24. Sturm's Problem of the Number of Roots

Find the number of real roots of a polynomial equation with real coefficients over a given interval.

This very important algebraic problem was solved in a surprisingly simple way in 1829 by the French mathematician Charles Sturm (1803-1855). The paper containing the famous Sturm theorem appeared in the eleventh volume of the *Bulletin des sciences de Férussac* and bears the title "Mémoire sur las résolution des équations numériques".

"With this major discovery", says Liousville, "Sturm at once simplified and perfected the elements of algebra, enriching them with new results".

SOLUTION: We consider two cases:

I. The real roots of the equation in question are all simple over the given interval.

II. The equation has multiple real roots over the interval.

We will first show that the second case leads back to the first.

Let the given polynomial equation be $F(x) = 0$ have distinct real roots $\alpha, \beta, \gamma, \ldots$ with multiplicities $a, b, c, \ldots,$ so that

$$F(x) = (x-\alpha)^a (x-\beta)^b (x-\gamma)^c \ldots$$

Then

$$\begin{aligned} \frac{F'(x)}{F(x)} &= \frac{a}{x-\alpha} + \frac{b}{x-\beta} + \frac{c}{x-\gamma} + \ldots \\ &= \frac{a(x-\beta)(x-\gamma)\ldots + b(x-\alpha)(x-\gamma)\ldots + c(x-\alpha)(x-\beta)(x-\delta)\ldots + \ldots}{(x-\alpha)(x-\beta)(x-\gamma)(x-\delta)\ldots} \end{aligned}$$

Let

$$\begin{aligned} p(x) &= a(x-\beta)(x-\gamma)\ldots + b(x-\alpha)(x-\gamma)\ldots + c(x-\alpha)(x-\beta)(x-\delta)\ldots + \ldots \\ q(x) &= (x-\alpha)(x-\beta)(x-\gamma)(x-\delta)\ldots \\ G(x) &= \frac{F(x)}{q(x)} = (x-\alpha)^{a-1}(x-\beta)^{b-1}(x-\gamma)^{c-1}\ldots \end{aligned}$$

Then

$$F(x) = G(x) \cdot q(x) \quad \text{and} \quad F'(x) = F(x)\frac{p(x)}{q(x)} = G(x) \cdot p(x).$$

Since $p(x)$ and $q(x)$ have no common (non-constant) divisor, $G(x)$ (or any constant multiple of it) is the greatest common divisor of $F(x)$ and $F'(x)$. This GCD can be easily found from the Euclidean algorithm, and can therefore be considered known, as a result of which $q(x) = \frac{F(x)}{G(x)}$ is also known (but not necessarily in factored form).

The equation $F(x) = 0$ then is equivalent to

$$q(x) = 0 \quad \text{and} \quad G(x) = 0,$$

the first of which has only simple roots, while the second can be further reduced in the same way that $F(x) = 0$ was. An equation with multiple roots can therefore always be transformed into equations (with know coefficients) having only simple roots. Thus it is sufficient to solve the problem for the first case.

Let $f(x) = 0$ be a polynomial equation all of whose roots are simple. The derivative $f'(x)$ of $f(x)$ then vanishes for none of these roots and the greatest common divisor of $f(x)$ and $f'(x)$ is a non-zero constant K. We can use the Euclidean algorithm to determine the GCD of $f(x)$ and $f'(x)$, writing $f_0(x)$ and $f_1(x)$ for $f(x)$ and $f'(x)$ for convenience, and calling the quotients and remainders in the divisions $q_0(x), q_1(x), q_2(x), \ldots,\ -f_2(x), -f_3(x), \ldots.$ Dropping x for brevity, we obtain

(0) $\qquad f_0 = q_0 f_1 - f_2$

(1) $\qquad f_1 = q_1 f_2 - f_3$

(2) $\qquad f_2 = q_2 f_3 - f_4$, etc.

Let the last non-zero constant remainder by $-f_s(x)$ (or K). This is where the Euclidean algorithm stops. The functions

$$f_0, f_1, f_2, \ldots, f_s$$

form a "*Sturm chain*", and in this setting are called *Sturm functions*. The Sturm functions have the following three properties:

1. Two neighboring functions do not vanish simultaneously at any point of the interval.
2. At a zero of a Sturm function, its two neighboring functions have different signs.
3. Within a sufficiently small interval about a zero of $f_0(x)$, $f_1(x)$ is everywhere greater than zero or everywhere less than zero.

Proof of 1 If for example f_2 and f_3 vanish at any point of an interval, f_4 [by (2)] also vanishes at this point, and consequently f_5 also, etc., so that finally f_s, also vanishes there, which however contradicts the fact that f_s is a non-zero constant.

Proof of 2. If the function f_3 vanishes at the point σ, for example, then it follows from (2) that $f_2(\sigma) = -f_4(\sigma)$.

Proof of 3. At α for example $f(x) = (x - \alpha)q(\alpha)$ where $q(\alpha) \neq 0$. Then $f'(\alpha) = q(\alpha)$ is non-zero. q is continuous, and hence has a constant sign (the sign of $q(\alpha)$) in some small interval about α.

We now choose any point x in the given interval, record the signs of $f_0(x), f_1(x), \ldots, f_s(x)$, and thus obtain a *Sturm sign chain*. (It is assumed here that none of the $s+1$ values is zero.) Two consecutive signs can be $++, --, +-$ and $-+,$ the latter two being called sign changes. We shall consider the number $Z(x)$ of sign changes undergone by $Z(x)$ when x passes through the interval. A change can occur only if one or more of the Sturm functions changes sign. We will thus study the effect produced on $Z(x)$ by the passage of a function $f_v(x)$ through zero. Note that $v < s,$ since $f_s(x)$ is a constant non-zero function.

Let k be a point at which $f_v(k) = 0,\ h < k < l$ so that on the interval $[h,l]$

1. $f_v(x) \neq 0,$ except at $k,$

2. f_{v+1} does not change sign, and

3. if $v > 0,\ f_{v-1}$ does not change sign.

In case $v > 0,$ we are concerned with the triplet $f_{v-1}, f_v, f_{v+1},$ and if $v > 0$ the pair $f_0 = f$ and $f_1 = f'$.

In the triplet case, f_{v-1} and f_{v+1} are each either positive or negative at all three points h,k,l but opposite in sign. Thus, whatever the sign of f_v at these points, the triplet f_{v-1}, f_v, f_{v+1} has just one change of sign for each of $h,k,l.$ The passage through a zero of f_v thus does not affect the number of sign changes in the Sturm chain.

In the pair case, f_1 is either positive or negative at all three points $h,k,l.$ If $f_1(h) > 0,\ f$ is increasing and $f(h) < f(k) = 0 < f(l).$ In the second case $f(h) > f(k) = 0 > f(l).$ In both cases a sign change is lost (since f_1 does not change sign on $[h,l]$).

Thus the Sturm sign chain undergoes a change in the number of sign changes $Z(x)$ only when x passes through a zero of $f(x)$; and specifically, the chain then loses exactly one sign change. Thus, if x passes through the interval (the ends of which are not zeros of $f(x)$) from left to right, the sign chain loses exactly as many sign changes as there are roots of $f(x)$ within the interval. We have

Sturm's Theorem: *The number of real roots of a polynomial equation with real coefficients all of whose roots are simple over an interval, the end points of which are not roots, is equal to the difference between the numbers of sign changes for the Sturm sign chains formed for the interval ends.*

Note. The same conclusion results if we multiply $f_0, f_1, f_2, \ldots, f_s$ by any positive constants; this series is then likewise called a Sturm chain. In the formation of the Sturm chain, fractions can be avoided (if the coefficients of f are all integers).

Example 1. Determine the number and location of the real roots of the equation $x^5 - 3x - 1 = 0.$

The Sturm chain is

$$f_0 = x^5 - 3x - 1,\quad f_1 = 5x^4 - 3,\quad f_2 = 12x + 5,\quad f_3 = 1,$$

and the signs for $x = -2, -1, 0, 1, 2$ are

x	f_0	f_1	f_2	f_3	$Z(x)$
-2	$-$	$+$	$-$	$+$	3
-1	$+$	$+$	$-$	$+$	2
0	$-$	$-$	$+$	$+$	1
1	$-$	$+$	$+$	$+$	1
2	$+$	$+$	$+$	$+$	0

The equation thus has three real roots: one between -2 and -1, one between -1 and 0, and one between 1 and 2. The other two roots are complex.

Example 2. Determine the number of real roots of the equation $x^5 - ax - b = 0$ when $a, b > 0$ and $4^4 a^5 > 5^5 b^4$.

The Sturm chain is

$$x^5 - ax - b, \quad 5x^4 - a, \quad 4ax + 5b, \quad 4^4 a^5 - 5^5 b^4,$$

and the sign changes from $-\infty$ to ∞ are

x	f_0	f_1	f_2	f_3	$Z(x)$
$-\infty$	$-$	$+$	$-$	$+$	3
∞	$+$	$+$	$+$	$+$	0

The equation has three real roots and two complex roots.