

THE GENERALIZATION of THE HURWITZ THEOREM

Ender DOGAN,
Dr. Vilademir TEN

We use polynomials as a relationship between an input and an output. To find the result we should make the polynomial more specific expression that is an equation. Solutions of equation are the roots. The aim is to find exact results. But sometimes it may not happen. Then we try to estimate an interval of solution by use of Hurwitz theorem.

The goal of this project is the solution of the following problem:

Let $f(z) = a_0 + a_1z + \dots + a_mz^m$ be a polynomial with real coefficients. Find the necessary and sufficient conditions which coefficients of the a_0, a_1, \dots, a_m of the polynomial $f(z)$ have to satisfy such that all roots of this polynomial to :

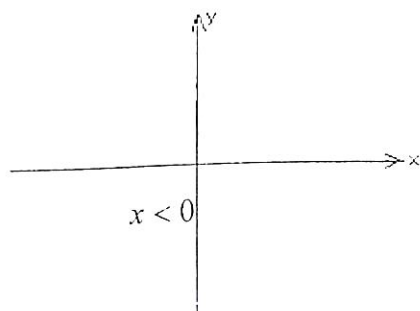
a). Semi-plane $x < r$.

b). Semi-plane $x > r$.

c). Domain $r_1 < x < r_2$.

of the complex plane $z = x + iy$, where r, r_1, r_2 are real numbers.

The problem we formulated is the generalization of classic Hurwitz theorem that was proved in 1895 year which gives us the necessary and sufficient conditions for coefficients of polynomials with real coefficients such that roots of this polynomial belong to the left semi-plane of the complex plane $z = x + iy$, i.e. $x < 0$.



Hurwitz theorem is closely connected with the problem of stability of solutions of differential equations. Project contains two parts. The first part contains some definitions and statements from the course of higher algebra which we need for the proof of the main result. They are:

a). Some facts about complex numbers and polynomials.

b). Some facts about matrices and determinants. Hurwitz theorem.

In the second part we prove the main theorem and offer the computer realization of our problem.

§1. Complex numbers.

Let us consider the equation $x^2 + 1 = 0$ (1)
it has no solution in \mathbf{R} , where \mathbf{R} is the field of real numbers. The goal of this paragraph to construct a field such that equation (1) has a solution in this field.

Now let's introduce the new number $\sqrt{-1}$ and denote it as i . And let's determine by definition:

a). Product βi of any real number β by the number i .

б). A sum $\alpha + \beta i$ for any real numbers α, β .

Definition 1.1. Let $z_1 = \alpha + \beta i$, $z_2 = \gamma + \delta i$. Then by definition $z_1 = z_2$ if and only if $\alpha = \gamma, \beta = \delta$.

§ 2. Polynomials and their roots.

Let F is a field, x is a variable. If, $\alpha_0, \alpha_1, \dots, \alpha_n \in F$ then an expression

$$f(x) = \alpha_0 + \alpha_1 x + \dots + \alpha_n x^n$$

is called a *polynomial* of one variable x with coefficients $\alpha_0, \alpha_1, \dots, \alpha_n$ of a field F .

We denote the set of all polynomials with coefficients from F as $F[x]$.

Definition 1.6 Let n be the greatest number that $\alpha_n \neq 0$, then n is called the degree of the polynomial $f(x)$ and denoted as $\deg f$.

Remark a). $\deg f = 0$ if and only if $f(x) = \alpha \in F$.

b). If $f(x) = 0$ then $f(x)$ has no degree.

Let $f(x), g(x) \in F[x]$ and $f(x) = \alpha_0 + \alpha_1 x + \dots + \alpha_n x^n$, $\alpha_n \neq 0$ (1)

$$g(x) = \beta_0 + \beta_1 x + \dots + \beta_m x^m, \quad \beta_m \neq 0 \quad (2)$$

Definition 1.7. $f(x) = g(x)$ if and only if $n = m$ and $\alpha_k = \beta_k$, $k = 0, 1, \dots, n$.

Definition 1.8. a). If $n \geq m$,

then $f(x) + g(x) = \gamma_0 + \gamma_1 x + \dots + \gamma_n x^n$, $\gamma_k = \alpha_k + \beta_k$, $k = 0, 1, \dots, n$

$f(x) + g(x)$ is called the sum of polynomials $f(x), g(x)$.

b). If $\alpha \in F$, then $\alpha \cdot f(x) = \alpha \alpha_0 + \alpha \alpha_1 x + \dots + \alpha \alpha_n x^n$.

c) $f(x)g(x) = \gamma_0 + \gamma_1 x + \dots + \gamma_{n+m} x^{n+m}$,

$$\gamma_k = \sum_{i+l=k} \alpha_i \beta_l = \alpha_0 \beta_k + \alpha_1 \beta_{k-1} + \dots + \alpha_k \beta_0, \quad k = 0, 1, \dots, n.$$

§3. Matrices and determinants. Hurwitz theorem.

Definition 1.16. Let F be a field. An $m \times n$ -matrix A is a rectangular array of scalars from F with m rows and n columns. Each scalar in the matrix is called an *element*

of the matrix A .

So if

$$A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \cdot & \cdot & \cdot \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \quad (1)$$

a_{ij} is the scalar in row i and column j of the matrix A . The short notation of (1) is $A = \|a_{ij}\|$. If $m = n$ then A is called a square matrix of the order n . Diagonal of a square matrix that is formed by elements $a_{11}, a_{22}, \dots, a_{nn}$ is called the main diagonal.

Let's consider the following matrix

$$\begin{pmatrix} a_1 & a_0 & 0 & 0 & \dots & 0 \\ a_3 & a_2 & a_1 & a_0 & \dots & 0 \\ a_5 & a_4 & a_3 & a_2 & \dots & 0 \\ \cdot & \cdot & \cdot & \cdot & \dots & \cdot \\ 0 & 0 & 0 & 0 & \dots & a_n \end{pmatrix} \quad (3)$$

Definition 1.18. Matrix (3) is called the Hurwitz matrix of polynomial (2).

Now let's consider the following determinants:

$$\Delta_1 = a_1, \quad \Delta_2 = \begin{vmatrix} a_1 & a_0 \\ a_3 & a_2 \end{vmatrix}, \quad \Delta_3 = \begin{vmatrix} a_1 & a_0 & 0 \\ a_3 & a_2 & a_1 \\ a_5 & a_4 & a_3 \end{vmatrix}, \dots,$$

$$\Delta_n = \begin{vmatrix} a_1 & a_0 & 0 & \dots & 0 \\ a_3 & a_2 & a_1 & \dots & 0 \\ a_5 & a_4 & a_3 & \dots & 0 \\ \cdot & \cdot & \cdot & \dots & \cdot \\ 0 & 0 & 0 & \dots & a_n \end{vmatrix}.$$

Theorem 1.13. (Hurwitz). Polynomial (2) is Hurwitz polynomial if and only if

$$\Delta_k > 0, \quad k = 1, 2, \dots, n.$$

Generalization of Hurwitz theorem.

The main theorem:

In this paragraph we consider the proof of the main theorem of our project.

Let

$$f(z) = a_0 + a_1 z + \dots + a_n z^n \quad (1)$$

be a polynomial with real coefficients. Let's consider the following polynomials

$$p_r(z) = f(z+r), \quad (2)$$

$$q_r(z) = f(-z+r). \quad (3)$$

Lemma 2.1.

a). Coefficients of polynomial (2) are connected with coefficients of polynomial (1) by the formulas:

$$b_k = \sum_{l=k}^n C_l^{l-k} r^{l-k} a_l, \quad k = 0, 1, \dots, n,$$

where $C_l^i = \frac{l!}{i!(l-i)!}$.

b). Coefficients of polynomial (3) are connected with coefficients of polynomial (1) by the formulas:

$$c_k = (-1)^k \sum_{l=k}^n C_l^{l-k} r^{l-k} a_l, \quad k = 0, 1, \dots, n.$$

Proof.

a). By definition

$$p_r(z) = f(z+r) = \sum_{l=0}^n a_l (z+r)^l = \sum_{l=0}^n a_l (z+r)^l. \quad (4)$$

Let's consider polynomials $(z+r)^l$. It is clear that if $l < k$ then the polynomial $(z+r)^l$ does not contain z^k , i.e. z^k the zero coefficient. So z^k has nonzero coefficients after opening brackets in $(z+r)^l$ if and only if $k \leq l \leq n$. By Newton binomial formula we have:

$$\begin{aligned} (z+r)^k &= C_k^0 z^k + k z^{k-1} r + \dots + r^k \\ \dots\dots\dots \\ (z+r)^l &= z^l + \dots + C_l^{l-k} r^{l-k} z^k + \dots + r^l \\ \dots\dots\dots \\ (z+r)^n &= z^n + \dots + C_n^{n-k} r^{n-k} z^k + \dots + r^n. \end{aligned}$$

So the corresponding coefficients of z^k will be

$$\begin{aligned} &C_k^0 \\ &\dots\dots\dots \\ &C_l^{l-k} r^{l-k} \\ &\dots\dots\dots \\ &C_n^{n-k} r^{n-k}. \end{aligned} \quad (5)$$

Then from (4) и (5) we have

$$b_k = \sum_{l=k}^n C_l^{l-k} r^{l-k} a_l.$$

b). Again by definition we have $q_r(z) = f(-z+r) = \sum_{l=0}^n a_l (-z+r)^l. \quad (6)$

If we repeat the explanations we use in a) we receive that coefficients of z^k in the polynomial $q_r(z)$ are equal to

$$c_k = (-1)^k C_l^{l-k} r^{l-k} a_l, \quad k = 0, 1, \dots, n.$$

Now we may formulate and prove the main theorem of our project.

Theorem 2.1.

Roots of the polynomial $f(z)$ belong to

a). The semi-plane $\operatorname{Re} z < r$ if and only if the polynomial $p_r(z)$ is the Hurwitz polynomial.

б). The semi-plane $\operatorname{Re} z > r$ if and only if the polynomial $q_r(z)$ is the Hurwitz polynomial.

в). The domain $r_1 < \operatorname{Re} z < r_2$ if and only if the polynomials $q_{r_1}(z)$ and $p_{r_2}(z)$ are the Hurwitz polynomials.

Proof.

а). Let α be a root of the polynomial $f(z)$, i.e. $f(\alpha) = 0$. Then no problem to see that $\alpha - r$ is the root of the polynomial $p_r(z) = f(z + r)$ as

$$p_r(\alpha - r) = f(\alpha - r + r) = f(\alpha) = 0.$$

So if $\alpha_1, \dots, \alpha_n$ are all roots of the polynomial $f(z)$ then $\alpha_1 - r, \dots, \alpha_n - r$ are all roots of the polynomial $p_r(z)$. So the polynomial $p_r(z)$ is the Hurwitz polynomial if and only if

$$\operatorname{Re}(\alpha_k - r) = \operatorname{Re} \alpha_k - r < 0, \quad k = 0, 1, \dots, n,$$

or

$$\operatorname{Re} \alpha_k < r, \quad k = 0, 1, \dots, n.$$

б). If α is a root of the polynomial $f(z)$ then $-\alpha + r$ is the root of the polynomial $q_r(z)$ as

$$q_r(-\alpha + r) = f(-(-\alpha + r) + r) = f(\alpha) = 0.$$

So if $\alpha_1, \dots, \alpha_n$ are all roots of the polynomial $f(z)$ then $-\alpha_1 + r, \dots, -\alpha_n + r$ are all roots of the polynomial $q_r(z)$. So $q_r(z)$ is the Hurwitz polynomial if and only if

$$\operatorname{Re}(-\alpha_k + r) = -\operatorname{Re} \alpha_k + r < 0, \quad k = 0, 1, \dots, n$$

or

$$\operatorname{Re} \alpha_k > r, \quad k = 0, 1, \dots, n.$$

в). The proof of this item immediately follows from items а) and б) of our theorem.

Now let's consider the following general problem:

Let $f(z)$ be a polynomial with real coefficients, $y = kx + b$ be a line in the complex plane $z = x + yi$. Find the necessary and sufficient conditions the coefficients of the polynomial $f(z)$ satisfy such that all roots of the polynomial belong to

а). In the semi-plane $y < kx + b$.

б). In the semi-plane $y > kx + b$.

The following example shows that in general case this problem has no solution.

Example. Let's consider the line $y = 0$. If $f(z)$ is any polynomial with real coefficients and α is a complex root of $f(z)$, i.e. $\operatorname{Im} \alpha \neq 0$ then by theorem 1.12. it follows that conjugate number $\bar{\alpha}$ is the root of $f(z)$ too. So all roots of any polynomial with real coefficients can not belong to the semi-plane $y < 0$ or to the semi-plane $y > 0$.

References:

1. А.И. Кострикин. Введение в алгебру. Наука, Москва, 1977.
2. А.Г. Курош. Курс высшей алгебры. Наука, Москва, 1971.

Резюме

В данной работе с помощью теоремы Гурвица, разработана связь между коэффициентами многочлена, $f(z) = a_0 + a_1 z + \dots + a_m z^m$, которые являются вещественными членами, и корнем многочлена, который определяется комплексными членами.