REAL ROOTS!

CHRISTOPHE DOCHE AND MICHEL MENDÈS FRANCE

To Jonas Kubilius on his 80^{th} birthday

1. The origins, Descartes' theorem

In 1637, R. Descartes wrote on page 373 of his "géométrie",
« On connoist aussi de cecy [équation polynomiale] combien il peut y avoir de vrayes racines, & combien de fausses en chasque Equation. A sçavoir il y en peut avoir autant de vrayes, que les signes + & − s'y trouve de fois estre changés; & autant de fausses qu'il s'y trouve de fois deux signes +, ou deux signes − qui s'entresuivent. »
This is the famous "Descartes Rule" which D. E. Smith and M. L. Lathan translate as follows in [De].
"An equation can have as many true roots (*i.e.* real positive) as it contains changes of sign from + to − or from − to +; and as many false roots (*i.e.* real negative) as the number of times two + signs or two − signs are found in succession."
Following John Wallis in his treatise of Algebra (1685), the translations of Descartes add in a footnote that the result was already known to Harriot as early as 1631. Not being able to verify this claim, we believe that Descartes is the discoverer of his Rule. A modern statement could read as follows.

Theorem 1. *Let $\lambda_1, \lambda_2, \ldots, \lambda_n$ be n positive increasing integers, and let $a_1, a_2, \ldots, a_n$ be n nonzero real numbers. Then the number of strictly positive real distinct zeros of*

$$P(X) = \sum_{j=1}^{n} a_j X^{\lambda_j}$$

is less than or equal to the number of sign changes in the sequence $a_1, a_2, \ldots, a_n$.

Changing X into $-X$, we see that the number of strictly negative distinct real zeros is less than or equal to the number of sign changes in the sequence $(-1)^{\lambda_1} a_1$, $(-1)^{\lambda_2} a_2, \ldots, (-1)^{\lambda_n} a_n$.

As the number of sign changes in a sequence of length n is at most $n-1$ and since $X = 0$ may well be a real zero of P (if $\lambda_1 > 0$), we get the following corollary which we will use later on.

Corollary 2. *Let $0 \leqslant \lambda_1 < \lambda_2 < \cdots < \lambda_n$ be given integers and $a_1, a_2, \ldots, a_n$ be real nonzero numbers. The polynomial*

$$P(X) = \sum_{j=1}^{n} a_j X^{\lambda_j}$$

has at most $2n-1$ distinct real zeros.

Date: on December 5, 2002.
Key words and phrases. Integral geometry, real roots.

The most precise result is due to Sturm (1829), see [AG] p. 22. Let P be a polynomial of degree n with nonzero discriminant. Consider the sequence P_k

$$P_0(X) = P(X),\ P_1(X) = P_0'(X)$$
$$P_{k-2} \equiv -P_k \pmod{P_{k-1}},\ k \geqslant 2.$$

The number of real roots of P between a and b is equal to the excess of sign changes of $P_0(a), P_1(a), \ldots, P_n(a)$ compared to these of $P_0(b), P_1(b), \ldots, P_n(b)$.

2. Averages

In the 1930's several authors started to investigate the problem of estimating the expectation on average of the number of real zeros of a random n degree real polynomial. See for example [BP] and [LO1, LO2]. By a very cunning trick these last authors manage to prove that the probability that $\sum \pm X^j$ has more than $25(\log n)^2$ real roots is at most $12 \log n/n$. The probability that it has fewer than $\alpha \log n/\log\log\log n$ is less than $A/\log n$. But the first real breakthrough was completed by Mark Kac [Kac1, Kac2] when he established that if the coefficients $a_0, a_1, \ldots, a_n$ are real independent random variables with standard normal distribution (Gaussian) then the expected number of real roots is

$$N(n) = \frac{2}{\pi} \log n + O(1)$$

as n increases to infinity. He actually gives an explicit formula

$$N(n) = \frac{1}{\pi} \int_{-\infty}^{+\infty} \sqrt{\frac{1}{(t^2-1)^2} - \frac{(n+1)^2 t^{2n}}{(t^{2n+2}-1)^2}}\, dt$$

which we shall discuss later when we survey the 1995 geometric proof of A. Edelman and E. Kostlan [EK]. In his papers, M. Kac also studies the expected average number of real roots when the coefficients of the random polynomial are distributed according to different probability laws. In particular, he notes that if they are uniformly distributed on $[-1, +1]$ or equally distributed on the finite set $\{-1, +1\}$ then again

$$N(n) \sim \frac{2}{\pi} \log n.$$

At this point one should mention the two articles of B. F. Logan and L. A. Shepp [LS1, LS2] in which it is established that if the characteristic function of the probability law is $\exp(-|x|^\alpha)$, $0 < \alpha \leqslant 2$, then the expectation of the number of real zeros is $\sim c(\alpha) \log n$ where $\frac{2}{\pi} \leqslant c(\alpha) < 1$ seems to be a decreasing function of α. In particular $c(2) = \frac{2}{\pi}$ (Kac's result) and if $\alpha = 1$, then the probability law is the Cauchy law

$$P(E) = \frac{1}{\pi} \int_E \frac{dx}{1+x^2}$$

and

$$c(1) = \frac{8}{\pi^2} \int_0^\infty \frac{x e^{-x}}{x - 1 + 2e^{-x}}\, dx = 0.7413\ldots$$

whereas $c(2) = \frac{2}{\pi} = 0.6366\ldots$ In [Kac3] the author draws the graph of the density of zeros for a given n

$$t \longmapsto \frac{1}{\pi} \sqrt{\frac{1}{(t^2-1)^2} - \frac{(n+1)^2 t^{2n}}{(t^{2n+2}-1)^2}} = p_n(t)$$

and shows that it has two peaks around $t = \pm 1$ of height $\frac{n}{2\pi\sqrt{3}}(1 + o(1))$ which illustrates the fact that as n increases to infinity the real zeros tend to concentrate around ± 1, see figure 1.

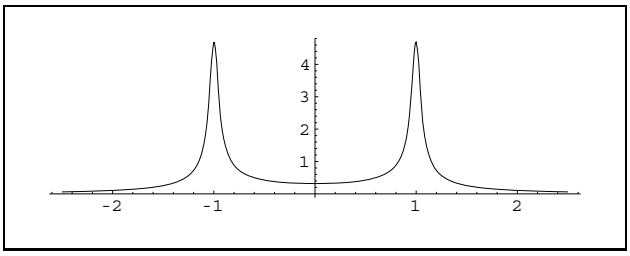

FIG. 1

This observation is in complete agreement with a result of P. Erdős and P. Turán [ET] according to which for large n the complex zeros of a large family of n degree polynomials are essentially located in the close vicinity of the circumference $|z| = 1$.

More precisely for all polynomials

$$P(X) = 1 + a_1 X + \cdots + a_n X^n$$

with $|a_j| \in (j^{-1}, j)$ all the zeros lie in the annulus $1 - \frac{1}{\sqrt{n}} \leqslant |z| \leqslant 1 + \frac{1}{\sqrt{n}}$ with the possible exception of $O(\sqrt{n \log n})$ zeros. Furthermore, within the annulus the arguments of the zeros are essentially uniformly distributed.

3. The Erdős-Offord paper

In a beautiful but difficult and highly technical article [EO], P. Erdős and A. C. Offord establish the following

Theorem 3. *Among the* 2^{n+1} *polynomials*

$$\sum_{j=0}^{n} \pm X^j$$

they all have

$$\frac{2}{\pi} \log n + o((\log n)^{2/3} \log\log n)$$

real zeros except for at most $o(2^n (\log\log n)^{-1/3})$ *polynomials.*

This is of course independent of Kac's theorem. Neither of these results imply the other.

4. The Edelman-Kostlan approach

In 1995, A. Edelman and E. Kostlan [EK] found a very simple proof of Kac's theorem : the main tool is an extension of Buffon's needle problem [Bu]. In this paragraph we propose to give the sketch of Kac's theorem following A. Edelman and E. Kostlan.

Theorem 4 (Kac)**.** *Let*

$$P(X) = \sum_{j=0}^{n} a_j X^j$$

be a real n degree polynomial with independent random coefficients according to the standard normal distribution law

$$\Pr\{a_j < \xi\} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\xi} e^{-\frac{1}{2}t^2}\, dt.$$

The expectation of the number of real zeros is

$$\begin{aligned} N(n) &= \frac{1}{\pi} \int_{-\infty}^{+\infty} \sqrt{\frac{1}{(t^2-1)^2} - \frac{(n+1)^2 t^{2n}}{(t^{2n+2}-1)^2}}\, dt \\ &= \frac{2}{\pi} \log n + 0.625\ldots + \frac{2}{\pi n} + O\left(\frac{1}{n^2}\right) \text{ [Edelman–Kostlan].} \end{aligned}$$

The main tool is the following theorem.

Theorem 5. *Let γ be a rectifiable curve drawn on the n-sphere*

$$\mathbb{S}^n = \left\{ (x_0, \ldots, x_n) \in \mathbb{R}^{n+1} \,\middle|\, \sum_{j=0}^{n} x_j^2 = 1 \right\}.$$

A random great circle

$$\sum_{j=0}^{n} a_j x_j = 0,\ \sum_{j=0}^{n} x_j^2 = 1$$

intersects γ on average in $\frac{1}{\pi}|\gamma|$ where $|\gamma|$ is the length of γ.

Proof. (Using Barbier's trick [Ba]) Suppose γ is a "small needle" of length ℓ drawn on $\mathbb{S}^n$, *i.e.* a small portion of a great circle. The expectation of the number of intersection points of γ with a random great circle is $\mathcal{E}(\gamma) = \varphi(\ell)$ where φ is an unknown function. Now if γ is the union of two needles $\gamma_1 \cup \gamma_2$ forming a larger needle, yet small again then

$$\varphi(\ell_1 + \ell_2) = \varphi(\ell_1) + \varphi(\ell_2)$$

since the expectation $\mathcal{E}(.)$ is linear. The above equation implies $\varphi(\ell) = c\ell$ for some constant c.

If γ is a broken line composed by k needles $\gamma_1, \ldots, \gamma_k$ then again by linearity

$$\mathcal{E}(\gamma) = \sum_{j=1}^{k} \varphi(\ell_j)$$

and therefore

$$\mathcal{E}(\gamma) = c \sum_{j=1}^{k} \ell_j = c|\gamma|.$$

By a limiting process (which is not so obvious!) we see that if γ is an arbitrary rectifiable curve on $\mathbb{S}^n$ then $\mathcal{E}(\gamma) = c|\gamma|$.

Choose in particular for γ a great circle. Clearly $\mathcal{E}(\gamma) = 2$ and $|\gamma| = 2\pi$ so that $c = 1/\pi$. This establishes Theorem 5.

Now we complete the proof of Kac's theorem. Consider the curve Γ in $\mathbb{R}^{n+1}$ defined by

$$\Gamma \begin{cases} x_0 = 1 \\ x_1 = t \\ x_2 = t^2 \\ \quad \vdots \\ x_n = t^n, \qquad t \in \mathbb{R}. \end{cases}$$

Projected on $\mathbb{S}^n$ we obtain

$$\gamma \begin{cases} x_0 = 1/\Big(\sum_{j=0}^{n} t^{2j}\Big)^{1/2} \\ \quad \vdots \\ x_n = t^n/\Big(\sum_{j=0}^{n} t^{2j}\Big)^{1/2}. \end{cases}$$

Intersecting γ by a random great circle is equivalent to counting the number of real zeros of

$$\sum_{j=0}^{n} a_j t^j.$$

By the above theorem we know that this is $\frac{1}{\pi}|\gamma|$. Kac's theorem is therefore reduced to computing the length of γ and this is trivial and leads to the integral representation of $N(n)$. It is then left to estimate the integral. □

The reader may have guessed that A. Edelman and E. Kostlan obtain many other results with their technique. For example, given $(n+1)$ rectifiable functions $f_0, f_1, \ldots, f_n$, the expectation of the number of zeros of

$$\sum_{j=0}^{n} a_j f_j(t)$$

is

$$\frac{1}{\pi} \int_{-\infty}^{+\infty} ||\gamma'(t)||\, dt$$

where

$$||\gamma'(t)||^2 = \sum_{j=0}^{n} \bar{f}_j'^{\,2}(t) \text{ and } \bar{f}_j(t) = \frac{f_j(t)}{\Big(\sum_{k=0}^{n} f_k^2(t)\Big)^{1/2}}$$

Remark due to M. Kac: The careful reader may have noticed that the statement of Kac's theorem involves independent standard normal distributions of the coefficients a_j:

$$\Pr\{a_j < \xi\} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\xi} e^{-\frac{1}{2}t^2}\, dt$$

whereas the proof uses the uniform probability on the sphere $\mathbb{S}^n$. In the first case the expectation of the number of real zeros is

$$N_1(n) = \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} Z(\underline{a}) e^{-\frac{1}{2}||\underline{a}||^2}\, da_1 \ldots da_n$$

where $Z(\underline{a})$ is the number of zeros of

$$\sum_{j=0}^{n} a_j X^j$$

and

$$||\underline{a}|| = \Big(\sum_{j=0}^{n} a_j^2\Big)^{1/2},$$

and in the second case the expectation is

$$N_2(n) = \frac{1}{\mathrm{vol}(\mathbb{S}^n)} \int_{\mathbb{S}^n} Z(\underline{a})\, d\sigma$$

where $d\sigma$ is the normalized uniform measure on $\mathbb{S}^n$. The claim is that $N_1(n) = N_2(n)$.

Indeed,

$$N_1(n) = \frac{1}{(2\pi)^{n/2}} \int_0^{+\infty} e^{-\frac{1}{2}r^2} \int_{\mathbb{S}_n(r)} Z(\underline{a})\, d\sigma_r dr$$

where $d\sigma_r = r^{n-1} d\sigma$ is the measure on the sphere $\sum_{j=0}^{n} x_j^2 = r^2$. Therefore

$$N_1(n) = \frac{1}{\mathrm{vol}(\mathbb{S}^n)} \int_{\mathbb{S}^n} Z(\underline{a})\, d\sigma = N_2(n)$$

since

$$\frac{\mathrm{vol}(\mathbb{S}^n)}{(2\pi)^{n/2}} \int_0^{+\infty} r^{n-1} e^{-\frac{1}{2}r^2}\, dr = 1.$$

□

5. Related problems

With Kac and Erdős-Offord we have a quite good knowledge of the average number of real zeros. But we are left with the difficult problem of estimating the optimal number of real zeros given that the coefficients of the polynomials are restricted to some conditions.

For example, D. Boyd [Bo] shows that real polynomials with ± 1 coefficients have at most $c \log^2 n / \log\log n$ zeros at 1.

I. Schur [Sc], G. Szegő [Sz] and P. Erdős and P. Turán [ET] establish that n degree polynomials with coefficients $0, \pm 1$ have at most $c\sqrt{n \log n}$ real zeros. See also [BE]. P. Borwein, T. Erdélyi and G. Kós have since shown in 1999 that the optimal bound is $c\sqrt{n}$ [BEK].

Many other problems come to mind concerning real zeros of real polynomials, some of which may be difficult. For example, find all k for which there is a polynomial of degree n

$$\sum_{j=0}^{n} \pm X^j$$

having exactly k real zeros.

6. Prefix polynomials

Let $\sum_{j=0}^{\infty} a_j X^j$ be a real series. We call n degree or n^{th} prefix polynomial the polynomial

$$\sum_{j=0}^{n} a_j X^j.$$

The number of real zeros of the n^{th} prefix polynomial will be denoted $Z(\underline{a}, n)$.

The results of Kac and Erdős-Offord may suggest the following definition. A real infinite sequence $\underline{a} = (a_n)$ is said to be MR (mimics randomness) if as n grows to infinity

$$Z(\underline{a}, n) \sim \frac{2}{\pi} \log n.$$

A less stringent condition would be

$$\frac{1}{N} \sum_{n=1}^{N} Z(\underline{a}, n) \sim \frac{2}{\pi} \log N$$

in which case we would have a weak MR sequence.

The problem we address here is to know whether MR (resp. weak MR) sequences exist. At the time of writing we only have very limited answers to the question.

A very first remark is that an ultimately periodic sequence is not MR neither is it weakly MR. Indeed, suppose p is the period of the tail of $\underline{a}$. For all sufficiently large n, say $n > n_0$, $a_{n+p} = a_n$. Let $n > n_0$. Then n has the unique representation $n = \ell + kp$ where $n_0 \leqslant \ell < n_0 + p$. The n^{th} prefix polynomial is then

$$\begin{aligned}
P_n(X) &= \sum_{j=0}^{\ell} a_j X^j + \sum_{j=\ell+1}^{\ell+kp} a_j X^j \\
&= A(X) + \sum_{j=\ell+1}^{\ell+p} a_j X^j + \sum_{j=\ell+p+1}^{\ell+2p} a_j X^j + \cdots + \sum_{j=\ell+(k-1)p+1}^{\ell+kp} a_j X^j \\
&= A(X) + \Big(\sum_{j=\ell+1}^{\ell+p} a_j X^j \Big) \Big(1 + X^p + \cdots + X^{kp} \Big) \\
&= A(X) + B(X) \frac{1 - X^{(k+1)p}}{1 - X^p} \\
&= \frac{A(X)(1 - X^p) + B(X)(1 - X^{(k+1)p})}{1 - X^p}.
\end{aligned}$$

The real zeros of P_n are the same as the numerator's apart maybe for $X = \pm 1$. The numerator is a polynomial which contains at most twice the number of nonzero terms of A and B; this number is bounded by $2\ell+2p+2 < 2n_0+4p+2$. The corollary of Descartes theorem then implies that the number of real zeros is $Z(\underline{a}, n) \leqslant 4n_0 + 8p + 3$. Therefore $Z(\underline{a}, n)$ is bounded independently of n. □

So, if MR sequences exist, they must be somehow more complex than ultimately periodic sequences.

The exceptional n degree polynomials of the Erdős-Offord Theorem (Theorem 3) are unfortunately too numerous to serve as the basis of a proof of the existence of MR sequences. Nevertheless, it does give some information. Indeed together with the Borel-Cantelli Theorem it is easy to see that there exists a sequence $n_K \sim \exp\exp\exp K^4$ such that for almost all sequences $\underline{a}$ in $\{-1,1\}^{\mathbb{N}}$, $Z(\underline{a}, n_K) \sim \frac{2}{\pi}\log n_K$. In particular,

$$\liminf_{n\to\infty} \frac{Z(\underline{a},n)}{\log n} \leqslant \frac{2}{\pi} \leqslant \limsup_{n\to\infty} \frac{Z(\underline{a},n)}{\log n}.$$

The absurd rate with which the sequence n_k increases can be improved by using result of Maslova [Mas]. Consider the measure

$$d\mu = \prod_{j=0}^{\infty} \left(\frac{e^{-\frac{1}{2}a_j^2}}{\sqrt{2\pi}}\, da_j \right)$$

defined on $\mathbb{R}^{\mathbb{N}}$. She computes the variance

$$\int_{\mathbb{R}^{\mathbb{N}}} |Z(\underline{a},n) - E(Z(\underline{a},n))|^2\, d\mu \sim \frac{4}{\pi}\left(1-\frac{2}{\pi}\right)\log n$$

from which it follows from the Beppo Levi principle that for all $\delta > 0$ and for all sequences $n_K \sim e^{K^{1+\delta}}$, μ-almost all $\underline{a}$ are such that $Z(\underline{a}, n_K) \sim \frac{2}{\pi}\log n_K$. Proving that $\liminf = \limsup$ would then establish the existence of MR sequences.

Not being able to obtain definite results neither by probabilistic methods nor by ergodic methods, one can try to construct a MR sequence. The difficulty is to control the number of real zeros of polynomials. The idea could then be to consider sequences $\underline{a}$ for which we do have some information concerning the zeros, and then to hope to prove that they satisfy $Z(\underline{a}, n) \sim \frac{2}{\pi}\log n$. The results we obtain are unfortunately not sufficiently precise. We shall nevertheless present them here [DMF], [Do].

To every integer $n \geqslant 0$ we associate the sum $s(n)$ of its binary digits. The Thue-Morse sequence is defined by $\varepsilon_n = (-1)^{s(n)}$ and it is an easy matter to verify that

$$\sum_{k=0}^{2^n-1} \varepsilon_k X^k = \prod_{j=0}^{n-1} \left(1 - X^{2^j}\right).$$

The real zeros of this polynomial are obvious, namely $+1$ with multiplicity n and -1 with multiplicity $n-1$. The number of real zeros is therefore $2n-1$. In other words, if $\underline{\varepsilon}$ is the Thue-Morse sequence

$$Z(\underline{\varepsilon}, 2^n - 1) = \frac{2}{\log 2}\log(2^n-1) + O(1)$$

and as a consequence

$$\limsup_{N\to\infty} \frac{Z(\underline{\varepsilon},N)}{\log N} \geqslant \frac{2}{\log 2}.$$

Unfortunately, $\frac{2}{\log 2} > \frac{2}{\pi}$ so that we can already conclude that $\underline{\varepsilon}$ is not MR, even though much more complex than an ultimately periodic sequence. It is actually not even almost-periodic in the sense of Bohr and Besicovitch. We shall come back to that later.

The Thue-Morse sequence is far from being MR. And indeed, it is possible, but not easy at all to compute the number of real zeros of the prefix polynomials of the Thue-Morse sequence [DMF].

Theorem 6. *Let $\underline{\varepsilon}$ be the Thue-Morse sequence.*

1. *If $n \in \mathbb{N}$ is even, $Z(\underline{\varepsilon}, n) \leqslant 2$. More precisely:*
If $n \equiv 0\ (4)$ and $\varepsilon_n = 1$ then $Z(\underline{\varepsilon}, n) = 0$.
If $n \equiv 2\ (4)$ and $\varepsilon_n = 1$ then $Z(\underline{\varepsilon}, n) = 2$ and both real zeros are negative.
If $n \equiv 0\ (2)$ and $\varepsilon_n = -1$ then $Z(\underline{\varepsilon}, n) = 2$ and the real zeros are of different sign.

2. *If $n \in \mathbb{N}$ is odd, let us note $\nu(n)$ the 2-adic valuation of $n+1$.*
Then $Z(\underline{\varepsilon}, n) = 2\nu(n) + (-1)^{\nu(n)+1}\varepsilon_n$.

We now state two simple consequences.

Corollary 7.

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} Z(\underline{\varepsilon}, k) = \frac{11}{4}\cdot$$

This result underlines once more how far this sequence is from being MR.

Corollary 8. *For every $\alpha \in [0, \frac{2}{\log 2}]$ there exists a subsequence $n_k = k^{(1-\frac{\alpha}{2}\log 2)^{-1}+o(1)}$ such that $Z(\underline{\varepsilon}, n_k) \sim \alpha \log n_k$. In particular, choosing $\alpha = 2/\pi$, $n_k \sim k^{(1-\frac{\log 2}{\pi})^{-1}+o(1)}$ and $Z(\underline{\varepsilon}, n_k) \sim \frac{2}{\pi}\log n_k$.*

In the search for a MR sequence we introduce the extended Thue-Morse sequences. Let $g_1, g_2, \ldots$ an infinite sequence of integers, $g_n \geqslant 2$. Define the sequence of polynomials

$$Q_n(X) = 1 + \delta_{1,n}X + \delta_{2,n}X^2 + \cdots + \delta_{g_{n+1}-1,n}X^{g_{n+1}-1}$$

where $\delta_{i,j} = \pm 1$. It is easily verified that

$$\prod_{n=0}^{\infty} Q_n(X^{g_1 g_2 \cdots g_n}) = \sum_{j=0}^{\infty} \varepsilon_j X^j$$

where $\varepsilon_j = \pm 1$. For the special choice $g_n = 2$ and $Q_n(X) = 1 - X$, $n = 1, 2, 3, \ldots$ we recover the Thue-Morse sequence. And incidentally, if the g_n's are arbitrary and all the $\delta_{i,j}$ are $+1$ then

$$\prod_{n=0}^{\infty} Q_n(X^{g_1 \cdots g_n}) = (1 - X)^{-1}.$$

Extending a result in [Do] and choosing for $Q_n(X)$ polynomials for which the number of real zeros is equivalent to $\frac{2}{\pi}\log(g_{n+1} - 1)$ (for $g_n \to \infty$ this is possible according to the Erdős-Offord Theorem 3) one would then be able to show that there exist infinite sequences $\underline{\varepsilon} \in \{-1, +1\}^{\mathbb{N}}$ such that

$$\frac{1}{\pi} \leqslant \liminf_{n\to\infty} \frac{Z(\underline{\varepsilon}, n)}{\log n} \leqslant \frac{2}{\pi} \leqslant \limsup_{n\to\infty} \frac{Z(\underline{\varepsilon}, n)}{\log n}\cdot$$

This is a slight improvement of a previous remark in that here one asserts that the lim inf is strictly positive.

All this is quite frustrating and in order to confirm one's feelings it may be useful to have some numerical evidence. We have arbitrarily chosen to test some

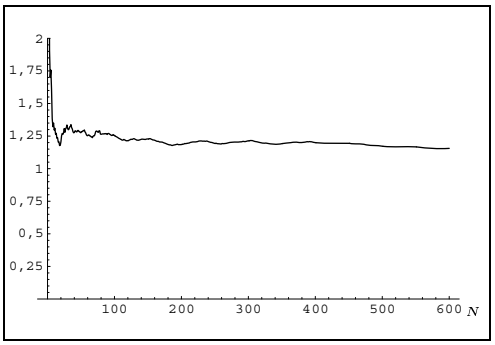

FIG. 2

well-known ± 1 infinite sequences. The first one is the sequence of odd primes p_n reduced (mod 4) and indeed up to the degree $n = 600$ it seems to behave à la MR.

The second sequence we tested is $(-1)^{\lfloor 10^i \pi \rfloor}$ which also seems quite convincing, see figures 2 and 3 where we plotted

$$\frac{1}{\frac{2N}{\pi} \log N} \sum_{k=1}^{N} Z(\underline{a}, k)$$

against $N \leqslant 600$.

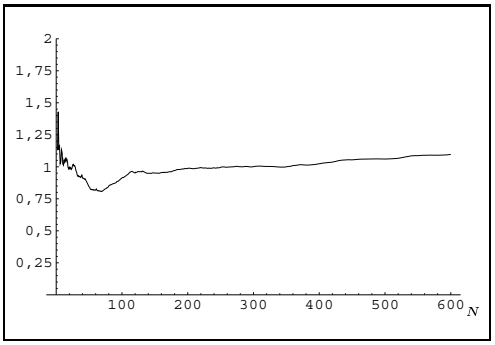

FIG. 3

7. A LAST REMARK CONCERNING MR SEQUENCES

If, as we guess, MR sequences really do mimic randomness, then we should expect that they behave as such, and that in particular their spectral properties should be close to that of random sequences.

Given a real or complex infinite sequence $\underline{a} = (a_n)$, we define the correlation sequence γ_a

$$\gamma_a(k) = \lim_{N \to \infty} \frac{1}{N} \sum_{n=0}^{N-1} \overline{a_n} a_{n+k}, \ k \in \mathbb{Z}$$

provided the limit exists. A beautiful result of Bochner's and Herglotz' [Kat] asserts the existence of a measure Λ_a (the spectral measure) such that

$$\gamma_a(k) = \int_0^1 e^{2i\pi kt}\, \Lambda_a(dt).$$

For a centered random sequence, $\Lambda_a(dt) = dt$ *i.e.* $\gamma_a(0) = ||\underline{a}||^2$ (the Marcinkiewicz semi-norm) and $\gamma_a(k) = 0$ for $k \neq 0$.

One would guess that this is the rule for a MR sequence. For an almost-periodic sequence, Λ_a is the sum of Dirac measures and so such a sequence is probably not MR. For the Thue-Morse sequence, Λ_ε is a continuous measure, purely singular (see [Mah], [Kak]) and so once again this underlines the distance the Thue-Morse is from being MR. As for the extended Thue-Morse sequences, the situation is more complicated: if the g_n's are uniformly bounded, Λ_ε is continuous purely singular except for 2 trivial cases where Λ_ε is the Dirac measure δ. If however $g_n \to \infty$, one could hope to get a spectral measure Λ_ε somehow smoother than in the case of bounded g_n.

Appendix. Complexity

F. Cucker and M. F. Roy discuss the problem of computing real algebraic numbers and their complexity [CR]. They show that the complexity depends on the "norm" of the polynomial $(\sum_0^n a_j^2)^{1/2}$ where the a_j are the coefficients, and on the number of real zeros, and of course on the degree of the polynomial. We shall not develop their theme here. See also the references in [CR].

Another possible way to define the complexity of a real polynomial could be as follows.

In previous articles (see for example [MF]) the second named author introduced the notion of entropy of a rectifiable finite curve Γ

$$\log \frac{2L}{C}$$

where L is its length and C, the so-called perimeter is the length of the boundary of the convex hull of Γ.

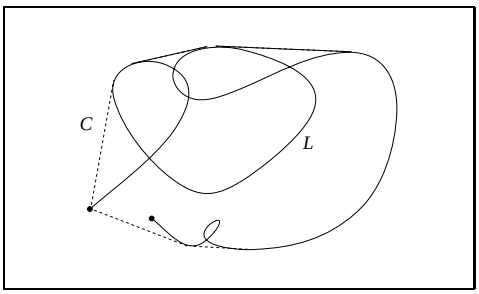

Fig. 4

Now consider a real polynomial $P(X)$ and its graph Γ. If A and B are two points on Γ, denote by $L(A,B)$ the length of the arc $\overset{\frown}{AB}$ on Γ and let $C(A,B)$ be the perimeter. We define the entropy

$$\gamma_1(P) = \sup_{A,B\in\Gamma} \log \frac{2L(A,B)}{C(A,B)}.$$

We now introduce another notion of entropy $\gamma_2(P)$.

If $Q(X)$ is any real polynomial of degree at least 2, let $Z^*(Q(X))$ be the number of distinct real zeros of $Q(X)$. Define

$$\gamma_2(P) = \sup_{a,b\in\mathbb{R}} \log Z^*(P(X) - aX - b).$$

$\gamma_1(P)$ describes the complexity in terms of the shape of the graph. $\gamma_2(P)$ depends on the number of distinct real zeros of a related polynomial.

Theorem 9. *On has*

$$\gamma_1(P) \leqslant \gamma_2(P).$$

The proof is actually quite trivial since by a theorem of H. Steinhaus $2L(A,B)/C(A,B)$ is the average number of intersecting points of the arc $\overset{\frown}{AB}$ with a "random" straight line [St, Sa], and since $\sup_{a,b\in\mathbb{R}} Z^*(P(X)-aX-b)$ is the maximum number of such intersections. It is easily seen that for a 2-degree polynomial the two entropies coincide.

This approach leads to a curious inequality which we now discuss. Let us consider an arbitrary rectifiable map $f : [\alpha,\beta] \to \mathbb{R}$. We have just seen that

$$\frac{2}{C(\alpha,\beta)} \int_\alpha^\beta \sqrt{1+f'^2(t)}\, dt \leqslant \sup_{a,b} Z^*(f(X) - aX - b)$$

where $C(\alpha,\beta)$ is the perimeter of the graph of f. Quite obviously

$$C(\alpha,\beta) \leqslant 2(\beta - \alpha + f_M - f_m)$$

where

$$\begin{aligned} f_M &= \sup_{x\in[\alpha,\beta]} f(x) \\ f_m &= \inf_{x\in[\alpha,\beta]} f(x). \end{aligned}$$

On the other hand,

$$\int_\alpha^\beta \sqrt{1+f'^2(t)}\, dt > \int_\alpha^\beta |f'(t)|\, dt$$

so that

$$\sup_{a,b} Z^*(f(X) - aX - b) \geqslant \frac{\int_\alpha^\beta |f'(t)|\, dt}{2(\beta - \alpha + f_M - f_m)}.$$

Replace f by λf where $\lambda > 0$ is a parameter which tends to infinity. The inequality now reads

$$\sup_{a,b} Z^*(f(X) - aX - b) \geqslant \frac{\int_\alpha^\beta |f'(t)|\, dt}{2(f_M - f_m)}.$$

Finally let us specialize f, α and β. We choose $\alpha = 0$, $\beta = N$ and f continuous on $[0,N]$ affine on each interval $]n, n+1[$, $n = 0, 1, \ldots, N-1$.

$Z^*(f(X) - aX - b)$ now represents the number of sign changes of the sequence $n \mapsto f(n) - an - b$. We then arrive at the following result

Theorem 10. *Given $f : \{0, 1, \dots, N\} \mapsto \mathbb{R}$. There exists a couple (a, b) such that the number of sign changes of the sequence $f(n) - an - b$ is larger than*

$$\frac{\sum_{n=0}^{N-1} |f(n+1) - f(n)|}{2 \max_{0 \leqslant k, \ell \leqslant N} |f(k) - f(\ell)|}.$$

We started out with Descartes' rule of signs and we conclude with a result on sign changes.

Acknowledgements

We wish to thank Martine Queffelec, Hugh Montgomery and Jean-Marc Deshouillers with whom we had many discussions.

References

[AG] A. Alesina, M. Galuzzi, *A new proof of Vincent's Theorem*, Enseignement Mathématique **44** (1998), 219–256.

[Ba] E. Barbier, *Note sur le problème de l'aiguille et le jeu du joint couvert*, J. Mathématiques Pures et Appliquées **5** (1860), 273–286.

[BE] P. Borwein, T. Erdélyi, *Polynomials and polynomial inequalities*, Springer-Verlag 1995.

[BEK] P. Borwein, T. Erdélyi, G. Kós, *Littlewood-type problems on* $[0, 1]$, Proc. London Math. Soc. **79** (1999), 22–46.

[Bo] D. Boyd, *On a problem of Byrnes concerning polynomials with restricted coefficients*, Math. Comp. **66** (1997), 1697–1703.

[BP] A. Bloch, G. Pólya, *On the roots of certain algebraic equations*, Proc. London Math. Soc. **33** (1932), 102–114.

[Bu] G. Buffon, *Essai d'arithmétique morale*, in Supplément à l'histoire naturelle, 1777.

[CR] F. Cucker, M. F. Roy, *A theorem on random polynomials and some consequences in average complexity* , J. Symbolic Comput. **10** (1990), 405–409.

[De] R. Descartes, *La géométrie* in Discours de la méthode, Ian Maire 1637 (Leyde) and its english version *The geometry of René Descartes* translated by D. E. Smith and M. L. Lathan, Dover 1954.

[DMF] C. Doche, M. Mendès France, *Integral geometry and real zeros of Thue-Morse polynomials*. Experiment. Math. **9** (2000), 339–350.

[Do] C. Doche, *On the real roots of generalized Thue-Morse polynomials*, Acta Arith. XCIX.4 (2001), 309–319.

[EK] A. Edelman, E. Kostlan, *How many zeros of a random polynomial are real?* Bull. (New Series) AMS **32** (1995), 1–37.

[EO] P. Erdős, A. C. Offord, *On the number of real roots of a random algebraic equation*, Proc. London Math. Soc. **6** (1956), 139–160.

[ET] P. Erdős, P. Turán, *On the distribution of roots of polynomials*, Ann. of Math. **51** (1950), 105–119.

[Kac1] M. Kac, *On the average number of real roots of a random algebraic equation*, Bull. Amer. Math. Soc. **49** (1943), 314–320 and 938.

[Kac2] M. Kac, *On the average number of real roots of a random algebraic equation (II)*, Proc. London Math. Soc. **50** (1949), 390–408.

[Kac3] M. Kac, *Probability and related topics in physical sciences*, Lectures in applied mathematics, first printing Interscience publishers 1960, Second printing A.M.S. 1976.

[Kak] S. Kakutani, *Strictly ergodic symbolic dynamical systems,* in Proceedings of the 6th Berkeley symposium on mathematical statistics and probability, [Berkeley 1970] vol. **2** (Berkeley, University of California Press, 1972), 319–326.

[Kat] Y. Katznelson, An introduction to harmonic analysis, Dover Publications Inc., New York 1976, xiv+264

[KR] D. A. Klain, G. C. Rota, *Introduction to geometric probability*, Lezioni lincei, Accademia nazionale dei lincei Cambridge Univ. Press, 1997.

[LO1] J. E. Littlewood, A. C. Offord, *On the number of real roots of a random algebraic equation*, J. London, Math. Soc. **13** (1938), 288–295.

[LO2] J. E. Littlewood, A. C. Offord, *On the number of real roots of a random algebraic equation II*, Proc. Cambridge, Philo. Soc. **35** (1939), 133–148.

[LS1] B. F. Logan and L. A. Shepp, *Real zeros of random polynomials. I*, Proc. London Math. Soc. **18** (1968), 29–35.

[LS2] B. F. Logan and L. A. Shepp, *Real zeros of random polynomials. II* Proc. London Math. Soc. **18** (1968), 308–314.

[Mah] K. Mahler, *On the translation properties of a simple class of arithmetical functions*, Jour. Math. and Phys. **6**, (1927), 158–163.

[Mas] N. B. Maslova, *The variance of the number of real roots of random polynomials*, Teor. Verojatnost. i Primenen. **19** (1974), 36–51.

[MF] M. Mendès France, *The Planck constant of a curve*, Fractal geometry and analysis (Montreal, PQ, 1989), Kluwer Acad. Publ., Dordrecht 1991, 325–366.

[Sa] L. A. Santaló, *Integral geometry and geometric probability*, Addison-Wesley (1976).

[Sc] I. Schur, *Untersuchungen über algebraische gleichungen*, Sitz. Preuss. Akad. Wiss. Phys. Math. Kl. (1933), 403–428.

[St] H. Steinhaus, *Length, shape and area*, Colloquium Mathem. (1955), 1–13.

[Sz] G. Szegö, *Bemerkungen zu einem Satz von E. Schmidt über algebraische gleichungen*, Sitz. Preuss. Akad. Wiss. Physics Math. Kl. (1934), 86–98.

Laboratoire d'Algorithmique Arithmétique, Université Bordeaux I, 351 cours de la Libération, F-33405 Talence Cedex France.

E-mail address: cdoche@math.u-bordeaux.fr

E-mail address: mmf@math.u-bordeaux.fr