On Squares of the Solutions for a Polynomial Equation

Ali Astaneh Ph.D(Lon), Vancouver, B.C.

Given a polynomial equation $x^n + a_{n-1}x^{n-1} + \ldots + a_1x + a_0 = 0$ with real or complex coefficients, the main objective of this article is to construct a new polynomial equation $X^n + A_{n-1}x^{n-1} + \ldots + A_1x + A_0 = 0$ whose n roots are "squares" of the solutions of the original equation. As expected the coefficients of the second equations are to be expressed in terms of the coefficients of the original one. For convenience (as well as a good reason!), I will refer to the second equation as the "transit" equation for the first one. One reason being , if you are interested to find those solutions of the first equations that are square root $\sqrt{k}$ of non-complete square positive integers k, you can find positive integer roots of the transition equation, and consider $\pm$ square root of them to be "square root" solutions of the original equation.

As for how I ended up with the odd looking sigma in the following theorem, I might add I simply tried to find a system of *n* linear equations in terms of the *n* unknown coefficients A_k's and solve the system in terms of the original coefficients a_k's. To this end, I used the popular set of *n* relations that hold between the *n* zeros of a polynomial of degree *n* and its *n* coefficients $a_k,\ k = 1, 2, \ldots, n-1.$ I simply grouped those relations in a way to enable me to solve forA_k's in terms of a_k's. This process will not appear in the proof, as I realized at the end that a proof of my formulation can be delivered by induction on *n* . Here is the formulation,

Theorem [A. Astaneh] Let $x^n + a_{n-1}x^{n-1} + \ldots + a_1x + a_0 = 0$ be a given polynomial equation with real or complex coefficients and set,

$A_0 = (-1)^n a_0^2,$
$A_k = (-1)^{n+k}[a_k^2 + 2\sum_{j=0}^{k-1}(-1)^{k-j}a_j a_{2k-j}],$ for $k = 1, 2, \ldots n-1.$

Then roots of the "transit" equation $X^n + A_{n-1}X^{n-1} + \ldots + A_1X + A_0 = 0$ will be exactly squares of the roots of the original polynomial.

Note that, whenever in manipulations in the proof the coefficient a_n occurs we will consider it to be the leading coefficient of the original equation which means $a_n = 1$. This, for example occurs as a term of the sigma when $k = n - 1$ and $j = k - 1$, as then the coefficient a_{2k-j} becomes $a_n = 1$.

Proof: We prove the claimed assertion by induction on the degree n of the polynomial .The case $n = 1$ is obvious, because the polynomial whose root is the square of the only solution of the polynomial $x + a_0 = 0$ is $X - a_0^2 = 0$, which means the relation $A_0 = (-1)^1 a_0^2$ holds.
Assume that the assertion holds for n, and consider an $n + 1$ degree polynomial equation

$$x^{n+1} + c_n x^n + c_{n-1} x^{n-1} + \ldots + c_1 x + c_0 = 0. \quad (1)$$

By the fundamental theorem of algebra, this equation has $n + 1$ real or complex roots. Therefore, in particular for some real or complex number r, the equation it can be factored as

$$(x - r)(x^n + a_{n-1} x^{n-1} + \ldots + a_1 x + a_0) = 0. \quad (2)$$

Next, by induction hypothesis, the roots of the following equation are exactly squares of the roots of the original equation

$$(X - r^2)(X^n + A_{n-1} x^{n-1} + \ldots + A_1 x + A_0) = 0, \quad (3)$$

where $A_0 = (-1)^1 a_0^2$, and $A_k = (-1)^{n+k} [\, a_k^2 + 2\sum_{i=0}^{k-1} (-1)^{k-i} a_i a_{2k-i}]$.
If we know assume the roots of the polynomial

$$X^{n+1} + C_n X^n + \ldots + C_1 X + C_0 = 0 \quad (4)$$

are squares of the solutions of equation (1) , then we must show that

$$C_0 = (-1)^{n+1} c_0^2, \quad (5)$$

$$C_k = (-1)^{n+1+k} [c_k^2 + 2 \sum_{j=0}^{k-1} (-1)^{k-j} c_j c_{2k-j}], \quad \text{for } k = 1, 2, \ldots n. \quad (6)$$

To settle (5), since equations (3) and (4) have the same roots and the same leading coefficients 1, all other coefficients must be the same, in particular

$$C_0 = -r^2 A_0 = -r^2 (-1)^n a_0^2 = (-1)^{n+1} r^2 a_0^2 = (-1)^{n+1} c_0^2$$

which settles (5). And finally to show (6), once again considering that not only equations because (3) and (4) have the same coefficients, but also (1) and (2) have the same coefficients we have the following two relations

$$C_k = A_{k-1} - r^2 A_k, \quad \text{for } k = 1, 2, \ldots n. \quad (7)$$

$$c_k = a_{k-1} - r a_k, \quad \text{for } k = 1, 2, \ldots n - 1. \quad (8)$$

The proof will be complete if we show that the right hand sides of (6) and (7) are equal. Since the right hand side of (7) is

$$A_{k-1} - r^2 A_k$$
$$= (-1)^{n+k-1}[a_{k-1}^2 + 2\sum_{i=0}^{k-2}(-1)^{k-1-i}a_i a_{2k-2-i}]$$
$$- r^2(-1)^{n+k}[a_k^2 + 2\sum_{i=0}^{k-1}(-1)^{k-i}a_i a_{2k-i}]$$

it is enough to show that for $k = 1,\ 2,\ \ldots, n$ we have,

$$(-1)^{n+1+k}\ [\ c_k^2\ + 2\ \sum_{i=0}^{k-1}(-1)^{k-i}c_i c_{2k-i}] =$$
$$(-1)^{n+k-1}[\ a_{k-1}^2 + 2\ \sum_{i=0}^{k-2}(-1)^{k-i-1}a_i a_{2k-2-i}]$$
$$- r^2\ (-1)^{n+k}[\ a_k^2\ + 2\ \sum_{i=0}^{k-1}(-1)^{k-i}a_i a_{2k-i}]$$

Or

$$c_k^2\ + 2\ \sum_{i=0}^{k-1}(-1)^{k-i}c_i c_{2k-i} = a_{k-1}^2 + 2\ \sum_{i=0}^{k-2}(-1)^{k-i-1}a_i a_{2k-2-i}$$
$$+\ r^2[\ a_k^2\ + 2\ \sum_{i=0}^{k-1}(-1)^{k-i}a_i a_{2k-i}] .$$

The above relation will be established by considering that by (1) and (2) $c_0 = -ra_0$,$c_n = a_{n-1} - r$, and $c_k = a_{k-1} - ra_k$. However, in order to avoid the appearance of an undefined coefficient of a_{-1} it is necessary to isolate the first term of the sigma on the left. Also for a preliminary cancellation it helps to isolate the first term of the last sigma on the right, to get

$$(a_{k-1} - ra_k)^2 + 2(-1)^k(-ra_0)(a_{2k-1} - ra_{2k})$$
$$+2\ \sum_{i=1}^{k-1}(-1)^{k-i}(a_{i-1} - ra_i)(a_{2k-i-1} - ra_{2k-i}) =$$
$$a_{k-1}^2 + 2\ \sum_{i=0}^{k-2}(-1)^{k-i-1}a_i a_{2k-2-i} +\ r^2[\ a_k^2\ + 2(-1)^k a_0 a_{2k}\ + 2\sum_{i=0}^{k-1}(-1)^{k-i}a_i a_{2k-i}]$$

The above equation is first simplified into

$$-ra_k a_{k-1} + (-1)^{k+1} ra_0 a_{2k-1}$$
$$+\sum_{i=1}^{k-1} (-1)^{k-i} (a_{i-1} - ra_i)(a_{2k-i-1} - ra_{2k-i}) = \quad .$$
$$+\sum_{i=0}^{k-2} (-1)^{k-i-1} a_i a_{2k-2-i} + r^2 \sum_{i=1}^{k-1} (-1)^{k-i} a_i a_{2k-i}$$

Then, it is further simplified into

$$-ra_k a_{k-1} + (-1)^{k+1} ra_0 a_{2k-1}$$
$$+\sum_{i=1}^{k-1} (-1)^{k-i} (a_{i-1} a_{2k-i-1} - ra_{i-1} a_{2k-i} - ra_i a_{2k-i-1}) =$$
$$+\sum_{i=0}^{k-2} (-1)^{k-i-1} a_i a_{2k-2-i}$$

Next, a change of index $j = i+1$ in the sigma on the right will give rise into

$$-ra_k a_{k-1} + (-1)^{k+1} ra_0 a_{2k-1}$$
$$+\sum_{i=1}^{k-1} (-1)^{k-i} (a_{i-1} a_{2k-i-1} - ra_{i-1} a_{2k-i} - ra_i a_{2k-i-1}) =$$
$$+\sum_{j=1}^{k-1} (-1)^{k-j} a_{j-1} a_{2k-1-j}$$

Therefore the equation is reduced to

$$-a_k a_{k-1} + (-1)^{k+1} a_0 a_{2k-1}$$
$$+\sum_{i=1}^{k-1} (-1)^{k-i+1} (a_{i-1} a_{2k-i} + a_i a_{2k-i-1}) = 0 \, .$$

Next we will split the sigma involved into two sigmas, and get,

$$-a_k a_{k-1} + (-1)^{k+1} a_0 a_{2k-1}$$
$$+\sum_{i=1}^{k-1} (-1)^{k-i+1} a_{i-1} a_{2k-i} + \sum_{i=1}^{k-1} (-1)^{k-i+1} a_i a_{2k-i-1} = 0$$

Again, upon change of index $j = i+1$ in the sigma on the right we get

$$-a_k a_{k-1} + (-1)^{k+1} a_0 a_{2k-1}$$
$$+\sum_{i=1}^{k-1} (-1)^{k-i+1} a_{i-1} a_{2k-i} + \sum_{j=2}^{k} (-1)^{k-j+2} a_{j-1} a_{2k-j} = 0$$

And finally we isolae the first term of the sigma on the left, and the last term of the sigma on the right , giving rise to difference of two identical sigmas that cancel. The equality to be shown will eventually turn into

$$-a_k a_{k-1} + (-1)^{k+1} a_0 a_{2k-1}$$
$$+ (-1)^k a_0 a_{2k} + (-1)^2 a_{k-1} a_k = 0$$

which is a true relation. Therefore the proof is complete.

Remark The above Theorem can reduce the problem of finding the 'square root' zeros of a given polynomial $x^n + a_{n-1}x^{n-1} + \ldots + a_1 x + a_0 = 0$ with real or complex coefficients to that of finding positive roots of the "transition" polynomial equation $X^n + A_{n-1}x^{n-1} + \ldots + A_1 x + A_0 = 0$, as seen in the following six examples.

Example 1 The transition equation for the quadratic equation $x^2 + a_1 x + a_0 = 0$ reads as $X^2 - (a_1^2 - 2a_0)X + a_0^2 = 0$.
Therefore, for the specific equation $x^2 - \sqrt{2}\, x - 4 = 0$ the transition equation is $X^2 - 10X + 16 = 0$. Since the transition equation has roots $X = 2,\ 8$ one can check that $x = -\sqrt{2},\ \text{and } 2\sqrt{2}$ are roots of the original equation.
Observe in the event that equation $x^2 + a_1 x + a_0 = 0$ has complex roots instead, the transition equation will have squares of those complex numbers as its roots. For example the roots of $x^2 - 6x + 13 = 0$ are $x = 3 \pm 2i$, and the roots of the transition equation $X^2 - 10X + 169 = 0$ are $X = 5 \pm 12i$, which are respective squares of the complex numbers $x = 3 \pm 2i$.

Example 2 The transition equation for the cubic equation $x^3 + a_2 x^2 + a_1 x + a_0 = 0$ reads as
$X^3 - (a_2^2 - 2a_1)X^2 + (a_1^2 - 2a_0 a_2)X - a_0^2 = 0$. To see a direct verification

of this assertion, we only need to observe the following three identities in terms of the roots x_1, x_2, x_3 of the original cubic equation,

$$x_1^2 + x_2^2 + x_3^2 = (x_1 + x_2 + x_3)^2 - 2\ (x_1x_2 + x_2x_3 + x_1x_3)$$

$$x_1^2x_2^2 + x_2^2x_3^2 + x_1^2x_3^2 = (x_1x_2 + x_2x_3 + x_1x_3)^2 - 2\ x_1x_2x_3(x_1 + x_2 + x_3)$$

$$x_1^2x_2^2x_3^2 = (x_1x_2x_3)^2$$

Therefore, for the specific equation $x^3 + (2 - \sqrt{2}\,)\,x^2 \ - 2\ (\sqrt{2}\ + 2)\ x\ - 8\ =\ 0$ the transition equation is $X^3 - 14\,X^2\ +\ 56\,X\ - 64 = 0$. Since the latter equation has roots $X = 2,\ 4,\ \text{and}\ 8$, one can check that $x = \sqrt{2},\ -2,\ \text{and}\ 2\sqrt{2}$ are roots of the original cubic equation.

Example 3 The transition equation for the quartic equation

$$x^4 + a_3x^3 + a_2\ x^2\ + a_1\ x\ +\ a_0 = 0$$

reads as

$$X^4 - (a_3^2 - 2a_2)\,X^3 + (a_2^2 - 2a_1a_3 + 2a_0)\,X^2 - (a_1^2 - 2\ a_0\ a_2)\ X\ + a_0^2 = 0.$$

Therefore, for the specific equation $x^4 - \sqrt{2}\ x^3 - 7\ x^2\ + 3\sqrt{2}\ x\ + 12 = 0$ the transition equation is $X^4 - 16\,X^3 + 85\,X^2 - 186\,X\ + 144 = 0$. Since the latter equation has roots $X = 2,\ 3,\ \text{and}\ 8$, one can see that $x = -\sqrt{2},\ 2\sqrt{2},\ \text{and}\ \pm\sqrt{3}$ are roots of the original quartic equation.

Example 4 The transition equation for the Hexatic equation

$$x^5 + a_4x^4 + a_3x^3 + a_2\ x^2\ + a_1\ x\ \ +\ a_0\ =\ 0$$

reads as,

$$X^5 - (a_4^2 - 2a_3)\,X^4 + (a_3^2 - 2a_2a_4 + 2a_1)\,X^3 - (a_2^2 - 2a_1a_3 + 2a_0a_4)\,X^2 + (a_1^2 - 2\ a_0\ a_2)\ X\ - a_0^2 = 0.$$

Therefore, for the specific equation $x^5 + 2x^4 - 5\,x^3 - 10\,x^2\ + 6\,x\ + 12 = 0$ the transition equation is $x^5 - 14x^4 + 77\,x^3 - 208\,x^2\ + 276\,x\ - 144 = 0$. Since the transition equation has roots $X = 2,\ 3$, and 4, one can check that $x = \pm\sqrt{2},\ -2,\ \text{and}\ \pm\sqrt{3}$ are roots of the original quintic equation.

Example 5 The transition equation for the quintic equation

$$x^6 + a_5x^5 + a_4x^4 + a_3x^3 + a_2\ x^2\ + a_1\ x\ \ +\ a_0\ =\ 0$$

reads as follows

$$X^6 - (a_5^2\ -\ 2a_4\ )\,X^5 + (a_4^2 - 2a_3a_5 + 2a_2)\,X^4 - (a_3^2 - 2a_2a_4 + 2a_1a_5 - 2a_0)\,X^3$$
$$+ (a_2^2 - 2a_1a_3 + 2a_0a_4)\,X^2 - (a_1^2 - 2\ a_0\ a_2)\ X\ + a_0^2 = 0.$$

Example 6 An application of theorem 2 will show that the transition equation for the quartic equation $a_4x^4 + a_3x^3 + a_2 x^2 + a_1 x + a_0 = 0$ is as follows

$$X^4 - (a_3^2 - 2a_2a_4)X^3 + a_4^2(a_2^2 - 2a_1a_3 + 2a_0a_4)X^2 - a_4^4(a_1^2 - 2a_0a_2)X + a_4^6a_0^2 = 0.$$

For example, in the case of the equation $2x^4 + 2x^3 - x^2 - 3x - 3 = 0$ the transition function will read as $X^4 - 8X^3 + 4X^2 - 48X + 576 = 0$. An application of the integral zero theorem will show that $X = 6$ is a (double) root of the transition equation, and therefore both $x = \pm\sqrt{6/4} = \pm\sqrt{6}/2$ are 'square root' zeros of the original quadratic equation.