Circulant Matrices and Polynomials

Dave Frank

December 13, 2002

Abstract

The goal of this paper is to provide a simple, unified approach to the solutions of quadratic, cubic, and quartic polynomials.

1. Introduction

Solution methods for solving polynomial equations have evolved for over 4000 years. As early as 2000 B.C., the Babylonians could solve systems of equations that yielded quadratic expressions. Euclid provided a geometric solution of the quadratic with positive coefficients. Later algebraic developments by sixteenth century mathematicians such as Scipione del Ferro, Niccolo Tartagla, and Lodovico Ferrari provided the solutions to the general cubic and quadratic polynomials. However, these solution methods are arduous, and require a great deal of complex induction.

However, this paper intends to provide an alternative method to the solutions of polynomials using circulant matrices. We begin by exploring the basic structure of circulant matrices and defining them mathematically. Then we provide a crucial connection between the eigenvalues of a circulant and the *nth* roots of unity. Finally, we will use the facts we develop to solve polynomial equations.

2. Circulant Matrices

We begin by defining a circulant matrix and demonstrating how they are formed. Ultimately, we will demonstrate that computing the eigenvalues of a circulant is actually quite trivial.

An $n \times n$ circulant matrix is formed by taking a vector with n components and making it the first row of a matrix C. Now, subsequent rows contain the entries of the previous row, shifted to the right. Below are the general 3×3 and 4×4 circulants that expose the basic structure of circulant matrices.

$$C = \begin{pmatrix} a & b & c \\ c & a & b \\ b & c & a \end{pmatrix} \qquad C = \begin{pmatrix} a & b & c & d \\ d & a & b & c \\ c & d & a & b \\ b & c & d & a \end{pmatrix}$$

Notice that the same entry runs along each diagonal.

2.1. Generating a Circulant

Now we intend to show how to construct a circulant mathematically. To do so, we define the generator matrix W. Of course, the dimension of a circulant determines the dimension of it's corresponding generator W. However, for our purpose, we are only concerned with circulants of dimension 2, 3, or 4. But in each case, W is defined to be a special permutation matrix. Therefore, every exponential of W is also a permutation matrix. For 2×2 circulants, the corresponding generator matrix W is defined as

$$W = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad \text{and} \qquad W^2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Notice that W^2 is the identity matrix.

For the 3×3 case, we have

$$W = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} \quad , \quad W^2 = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \quad \text{and} \quad W^3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

In this case, W^3 is the identity.

Now, in order to generate a circulant C, we define any polynomial q, whose degree is one less than the dimension of C, and evaluate $q(W)$. For example, to generate a 3×3 circulant, let

$$q(t) = a + bt + ct^2$$

Now we evaluate

$$\begin{aligned} C &= q(W) \\ &= aI + bW + cW^2 \\ &= a \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + b \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} + c \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \\ &= \begin{pmatrix} a & 0 & 0 \\ 0 & a & 0 \\ 0 & 0 & a \end{pmatrix} + \begin{pmatrix} 0 & b & 0 \\ 0 & 0 & b \\ b & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & c \\ c & 0 & 0 \\ 0 & c & 0 \end{pmatrix} \\ &= \begin{pmatrix} a & b & c \\ c & a & b \\ b & c & a \end{pmatrix} \end{aligned}$$

An important point is that the coefficients of q are now the first row of C. But, notice that q is written in ascending order. So, suppose now we want to generate a 4×4 circulant with $[a, b, c, d]$ for its first row. So we evaluate

$$q(t) = a + bt + ct^2 + dt^3$$

at W. But of course, the generating matrices are now 4×4

$$W = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix}, \ W^2 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}, \ W^3 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

And now W^4 is the identity. So we have

$$\begin{aligned} C &= q(W) \\ &= aI + bW + cW^2 + dW^3 \\ &= \begin{pmatrix} a & b & c & d \\ d & a & b & c \\ c & d & a & b \\ b & c & d & a \end{pmatrix} \end{aligned}$$

So we now have a mathematically consistent way of defining a circulant matrix in terms of W. However, "what concerns us most about circulant matrices is the simple computation of their eigenvalues using nth roots of unity" (Kalman 821). That is, provided we are familiar with the nth roots of unity.

2.2. The nth Roots of Unity

When we say "the nth roots of unity", we mean that we want the solutions to

$$z^n = 1.$$

And we know that there must be exactly n solutions. But, again, for our purpose, we are only concerned with the square, cube, and fourth roots of unity.

To find the square roots of unity, we solve $z^2 = 1$, and we know that the two roots are ± 1. However, for exponents greater than 2, there will be complex solutions. And recall

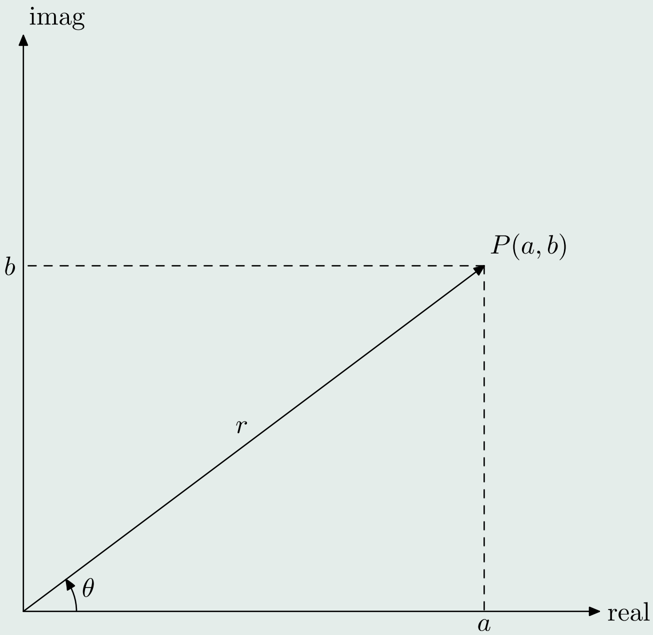

Figure 1: The Complex Plane

that any complex number $z = a + bi$ can be represented as a vector on the complex plane with coordinates (a, b). This is demonstrated in Figure 1.

When we want the cube roots of unity, we solve $z^3 = 1$, and letting $z = re^{i\theta}$

$$(re^{i\theta})^3 = e^{i2k\pi}$$
$$r^3 e^{i3\theta} = e^{i2k\pi}$$

Thus,

$$r^3 = 1 \qquad 3\theta = 2k\pi$$
$$r = 1 \qquad \theta = \frac{2k\pi}{3}$$

Substituting in $z = re^{i\theta}$,

$$z = e^{i2k\pi/3}$$

Now our three roots z_0, z_1, and z_2 are given by $k = 0$, 1, and 2.

$$z_0 = e^0 = 1$$
$$z_1 = e^{i2\pi/3} = -\frac{1}{2} + i\frac{\sqrt{3}}{2}$$
$$z_2 = e^{i4\pi/3} = -\frac{1}{2} - i\frac{\sqrt{3}}{2}$$

Notice that this is the complete solution to $z^3 = 1$, since letting $k = 3$ gives $z = z_0$, and we begin a second cycle through our given roots.

When we plot our three roots on the complex plane, we find that they all lie on the unit circle (since $r = 1$), as shown in Figure 2. Notice that each vector splits the circle into three equal parts. Now we must make one last observation. If we let $\omega = z_1$, then $\omega^2 = z_2$, and ω^3 cycles back to unity. This allows us to have a convenient method for representing the cube roots of unity: 1, ω, and ω^2.

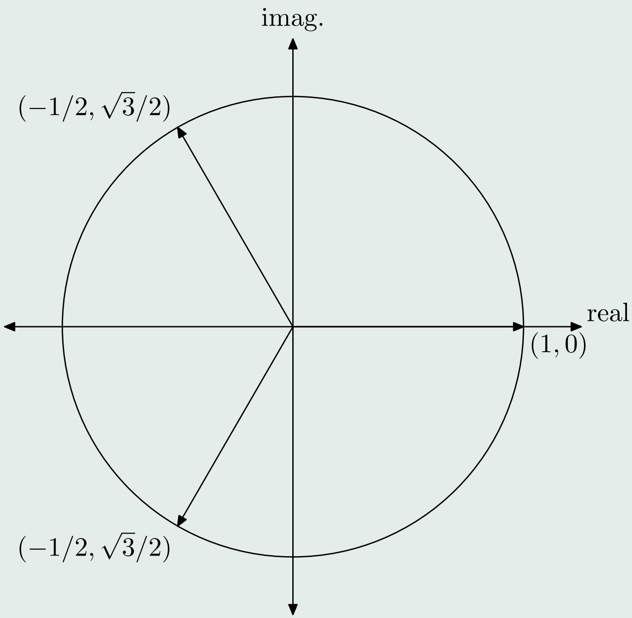

Figure 2: The cube roots of unity

We are also interested in the fourth roots of unity, so we solve $z^4 = 1$ in the same manner as before.

$$(re^{i\theta})^4 = e^{i2k\pi}$$
$$r^4 e^{i4\theta} = e^{i2k\pi}$$

Thus,

$$r^4 = 1 \qquad 4\theta = 2k\pi$$
$$r = 1 \qquad \theta = \frac{k\pi}{2}$$

Substituting in $z = re^{i\theta}$,

$$z = e^{ik\pi/2}$$

Our four solutions are then given by $k = 0$, 1, 2, and 3.

$$z_0 = e^0 = 1$$
$$z_1 = e^{i\pi/2} = i = \omega$$
$$z_2 = e^{i\pi} = -1 = \omega^2$$
$$z_3 = e^{i3\pi/2} = -i = \omega^3$$

The solution vectors again all lie on the unit circle in the complex plane as shown in Figure 3. And in this case, ω^4 cycles back to unity. We now conclude our analysis of the *nth* roots of unity and proceed with the more interesting aspects of our discussion.

2.3. Eigenvalues of a Circulant

We know if you have a matrix A with eigenvalues λ

$$A\mathbf{x} = \lambda\mathbf{x}$$

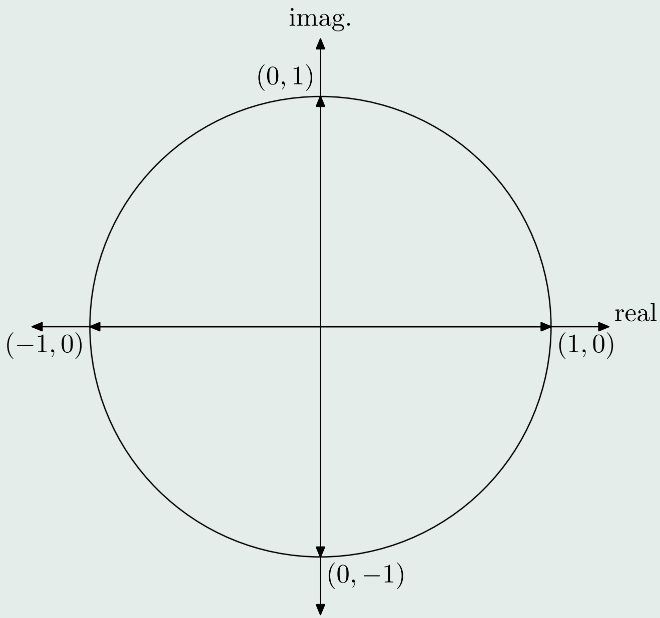

Figure 3: The fourth roots of unity

But the interesting implication is that

$$q(A)\mathbf{x} = q(\lambda)\mathbf{x}$$

for any polynomial q. That is, the number $q(\lambda)$ must be an eigenvalue of the matrix $q(A)$. A simple 2×2 example should demonstrate this. For simplicity, we let A be an upper triangular matrix.

$$A = \begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix}$$

The eigenvalues of A are $\lambda_1 = 1$ and $\lambda_2 = 3$. Now, let

$$q(t) = 1 + 2t + t^2.$$

Now we evaluate

$$\begin{aligned} q(A) &= I + 2A + A^2 \\ &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + 2\begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix} + \begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix}^2 \\ &= \begin{pmatrix} 4 & 14 \\ 0 & 16 \end{pmatrix} \end{aligned}$$

Note that the eigenvalues are $\lambda = 4$ and $\lambda = 16$. Now, in order to confirm our premise, we evaluate $q(1)$ and $q(3)$.

$$\begin{aligned} q(1) &= 1 + 2(1) + (1)^2 = 4 \\ q(3) &= 1 + 2(3) + (3)^2 = 16 \end{aligned}$$

For clarification, we now summarize our results:

- The eigenvalues of A are $\lambda = 1$ and $\lambda = 3$.

- The eigenvalues of $q(A)$, by inspection, are $\lambda = 4$ and $\lambda = 16$.
- However, the eigenvalues of $q(A)$ are also given by $q(1)$ and $q(3)$.

But, again, we are most concerned with the eigenvalues of a given circulant $C = q(W)$. But we now know that if λ is an eigenvalue of W, then $q(\lambda)$ is an eigenvalue of $q(W)$, for any polynomial q. But, we need to find the eigenvalues of W. So, first, let us solve for λ when W is the 3×3 generator matrix.

$$\begin{aligned}
\det(W - \lambda I) &= 0 \\
\left| \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} - \begin{pmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{pmatrix} \right| &= 0 \\
\begin{vmatrix} -\lambda & 1 & 0 \\ 0 & -\lambda & 1 \\ 1 & 0 & -\lambda \end{vmatrix} &= 0 \\
\begin{vmatrix} -\lambda & 1 & 0 \\ 0 & -\lambda & 1 \\ 0 & 0 & -\lambda + \frac{1}{\lambda^2} \end{vmatrix} &= 0 \\
\lambda^2(-\lambda + \frac{1}{\lambda^2}) &= 0 \\
\lambda^3 &= 1
\end{aligned}$$

Therefore, the eigenvalues are precisely the cube roots of unity. And we make a similar analysis to find the eigenvalues of the 4×4 generator W.

$$\det(W - \lambda I) = 0$$

$$\left| \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix} - \begin{pmatrix} \lambda & 0 & 0 & 0 \\ 0 & \lambda & 0 & 0 \\ 0 & 0 & \lambda & 0 \\ 0 & 0 & 0 & \lambda \end{pmatrix} \right| = 0$$

$$\begin{vmatrix} -\lambda & 1 & 0 & 0 \\ 0 & -\lambda & 1 & 0 \\ 0 & 0 & -\lambda & 1 \\ 1 & 0 & 0 & -\lambda \end{vmatrix} = 0$$

$$-\lambda \begin{vmatrix} -\lambda & 1 & 0 \\ 0 & -\lambda & 1 \\ 0 & 0 & -\lambda \end{vmatrix} + \begin{vmatrix} 0 & 1 & 0 \\ 0 & -\lambda & 1 \\ 1 & 0 & -\lambda \end{vmatrix} = 0$$

$$\lambda^4 + \begin{vmatrix} 0 & 1 \\ 1 & -\lambda \end{vmatrix} = 0$$

$$\lambda^4 = 1$$

Therefore, for the 4×4 case, the eigenvalues are the fourth roots of unity.

So, if we let ω_n be the nth roots of unity, we can say that the values of $q(\omega_n)$ are the eigenvalues of a corresponding circulant C. But, as we've seen, q is defined by the first row of C. Thus, we now have a simple method for calculating the eigenvalues of any given $n \times n$ circulant matrix, C.

- Write down the polynomial q, in ascending order, defined by the first row of C.
- Evaluate $q(\omega_n)$.
- The values of $q(\omega_n)$ are now the eigenvalues of C.

As an example, consider the given 4×4 circulant matrix.

$$C = \begin{pmatrix} 1 & 2 & 1 & 3 \\ 3 & 1 & 2 & 1 \\ 1 & 3 & 1 & 2 \\ 2 & 1 & 3 & 1 \end{pmatrix}$$

The first row of C defines the polynomial q in ascending order.

$$q(t) = 1 + 2t + t^2 + 3t^3$$

The eigenvalues of C are now $q(\omega)$, where ω, in this case, is the fourth roots of unity.

$$\begin{aligned} q(1) &= 1 + 2(1) + (1)^2 + 3(1)^3 = 7 \\ q(-1) &= 1 + 2(-1) + (-1)^2 + 3(-1) = -3 \\ q(i) &= 1 + 2i + i^2 + 3i^3 = -i \\ q(-i) &= 1 + 2(-i) + (-i)^2 + 3(-i)^3 = i \end{aligned}$$

This demonstration has shown that the computation of a circulant's eigenvalues is actually quite trivial. We will use this tool in the next section to solve polynomials.

3. Solving Polynomials

Up until now, we have started with a given circulant and extracted its eigenvalues, which are the roots of its characteristic polynomial p. But we normally want to work in the reverse order, where we are given a polynomial p, whose roots we want to calculate. Now, the trick is to find a circulant matrix whose characteristic polynomial *is* p. But once we've done that, the roots of p are simply the eigenvalues $q(\omega_n)$.

3.1. A Quadratic Example

We are given a polynomial p whose roots we want to calculate.

$$p(x) = x^2 - 2x - 3$$

First, we want to find a 2×2 circulant matrix whose characteristic polynomial is p. So we must consider the general 2×2 circulant.

$$C = \begin{pmatrix} a & b \\ b & a \end{pmatrix}$$

The characteristic polynomial of C is then given by

$$\det(xI - C) = \det \begin{pmatrix} x - a & b \\ b & x - a \end{pmatrix} = x^2 - 2ax + a^2 - b^2$$

By inspection we can make the following relationships:

$$\begin{aligned} -2a &= -2 \\ a^2 - b^2 &= -3 \end{aligned}$$

Solving this system we get $a = 1$ and $b = \pm 2$. For simplicity we take $b = 2$ and we have

$$C = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$$

The first row of C now defines q

$$q(t) = 1 + 2t$$

Finally, evaluating q at the two square roots of unity, gives the roots of p.

$$\begin{aligned} q(1) &= 1 + 2(1) = 3 \\ q(-1) &= 1 + 2(-1) = -1 \end{aligned}$$

3.2. The Quadratic Formula

Now we extend the same method to the solution of the general quadratic polynomial.

$$p(x) = x^2 + \alpha x + \beta$$

So, again we find the characteristic polynomial of the general 2×2 circulant.

$$\det(xI - C) = x^2 - 2ax + a^2 - b^2$$

Now we need to find a and b so that our above expression equals p.

$$\begin{aligned} -2a &= \alpha \\ a^2 - b^2 &= \beta \end{aligned}$$

Thus,

$$a = -\frac{\alpha}{2} \qquad \text{and} \qquad b = \pm\sqrt{\frac{\alpha^2}{4} - \beta}$$

Substituting a and b into the general 2×2 circulant C, we get

$$C = \begin{pmatrix} -\frac{\alpha}{2} & \sqrt{\frac{\alpha^2}{4} - \beta} \\ \sqrt{\frac{\alpha^2}{4} - \beta} & -\frac{\alpha}{2} \end{pmatrix}$$

And by inspection we have

$$q(t) = -\frac{\alpha}{2} + t\sqrt{\frac{\alpha^2}{4} - \beta}$$

So, the roots of p are now the values of $q(\omega)$. Where ω in this case is the two square roots of unity.

$$q(1) = -\frac{\alpha}{2} + \sqrt{\frac{\alpha^2}{4} - \beta}$$
$$q(-1) = -\frac{\alpha}{2} - \sqrt{\frac{\alpha^2}{4} - \beta}$$

These are now the formulas which give the roots of any second degree polynomial, and when the two are combined, they yield the familiar quadratic formula.

3.3. The Cubic

The same method extends naturally to the solution of the cubic. In general, we want the roots of

$$p(x) = x^3 + \alpha x^2 + \beta x + \gamma.$$

Now consider the general 3×3 circulant.

$$C = \begin{pmatrix} a & b & c \\ c & a & b \\ b & c & a \end{pmatrix}$$

The characteristic polynomial of C is then

$$\det(xI - C) = \begin{vmatrix} x-a & -b & -c \\ -c & x-a & -b \\ -b & -c & x-a \end{vmatrix} = (x-a)^3 - b^3 - c^3 - 3bc(x-a)$$

We want to find a, b, and c such that p is the characteristic polynomial of C. To do so, we essentially must make a change of variables, letting $y = x - a$. This will eliminate the quadratic term. It is interesting to note that this change of variables is a similar strategy used in the traditional solution method.

But the point here is that we want to eliminate the quadratic term in p. Traditionally, this is accomplished by letting $x = y - \frac{\alpha}{n}$ where n is the polynomial's degree, and α is the coefficient of the $n-1$ term. So, we really want to solve a modified polynomial p after making the appropriate change of variables.

$$p(y) = y^3 + \beta y + \gamma$$

But it is important to note that β is now a new coefficient and γ is a new constant.

Now we want to find a 3×3 circulant whose characteristic polynomial is $p(y)$. But let us first consider some interesting facts. We know that the sum of the roots of a polynomial of degree n is equal to the negative of the coefficient of the $n-1$ term. But since we have eliminated that term, the sum of the roots of $p(y)$ is now zero. It follows directly that the sum of the eigenvalues of its corresponding circulant must be zero, and thus, its trace must also be zero. But in the domain of circulant matrices, the main diagonal is constant. So the important conclusion here is that the circulant C, whose characteristic polynomial is $p(y)$, must have all zeros along its main diagonal. This is known as a *traceless circulant.*

$$C = \begin{pmatrix} 0 & b & c \\ c & 0 & b \\ b & c & 0 \end{pmatrix}$$

The characteristic polynomial of C is now

$$\det(yI - C) = \begin{pmatrix} y & -b & -c \\ -c & y & -b \\ -b & -c & y \end{pmatrix} = y^3 - b^3 - c^3 - 3bcy$$

which equals $p(y)$ if and only if:

$$-3bc = \beta$$
$$-b^3 - c^3 = \gamma$$

To solve the system, think of the unknowns as b^3 and c^3. So, dividing the first equation by -3 and cubing gives us

$$b^3c^3 = -\frac{\beta^3}{27}$$
$$b^3 + c^3 = -\gamma$$

By substitution, we can obtain a quadratic in b and c.

$$-b^6 - \gamma b^2 + \frac{\beta^3}{27} = 0$$

Applying the quadratic formula we have

$$b^3 = \frac{\gamma \pm \sqrt{\gamma^2 + \frac{4\beta^3}{27}}}{-2}$$
$$b = \left[\frac{\gamma \pm \sqrt{\gamma^2 + \frac{4\beta^3}{27}}}{-2}\right]^{1/3}$$

We can now define b using *any* of the possible values of the square and cube roots. All choices for b yield the same roots. Back substitution then gives $c = -\beta/(3b)$. So we can now define the polynomial q.

$$q(t) = bt - \frac{\beta}{3b}t^2$$

And finally, we obtain the roots of $p(y)$ by evaluating

$$q(1) = b - \frac{\beta}{3b}$$
$$q(\omega) = b\omega - \frac{\beta}{3b}\omega^2$$
$$q(\omega^2) = b\omega^2 - \frac{\beta\omega}{3b}$$

4. Conclusion

A similar analysis leads to the solution of the quartic. One would begin by eliminating the x^3 term and following the same logical process. Although recall that in the previous section, solving the cubic led to solving a quadratic for b and c. But in this case, we are forced to solve a cubic for the constants b, c, and d, which define q. This becomes an arduous process. Nevertheless, the method remains consistent.

However, when applied to general polynomials of degree greater than four, the circulant method fails. But keep in mind, that's only because the roots of these higher degree polynomials cannot necessarily be expressed in terms of pure radicals. Therefore we are left with a simple, unified approach to the solutions of all polynomials through degree 4.

References

[1] John Stillwell. *Mathematics and its History.*

[2] Dan Kalman and James E. White. *Polynomial Equations and Circulant Matrices.*

[3] Gilbert Strang. *Introduction to Linear Algebra.*