 0$. This has been represented in Fig. 1b, and the equivalent of Eq. (12) is entered in Table 1. Again, the inverse of the expression for the root would, itself, be a root under the circumstances in which the roles of $b = p/r^{1/3}$ and $c = q/r^{2/3}$ were interchanged.

#### Special Cubic Solution Function

A recent note[32] has pointed out the advantages secured by a combination of the techniques of the reduced and normalized equations. The substitution $s = y/[r + 2p^3/27 - pq/3]^{1/3}$ in Eq. (5) yields the normalized, reduced equation

$$s^3 + ks + (1)\operatorname{sgn} h = 0 \tag{13}$$

where

$$k = (q - p^2/3)/h^{2/3} \qquad h = (2p^2 - 9q)\frac{p}{27} + r$$

The solution $s_I = G(k)$ of Eq. (13) for the isolated real root $-1/T$ is the solution of Eq. (8) with $b = 0$. In other words, it is the root along the *c* axis in Fig. 1a. This solution has remarkable overlapping asymptotic properties; and, because the root $s_I$ is always isolated, Newton's method for refining a starting estimate provided by asymptotic approximations allows one to achieve great accuracy with very little effort. There is no difficulty in the vicinity of double or triple roots. The method fails only when $h = 0$; but then $x_I = -p/3$ is known as a solution of Eq. (1). Otherwise

$$x_I = s_I h^{1/3} - p/3 \tag{14}$$

The solution of Eq. (1) with arbitrary coefficients by this method is reasonably convenient and very reliable. Asymptotic starting estimates for the solution of Eq. (13) are displayed in Table 2.

## Quartic Equations

We may take the complete, standard form, quartic equation with real coefficients to be

$$\lambda^4 + B\lambda^3 + C\lambda^2 + D\lambda + E = [\lambda^2 + l\lambda + m][\lambda^2 + L\lambda + M]$$

$$= [(\lambda + \alpha_1)^2 + \beta_1^2][(\lambda + \alpha_2)^2 + \beta_2^2] = 0 \tag{15}$$

Either or both of the quadratic factors may, on occasion, be split into two first-degree factors. The factors separately set equal to zero yield the roots of the equation. These three comprise four reals, two reals and a complex pair, or two complex pairs.

Multiplication of the indicated factors shows

$$L + l = B \tag{16}$$

$$(M + m) + Ll = C \tag{17}$$

$$Lm + Ml = D \tag{18}$$

$$Mm = E \tag{19}$$

**Table 2 Asymptotic approximations to the special cubic solution function, $G(k)$**

| Range | Approximation |
|---|---|
| $k > 1$ | $s_I \doteq -\left(\frac{1}{k}\right) \cfrac{1}{1 + \cfrac{1/k^3}{1 + \cfrac{1/k^3}{1 + \ddots}}}$ |
| | $\cong -\left(\frac{1}{k}\right)\left(\frac{1 + k^3}{2 + k^3}\right)$ |
| $-2 < k \leq 1$ | $s_I \doteq -1 + k/3$ |
| $k \leq -2$ | $s_I \doteq -\sqrt{-k} + (1/2k)$ |

The methods[2] associated with the names of Ferrari and Lagrange, Descartes and Euler, and Duncan[28] and Babister,[33] in effect, depend on the substitution of an auxiliary variable for a simple combination of the (unknown) terms in Eqs. (16-19).

Thus, the substitution, $t = M + m$, yields

$$L,l = \frac{B}{2} \pm \frac{Bt - 2D}{2\sqrt{t^2 - 4E}} \tag{20}$$

$$M,m = \frac{t}{2} \pm \frac{\sqrt{t^2 - 4E}}{2} \tag{21}$$

and the resolvent cubic

$$t^3 - Ct^2 + (BD - 4E)t - (D^2 + B^2E - 4CE) = 0 \tag{22}$$

When an appropriate $(t > 2\sqrt{E})$ real solution of Eq. (22) is found, the factors of Eq. (15) are then known. There is, however, some difficulty where $t = t_{crit} = 2\sqrt{E}$. This latter is the case in which $M = m$. (This implies equal "frequencies" in physical problems.)

Similarly, the substitution $\$ L - l$, yields

$$L,l = (B \pm \$)/2 \tag{23}$$

$$M,m = \frac{\$^3 + (4C - B^2)\$ \pm [B\$^2 + (4BC - B^3 - 8D)]}{8\$} \tag{24}$$

and the resolvent cubic in $\$^2$

$$\$^6 - (3B^2 - 8C)\$^4 + [(B^2 - 4C)(3B^2 - 4C) + 16(BD - 4E)]\$^2 - (B^3 - 4BC + 8D)^2 = 0 \tag{25}$$

Obviously, the method will fail when $\$ = \$_{crit} = 0$. This is the case of equal real parts or "damping factors."

Further, the substitution $\chi = Ll$, yields

$$L,l = \frac{B}{2} \pm \frac{\sqrt{B^2 - 4\chi}}{2} \tag{26}$$

$$M,m = \frac{C - \chi}{2} \pm \frac{B(C - \chi) - 2D}{2\sqrt{B^2 - 4\chi}} \tag{27}$$

and the resolvent cubic

$$\chi^3 - 2C\chi^2 + (C^2 + BD - 4E)\chi - (BCD - B^2E - D^2) = 0 \tag{28}$$

Note that $\chi = C - t$. This equation, suggested for the purpose by Babister,[33] has the evident advantage that the trailing term is the well known (and often physically significant) critical Routh test function.[27,28] Clearly, this method too will fail when $\chi = \chi_{crit} = B^2/4$. Again this is the case of equal real parts or damping factors.

For reasons shortly to become clear, we might wish that there were also a resolvent cubic in the auxiliary variable, $w = M - m$. This dream, however, becomes an algebraic nightmare. Nevertheless, some progress can be made by means of this substitution; and we readily can get as far as the simultaneous equations

$$w^3 + (2m)w^2 + B(D - Bm)w - (D - Bm)^2 = 0$$

$$m^2 + mw - E = 0 \tag{29}$$

Now, however, when $w$ is very small: $w \doteq (D/B - m)$, and $m \doteq BE/D$. (When $w_{crit} \equiv 0$, $M = m\sqrt{E}$. The method fails. This once more is the case of equal frequencies.) In any event,

$$L,l = \frac{D - B\left[\sqrt{E + \left(\frac{w}{2}\right)^2} \pm \frac{w}{2}\right]}{w} \tag{30}$$

$$M,m = \sqrt{E + \left(\frac{w}{2}\right)^2} \pm \frac{w}{2} \tag{31}$$

On the other hand, if $w = M - m$ should not be small, then Bairstow's[13] starting assumption appears justified. Such matters are taken up again further on.

### Exact and Approximate Factors for Quartics

Neumark[2] has discussed a number of special cases for quartics in which the factor coefficients may be determined directly. Now logically, to these we may make a minor addition. Excluding several somewhat precious cases of double, triple, and quadruple roots (see Ref. 2); the most interesting special cases are summarized in Fig. 2. Note that in the first case (perfect square) every method alluded to so far will fail and should not be attempted. In other cases which may approximate the perfect square one or another or several of the resolvent cubic methods as well as a number of popular iterative methods likely will show signs of distress.

Somewhat to the contrary, however, when the actual factors are very nearly the critical ones involving equal damping factors, equal frequencies, or both, these are precisely the circumstances in which the small root approximation for one or another of the resolvent cubics will

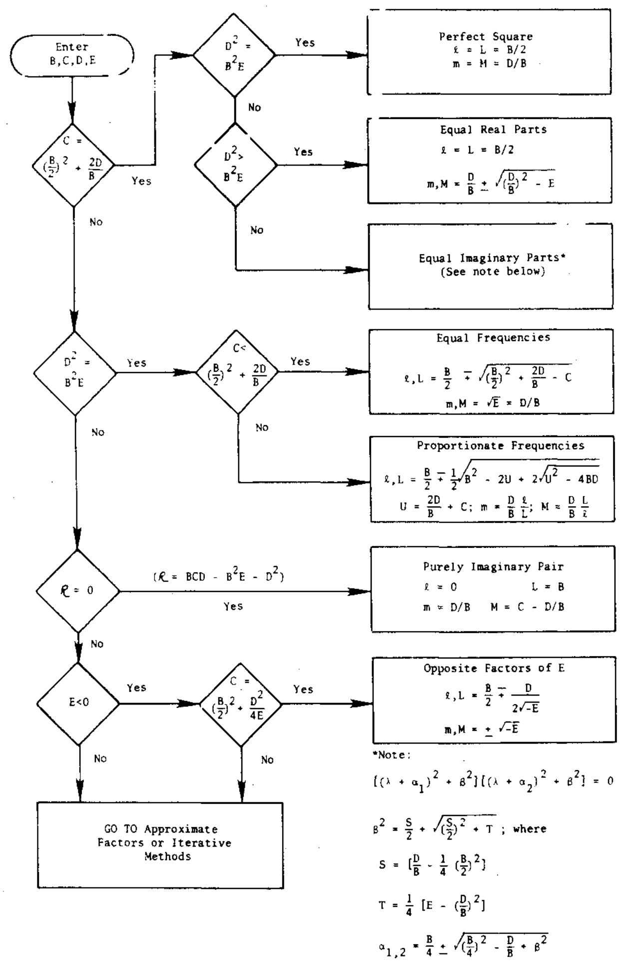


**Fig. 2 Exact factors for the quartic equation, $\lambda^4 + B\lambda^3 + C\lambda^2 + D\lambda + E = (\lambda^2 + l\lambda + m)(\lambda^2 + L\lambda + M) = 0$**

**Table 3 Approximate factors for quartics**

| | Case | Factor coefficients | Remarks[a] |
|---|---|---|---|
| 1a) | $l=\frac{(CD-BE)}{C^2}$<br>$m\doteq\frac{E}{C}$ | $L=B-\frac{(CD-BE)}{C^2}$<br>$M=C-\frac{E}{C}-\frac{BD}{C}+\frac{B^2E+D^2}{C^2}-\left(\frac{B^2E}{C^2}\right)^2$ | Small complex pair (airplane longitudinal stability quartic) |
| 1b) | $l\doteq\frac{D}{C}$<br>$m=\frac{\left[\frac{BD^2}{C}-\left(\frac{D}{C}\right)^3\right]}{(B-2D/C)}$ | $L=B-\frac{D}{C}$<br>$M=\frac{E(B-2D/C)}{\left[\frac{BD^2}{C}-\left(\frac{D}{C}\right)^3\right]}$ | Small complex pair (airplane longitudinal stability quartic) |
| 2a) | $l\doteq\frac{C}{B}$<br>$m=\frac{D-(C/B)^3}{B-2C/B}$ | $L=B-\frac{C}{B}$<br>$M=\frac{E(B-2C/B)}{D-(C/B)^3}$ | Large and small real roots (airplane lateral stability quartic) |
| 2b) | $l\doteq(CD-BE)/BD$<br>$m=\frac{D-l[C-l(B-l)]}{B-2l}$ | $L=B-(CD-BE)/BD$<br>$M=\frac{E(B-2l)}{D-l[C-l(B-l)]}$ | Large and small real roots (airplane lateral stability quartic) |
| 2c) | $l\doteq 0$<br>$m=D/B$ | $L=B$<br>$M\doteq EB/D$ | One pair nearly undamped (lateral stability quartic) |
| 3) | $l\doteq B/4$<br>$m=\frac{2D}{B}-\frac{C}{2}+\frac{3}{8}\left(\frac{B}{2}\right)^2$ | $L=3B/4$<br>$M=\frac{E}{\frac{2D}{B}-\frac{C}{2}+\frac{3}{8}\left(\frac{B}{2}\right)^2}$ | One damping factor three times the other (other ratios are easily worked out) |
| 4) | $l=\frac{[D-B\sqrt{E}/2]}{3\sqrt{E}/2}$<br>$m\doteq\sqrt{E}/2$ | $L=B-\frac{[D-B\sqrt{E}/2]}{3\sqrt{E}/2}$<br>$M=2\sqrt{E}$ | One frequency twice the other (other ratios are easily worked out) |
| 5) | $l,L=\frac{1}{2}\left[B\mp\frac{Bt-2D}{\sqrt{t^2-4E}}\right]$ | $m,M=\frac{1}{2}[t\mp\sqrt{t^2-4E}]$<br>where $t\doteq C-(BD-4E)/C>2\sqrt{E}$ | Both frequencies moderate |
| 6) | $l,L=\frac{1}{2}[B\mp\mathcal{S}]$ | $m=\frac{1}{\mathcal{S}}\left[D-BC/2+C\mathcal{S}/2+\frac{1}{8}(B^3-B^2\mathcal{S}-B\mathcal{S}^2+\mathcal{S}^3)\right]$<br>$M=E/m$ | Damping factors nearly equal |
| | where $\mathcal{S}\doteq\frac{(B^3-4BC+8D)}{\sqrt{(B^2-4C)(3B^2-4C)+16(BD-4E)}}\neq 0$ | | |
| 7) | $l,L=\frac{1}{2}[B\mp\sqrt{B^2-4\chi}]$ | $m,M=\frac{1}{2}\left[C-\chi\mp\frac{B(C-\chi)-2D}{\sqrt{B^2-4\chi}}\right]$ | One damping factor only moderate |
| | where $\chi\doteq\frac{BCD-B^2E-D^2}{C^2+BD-4E}<\frac{B^2}{4}$ | | |

[a] In principle, the conditions of validity could be expressed algebraically. These, for the most part, however, are very complex and unedifying. Instead, the basic approximation in each set of factor coefficients is shown by the symbol $\doteq$; and the remarks indicate typical features of the zero constellations. For traditional reasons, item 2c is repeated from Fig. 2.

give good results. The effect is that we are equipped with an armament of useful, approximate factors. Together with previously well known ones, these are summarized in Table 3. One may remark that since the approximate solutions to the resolvent cubics are often (under appropriate circumstances) exact (recall Table 1), the number of exact factors of quartics which may, in principal, be "known" is very large indeed. Only the most readily apparent ones are displayed here.

### Iterative Procedures for Quartics

In connection with numerical problems, one might choose to solve one or another of the resolvent cubics; or one might prefer to proceed with an iterative solution of the quartic itself. Exercising the second option is likely to give quick, accurate answers only where a great deal is known about the solution to start with. For practical work, however, this is often the case.

Although there are many others[20,34-36] with merit, a discriminatory choice is made here of three elementary methods which (experience shows) are appropriately likely to make short work of the numerical solution of quartic equations. These are the methods of Bairstow[13] and Lin,[16] of Graham,[17] and of Porter and Mack.[18]

In effect, the method of Bairstow and Lin involves the simultaneous particular initial choice of $L$ and $M$ and the use of Eqs. (16-19) to compute new values of $l$ and $m$ which lead, in turn, to a revision of the values of $L$ and $M$, and so forth. The starting estimate (corresponding to $m=l=0$) is $L=B$, $M=C$. The procedure is appropriately represented in Table 4. On each line the table is filled out from left to right with calculated numbers. The unprimed symbols refer to values from the previous line, while the primed symbols represent numerical values from the same (current) trial (or line). When $m \ll M$ the procedure converges rapidly, provided no errors are make in the calculation. On the other hand, if $m=\mathrm{O}(M)$, the calculations are extraordinarily tedious and one is tempted to jump ahead. This will not do. The process must be carried out exactly as prescribed until the test function in the last column becomes acceptably small. Limited experience makes it appear that $|E|<(C/5)^2$ is an indication of easy success.

A valuable alternative,[17,18] particularly when (as is often thp case) all the roots have real parts with the same sign, is to start with an estimate of one damping factor, say $l_i$. Then, of course, the estimates of the other factor coefficients can be calculated from Eqs. (16-19) according to Table 5. When the test shows zero, the estimate is exact.

Unfortunately, the only realistic method of formulating a next estimate of the factor coefficient is by linear interpolation (or extrapolation). This must be based on two previous values and on the value of the test function in the last column. Not only that, but also the test is a strongly nonlinear function of the estimate. This signifies that (for good results) we actually require two estimates close to the actual value.

Indeed, however, in the case most often of concern all the roots of the quartic have negative real parts. (All the coefficients have the same sign and $\mathcal{R}=BCD-B^2E-D^2>0$.) In that event, excluding the case of equal real parts, $0<l<B/2$, and the test function will change sign in this range. A trial, $l=B/4$, may yield a test function which is negative. Then $0<l<B/4$, and an appropriate new estimate is

$$l=\frac{1}{2}\left[B-\sqrt{B^2-4\chi}\right]$$

where $\chi \doteq \mathcal{R}/(C^2+BD-4E)<B^2/4$ (see item 7, Table 3). Contrarily, if the trial, $l=B/4$, yields a test function which is positive, then $B/4<l<B/2$, and an appropriate new estimate is $l=\frac{1}{2}B-\$$, where

$$\$ \doteq (B^3-4BC+8D)/\sqrt{(B^2-4C)(3B^2-4C)+16(BD-4E)}>0$$

(see item 6, Table 3). In either event we have two estimates which are (more or less) "close" to exact, and can proceed with linear interpolation, which may need to be repeated.

**Table 4 Method of Bairstow and Lin**

| Trial | $m'=E/M$ | $l'=(D-Lm')/M$ | $L'=B-l'$ | $M'=C-m'-L'l'$ | Test $M-M'$ |
|---|---|---|---|---|---|
| 0 | 0 | 0 | $B$ | $C$ | – |
| 1 | $E/C$ | $\frac{D}{C}-\frac{BE}{C^2}$ | $B-\frac{D}{C}+\frac{BE}{C^2}$ | $C-\frac{E}{C}\frac{B^2E+D^2}{C^2}$ | $+\frac{E}{C}-\frac{B^2E+D^2}{C^2}$ |
| | | | | $-\frac{BD}{C}-\left(\frac{B^2E}{C^2}\right)^2$ | $+\frac{BD}{C}+\left(\frac{B^2E}{C^2}\right)^2$ |
| 2 | ... | ... | ... | ... | ... |

**Table 5 Method of Graham and method A of Porter and Mack**

| ① Estimate | ② $L_i$ | ③ $L_i-l_i$ | ④ $D-l_i(C-l_iL_i)$ | ⑤ $m_i$ | ⑥ $M_i$ | ⑦ Test |
|---|---|---|---|---|---|---|
| $l_i$ | $B-$① | $B-2\times$① | $D-$①$\times(C-$①$\times$②$)$ | ④/③ | $C-$⑤$-($①$\times$②$)$ | $($⑤$\times$⑥$)-E$ |

**Table 6 Method B of Porter and Mack**

| ① Estimate | ② $M_i$ | ③ $l_i$ | ④ $L_i$ | ⑤ Test, $g$ |
|---|---|---|---|---|
| $m_i$ | $E/$① | $(B\times$①$-D)/($①$-$②$)$ | $B-$③ | ①$+$②$+($③$\times$④$)-C$ |

Choosing an initial frequency squared estimate is a complementary procedure.[18] The solution method according to Eqs. (16-19) then appears in Table 6.

Again, however, we must rely on linear interpolation (or extrapolation) for an iterative estimate, and the test, once more, is nonlinear. In fact, the behavior of the test function often seems worse than in the previous case, but we have a variety of trial values, which we may use in a systematic way.

If $E \equiv 0$, one of the roots of the quartic is zero, and we only have to do with a reduced equation which is a cubic. If $E>0$, excluding the case of equal frequencies, there is a positive value such that, $0<m<\sqrt{E}$. When $E<0$; $0<m<+\sqrt{-E}$ if $H=(B/2)^2+D^2/4E-C>0$; otherwise, $-\sqrt{-E}<m<0$, when $H<0$. (Recall from Fig. 2, that, if $H \equiv 0$, $m,M=\pm\sqrt{-E}$.) In either event, a trial with a suitably signed magnitude, $|m|=\sqrt{-E/2}$ will section the range. Then, if the difference, $w=M-m$, should be small in absolute magnitude, $m=D/B$ is a suitable choice (item 2c, Table 3); while if the difference should be large, $m=E/C$ is an obvious possibility (see item 1a, Table 3). Furthermore, $m=[t-\sqrt{t^2-4E}]/2$, where $t \doteq C-(BD-4E)/C>2\sqrt{|E|}$ is then also appropriate (see item 5, Table 3). Between the four choices, it cannot fail that two of them (possibly presciently chosen first) will be reasonably close to an exact value so that one may then confidently proceed with interpolation (or extrapolation). This, of course, may have to be repeated several times (using pairs of trials which yield small absolute values of the test function $g$).

Because the initial trial values themselves are readily calculated, and because filling in Table 6 involves a minimum amount of calculation; because the first several trials need not be made with precision, and the procedure allows for recovery from mistakes; this method B of Porter and Mack's is to be recommended over others. In my opinion, it has not received the attention and appreciation it deserves.

## Suggestions for Quintics

Although the methods which have been discussed to this point can be (and have been[16,18]) extended to quintics, sextics, and equations of even higher degree, a comparative account of such matters is beyond the scope attempted here. Nevertheless, it is conventionally perceived that one way to treat a quintic is to, in effect, "divide out" a real root via synthetic division. (There must be at least one.) Then, the depressed equation is a quartic whose competent solution has been the subject of our attention.

By analogy to the techniques employed for cubics we can readily formulate real root estimates which section the possibilities for quintics. For example, consideration of the three leading terms yields large root and improved large root estimates; while consideration of only the three trailing terms (in standard form) yields small root and improved small root estimates. Further, intermediate root estimates comprise both the fifth root of the constant term and the nonzero result of equating only the (standard form) middle two terms to zero. In every circumstance, such estimates may be improved by one or another of the methods presented in the Appendix. By these means, quintics readily are soluble in numerical terms. For reasons far better explained elsewhere,[37,38] it is seldom necessary for an engineer to give detailed consideration to polynomial equations of more than the fifth degree.

## Conclusion

A comparison of methods for the solution of cubic and quartic polynomial equations shows that no single method is equally effective and efficient for either type. On the other hand, research in a scattered literature, together with some novel algebraic results, has revealed a large number of exact or asymptotic approximations appropriate to improved starting estimates for iterative calculations; and (in literal terms) for facilitating the understanding of designers. Extension of the considerations appropriate to cubic equations makes the numerical solution of quintic equations convenient.

## Appendix: Iteration for Real Roots

There are three more or less generally applicable methods for the iterative improvement of real root estimates in connection with the numerical solution of polynomial equations. These are Newton's method, linear interpolation (or extrapolation), and the "cross-divide" process. They are shown here for a cubic equation. Extension to equations of higher degree is obvious.

### Newton's Method

Synthetic division with a root estimate $x_j$ of the polynomial equation, $x^3+px^2+qx+r=0$, appears as follows.

$$\begin{array}{ccccl} 1 & p & q & r & \lfloor x_j \\ & x_j & x_j(p+x_j) & x_j[q+x_j(p+x_j)] & \\ \hline 1 & p+x_j & q+x_j(p+x_j) & x_j^3+px_j^2+qx_j+r & \\ & & & =R(x_j) & \end{array}$$

By Newton's method, the next root estimate is $x_{j+1}=x_j-R(x_j)/[\mathrm{d}R(x_j)/\mathrm{d}x_j]$. The remainder $R(x_j)$ is already available. Its numerical derivative is readily obtained by "continuing" the synthetic division with the first three terms in the bottom line. Thus,

$$\begin{array}{ccc} 1 & p+x_j & q+x_j(p+x_j) \\ & x_j & x_j(p+2x_j) \\ \hline 1 & p+2x_j & q+p(2x_j)+3x_j^2 \\ & & =\dfrac{\mathrm{d}R(x_j)}{\mathrm{d}x_j} \end{array}$$

### Linear Interpolation (or Extrapolation)

If there should be a function $F(a)$, such as the remainder on synthetic division, or a test function as described in the text, and this function has been evaluated for two values of the independent variable, say $a_1$ and $a_2$, an estimate of the value $a_3$ which will make the function zero is

$$a_3=a_1+F_1(a_2-a_1)/(F_1-F_2)$$

### Cross Divide

When a real root of a polynomial equation is small in magnitude and negative, a simple improved estimate can be formed from the preceding results of synthetic division as $x_{j+1}=-r/[q+x_j(p+x_j)]$. Here $q+x_j(p+x_j)$ is the term under the solidus in the third column. Repeated trials will converge to the exact result if the root is indeed small and negative. The Lipschitz condition may be used to show precisely when this is expected to be the case, but the general expression is so complicated as to be unedifying.

## References

[1] Struik, D.J., *A Concise History of Mathematics,* Dover, New York, 1948.

[2] Neumark, J., *Solution of Cubic and Quartic Equations,* Pergamon Press, Oxford, 1965, pp. 1-30, 37, 38.

[3] Zimmerman, C.H., "An Analysis of Lateral Stability in Power-Off Flight with Charts for Use in Design," NACA TR-589, 1937.

[4] Liu, Y.J., "Servomechanisms, Charts for Finding Their Stability and for Finding the Roots of Their Third and Fourth Degree Characteristic Equations," Massachusetts Institute of Technology, Cambridge, Mass., Oct. 1941.

[5] Kochenburger, R.J., unpublished, date uncertain; see Smith, E.S., *Automatic Control Engineering,* McGraw-Hill, New York, 1944, pp. 244-246.

[6]Trimmer, J.D., *Response of Physical Systems,* John Wiley & Sons, New York, 1950, pp. 112, 113.

[7]Jahnke, E. and Emde, F., *Tables of Functions,* 4th ed., Dover, New York, 1945, App., pp. 20-30.

[8]Ashkenas, I.L. and McRuer, D.T., "Approximate Airframe Transfer Functions and Application to Single Sensor Control Systems," Wright-Patterson Air Force Base, Ohio, WADC TR 58-82, June 1958, pp. 191-210.

[9]Lyon, W.V., "Note on a Method of Evaluating the Complex Roots of Quartic Equations," *Journal of Math and Physics,* Vol. 3, April 1924, pp. 188-190.

[10]Nogrady, H.A., *A New Method for the Solution of Cubic Equations,* Edwards Brothers, Ann Arbor, Mich., 1936.

[11]Zavrotsky, A., *Tablas para la Resolución de las Ecuaciones Cúbicas,* Editorial Standard, Caracas, 1945.

[12]Salzer, H.E., Richard, C.H., and Abraham, I., *Table for the Solution of Cubic Equations,* McGraw-Hill, New York, 1958.

[13]Bairstow, L., *Applied Aerodynamics,* Longmans Green, London, 2nd ed., 1939.

[14]Jones, B.M., "Dynamics of the Airplane," *Aerodynamic Theory,* Vol. V, edited by W.F. Durand, Durand Reprinting Committee, California Institute of Technology, Pasadena, Calif., 1943, pp. 169, 170, 199-201.

[15]McRuer, D.T., "Unified Analysis of Linear Feedback Systems," Wright-Patterson Air Force Base, Ohio, ASD-TR-61-118, July 1961.

[16]Lin, S-N., "A Method of Successive Approximations of Evaluating the Real and Complex Roots of Cubic and Higher Order Equations," *Journal of Math and Physics,* Vol. 20, No. 3, Aug. 1941, pp. 231-242.

[17]Graham, D., "Automatic Control and the Lateral Oscillatory Motion," Cornell Aeronautical Lab., Buffalo, N.Y., Rept. TB-574-F-1, April 1, 1949, pp. 73-75.

[18]Porter, A. and Mack, C., "New Methods for the Solution of Algebraic Equations," *The Philosophical Journal,* Ser. 7, Vol. 40, May 1949, pp. 578-585.

[19]Luke, Y.L. and Ufford, D., "On the Roots of Algebraic Equations," *Journal of Math and Physics,* Vol. 30, No. 2, July 1951, pp. 94-101.

[20]Siljak, D.D., *Nonlinear Systems,* John Wiley and Sons, New York, 1969, pp. 9-15, 421-423.

[21]Dickson, L.E., *New First Course in the Theory of Equations,* John Wiley & Sons, New York, 1939, pp. 42-54.

[22]Wischnegradski, I.A., "Sur la Théorie Générale des Régulateurs," *Comptes Rendus de L'Acadèmie des Sciences,* Vol. 85, Gautiers-Villars, Paris, 1877, pp. 318-321.

[23]Bryant, L.W. and Gandy, R.W., "An Investigation of Lateral Stability of Aeroplanes with Rudder Free," National Physical Lab., Great Britain, Stability and Control, 4304, 1939.

[24]Jones, R.T. and Cohen, D., "An Analysis of the Stability of an Aircraft with Free Controls," NACA TR 709, 1941.

[25]Brown, W.S., "A Simple Method of Constructing Stability Diagrams," Aeronautical Research Council, Great Britain, R&M 1905, 1942.

[26]Whatham, J. and Lyon, H.M., "A Theoretical Investigation of Dynamic Stability with Free Elevators," Aeronautical Research Council, Great Britain, R&M 1980, 1943.

[27]Routh, E.J., *Stability of a Given State of Motion,* Macmillan, London, 1877.

[28]Frazer, R.A. and Duncan, W.J., "On the Criteria for the Stability of Small Motions," *Proceedings of the Royal Society,* London, Ser. A., Vol. 124, 1929, p. 642.

[29]Graham, D. and Lathrop, R.C., "The Synthesis of 'Optimum' Transient Response: Criteria and Standard Forms," *Transactions of the AIEE,* Vol. 72, Nov. 1953, pp. 273-288.

[30]Turnbull, H.W., *Theory of Equations,* 5th ed., Oliver and Boyd, Edinburgh, 1952, p. 156.

[31]Graham, D., "Exact 'Approximate' Solutions of the Cubic Equation," Memorandum, Princeton Univ., Princeton, N.J., July 7, 1979.

[32]Graham, D., "Special Cubic Solution Function," *Journal of Guidance and Control,* Vol. 4, Jan.-Feb. 1981, pp. 90-92.

[33]Babister, A.W., *Aircraft Stability and Control,* Pergamon Press, Oxford, 1961, pp. 672-678.

[34]Crouse, J.E. and Putt, C.W., "Roots of Polynomials by Ratio of Successive Derivatives," NASA TN D-6793, May 1972.

[35]Branin, F.H. Jr., "Poles and Zeroes, Eigenvalues of Matrices, and Roots of Polynomials by the Method of Signatures," *Proceedings of the 1977 IEEE Intenational Symposium on Circuits and Systems,* Phoenix, Ariz., April 1977, pp. 89-94.

[36]Hitchcock, F.L., "Finding Complex Roots of Algebraic Equations," *Journal of Math and Physics,* Vol. 17, No. 2, June 1938, pp. 55-58.

[37]Chen, K., "A Quick Method for Estimating Closed-Loop Poles of Control Systems," *Applications and Industry,* No. 30, AIEE, New York, May 1957, pp. 80-87.

[38]McRuer, D., Ashkenas, I., and Graham, D., *Aircraft Dynamics and Automatic Control,* Princeton Univ. Press, Princeton, N.J., 1973, pp. 155-160.