Viewing Polynomial Roots With Matrix Eyes:

Part 1: Irrational Roots via Matrix Dynamics
Part 2: Solving Cubics and Quartics via Circulant Matrices

Dan Kalman, American University

Related papers available on line in PDF (acrobat reader) format
at http://www.cas.american.edu/∼kalman/pdffiles

Part 1: Dynamics and Irrational Roots

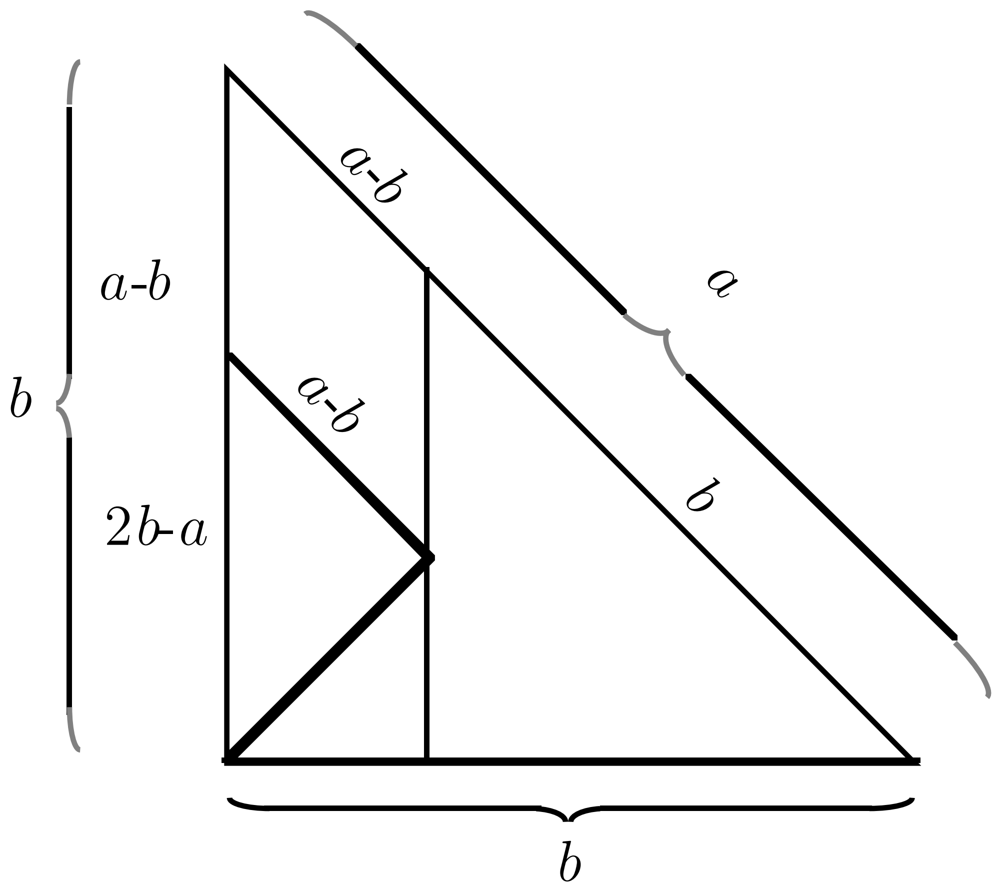
Joint Work With
Robert Mena, California State University Long Beach
Shahriar Shahriari, Pomona College

Our article:
Variations on an Irrational Theme – Geometry, Dynamics, Algebra. **Mathematics Magazine,** Volume 70, Number 2, April 1997, pp 93 – 104.

On the web page: irrat.pdf

Outline

- Incommensurability Arguments
- Matrix Formulation
- Dynamic Viewpoint
- Generalizations



$$\frac{a}{b} = \frac{2b-a}{a-b}$$

Algebraic Approach

$$\begin{aligned} a^2 &= 2b^2 \\ a^2 - ab &= 2b^2 - ab \\ a(a-b) &= b(2b-a) \\ \frac{a}{b} &= \frac{2b-a}{a-b} \end{aligned}$$

$$\begin{array}{ccccc} 0 & < & 2b-a & < & a \\ 0 & < & b-a & < & b \end{array}$$

Matrix Approach

$$\begin{bmatrix} -1 & 2 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 2b - a \\ a - b \end{bmatrix}$$

- Call the matrix A
- If $\mathbf{v}$ has integer coordinates, so does $A\mathbf{v}$
- If $\mathbf{v}$ is on the line with slope $1/\sqrt{2}$, so is $A\mathbf{v}$

Eigenvectors and Things

$$\begin{bmatrix} -1 & 2 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} \sqrt{2} \\ 1 \end{bmatrix} = \begin{bmatrix} 2-\sqrt{2} \\ \sqrt{2}-1 \end{bmatrix}$$

$$= \begin{bmatrix} \sqrt{2}(\sqrt{2}-1) \\ \sqrt{2}-1 \end{bmatrix}$$

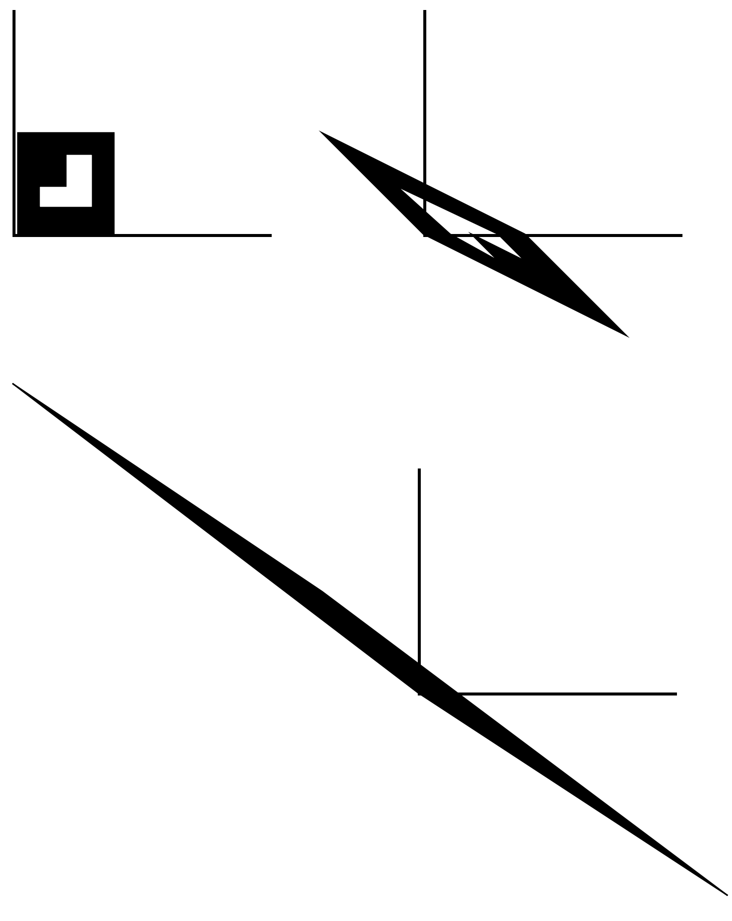
$$= (\sqrt{2}-1) \begin{bmatrix} \sqrt{2} \\ 1 \end{bmatrix}$$

- $\lambda = \sqrt{2} - 1$ is an eigenvalue
- $0 < \lambda < 1$
- $A\mathbf{v} = \lambda \mathbf{v} \Rightarrow A^n \mathbf{v} \downarrow 0$
- If $\mathbf{v}$ has integer entries, $A^n\mathbf{v}$ is a sequence of lattice points that converges to 0

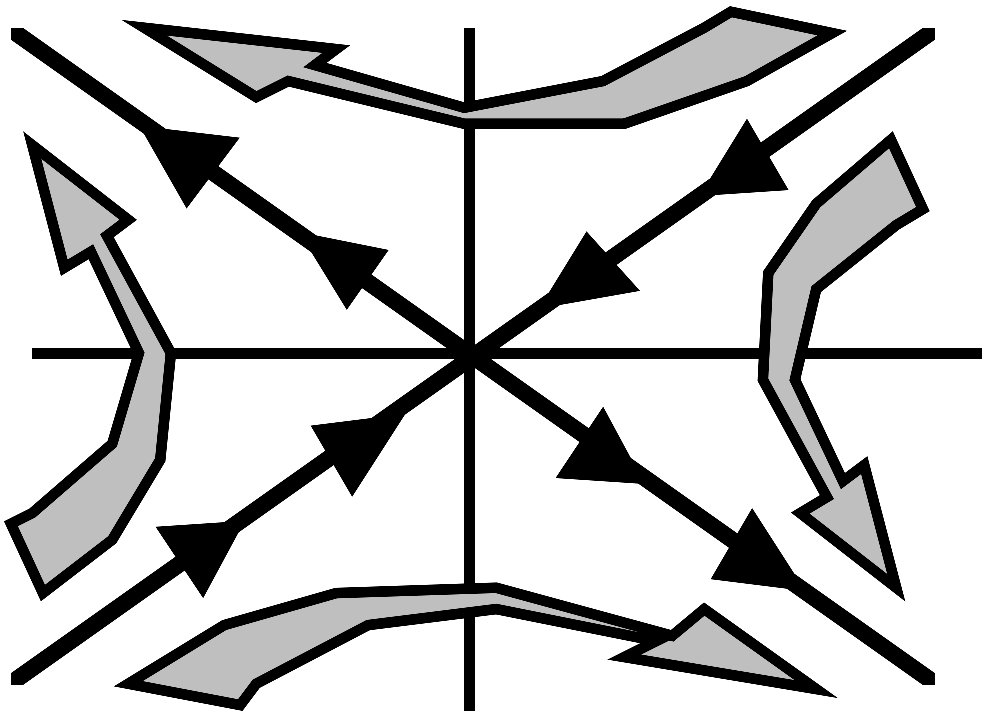
Dynamical View

- Consider the orbits under repeated application of A starting from various points in the plane
- All the points on the eigenline L (with slope $1/\sqrt{2}$) contract to the origin
- Other eigenvalue is $\mu = -1 - \sqrt{2}$ which has magnitude > 1
- All the points (including all the lattice points) off L escape to infinity, oscilating along the eigenline for μ

Dynamics of A



Dynamics of A^2



Inverse Iteration

- $A^{-1} = \begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix}$
- Same eigenlines, reciprocal eigenvalues, integer entries
- Orbits from lattice points all converge to the line L
- Let $\mathbf{v}$ be any integer vector. Let $\mathbf{v_n} = A^{-n}\mathbf{v} = [a_n \quad b_n]^T$. Then $a_n/b_n \to \sqrt{2}$.

Example

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad \begin{bmatrix} 3 \\ 2 \end{bmatrix} \quad \begin{bmatrix} 7 \\ 5 \end{bmatrix} \quad \begin{bmatrix} 17 \\ 12 \end{bmatrix}$$

$$\begin{bmatrix} 41 \\ 29 \end{bmatrix} \quad \begin{bmatrix} 99 \\ 70 \end{bmatrix} \quad \begin{bmatrix} 239 \\ 169 \end{bmatrix} \quad \begin{bmatrix} 577 \\ 408 \end{bmatrix} \quad \begin{bmatrix} 1393 \\ 985 \end{bmatrix}$$

$$\begin{bmatrix} 3363 \\ 2378 \end{bmatrix} \quad \begin{bmatrix} 8119 \\ 5741 \end{bmatrix} \quad \begin{bmatrix} 19601 \\ 13860 \end{bmatrix} \quad \begin{bmatrix} 47321 \\ 33461 \end{bmatrix} \quad \begin{bmatrix} 114243 \\ 80782 \end{bmatrix}$$

The last vector gives us 114243/80782 as a rational approximation of $\sqrt{2}$. This approximation is correct to 10 decimal places.

From 2 to n

- Let $A = \begin{bmatrix} -k & n \\ 1 & -k \end{bmatrix}$
- Eigenvectors: $\begin{bmatrix} \pm\sqrt{n} \\ 1 \end{bmatrix}$
- Eigenvalues: $\lambda = \sqrt{n} - k$ and $\mu = -(\sqrt{n} + k)$
- Choose $k = \lfloor \sqrt{n} \rfloor$
- If n is not a perfect square, $0 < \lambda < 1$ and we get the same dynamics as before. No integer eigenvectors for λ. This shows that $\sqrt{n}$ is irrational.
- If n is a perfect square, $\lambda = 0$. In this case there are lattice points on line L and they collapse to the origin in one jump

More Generally

Theorem: An eigenvalue of an integer matrix is either an integer or irrational.

Proof Outline:

- Let A be an integer matrix with rational eigenvalue λ that is not an integer
- WLOG, $0 < \lambda < 1$ for otherwise, replace A with $A - \lfloor \lambda \rfloor I$
- There is an integer eigenvector $\mathbf{v}$ associated with λ
- $A^k \mathbf{v} \to 0$ landing only at lattice points

Corollary

A polynomial with integer coefficients has roots that are either integers or irrational.

Proof: Given the polynomial p, construct the companion matrix A: an integer matrix with p as the characteristic polynomial. The roots of p are all eigenvalues of the matrix A, so must be integers or irrational.

Gauss's Lemma

- $f = gh$ is a polynomial factorization
- Assume f, g, h monic
- Assume coefficients of f are integers, of g, h rational
- Conclusion: coefficients of g, h are integers
- For linear g and h this is same as previous result that rational roots must be integers

Algebraic Integers

- An *Algebraic Integer* is a root of a monic polynomial with integer coefficients
- Earlier theorem: $\{\text{Algebraic Integers}\} \cap \{\text{Rationals}\} \subseteq \{\text{Integers}\}$
- Earlier theorem: r algebraic integer $\Leftrightarrow$ r eigenvalue of an integer matrix
- Lemma: Sums and products of algebraic integers are algebraic integers
- Proof of Gauss's theorem: roots of g and h are roots of f hence algbraic integers. Coefficients of g and h are made up of sums and products of the roots, hence algebraic integers and rational, hence integers

Kronecker Products

- $A \otimes B$ defined for any matrices A and B
- $\begin{bmatrix} a & b \\ c & d \end{bmatrix} \otimes B = \begin{bmatrix} aB & bB \\ cB & dB \end{bmatrix}$
- $(A \otimes B)(C \otimes D) = AC \otimes BD$
- $Av = \lambda v$ and $Bw = \mu w \quad \Rightarrow (A \otimes B)(v \otimes w) = Av \otimes Bw = \lambda\mu v \otimes w$
- $(A \otimes I + I \otimes B)(v \otimes w) = Av \otimes w + v \otimes Bw = (\lambda + \mu)v \otimes w$

Part 2: Solving Cubics and Quartics by Completing the Circulant

Joint work with James White, Mathwright Library
Closely related to results by A. Pen-Tung Sah,
A uniform method of solving cubics and quartics,
Monthly, vol 52, pp202-206, 1945.

Outline

- Circulant Matrices
- Special Case: Trace $= 0$
- Solving the Cubic
- Solving the Quartic
- Other Aspects

Circulant Matrices

- Generator W: Identity matrix with top row shifted to the bottom

$$\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$$

- General circulant: $q(W)$ where q is a polynomial of degree one less than the dimension of W.

- For 3×3 $W : aI + bW + cW^2$

$$\begin{bmatrix} a & b & c \\ c & a & b \\ b & c & a \end{bmatrix}$$

Properties of Circulants

- Minimal/characteristic polynomial for $W : W^n - 1$
- Eigenvalues of W : n^{th} roots of unity ω
- Eigenvalues of $q(W)$ are $q(\omega)$

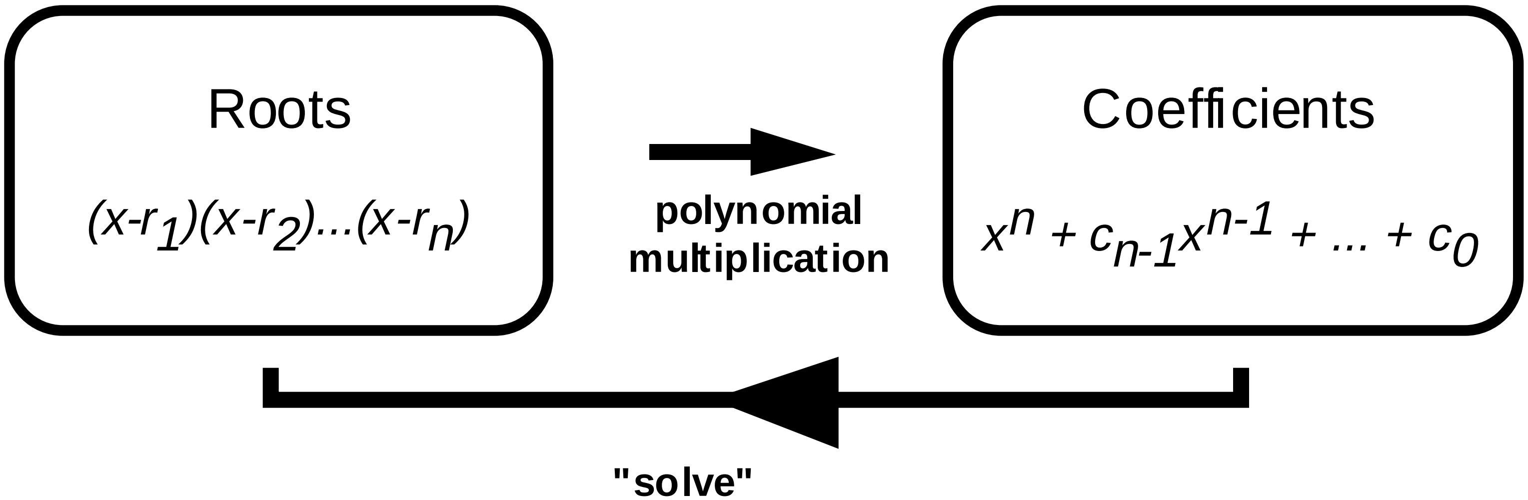
$$W\mathbf{v} = \omega\mathbf{v} \Rightarrow q(W)\mathbf{v} = q(\omega)\mathbf{v}$$

- Eigenvalues of $aI + bW + cW^2$ are $a + b\omega + c\omega^2$ where ω is any 3^{rd} root of unity

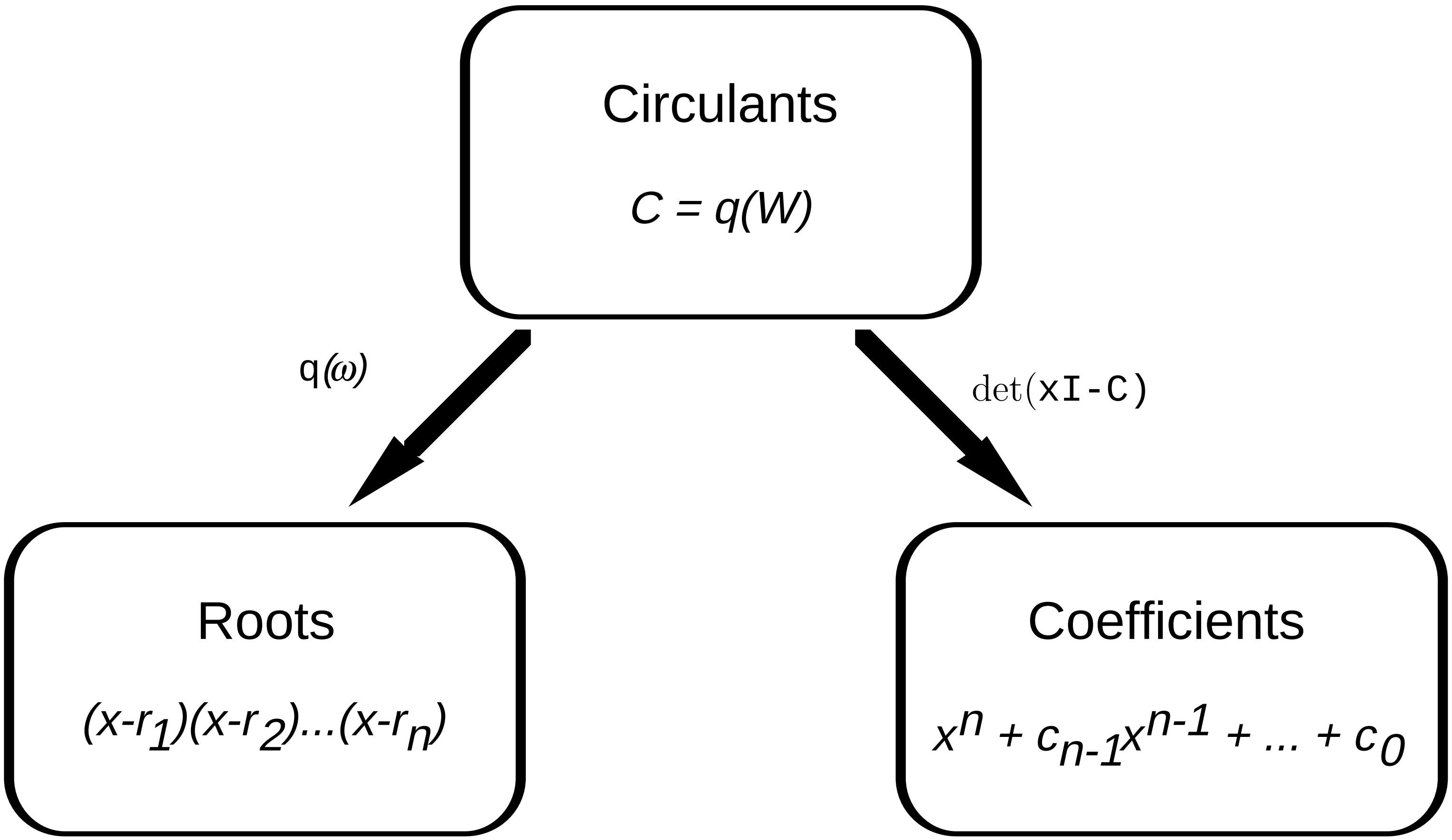
Solving Polynomials with Circulants

- Usual notion of solving a polynomial: given coefficients, find roots
- Circulants give us a rich set of polynomials with known roots
- New approach to solving a polynomial: given p (defined in terms of coefficients), find a circulant matrix $C = q(W)$ for which p is the characterisic polynomial. The eigenvalues $q(\omega)$ of C are then the roots of p

Usual Method



Circulant Method



Traceless Circulants

- WLOG the degree $n-1$ term of p vanishes
- Equivalently, sum of roots of p vanishes
- Equivalently, sum of eigenvalues of $q(W)$ vanishes
- Equivalently, trace of $q(W)$ vanishes
- Equivalently, diagonal of $q(W)$ vanishes

Order 3 Characteristic Polynomial

The characteristic polynomial of $M = q(W) = q_1 W + q_2 W^2$:

$$\begin{aligned}
\det\left(xI - \begin{bmatrix} 0 & q_1 & q_2 \\ q_2 & 0 & q_1 \\ q_1 & q_2 & 0 \end{bmatrix}\right) &= \det\begin{bmatrix} x & -q_1 & -q_2 \\ -q_2 & x & -q_1 \\ -q_1 & -q_2 & x \end{bmatrix} \\
&= x^3 - q_1^3 - q_2^3 - 3q_1 q_2 x \\
&= x^3 - 3q_1 q_2 x - (q_1^3 + q_2^3)
\end{aligned}$$

Solving the Cubic

- Given $p(x) = x^3 + ax + b$ find $q(W) = q_1 W + q_2 W^2$
- Characteristic polynomial of $q(W)$ is $x^3 - 3q_1 q_2 x - (q_1^3 + q_2^3)$
- Must solve the system

$$\begin{aligned} q_1^3 + q_2^3 &= -b \\ q_1 q_2 &= -a/3 \end{aligned}$$

- Roots of p will be $q(1) = q_1 + q_2$, $q(\omega) = q_1\omega + q_2\omega^2$, and $q(\omega^2) = q_1\omega^2 + q_2\omega$ where $\omega = \dfrac{-1 + i\sqrt{3}}{2}$

Finding q

$$\left\{ \begin{array}{ccc} q_1^3 + q_2^3 & = & -b \\ q_1 q_2 & = & -a/3 \end{array} \right\} \Rightarrow \left\{ \begin{array}{ccc} q_1^3 + q_2^3 & = & -b \\ q_1^3 q_2^3 & = & -a^3/27 \end{array} \right\}$$

q_1^3 and q_2^3 are roots of $x^2 + bx - a^3/27 = 0$

Thus, q_1 and q_2 are given by

$$\left\{ \frac{-b \pm \sqrt{b^2 + 4a^3/27}}{2} \right\}^{1/3} \quad \text{or} \quad \left\{ \frac{-b}{2} \pm \sqrt{\frac{b^2}{4} + \frac{a^3}{27}} \right\}^{1/3}$$

Solving the Quartic

- $p(x) = x^4 + ax^2 + bx + c$
- Characteristic polynomial of $q(W) = q_1 W + q_2 W^2 + q_3 W^3$

  $x^4 - (4q_3q_1 + 2q_2^2)x^2 - 4q_2(q_1^2 + q_3^2)x + (q_2^4 - q_1^4 - q_3^4 - 4q_1q_3q_2^2 + 2q_1^2q_3^2) = 0$

- Elimination leads to

$$q_2^6 + \frac{a}{2}q_2^4 + \left(\frac{a^2}{16} - \frac{c}{4}\right)q_2^2 - \frac{b^2}{64} = 0$$

- 4th roots of unity: $\pm 1, \pm i$
- Roots of p: $\pm q_1 + q_2 \pm q_3$, and $\pm i q_1 - q_2 \mp i q_3$.

Degree n: Existence of $q(W)$

- Given monic p of degree n, find a circulant matrix $C(p)$ for which the characteristic polynomial is p
- $C(p) = q(W)$ for an appropriate polynomial q of degree $n-1$
- Existence: if the roots of p are r_k, and the n^{th} roots of unity are ω_k, it suffices to have $q(\omega_k) = r_k$. Existence of q assured by polynomial interpolation theory.

Another View

Let $q(x) = q_0 + q_1 x + \cdots + q_{n-1} x^{n-1}$. Let the roots of p be r_k for $1 \le k \le n$. Let $\omega = e^{2\pi i/n}$, so that the powers of ω are the n^{th} roots of unity. Then q maps the roots of unity to the r_k if

$$\begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & \omega & \omega^2 & \cdots & \omega^{n-1} \\ 1 & \omega^2 & \omega^4 & \cdots & \omega^{2(n-1)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \omega^{n-1} & \omega^{2(n-1)} & \cdots & \omega^{(n-1)(n-1)} \end{bmatrix} \begin{bmatrix} q_0 \\ q_1 \\ q_2 \\ \vdots \\ q_{n-1} \end{bmatrix} = \begin{bmatrix} r_1 \\ r_2 \\ r_3 \\ \vdots \\ r_n \end{bmatrix}$$

This system is easily seen to be solvable for any choice of the r_k.
Side Comment: The column of r_k is the discrete fourier transform of the column of q_j.

Other Aspects

- Roots real iff $q(W)$ is Hermitian
- Sufficient condition for roots to be rational
- Newton's identities related to trace of $(q(W))^k$