

When are Complex Zeroes on a Roll(e)?

B Sury



After spending 1981–1999 in TIFR Bombay, B Sury moved to ISI Bangalore in 1999, tempted by the opportunity to teach in the soon-to-start BMath Hons Programme.

*At endpoints of an interval,
if f has value null.
It will then be seen
somewhere in-between
to have a tangent horizontal!*

With Rolle's theorem as a starting point, we discuss some geometric and other aspects of the distribution of the zeroes of polynomials.

Geometry of Polynomial Zeroes

Finding the zeroes (another name for 'roots') of a polynomial becomes harder as the degree increases and it is useful if one is able to relate the zeroes of a polynomial to those of its derivative (which has one degree less). We will start the discussion¹ by recalling a familiar theorem proved by Michel Rolle in his book *Traite d'Algebre* in 1690. Rolle's theorem holds not only for polynomials but more generally as follows:

Rolle's Theorem. *If $f : [a, b] \rightarrow \mathbf{R}$ is continuous, and f is differentiable on (a, b) , such that $f(a) = 0 = f(b)$, then $\exists c \in (a, b)$ so that $f'(c) = 0$.*

An immediate application of this is:

Mean-Value Theorem. *If $f : [a, b] \rightarrow \mathbf{R}$ is continuous and f is differentiable on (a, b) , then $\exists c \in (a, b)$ so that $\frac{f(b)-f(a)}{b-a} = f'(c)$.*

The basic reason behind the validity of Rolle's theorem is that a continuous function assumes its maximum and minimum values on a closed, bounded interval, a fact which is a consequence of the Heine–Borel property. It

¹ Expanded version of talks in ISI Bangalore for BMath Hons entrants and at St.Xavier's College, Aluwa, Kerala in 2006.

Keywords

Rolle's theorem, Gauss–Lucas theorem, Rouche's theorem, fundamental theorem of algebra.



is important to note that the hypotheses are necessary; if we apply these theorems blindly, we may be in trouble. For example, suppose we were to apply the mean-value theorem to the function $\tan(x)$ on an interval $[a, b]$. Then,

$$\tan(b) - \tan(a) = (b - a)\sec^2(c)$$

for some point $c \in (a, b)$. But, $\cos^2(c) \leq 1$ gives $\tan(b) - \tan(a) \geq b - a$. In particular, if this were valid in the interval $[0, \pi]$, we would have $\tan(\pi) - \tan(0) \geq \pi$, a manifest absurdity as the left side is zero! So, what is wrong? Student readers are encouraged to think about it.

Rolle's theorem can be rephrased as saying that if $a < b$ are zeroes of a polynomial $f(x)$ with real coefficients (or more generally a nice enough function), then the line segment joining a and b contains a zero of the derivative function $f'(x)$. The first question which arises is whether a similar thing is true when we go to \mathbf{C} . We do not assume any knowledge of complex analysis. So, let us restrict ourselves to polynomials $f(x)$ which have complex coefficients.

A moment's thought tells us that the rephrased statement of Rolle does not hold good. For example, look at the polynomial $f(z) = z^3 - 1$, where z is allowed to take complex values. Its zeroes are $1, z, z^2$, where $z = e^{2i\pi/3}$. But $f'(z) = 3z^2$ is zero only at $z = 0$, which is not on any of the 3 line segments $\overline{1, z}, \overline{z, z^2}$ or $\overline{1, z^2}$ as can be seen from *Figure 1*.

Thus Rolle's theorem in this form fails. It is also clear that the mean-value theorem too fails. We shall be discussing some results for complex polynomials including a complex analogue of Rolle's theorem.

A basic idea we will use repeatedly is the *triangle inequality for complex numbers*

$$|z_1 + z_2| \leq |z_1| + |z_2|.$$

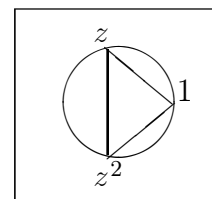
We will use this also in the form

$$|z_1 - z_2| \geq |z_1| - |z_2|$$

which follows from the first form by writing $z_1 = (z_1 - z_2) + z_2$.

The natural analogues in \mathbf{C} of closed and open intervals in \mathbf{R} are closed discs $\bar{B}(z_0, r) = \{z \in \mathbf{C} : |z - z_0| \leq R\}$

Figure 1.



and open discs $B(z_0, r) = \{z \in \mathbf{C} : |z - z_0| < R\}$ which are centred at a point z_0 and have radius r .

Before proceeding, we mention a motivating factor for the great Gauss to study these aspects. Basically, one is interested in locating the zeroes of a polynomial p in the complex plane. If p is the derivative of a polynomial f and if we can locate the zeroes of f easily, Gauss discovered that one could locate or at least find a region where all zeroes of p are located. I ask you only to note in the previous example that the only zero $z = 0$ of f' is inside the Δ spanned by the zeroes of f . That this is not an accident is shown by the following beautiful discovery:

Theorem (Gauss–Lucas). *Zeroes of $f'(z)$ lie in the convex hull of the zeroes of $f(z)$. Equivalently $\max\{|w| : f'(w) = 0\} \leq \max\{|z| : f(z) = 0\} \forall f$.*

This theorem can be thought of as a generalization of Rolle’s theorem for polynomials over \mathbf{C} . It has been generalized by Schmeisser who showed:

$$\sum_1^k |w_i| \leq \sum_1^k |z_i| \forall k \leq n - 1, \text{ where } |w_1| \geq \dots \geq |w_{n-1}| \text{ and } |z_1| \geq \dots \geq |z_n|.$$

Proof of the theorem. Let us rewrite the polynomial f as $f(z) = a_n \prod_{i=1}^r (z - z_i)^{m_i}$, where z_i ’s are distinct. Then $\frac{f'}{f} = \sum_{i=1}^r \frac{m_i}{z - z_i} = \sum_{i=1}^r \frac{m_i(\bar{z} - \bar{z}_i)}{|z - z_i|^2}$. Take a possible zero of f' , different from the z_i ’s, say a . So $0 = \frac{f'(a)}{f(a)} = \sum_{i=1}^r \frac{m_i(\bar{a} - \bar{z}_i)}{|a - z_i|^2}$. We have $\bar{a} = \sum_{i=1}^r \frac{m_i/|a - z_i|^2}{\sum_{j=1}^r \frac{m_j}{|a - z_j|^2}} \cdot \bar{z}_i$. Take conjugates to get $a = \sum_{i=1}^r \frac{\frac{m_i}{|a - z_i|^2}}{\sum_{j=1}^r \frac{m_j}{|a - z_j|^2}} \cdot z_i$. Clearly, this is a convex combination of the z_i ’s and proves the first statement.

The equivalence to the statement

$$\max\{|w| : f'(w) = 0\} \leq \max\{|z| : f(z) = 0\} \tag{1}$$

for all polynomials $f(z)$ can be shown as follows.

Firstly, evidently Gauss–Lucas theorem implies the above inequality for any polynomial $f(z)$. Now, for the converse implication, we will use the observation that the convex hull of a set S necessarily coincides with the intersection of all the closed discs containing S in their interiors. Suppose the Gauss–Lucas theorem were false for a certain polynomial $g(z)$. Therefore, there is a complex number w such that $g'(w) = 0$ but w is not in the convex hull \mathcal{C} of the set of zeroes of



$g(z)$. By the above observation, there is a closed disc $\bar{B}(c, r)$ containing \mathcal{C} in its interior such that $w \notin B(c, r)$. Consider the polynomial $f(z) = g(z + c)$; note $f'(z) = g'(z)$. Clearly, the zeroes of $f(z)$ are $z_i - c$, where z_i runs through the zeroes of $g(z)$. So, the absolute values $|z_i - c|$ of the zeroes of $f(z)$ are all bounded by r while the root w of $f'(z)$ has an absolute value bigger than r . This shows that the inequality (1) does not hold for the polynomial $f(z)$ here. Hence, we have shown the equivalence.

The above theorem was beautifully put in the following words by Gauss. (Recall that critical points of a polynomial are the zeroes of the derivative.)

'The critical points of a polynomial f which are not multiple zeroes of f are located at the equilibrium positions in a certain field of force. This field is one due to a particle placed at each zero of f , having a mass equal to the multiple of the zero and attracting to the inverse distance law.'

Here are two closer analogues of Rolle's theorem which we state without proof:

Theorem (Grace 1902; rediscovered by Heawood 1907). *If $z_1 \neq z_2$ are zeroes of a polynomial f , of degree n , then the disc centered at $(z_1 + z_2)/2$ and with radius $|z_1 - z_2|\cot(\pi/n)/2$ contains a zero of f' .*

Note the case $n = 2$!

Theorem (Alexander). *If $\bar{B}(0, M)$ contains zeroes $z_1 \neq z_2$ of polynomial f , of degree n , then*

$\bar{B}(0, M\operatorname{cosec}\frac{\pi}{n})$ contains a zero of f' .

Conjecture (Sendov 1962). *If all the zeroes of a polynomial f of degree n are in the unit disc $\bar{B}(0, 1)$, then each zero of f is within unit distance of a zero of f' .*

This is still open! Here is a way to get a disc around the origin containing all the zeroes of an arbitrary fixed polynomial. Note first that an arbitrary non-zero polynomial can be written as $az^r f(z)$ where $f(z)$ is a polynomial with non-zero constant term and top coefficient equal to 1. Hence the next result gives in reality a region where the zeroes of an arbitrary non-zero polynomial can be .

Theorem. *Let $n \geq 1, a_0 \neq 0$. Then, $z^n - |a_{n-1}|z^{n-1} - \dots - |a_1|z - |a_0|$ has a unique real, positive zero M . Moreover, the polynomial $z^n + a_{n-1}z^{n-1} + \dots + a_0$ has all its zeroes in $\bar{B}(0, M)$.*



Before proceeding to prove this theorem, we recall a very useful result due to Descartes:

Descartes’s Rule of Signs

If $f(x) = c_0 + c_1x + \dots + c_nx^n$ is a polynomial with real coefficients, then the number s of sign changes as we go along the sequence c_0, \dots, c_n and the number r of positive real roots of $f(x)$ are related by $s \geq r$ and $s - r$ is even.

Proof of above theorem. Note first that the polynomial $z^n - |a_{n-1}|z^{n-1} - \dots - |a_1|z - |a_0|$ has at most one positive zero as can be seen from Descartes’s rule of signs. Indeed, as there is one sign change, the number of real, positive zeroes is at most 1. But, clearly this polynomial is negative at $z = 0$ and positive for $|z|$ large. Therefore \exists a unique positive zero M of $z^n - |a_{n-1}|z^{n-1} - \dots - |a_0|$ and the value at any real $z > M$ is positive. Therefore, if $|z| > M$, we have

$$\begin{aligned} |z^n + a_{n-1}z^{n-1} + \dots + a_0| &\geq |z|^n - |a_{n-1}||z|^{n-1} - \dots - |a_0| \\ &> 0. \end{aligned}$$

Here is a result which gives, under some condition, a disc around the origin which does not contain any zero!

Lemma. Let $M > 0$ be such that $|a_i| < M|a_0| \forall i = 1, \dots, n$. Then $f = \sum_{i=0}^n a_i z^i$ has all its roots in $|z| \geq \frac{1}{M+1}$.

Proof. If $|z_0| < \frac{1}{M+1}$, then

$$\begin{aligned} |f(z_0)| &\geq |a_0| \left(1 - \frac{|a_1|}{|a_0|}|z_0| - \frac{|a_2|}{|a_0|}|z_0|^2 - \dots - \frac{|a_n|}{|a_0|}|z_0|^n \right) \\ &\geq |a_0|(1 - M((M + 1)^{-1} + \dots + (M + 1)^{-n})) > 0. \end{aligned}$$

Remarks. If f is a real polynomial with all zeroes being real and simple (i.e., of multiplicity 1), then the interval containing two consecutive zeroes contains a unique zero of f' .

However, this can fail when we consider functions which satisfy Rolle’s theorem but are not polynomials. For instance, look at $F = (X^2 - 4)e^{X^2/3}$. Then $F' = \frac{2}{3}X(X^2 - 1)e^{X^2/3}$. The zeroes of F are ± 2 and all the zeroes of F' are in between!

Let us say that a differentiable function f satisfies *Rolle’s property* if the line segment containing two zeroes of f contains at least one zero of f' . We started



with a simple example to show that there are complex polynomials which do not have Rolle's property. However, the question remains as to which complex polynomials do satisfy Rolle's property. This is the question implicitly mentioned in the title. The answer is provided by the following beautiful result:

Theorem (Dotson). *A polynomial f with complex coefficients has Rolle's property if, and only if, its zeroes are collinear.*

Proof. The sufficiency assertion for Rolle's property is easier; we prove it first. Let all zeroes of f lie on a line k . By translating and rotating (which amounts to changing z to $az + b$ for some $|a| = 1$), we may assume that the line is the X -axis and that all the zeroes of f are real. I leave it as an exercise to check that Rolle's property for real polynomials (applied to $c \cdot f(\frac{z-b}{a})$ for some $c \in \mathbf{C}$) implies that for f .

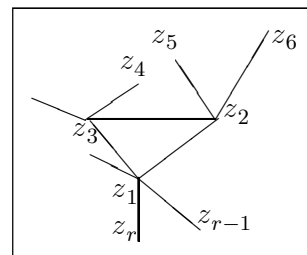
Let us look at the converse. Suppose f has Rolle's property. Suppose that not all zeroes are collinear. Let z_1, z_2, z_3 be zeroes of f forming a triangle of smallest area among all such triangles. Write z_4, \dots, z_r for the other *distinct* zeroes of f . It is clear that none of the z_4, \dots, z_r can lie on $\overline{\Delta z_1 z_2 z_3}$. Here, we have written $\overline{\Delta z_1 z_2 z_3}$ to denote the closed region including the interior of the triangle. Join all the points z_4, \dots, z_r to one of the vertices z_1, z_2, z_3 in such a way that all the segments are non-intersecting and, each segment intersects $\overline{\Delta z_1 z_2 z_3}$ only in that particular vertex of the triangle to which it is connected (see Figure 2).

By the hypothesis that Rolle's property holds for f , there are at least $r - 3$ zeroes of f' which are outside Δ and are different from the zeroes of f . Also, applying Rolle's property for the 3 edges of the triangle, there are 3 more zeroes of f' different from z 's. If the zeroes z_1, \dots, z_r have multiplicities $m_1, \dots, m_r (\geq 1)$, there are at least $\sum_{i=1}^r (m_i - 1) + r (= \sum_{i=1}^r m_i = \text{deg } f)$ zeroes (counted with multiplicity) of f' , which is a contradiction. Thus, no triangle can be formed and all zeroes must be collinear.

Schoenberg's Conjecture – Pereira's Theorem

The Gauss–Lucas theorem showed that the zeroes of f' are contained in the convex hull of the zeroes of f . In fact, it is trivial to see that the centroid $\frac{\sum_{i=1}^n z_i}{n}$ of the set $\{z_1, \dots, z_n\}$ of zeroes of f and the centroid $\frac{\sum_{i=1}^{n-1} w_i}{n-1}$ of the set $\{w_1, \dots, w_{n-1}\}$ of zeroes of f' coincide. Another

Figure 2.



simple computation shows $\sum_1^{n-1} w_i^2 = C^2 + \frac{n-2}{n} \sum_1^n z_i^2$. Note that for any rotation $z \mapsto e^{i\theta} z$, the polynomial $f(e^{-i\theta} z)$ has its roots to be $e^{i\theta} z_1, \dots, e^{i\theta} z_n$. Thus, if all zeroes $\{z_i\}_1^n$ of f were collinear, then all the $\{w_j\}_1^{n-1}$ are also on the same line as well (by the usual ‘real’ Rolle theorem). If this happens, then the above formula reduces to $\sum_{j=1}^{n-1} |w_j|^2 = |C|^2 + \frac{n-2}{n} \sum_{k=1}^n |z_k|^2$.

Schoenberg observed this and conjectured in 1984–85 that whenever $\{z_i\}_1^n$ is the zero-set of a polynomial f , and $\{w_i\}^{n-1}$, the corresponding zero set of f' , one has:

$$\sum_{j=1}^{n-1} |w_j|^2 \leq \frac{n-2}{n} \sum_{k=1}^n |z_k|^2 + |C|^2$$

with equality if and only if the z_j 's are collinear (and w_k 's are on the same line). Amazingly, it was only in 2003 that Rajesh Pereira proved this conjecture using ideas from linear algebra and Hilbert spaces.

Some Other Aspects of Polynomials in One Variable

The fundamental theorem of algebra says that every non-constant complex polynomial has a complex root. But the proof requires rather advanced techniques. The first correct proof was given by A Cauchy in his thesis. One result whose statement one can understand easily is:

Rouché’s Theorem. *Let f, g be two complex polynomials satisfying $|f(z) + g(z)| < |f(z)| + |g(z)| \forall z$ on the circle $|z| = R$. Then, inside the open disc $B(0, R)$, the polynomials f and g have the same number of zeroes.*

This wonderful theorem has so many applications that it is impossible in our discussion to even mention all the areas or questions to which it can be applied. One standard application of it is the fundamental theorem of algebra. Let us see how.

If $f(z) = \sum_{i=0}^n a_i z^i$ with $a_n \neq 0$ and $n \geq 1$ and $g(z) = -a_n z^n$, take R so large that $\left| \sum_{i=0}^{n-1} a_i z^i \right| < |a_n z^n|$ for $|z| = R$. Clearly, by Rouché’s theorem, (since $|f(z) + g(z)| < |g(z)| \forall |z| = R$), f and g have the same number n of zeroes in $|z| < R$.

Incidentally, Rouché’s theorem is valid for more general classes of functions than polynomials and more general regions. What is remarkable is that one can actu-



ally deduce Rouché's theorem for polynomials (the statement made above) from the fundamental theorem of algebra; in other words, both are equivalent. We prove this now considering the case of the unit circle:

Theorem (M Filaseta). *Suppose f, g are complex polynomials with $|f(z) + g(z)| < |f(z)| + |g(z)| \forall z$ with $|z| = 1$. Then f and g have the same number of zeroes in $B(0, 1)$.*

Proof. Note that the inequality of the hypothesis implies that neither f nor g have zeroes on the unit circle. If r and s denote the numbers of zeroes of f and g inside $B(0, 1)$, we will show that $r \leq s$ (this suffices, by symmetry).

Suppose, if possible, that $r > s$. It suffices to get $\theta \in [0, 2\pi]$ with $|f(e^{i\theta}) + g(e^{i\theta})| = |f(e^{i\theta})| + |g(e^{i\theta})|$. For this again, it is enough to find a θ with the property $\arg f(e^{i\theta}) = \arg g(e^{i\theta})$, where the argument takes values in $[-\pi, \pi]$. For $\alpha \in \mathbf{C}$ with $|\alpha| \neq 1$, we define a continuous function $w(\alpha) : \mathbf{R} \rightarrow \mathbf{R}$ as follows. Define

$$w(\alpha)(0) = \arg(1 - \alpha) ;$$

$$w(\alpha)(\theta) = 2t(\theta)\pi + \arg(e^{i\theta} - \alpha) \forall \theta \neq 0,$$

where $t(\theta) \in \mathbf{Z}$ is so chosen that $w(\alpha)$ is continuous. In other words, $w(\alpha)(\theta) - w(\alpha)(0)$ continuously represents the angle with vertex α and rays through 1 and $e^{i\theta}$.

Hence

$$w(\alpha)(2\pi) - w(\alpha)(0) = \left\{ \begin{array}{ll} 2\pi, & \text{if } |\alpha| < 1 \\ 0, & \text{if } |\alpha| > 1 \end{array} \right\}. \tag{2}$$

By the fundamental theorem of algebra, we may write

$$f = a \prod_1^n (z - \alpha_i) , \quad g = b \prod_1^m (z - \beta_j) . \tag{3}$$

Note that $|\alpha_i| \neq 1 \neq |\beta_j| \forall i, j$ by the hypothesis of the theorem.

Define $F(\theta) = \arg(a) + \sum_{i=1}^n w(\alpha_i)(\theta) + 2u\pi$, where $u \in \mathbf{Z}$ is chosen so that $F(0) \in (-2\pi, 0]$. Similarly, define $G(\theta) = \arg(b) + \sum_{j=1}^m w(\beta_j)(\theta) + 2v\pi$, where $v \in \mathbf{Z}$ is so chosen that $G(0) \in [F(0), F(0) + 2\pi]$.

Hence F, G are continuous.



Notice from (3) that $\arg f(e^{i\theta}) \equiv F(\theta)$ and $\arg g(e^{i\theta}) \equiv G(\theta) \pmod{2\pi}$, and $F(2\pi) - F(0) = 2r\pi, G(2\pi) - G(0) = 2s\pi$ (where r, s are the number of zeroes of f and g inside the unit disc).

Consider $H = F - G$. One has $H(0) = F(0) - G(0) \leq 0$ and $H(2\pi) = F(2\pi) - G(2\pi) = 2(r - s)\pi + F(0) - G(0) > 0$ (as first term is $\geq 2\pi$ and second term is in $[0, 2\pi)$). By the intermediate-value theorem, $H(\theta) = 0$ for some $\theta \in [0, 2\pi)$. So $\arg f(e^{i\theta}) = \arg g(e^{i\theta})$. This completes the proof of the theorem.

Common Divisors of $f^n - 1, g^n - 1$

Polynomials in one variable have many properties similar to integers. For example, the Euclidean division algorithm for natural numbers is valid for polynomials with complex coefficients. In fact, often polynomials are easier to study because there is the in-built notion of derivative which has no analogue for integers! We illustrate this theme by one beautiful conjecture which is wide open for integers but is a proved theorem for polynomials! We also mention in passing that the analogue of Fermat's last theorem for polynomials is an easy exercise.

If a, b are integers $\neq 0$ such that $\forall n$, the prime divisors of $a^n - 1$ are also divisors of $b^n - 1$. Erdős asked if it is then true that b is a power of a . This problem, known as the support problem was solved only about 10 years back. More generally, one has the following open conjecture:

Conjecture. Let $a, b \neq 0$ be such that $(a - 1, b - 1) = 1$ and $a^u \neq b^v \forall u, v \geq 1$ (i.e., a, b are multiplicatively independent). Then \exists infinitely many $n \geq 1$ such that the GCD $(a^n - 1, b^n - 1) = 1$.

Let us show now that integers are far more difficult to handle in comparison with polynomials by proving the analogue of the above conjecture for polynomials.

Theorem (Ailon and Rudnick – 2004). *Let f, g be complex polynomials which are multiplicatively independent. Then \exists a polynomial h such that $\text{GCD}(f^n - 1, g^n - 1)$ divides $h \forall n \geq 1$. Further, if $\text{GCD}(f - 1, g - 1) = 1$, then \exists a finite union $\bigcup_{i=1}^r d_i \mathbf{N}$ with $d_i \geq 2$ such that*

$$\forall n \notin \bigcup_{i=1}^r d_i \mathbf{N}, \text{GCD}(f^n - 1, g^n - 1) = 1.$$

Proof. One can give an easily-understandable proof of this assuming a certain fact called Lang's conjecture which has been proved now. This is just the simple-



looking statement (although it is a deep result requiring complicated mathematics):

When f, g are multiplicatively independent, there are only finitely many z for which both $f(z)$ and $g(z)$ are simultaneously roots of unity.

If S is this possible finite set, then $\text{GCD}(f^n - 1, g^n - 1)$ can only have linear factors from $\{z - s; s \in S\}$. Denote by ζ_n a primitive n -th root of unity. As $f(z)^n - 1 = \prod_{j=0}^{n-1} (f(z) - \zeta_n^j)$ and all factors are coprime, any $z - s$ can divide at most one of them; and has multiple $\leq \deg f$. Thus, $h(z) = \prod_{s \in S} (z - s)^{\min(\deg f, \deg g)}$ works. Further, if $\text{GCD}(f - 1, g - 1) = 1$, then $\forall s \in S$, the least positive integer d_s such that $(z - s)$ divides $\text{GCD}(f(z)^{d_s - 1}, g(z)^{d_s} - 1)$ is ≥ 2 . Also, $\forall n \notin \bigcup_{s \in S} d_s \mathbf{N}$, $(z - s)$ does not divide $\text{GCD}(f^n - 1, g^n - 1)$.

The Jacobian Conjecture – A Question about Polynomials in Several Variables

Until now we have discussed polynomials in a single variable. We end with a very concrete open problem about polynomials in 2 or more variables.

Let $f, g \in \mathbf{C}[x, y]$ be such that $\frac{\partial f}{\partial x} \frac{\partial g}{\partial y} - \frac{\partial f}{\partial y} \frac{\partial g}{\partial x}$ is a non-zero complex number (that is, do not involve x, y). Then x, y can be expressed as polynomials in f, g over \mathbf{C} .

This is the so-called Jacobian conjecture which is open and many famous mathematicians have given faulty proofs! Try your luck!

Address for Correspondence: B Sury, Indian Statistical Institute, 8th Mile Mysore Road, Bangalore 560 059, India.
Email:sury@isibang.ac.in

