STURM'S METHOD for the number of real roots of a real polynomial. 1

by K. N. RAGHAVAN, IMSc, Chennai

PRELUDE HINDSIGHT. Understanding of a situation or event *after* it has happened or developed.

~~Some~~ Many things in mathematics make sense only in hindsight. But to have hindsight, one must move ahead.

UPSHOT: As teachers, we must strive to know more material than we actually teach.

- Material of this lecture is not something that you are likely to teach but which nevertheless is good to know. ∵ it provides hindsight.
- Applying the same logic to this lecture itself, there is more material in the printed notes than we will cover in the lecture.

Polynomial $p(x) = a_n x^n + a_{n-1} x^{n-1} + \ldots + a_0$; $n \geq 0$, a_i real, $a_n \neq 0$.

Case 0. $p(x)$ has degree 0. Then it is a non-zero constant and has no roots

Case 1. $\deg p(x) = 1$. E.g. $3x-7$. Has exactly one root, viz., $7/3$

- There exists unique real root.
- It is rational if the coeffs of $p(x)$ are rational.

Case 2 $\deg p(x) = 2$. E.g. x^2+6x-2. Method of "Completion of Squares"

to find the roots. $x^2+6x-2=0 \overset{\text{rewrite}}{\rightsquigarrow} x^2+6x+9 = 2+9$

$\rightsquigarrow (x+3)^2 = 11 \rightsquigarrow x+3 = \pm\sqrt{11} \rightsquigarrow x = -3 \pm \sqrt{11}$.

Called SRIDHARACHARYA's method. Known for at least a 1000 years now.

General formula for the roots of ax^2+bx+c: $(-b \pm \sqrt{b^2-4ac})/2a$
$a \neq 0$

In particular, the roots are real iff $b^2-4ac \geq 0$.

Case 3 $\deg p(x) = 3$. ~~We now~~ General formula for the roots found in 16^{th} century.

Following is an illustration of a variation of the method of TARTAGLIA-CARDANO:

E.g. $3x^3 - 9x^2 + 6x - 1 = 0$. Idea: convert to the form $4z^3 - 3z = K$, constant

Put $z = \cos\theta$, so that lhs becomes $\cos 3\theta = K$. $\rightsquigarrow 3\theta = \cos^{-1}K \rightsquigarrow \theta = \frac{1}{3}\cos^{-1}K$

$\rightsquigarrow z = \cos\left(\frac{1}{3}\cos^{-1}K\right) \rightsquigarrow$ analogous to extracting ~~sq~~ cube root.

Put $x = (y+1)$. We get $3y^3 - 3y - 1 = 0$. Put $y = \frac{2}{\sqrt{3}} z$ and multiply by $\frac{\sqrt{3}}{2}$

$4z^3 - 3z = \frac{\sqrt{3}}{2}$. $z = \cos\left(\frac{1}{3}\cos^{-1}\frac{\sqrt{3}}{2}\right)$. So $z = \cos 10^\circ, \cos 130^\circ, \cos 250^\circ$

$y = \frac{2}{\sqrt{3}}\cos 10^\circ, \frac{2}{\sqrt{3}}\cos 130^\circ, \frac{2}{\sqrt{3}}\cos 250^\circ$

$x = 1 + \frac{2}{\sqrt{3}}\cos 10^\circ, 1 + \frac{2}{\sqrt{3}}\cos 130^\circ, 1 + \frac{2}{\sqrt{3}}\cos 250^\circ$

Case 4 $\deg p(x) = 4$. General formula for roots found in 18^{th} Century. LAGRANGE

ABEL-RUFFINI (early 19^{th} Century) No general formula exists for degrees ≥ 5.

ROOTS & THEIR ORDERS (MULTIPLICITIES)

$\alpha \in \mathbb{C}$ is called a <u>root</u> of $p(x)$ if $p(\alpha) = 0$, or, equivalently if $x - \alpha$ divides $p(x)$.

Proof of this equivalence:

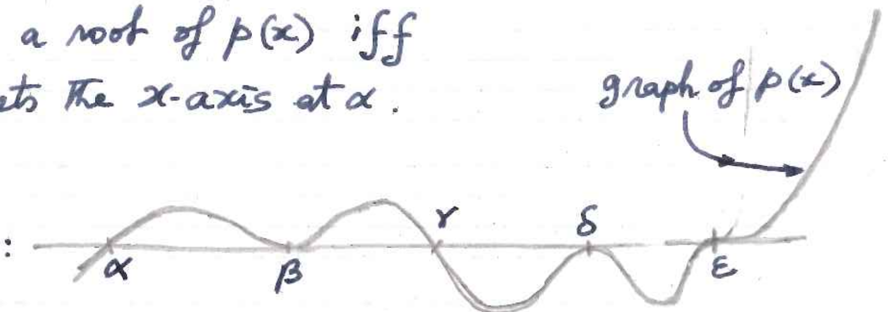
By the division algorithm: $p(x) = (x-\alpha)\, q(x) + r$ ← constant

So: $\boxed{p(\alpha) = r}$ ← This equality is called REMAINDER THEOREM

Clearly, $(x-\alpha)$ divides $p(x)$ iff $r = 0$ but $r = p(\alpha)$. ∎

Geometrically, $\alpha \in \mathbb{R}$ is a root of $p(x)$ iff the graph of $p(x)$ meets the x-axis at α.

$\alpha, \beta, \gamma, \delta$, and ε are the roots of $p(x)$ in the picture:



Multiplicity (or Order) of a root: $\alpha \in \mathbb{C}$ is a <u>root of order n</u> of $p(x)$ if n is the largest among $0, 1, 2, 3, \ldots$ s/t $(x-\alpha)^n$ divides $p(x)$.

- This order cannot exceed the degree of $p(x)$
- A root of order 0 means not a root at all
- <u>Simple root</u>: root of order 1
- <u>repeated root</u>: root of order ≥ 2

<u>PROPOSITION</u>: If α is a root of order $n \geq 1$ of $p(x)$, then α is a root of order $n-1$ of the derivative $p'(x)$.

<u>CAUTION</u>: α could be a root (even of high order) of $p'(x)$ without it being a root of $p(x)$.

<u>PROOF</u>: Write $p(x) = (x-\alpha)^n q(x)$ with $q(\alpha) \neq 0$.

$$p'(x) = n(x-\alpha)^{n-1} q(x) + (x-\alpha)^n q'(x)$$
$$= (x-\alpha)^{n-1} s(x), \quad \text{where } s(x) = n\,q(x) + (x-\alpha)\,q'(x).$$

Note $s(\alpha) = n\,q(\alpha) \neq 0$. ∎

Geometrically, $\alpha \in \mathbb{R}$ is a repeated root of $p(x)$ iff the tangent to the graph of $p(x)$ at $x = \alpha$ is the x-axis.

E.g. in the picture above, β, δ, and ε are repeated roots

α and γ are simple roots

REPEATED ROOTS & THE G.C.D. OF $p(x)$ and $p'(x)$

Set $d(x) :=$ G.C.D. of $p(x)$ and $p'(x)$.

COROLLARY: The roots of $d(x)$ are precisely the repeated roots of $p(x)$.
More precisely, if α is a root of order $n \geq 1$ of $p(x)$, then it is a root of order $n-1$ of $d(x)$.

In particular:
- $p(x)/d(x)$ has no repeated roots
- $p(x)$ has no repeated real roots iff $d(x)$ has no real roots.

$d(x)$ can be computed READILY by Euclid's algorithm
If $p(x)$ has rational coeffs, then all polynomials in the algorithm also have rational coeffs.

Unlike for polynomials, there is no READY algorithm known to determine whether a given integer is square-free (factorization is impractical at least as of now). For details, see IMSc Youtube video lecture "Some simple open problems in Mathematics" by OESTERLE.

CANONICAL SEQUENCE associated to $p(x)$:
We define inductively this sequence $p_0(x), \ldots, p_m(x)$ of polynomials with $p_0(x) = p(x)$.
Set $p_0(x) := p(x)$. If $p(x) = \text{const}$, then we stop with $p_0(x)$. Otherwise set $p_1(x) = p'(x)$
For $i \geq 2$, define $p_i(x)$ inductively as follows: if $p_{i-2}(x)$ divides $p_{i-1}(x)$, then $p_i(x)$ is <u>not</u> defined. Otherwise, $\boxed{p_i(x) := -\text{remainder when } p_{i-2}(x) \text{ is divided by } p_{i-1}(x)}$

EXAMPLE: $p_0(x) = p(x) = x^5 + x^2 + 1$, $p_1(x) = p'(x) = 5x^4 + 2x$

$$5x^4+2x \,\big)\, x^5 + x^2 + 1 \,\big(\, \tfrac{1}{5}x$$
$$x^5 + \tfrac{2}{5}x^2$$
$$\tfrac{3}{5}x^2 + 1$$

$p_2(x) = -\frac{3}{5}x^2 - 1$

$p_3(x) = -2x - \frac{125}{9}$

$p_4(x) = +\frac{125 \times 25}{9 \times 12} + 1$

Stops here.

$$-\tfrac{3}{5}x^2 - 1 \,\big)\, 5x^4 + 2x \,\big(\, -\tfrac{25}{3}x^2 + \tfrac{125}{9}$$
$$5x^4 + \tfrac{25}{3}x^2$$
$$-\tfrac{25}{3}x^2 + 2x$$
$$-\tfrac{25}{3}x^2 - \tfrac{125}{9}$$
$$2x + \tfrac{125}{9}$$

$$-2x - \tfrac{125}{9} \,\big)\, -\tfrac{3}{5}x^2 - 1 \,\big(\, \tfrac{3}{10}x - \tfrac{25}{12}$$
$$-\tfrac{3}{5}x^2 - \tfrac{25}{6}x$$
$$\tfrac{25}{6}x - 1$$
$$\tfrac{25}{6}x + \tfrac{125 \times 25}{9 \times 12}$$
$$-\tfrac{125 \times 25}{9 \times 12} - 1$$

<u>Example of canonical sequence</u> (refer to calculation above)

$p_0(x) = p(x) = x^5 + x^2 + 1$, $p_1(x) = p'(x) = 5x^4 + 2x$, $p_2(x) = -\frac{3}{5}x^2 - 1$,

$p_3(x) = -2x - \frac{125}{9}$, $p_4(x) = \frac{125 \times 25 + 9 \times 12}{9 \times 12}$

<u>Sturm's theorem:</u> Suppose that $p(x)$ has no repeated root.

Consider its canonical sequence $p_0(x), p_1(x), \ldots, p_m(x)$.

As $x \to -\infty$, each $p_i(x)$ tends to either $+\infty$ or $-\infty$ or some (non-zero) constant

Note down only the signs of these limits.

Eg In the example above, those are $-, +, -, +, +$

Denote by $\sigma(-\infty)$ the number of sign changes. E.g. it is 3 above

Iy As $x \to \infty$, each $p_i(x)$ tends to either $+\infty$ or $-\infty$ or some (non-zero) constant.

Note down only the signs of these limits.

In the example above, these are $+, +, -, -, +$

Denote by $\sigma(\infty)$ the number of signchanges. E.g. it is 2 above.

$\boxed{\text{The \# of real roots of } p(x) = \sigma(-\infty) - \sigma(\infty)}$ E.g. it is 3-2=1 above.

<u>Proof:</u> We track the "sign-change-number-function $\sigma(x)$" as x moves from $-\infty$ to ∞.

Definition of $\sigma(x)$: Given $x \in \mathbb{R}$, note down whether each $p_i(x)$ is +ve, −ve, or 0.

In the above example, for $x = -2$ we get $-, +, -, -, +$

for $x = 0$ we get $+, 0, -, -, +$

for $x = 1$ we get $+, +, -, -, +$

$\sigma(x)$ is the number of sign changes in the sequence ignoring zeros: $\sigma(-2) = 3$, $\sigma(0) = 2$, $\sigma(1) = 2$

Evidently, for $\sigma(x)$ to change, we must pass a root α of one of the $p_i(x)$.

By construction, we have $\boxed{p_{j-1}(x) = p_j(x)\, q(x) - p_{j+1}(x)}$ —— $(\star)$

<u>Claim:</u> No real α can be a root of two consecutive polynomials, say $p_i(x)$ & $p_{i+1}(x)$.

<u>Pf of claim:</u> Suppose $p_i(\alpha) = p_{i+1}(\alpha) = 0$. Then, by $(\star)$ with $j = i$, we conclude $p_{i-1}(\alpha) = 0$.

Again by $(\star)$ with $j = i-1$, we conclude $p_{i-2}(\alpha) = 0$. So continuing, we get $p_1(\alpha) = p_0(\alpha) = 0$.

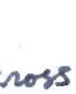
But $p_0(x) = p(x)$ has no repeated root by assumption and $p_1(x) = p'(x)$. $\Rightarrow\Leftarrow$.

<u>Case 1:</u> $p_i(\alpha) = 0$ for some $i \geq 1$. Then, by $(\star)$ with $j = i$, $p_{i-1}(\alpha)$ & $p_{i+1}(\alpha)$ have opposite signs.

Close to α, the signs look like:

	$p_{i-1}(x)$	$p_i(x)$	$p_{i+1}(x)$
either	+	?	−
or	−	?	+

Whatever the behaviour of $p_i(x)$ near α, it has no effect on $\sigma(x)$ across α.

<u>Case 2:</u> Suppose $p_0(\alpha) = p(\alpha) = 0$. Then, close to α, the graph of $p(x)$ looks like  or

	$p_0(x)$	$p_1(x)$	$p_0(x)$	$p_1(x)$
Correspondingly the signs of $p_0(x)$ and $p_1(x)$ are: $x < \alpha$:	−	+	+	−
$x = \alpha$:	0	+	0	−
$x > \alpha$:	+	+	−	−

In either case, $\sigma(x)$ decreases precisely by 1 as x moves across α from left to right. ▨