STURM'S METHOD FOR THE NUMBER OF REAL ROOTS OF A REAL POLYNOMIAL

K. N. RAGHAVAN

Let $p(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$ be a non-zero polynomial of degree $n \geq 0$ with real coefficients (i.e., n denotes a non-negative integer, the coefficients a_i are all real, and the leading coefficient a_n is non-zero). We describe a method, due to Sturm, dating from the 1820s, to determine the number of real roots of such a polynomial in any given interval on the real line. If the coefficients of the given polynomial and the end points of the given interval are rational, then the required calculations can all be carried out over the rational field.

Sturm's method and others related to it are explained and commented upon in many places, e.g., on wikipedia, but due diligence is advised when using these resources. In particular, one should be sensitive to variations in terminology.

Reduction of the problem to the case when $p(x)$ admits no repeated root. To solve our problem of determining the number of real roots of $p(x)$ in a given interval, we may assume without loss of generality that $p(x)$ has no repreated roots. Indeed, letting $g(x)$ be the g.c.d. of the given polynomial $p(x)$ and its derivative $p'(x)$, we observe that $p(x)/g(x)$ has no repeated roots: in fact, each root of $p(x)$ whether simple or repeated appears as a root of $p(x)/g(x)$ precisely once. If we are interested only in the roots of $p(x)$ without regard to their multiplicity, we may thus replace $p(x)$ by $p(x)/g(x)$ and proceed. If, on the other hand, we are interested in counting roots with multiplicity, then we can put the information for $p(x)/g(x)$ together with that for $g(x)$ to obtain the required information for $p(x)$: since $g(x)$ has smaller degree than $p(x)$, we may suppose that we can handle it (by an inductive argument).

The canonical sequence. We define a finite sequence of polynomials starting with the given non-zero polynomial $p(x)$. (We allow $p(x)$ to have repeated roots in this subsection.) Put $p_0(x) = p(x)$. If $p(x)$ is a constant polynomial, then the sequence stops right here. Othewise the derivative $p'(x)$ of $p(x)$ is non-zero, and we put $p_1(x) = p'(x)$. For $i \geq 2$, we define polynomials $p_i(x)$ inductively as follows. They are, up to sign, the (non-zero) remainders that occur in Euclid's algorithm for determing the gcd of $p(x)$ and $p'(x)$. If $p_{i-1}(x)$ divides $p_{i-2}(x)$, then $p_i(x)$ is not defined and the sequence stops

These are notes of a lecture given in September 2016 at a workshop at IMSc for school mathematics teachers. Comments are solicited and may please be sent to the author at knr.imsc@gmail.com or knr@imsc.res.in. Up to date version of these notes is permanently available at http://www.imsc.res.in/∼knr/past/sturm.

at $p_{i-1}(x)$. Otherwise, define:

$$p_i(x) = -\mathbf{rem}(p_{i-2}(x), p_{i-1}(x)) \tag{1}$$

where, for polynomials $a(x)$ and $b(x) \neq 0$, we denote by $\mathbf{rem}(a(x), b(x))$ the remainder when $a(x)$ is divided by $b(x)$.

Observe that $p'(x)$ has degree less than $p(x)$ (in fact, it has degree precisely one less), and that $p_i(x)$, if defined, has smaller degree than $p_{i-1}(x)$. Thus the sequence defined above stops in less than n steps (where n is the degree of $p(x)$):

$$p(x) = p_0(x), \quad p_1(x) = p'(x), \quad p_2(x), \quad \ldots, \quad p_m(x) \quad m \leq n = \deg p(x) \tag{2}$$

The sequence (2) is called the *canonical sequence* associated to the polynomial $p(x)$. The last term $p_m(x)$ of the sequence is, up to sign, the gcd of $p(x)$ and $p'(x)$. It divides all the terms of the sequence.

Sturm sequences. The canonical sequence defined above has many good properties—see (4) and (7) below—that we want to exploit to extract information about the roots of $p(x)$. The definition we now give of Sturm sequences is an attempt to isolate the relevant properties and exclude irrelevant ones.

Suppose we have a (non-empty) finite sequence of polynomials:

$$p_0(x), \quad p_1(x), \quad \ldots, \quad p_m(x) \tag{3}$$

where the initial term $p_0(x)$ is not identically zero. Such a sequence is called a *Sturm sequence associated to* $p_0(x)$ if it has the following properties (a)–(d):

(a) The last term $p_m(x)$ of the sequence is either always positive or always negative on the real line.

(b) No two consecutive $p_i(x)$ are simultaneouly zero for x a real number.

(c) Suppose that α is a real root of $p_i(x)$, for some i with $0 < i < m$. Then $p_{i-1}(\alpha)$ and $p_{i+1}(\alpha)$ have opposite signs (note that neither is zero by (b) above).

(d) At any real root α of $p_0(x)$, the graph of $p(x)$ "crosses" the x-axis at α, or, in other words, the values of $p(x)$ close to α and on either side of α it are of opposite sign.[1] Furthermore, depending upon whether $p_1(\alpha)$ is $+$ve or $-$ve—observe that it is not zero by (b)—the movement of $p(x)$ across α is from $-$ve to $+$ve or vice-versa.

We now prove:

the canonical sequence associated to a polynomial without repeated real roots is a Sturm sequence. (4)

PROOF: Let $p_0(x), \ldots, p_m(x)$ be the canonical sequence associated to a polynomial $p(x)$ without repeated real roots. We need to show that this sequence satisfies conditions (a)–(d) in the definition above.

[1] The graph of $p(x)$ is allowed to be tangential to the x-axis at α, like that of x^3 at 0, although this does not happen in the case when α is not a repeated root of $p(x)$.

The last term $p_m(x)$ being the gcd of $p(x)$ and $p'(x)$, it follows from our assumption about $p(x)$ that $p_m(x)$ has no real root and so is constant in sign. This proves (a).

It follows from (1) that consecutive terms of the sequence (2) are related thus:

$$p_{i-1}(x) = q(x)p_i(x) - p_{i+1}(x) \tag{5}$$

where $q(x)$ is the quotient when $p_{i-1}(x)$ is divided by $p_i(x)$. Thus, if α is a root of two consecutive polynomials, then it is a common root of all the polynomials. But this is not possible if α is real because of our assumption on $p(x)$ (by definition $p_1(x)$ is the derivative of $p_0(x)$ and as such has no real root in common with $p_0(x)$). This proves (b).

We now prove (c). It follows from (b) that neither $p_{i-1}(\alpha)$ nor $p_{i-1}(\alpha)$ is zero. Putting $x = \alpha$ in (5), we see that $p_{i-1}(\alpha) = -p_{i+1}(\alpha)$.

We now prove (d). If α is a real root of $p(x)$, then it is a simple root, and so the graph of $p(x)$ crosses—in fact, "cuts"—the x-axis at α. Since $p_1(x)$ is the derivative of $p(x)$, the rest follows from properties of the derivative. □

Another example of a Sturm sequence. Let $p(x)$ be a real polynomial of degree d with d distinct real roots. Then of course $p(x)$ and its derivative $p'(x)$ have no common roots. Let $\alpha_1, \ldots, \alpha_d$ be the roots of $p(x)$ arranged increasingly. By Rolle's theorem, the derivative $p'(x)$ has a root strictly between α_i and α_{i+1} for every i, $1 \leq i < d$. Since $p'(x)$ has degree $d-1$, it follows that all roots of $p'(x)$ are real and distinct. If $\beta_1, \ldots, \beta_{d-1}$ be the roots of $p'(x)$ arranged increasingly, then we have the "interlacing property": $\alpha_i < \beta_i < \alpha_{i+1}$ for all i, $1 \leq i < d$.

At a root β of $p'(x)$, the value $p''(\beta)$ of the double derivative is positive or negative accordingly as $p(\beta)$ is negative or positive. It is now clear that $p(x)$, $p^{(1)}(x)$, $p^{(2)}(x)$, ..., $p^{(d)}(x)$ is a Sturm sequence: here $p^{(i)}(x)$ denotes the i^{th} derivative of $p(x)$.

More examples of Sturm sequences. The constant polynomials 2, 0, and -1, for example, form a Sturm sequence, as can be readily checked. In particular, Sturm sequences could contain identically vanishing polynomials. For future use, let us record:

If we remove identically zero polynomials from a Sturm sequence the result is also a Sturm sequence. (6)

For the proof, observe that the neighbors in a Sturm sequence of an identically zero polynomial are non-zero constants (of opposite sign, but that is irrelevant for now). Thus every non-constant polynomial in the resulting sequence has the same neighbors as in the original. This proves (b), (c), and (d). The first and last terms of the original sequence being non-zero by definition, they continue to be the first and last terms of the resulting sequence. In particular, (a) holds. □

The following statement too will be used in the sequel:

If we divide all terms in the canonical sequence of a non-zero polynomial $p(x)$ by its last term $p_m(x)$, the result is a Sturm sequence for $p(x)/p_m(x)$. (7)

For the proof, put $\tilde{p}_i(x) := p_i(x)/p_m(x)$. Property (a) holds since $\tilde{p}_m(x) = 1$. From the construction of the canonical sequence, we have (see (5) above):

$$p_{i-1}(x) = p_i(x)q(x) - p_{i+1}(x), \quad \text{so} \quad \tilde{p}_{i-1}(x) = \tilde{p}_i(x)q(x) - \tilde{p}_{i+1}(x) \quad \text{for } 1 < i < m \tag{8}$$

where $q(x)$ is the quotient when $p_{i-1}(x)$ is divided by $p_i(x)$. This shows that if two consecutive $\tilde{p}_i(x)$ had a common root, then that root would be shared by all $\tilde{p}_i(x)$, but that is absurd since $\tilde{p}_m(x) = 1$. Thus (b) is proved.

If $\tilde{p}_i(\alpha) = 0$ for some real number α, then, from the second equation in (8) above, we have $\tilde{p}_{i-1}(\alpha) = -\tilde{p}_{i+1}(\alpha)$, so $\tilde{p}_{i-1}(\alpha)$ and $p_{i+1}(\alpha)$ have opposite sign. (That neither is zero follows from (b).) Thus (c) is proved.

We now prove (d). Since $p(x)/p_m(x)$ has no multiple roots, it follows from (4) above that (d) holds if we put $p_0(x) = p(x)/p_m(x)$ and $p_1(x) = (p/p_m)'(x)$. It is therefore enough to show that, for any root α of $p(x)$ of multiplicity e (note that the roots of $p(x)/p_m(x)$ are precisely those of $p(x)$ except that none of them is repeated), we have $(p'/p_m)(\alpha) = e(p/p_m)'(\alpha)$.

Put $p(x) = (x-\alpha)^e q(x)$ with $q(\alpha) \neq 0$ and $e \geq 1$. Then $p_m(x) = (x-\alpha)^{e-1}r(x)$ with $r(\alpha) \neq 0$. We get $(p'/p_m)(x) = eq(x)/r(x) + (x-\alpha)q'(x)/r(x)$ and so $(p'/p_m)(\alpha) = eq(\alpha)/r(\alpha)$. Also $(p/p_m)(x) = (x-\alpha)q(x)/r(x)$, so that $(p/p_m)'(\alpha) = q(\alpha)/r(\alpha)$. This finishes the proof of (7). □

The proof of the following observation is obvious:

If a polynomial in a Sturm sequence is nowhere vanishing on the real line, then we may omit the rest of the sequence to obtain a shorter Sturm sequence. □ (9)

The "sign-change-number" function σ. Let $p_0(x), \ldots, p_m(x)$ be an arbitrary sequence of polynomials (no conditions as Sturmness are imposed upon it). Given a real number α, we record sequentially whether $p_i(\alpha)$ is +ve, −ve, or zero. We define $\sigma(\alpha)$ to be the number of sign changes in this sequence, ignoring the zeros. The dependence of σ on the sequence is suppressed in notation.

For example, from the sequence

$$x^3, \quad x^2-1, \quad x-4, \quad 10$$

we obtain the following sequence when $\alpha = 1$:

$$+, \quad 0, \quad -, \quad +$$

And so $\sigma(1) = 2$.

We define $\sigma(\infty)$ and $\sigma(-\infty)$ in a similar way, by taking limits instead of values. The signs of the limits of the polynomials in the above sequence at $-\infty$ and ∞ respectively are:

$$-, \quad +, \quad -, \quad + \qquad\qquad +, \quad +, \quad +, \quad +$$

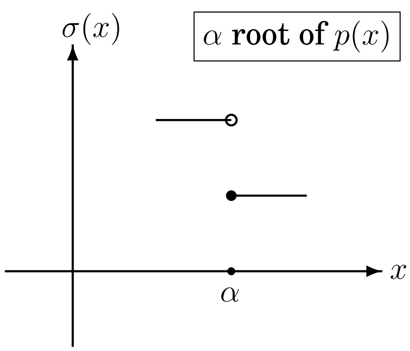
And so $\sigma(-\infty) = 3$ and $\sigma(\infty) = 0$.

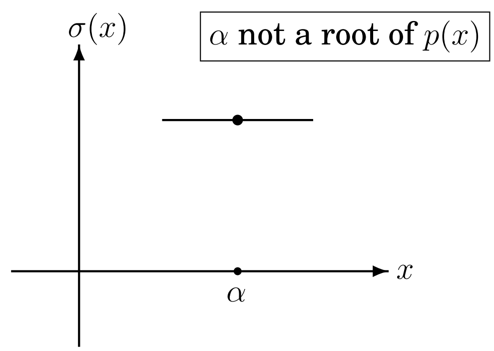
Sturm's theorem (first version: for Sturm sequences). *Let $p(x)$ be a non-zero polynomial with real coefficients. The number of distinct real roots (counted without multiplicity) of $p(x)$ in an interval $(a, b]$ of the real line (where we allow $a = -\infty$ or $b = \infty$ or both) is given by $\sigma(a) - \sigma(b)$, the difference in the sign-change-number function at the end points a and b, with respect to any Sturm sequence associated to the given polynomial $p(x)$.* [2]

Proof of the first version of Sturm's theorem. We first prune the Sturm sequence by deleting all the identically zero polynomials that it may contain: this does not affect $\sigma(x)$, for zeros are ignored in the calculation of σ; neither is the Sturmness of the sequence destroyed (see (6) above). We now track what happens to $\sigma(x)$ as x moves from left to right across the real line. It is evident that, for $\sigma(x)$ to change, we must cross a zero—call it α—of one of the polynomials $p_i(x)$ in the Sturm sequence. We focus on a neighborhood of α where none of the polynomials in the sequence attains a zero except possibly at α.

The theorem follows once we prove the following claim: *as we move across α from left to right, $\sigma(x)$ drops by 1 if α is a root of $p(x)$ and does not change otherwise:* see the picture below.

We now prove the claim. Suppose that α is a root (possibly repeated) of one of the $p_i(x)$ for $i > 0$. Then $i < m$ (by property (a)) and $p_{i-1}(\alpha)$, $p_{i+1}(\alpha)$ are non-zero and have opposite signs (by property (c)). This means that, in the neighborhood of α that we are focusing on, the polynomials $p_{i-1}(x)$ and $p_{i+1}(x)$ have no roots, and so do not change sign. Thus no matter what the behaviour of $p_i(x)$ is in this neighborhood, it has no effect on $\sigma(x)$.





Now suppose that $p_0(\alpha) = p(\alpha) = 0$. Then $p_1(\alpha) \neq 0$ because of (b). By (d), the signs of $p_0(x)$, $p_1(x)$ change across α either from $+, -$ to $-, -$ or from $-, +$ to $+, +$. In either case there is a drop by 1 of $\sigma(x)$ at α. $\square$

[2] As the proof shows, it is sufficient for the sequence to be Sturm "over the interval $(a, b]$", i.e., the properties (a)–(d) in the definition above of a Sturm sequence hold in $(a, b]$.

Sturm's theorem (second version: for the canonical sequence). *Let $p(x)$ be a non-zero polynomial with real coefficients. The number of real roots of $p(x)$ in an interval $(a, b]$ of the real line, where we allow $a = -\infty$ or $b = \infty$ (or both) and the roots are counted without multiplicity, is given by $\sigma(a) - \sigma(b)$, the difference in the sign-change-number function at the end points a and b, with respect to the canonical sequence associated to the given polynomial $p(x)$, provided that neither a nor b is a repeated root of $p(x)$ (they could be simple roots).*

Proof of the second version of Sturm's theorem. Let $p_0(x), \ldots, p_m(x)$ be the canonical sequence associated to $p(x)$. Put $\tilde{p}_i(x) := p_i(x)/p_m(x)$ and consider the sequence $\tilde{p}_0(x), \ldots, \tilde{p}_m(x)$. The latter sequence is a Sturm sequence by (7). Since $p_m(x)$ is the g.c.d. of $p(x)$ and its derivative $p'(x)$, it follows that $\tilde{p}_0(x)$ has no repeated root, and so, by the first version of Sturm's theorem, the number of roots of $\tilde{p}_0(x)$ in the interval $(a, b]$ is given by $\sigma(a) - \sigma(b)$, where σ is calculated with respect to the Sturm sequence $\tilde{p}_i(x)$ above.

The roots of $p(x)$ and those of $\tilde{p}_0(x)$ are identical except for multiplicity, so $\sigma(a) - \sigma(b)$ as in the previous paragraph also equals the number of distinct roots of $p(x)$ in the interval $(a, b]$.

Finally, we observe that $\sigma(a)$ (and similarly $\sigma(b)$) is the same whether calculated for the sequence $p_i(x)$ or for the sequence $\tilde{p}_i(x)$. Indeed, $p_m(a) \neq 0$, for a is not a repeated root of $p(x)$; and the sequence $p_0(a), \ldots, p_m(a)$ differs term-by-term from the sequence $\tilde{p}_0(a), \ldots, \tilde{p}_m(a)$ exactly by the non-zero factor $p_m(a)$. $\square$

Complements

A consequence of the proof. We have:

> *Suppose that properties (a)–(c) in the definition of a Sturm sequence hold for a sequence $p_0(x), \ldots, p_m(x)$ of real polynomials (with $m \geq 0$). Then they hold also for any sequence of the form $p_i(x), p_{i+1}(x), \ldots, p_m(x)$ obtained by omitting some number of initial terms. Further, in any interval $(a, b]$ of the real line (where we allow $a = -\infty$ or $b = \infty$ or both), the number of distinct real roots of $p_i(x)$ (i.e., counted without multiplicity) is at least $\sigma(a) - \sigma(b) - i$, where $\sigma(x)$ is the sign-change-number function with respect to the sequence $p_0(x), \ldots, p_m(x)$.*

It is evident that the first assertion holds. For the second, we first prove it for the case $i = 0$. If $p_0(x)$ is identically zero or if $\sigma(a) - \sigma(b) \leq 0$, then there is nothing to prove. Otherwise, we follow the proof of the first version of Sturm's theorem. We observe $\sigma(x)$ as x moves from left to right from a to b on the real line. The total decrease in $\sigma(x)$ over this range is $\sigma(a) - \sigma(b)$. We have $\sigma(a) \leq \sigma(a')$ for a' close to a and to its right. The decrease in $\sigma(x)$ can occur only when x crosses a root of $p_0(x)$. Moreover at each such root α, the decrease is at most 1. More precisely, here are all the possibilities for $\sigma(x)$ as it crosses α: it could decrease by 1 at α, it could increase by 1 just to the right of α, it could stay the same across α, or it could drop momentarily by 1 at α and come back up to its previous value. Thus there must exist at least $\sigma(a) - \sigma(b)$ distinct roots of $p_0(x)$ in $(a, b]$.

Let $\sigma_i(x)$ denote the sign-change-number function for $p_i(x), \dots, p_m(x)$. Then, clearly, $\sigma_0(x) = \sigma(x)$. Since the properties (a)–(c) hold for $p_i(x), \dots, p_m(x)$, and since $\sigma_i(a) - \sigma_i(b) \geq \sigma_0(a) - \sigma_0(b) - i$, the general case of the second assertion follows by applying the special case $i = 0$ to $p_i(x), \dots, p_m(x)$.

Interlacing of roots $p_0(x)$ and $p_1(x)$. Let $p_0(x), p_1(x), \dots, p_m(x)$ be a Sturm sequence. Between consecutive real roots α and α' of $p_0(x)$ (when the real roots are ordered increasingly), there exists a real root of $p_1(x)$: this follows from property (d) by applying to $p_1(x)$ the intermediate value theorem. In particular, if the number of real roots of $p_1(x)$ is one less than the number of roots of $p_0(x)$, then the roots are interlaced: that is, if $\alpha_1, \dots, \alpha_{r+1}$ are the real roots of $p_0(x)$ arranged increasingly and $\beta_1, \dots, \beta_r$ those of $p_1(x)$ arranged also increasingly, then $\alpha_i < \beta_i < \alpha_{i+1}$ for all $1 \leq i \leq r$.

Interlacing of roots. Let $p_0(x), \dots, p_m(x)$ be a Sturm sequence. Suppose that $p_0(x)$ has d distinct real roots (counted without multiplicity). Then of course $m \geq d$. Indeed, on the one hand, $d = \sigma(-\infty) - \sigma(\infty)$ by the first verstion of Sturm's theorem, and on the other, clearly, $m \geq \sigma(-\infty) \geq \sigma(-\infty) - \sigma(\infty)$.

Now let us suppose that $m = d$. Then $\sigma(-\infty) = m = d$ and $\sigma(\infty) = 0$. (To say $\sigma(\infty) = 0$ is the same as to say that the leading coefficients of $p_0(x), \dots, p_m(x)$ are all of the same sign.) Further, letting $\sigma_i(x)$ denote the sign-change-number function for the sequence $p_i(x), \dots, p_m(x)$, we have $\sigma_i(\infty) = 0$ and $\sigma_i(-\infty) = d - i$: the first equality is clear since $\sigma_i(\infty) \leq \sigma(\infty)$, and the second follows since the limits at $-\infty$ of the sequence $p_0(x), \dots, p_m(x)$ must strictly alternate in sign (for, that is the only way in which $\sigma(-\infty) = m$).

The sequences $p_i(x), \dots, p_m(x)$ (for i, $1 \leq i \leq m$) clearly inherit properties (a)–(c) in the definition of Sturm sequence. It follows from the result in the subsection "A consequence of the proof" that the number of distinct real roots of $p_i(x)$ is at least $d - i$.

Now suppose that the number of distinct real roots of each $p_i(x)$ is exactly $d - i$ (as for example, by some condition limiting their degrees). Then we claim:

$$p_i(x), \dots, p_m(x) \text{ is a Sturm sequence for each } i,\ 0 \leq i \leq m$$

We need only show that property (d) holds. As x moves from the far left to the far right on the x-axis, the function $\sigma_i(x)$ drops by $d - i$ in value. The value of this function can only change at roots of $p_i(x)$ and at each of these it decreases by at most 1: see the proof of the result in the subsection "A consequence of the proof". Since there are only $d - i$ roots, it follows that it drops for good by exactly 1 at each of these roots.

This forces (d). Indeed, if $p_i(x)$ does not cross the x-axis at one of its roots α, then $\sigma_i(x)$ is either constant across α or momentarily goes down by 1 at α before coming back up to its previous value, neither of which is possible. Suppose that $p_i(x)$ changes over from +ve to $-$ve at α and that $p_{i+1}(\alpha)$ is +ve. Then $\sigma_i(x)$ increases by 1 at α, which is not possible. Thus $p_{i+1}(\alpha)$ is $-$ve in this case. A similar argument shows that $p_{i+1}(\alpha)$ is +ve if $p_i(x)$ changes over from being $-$ve to +ve across α.

The preceding arguments clearly apply to the canonical sequence of a real polynomial $p(x)$ of degree d with d distinct real roots and so we have:

> *Let $p(x)$ be a real polynomial of degree d with d distinct real roots. Let $p_0(x)$, …, $p_m(x)$ be the canonical sequence associated to $p(x)$. Then*
> *(1) $m = d$*
> *(2) The signs of the leading coefficients of $p_i(x)$ are all the same.*
> *(3) $p_i(x)$ has degree $d - i$ and $d - i$ distinct real roots.*
> *(4) $p_i(x)$, …, $p_m(x)$ is a Sturm sequence, for any i, $0 \leq i \leq m$.*
> *(5) The roots of $p_i(x)$ and $p_{i+1}(x)$ are interlaced for $0 \leq i < m$.*

Descartes's rule of signs. Let $p(x)$ be a non-zero polynomial with real coefficients. We list the signs of the non-zero coefficients in order and count the number of changes in the resulting sequence. For example, for $x^6 - 3x^4 + 3X^3 + X^2 - 1$, we get the sequence $+ - + + -$, which has three changes. The rule states the following:

> *The number of positive real roots counted with multiplicity of a real polynomial with non-zero constant term has the same paririty as the number of sign changes in the coefficients and is less (by an even number) than that number.*

For the proof, proceed by induction on the degree of the polynomial. If the polynomial has degree zero, or, in other words, if it is a non-zero constant, then of course it has no sign changes and no roots, so the rule holds.

Now suppose that the polynomial is of degree at least one. We may assume that the polynomial is monic by dividing it by its leading coefficient: the non-zeroness of the constant term, the roots (and in particular the number of positive roots) and the number of sign changes are all preserved under this operation. Now the limit of $p(x)$ as x goes to infinity is infinity. By a continuity argument, it follows that $p(x)$ has an even or odd number of positive roots accordingly as its constant term is positive or negative. On the other hand, it is also evident that the number of sign changes is positive or negative accordingly as the constant term is positive or negative. This proves the parity part of the rule.

For the latter part, there is nothing to prove if $p(x)$ has no positive real roots. So let us assume that it has a positive real root α. Write $p(x) = (x - \alpha)q(x)$. By induction, the rule holds for $q(x)$: i.e., the number of its positive roots is an even number less than the number of its sign changes. Since $p(x)$ evidently has one positive root more than $q(x)$, it follows that the rule holds for $p(x)$ as well, provided that we can show that the number of sign changes in $p(x)$ is an odd number more than the number of those in $q(x)$. We now prove this last assertion, which is the main part of the proof.

In the following calculation, we keep track of the signs in the multiplication of $q(x)$ by $x-\alpha$ to produce $p(x)$. We have assumed that the constant term of $q(x)$ is positive, but a similar argument should work if it were negative.

$$\begin{array}{rl}
x \cdot q(x): & \boxed{+} + \ldots + \boxed{-} - \ldots - \boxed{+} + \ldots + \boxed{-} - \ldots \;\ldots\; \ldots \boxed{+} + \ldots + \\
-\alpha \cdot q(x): & \quad - \ldots - \boxed{-} + \ldots + \boxed{+} - \ldots - \boxed{-} + \ldots \;\ldots\; \ldots \boxed{+} - \ldots - \boxed{-} \\
\hline
p(x): & \boxed{+} \; ? \ldots ? \boxed{-} \; ? \ldots ? \; \boxed{+} \; ? \ldots ? \; \boxed{-} \; ? \ldots \;\ldots\; \ldots \boxed{+} \; ? \ldots ? \; \boxed{-}
\end{array}$$

In the first two rows, the boxed terms mark areas of constant sign: in the first row, they are the first terms after a sign has changed or first arisen; in the second row, they are the last terms of a given sign before the sign changes or forever vanishes. The two rows have the same number of boxes and the same number of sign changes. And the number of sign changes is one less than the number of boxes.

In the last row, the number of boxes is one more than in either the first or second row. The question marks in between the boxes indicate the possibility that those terms that could go any which way: they could be positive or negative or zero. At any rate, the number of sign changes between boxed terms in the third row is one or some other odd number.

Suppose that the number of sign changes in $q(x)$ is c. Then the number of boxes in either of the first two rows is $c+1$, the number of boxes in the third row is $c+2$, and, finally, the number of sign changes in the third row is a sum of $c+1$ odd numbers, which is easily seen to be an odd number more than c. $\square$

THE INSTITUTE OF MATHEMATICAL SCIENCES, CHENNAI
E-mail address: knr@imsc.res.in