THE SIGNUM EQUATION FOR ERDŐS-SURÁNYI SEQUENCES

Dorin Andrica
Faculty of Mathematics and Computer Science, Babeş-Bolyai University, Cluj-Napoca, Romania
dandrica@math.ubbcluj.ro

Eugen J. Ionascu[1]
Department of Mathematics, Columbus State University, Columbus, Georgia
ionascu@columbusstate.edu

Received: 3/1/14, Revised: 3/23/15, Accepted: 5/9/15, Published: 6/15/15

Abstract

For an Erdős-Surányi sequence it is customary to consider its signum equation. Based on some classical heuristic arguments, we conjecture the asymptotic behavior for the number of solutions of this signum equation in the case of the sequence $\{n^k\}_n$ ($k \geq 2$) and the sequence of primes. Surprisingly, we show that this method does not apply at all for the Fibonacci sequence. By computing the precise number of solutions, in this case, we obtain an exponential growth, which shows, in particular, the limitations of such an intuition.

1. INTRODUCTION

We recall (see [6] for $a_n = n^2$) that a sequence of distinct positive integers $\{a_m\}_{m\geq 1}$ is an *Erdős-Surányi sequence* if every integer k can be written in the form

$$k = \pm a_1 \pm a_2 \pm \cdots \pm a_n$$

for some choices of signs $+$ and $-$, in infinitely many ways. Using an induction argument, it is not difficult to see that $\{m\}_{m\geq 1}$, or even $\{m^2\}_{m\geq 1}$, is an Erdős-Surányi sequence. In 1979, J. Mitek proved in [8] that in fact, for an arbitrary positive integer s, $\{m^s\}_{m\geq 1}$ is an Erdős-Surányi sequence.

M. O. Drimbe gives ([4] and [5]) sufficient conditions for a sequence to be an Erdős-Surányi sequence.

Theorem 1.1 *Let $\{a_m\}_{m\geq 1}$ be a sequence of distinct positive integers such that $a_1 = 1$ and for every $n \geq 1$, $a_{n+1} \leq a_1 + \cdots + a_n + 1$. If the sequence $\{a_m\}_{m\geq 1}$ contains infinitely many odd integers, then it is an Erdős-Surányi sequence.*

[1]Honorific Member of the Romanian Institute of Mathematics "Simion Stoilow"

As an application, one can check that the Fibonacci sequence satisfies the hypothesis of this theorem, since $F_0 + F_1 + \cdots + F_n = F_{n+2} - 1$, $n \geq 1$. Since $F_n + F_{n+1} - F_{n+2} = 0$ for $n \geq 0$, it is not surprising that $k = \pm F_1 \pm F_2 \pm \cdots \pm F_n$ has infinitely many representations if it has one. The existence of at least one representation is related to Zeckendorf's theorem (see [11]) and as a curiosity we include one such representation for $k = 2015$:

$$2015 = \sum_{j=1}^{19} \epsilon_j F_j, \text{ with } \epsilon_j = 1, \text{ if and only if } j \in \{4, 11, 15, 17, 19\}.$$

Theorem 1.2 *Let $f \in \mathbb{Q}[X]$ be a polynomial such that for any $n \in \mathbb{Z}$, $f(n)$ is an integer. If the greatest common factor of the terms of the sequence $\{f(n)\}_{n\geq 1}$ is equal to 1, then $\{f(n)\}_{n\geq 1}$ is an Erdős-Surányi sequence.*

Using this result, one can obtain many Erdős-Surányi sequences: $\{(an-1)^k\}_n$ for fixed $a, k \in \mathbb{N}$, or $\{\binom{n+s}{s}\}_n$ for fixed $s \geq 2$, are just some possible examples. We found that it is difficult to use an induction type argument to show that these are Erdős-Surányi sequences.

In this paper, given an Erdős-Surányi sequence $\mathbf{a} = \{a_m\}_{m\geq 1}$, we are interested in the *signum equation* of $\mathbf{a}$ at level n, that is

$$\pm a_1 \pm a_2 \pm \cdots \pm a_n = 0. \tag{1}$$

For a fixed integer n, a *solution* to the signum equation is a choice of signs $+$ and $-$ that makes (1) true. We denote by $S_{\mathbf{a}}(n)$ the number of solutions of Equation (1). Clearly, if 2 does not divide u_n, where $u_n = a_1 + \cdots + a_n$, then we have $S_{\mathbf{a}}(n) = 0$.

We include next a few other properties of the sequence $S_{\mathbf{a}}(n)$ (see [1], [2]).

(i) $S_{\mathbf{a}}(n)/2^n$ is the unique real number α having the property that the function $f : \mathbb{R} \to \mathbb{R}$, defined by

$$f(x) = \begin{cases} \cos\frac{a_1}{x}\cos\frac{a_2}{x}\cdots\cos\frac{a_n}{x} & \text{if } x \neq 0, \\ \alpha \ \text{ if } \ x = 0, \end{cases}$$

is a derivative, i.e., it is the derivative of a differentiable function on $\mathbb{R}$.

(ii) $S_{\mathbf{a}}(n)$ is the term not depending on z in the expansion of

$$(z^{a_1} + \frac{1}{z^{a_1}})(z^{a_2} + \frac{1}{z^{a_2}})\cdots(z^{a_n} + \frac{1}{z^{a_n}}).$$

(iii) $S_{\mathbf{a}}(n)$ is the coefficient of $z^{u_n/2}$ in the polynomial

$$(1 + z^{a_1})(1 + z^{a_2})\cdots(1 + z^{a_n}).$$

(iv) $S_{\mathbf{a}}(n)$ can be calculated using the following integral formula

$$S_{\mathbf{a}}(n) = \frac{2^{n-1}}{\pi} \int_0^{2\pi} \cos(a_1 t)\cos(a_2 t)\cdots\cos(a_n t)dt. \tag{2}$$

(v) $S_{\mathbf{a}}(n)$ is the number of ordered bipartitions of the set $\{a_1, a_2, \cdots, a_n\}$ into two classes having equal sums.

(vi) $S_{\mathbf{a}}(n)$ is the number of distinct subsets of $\{a_1, a_2, \cdots, a_n\}$ whose elements sum to $u_n/2$ if 2 divides u_n, and $S_{\mathbf{a}}(n) = 0$ otherwise.

(vii) $S_{\mathbf{a}}(n) = 2N_{\mathbf{a}}(n)$, where $N_{\mathbf{a}}(n)$ is the number of representations of a_n as

$$a_n = \pm a_1 \pm a_2 \pm \cdots \pm a_{n-1}.$$

In the next section, we are interested in the asymptotic behavior of $S_{\mathbf{a}}(n)$, when $n \to \infty$.

2. Heuristic Asymptotic Behavior for $S_{\mathbf{a}}(n)$

S. Finch [7] has a very interesting probabilistic but heuristic approach to arrive at asymptotic results for $S_{\mathbf{a}}(n)$. In this section, we are going to describe it in detail starting with an Erdős-Surányi sequence $\mathbf{a}$. For a positive integer j, we consider the random variable E_j taking two values: a_j or $-a_j$ with probability $1/2$ each. We assume that for $i \neq j$, each E_i is independent of E_j. We then let the random variable

$$S_n := E_1 + E_2 + \cdots + E_n. \tag{3}$$

Since each of the E_j has expectation zero, we see that the expectation of S_n is $E(S_n) = \sum_{j=1}^n E(E_j) = 0$. Because each of the E_j has variation σ_j with $\sigma_j^2 = \mathrm{Var}(E_j) = E(|E_j - E(E_j)|^2) = a_j^2$, and the random variables are mutually independent, we conclude that

$$s_n^2 = \mathrm{Var}(S_n) = \sum_{j=1}^n \mathrm{Var}(E_j) = \sum_{j=1}^n a_j^2.$$

We would like to use the Berry-Essen Theorem ([3]), which is a very good version of the Central Limit Theorem for this situation.

Theorem 2.1 *Let X_1, X_2, ..., be independent random variables with $E(X_j) = 0$, $E(X_j^2) = \sigma_j^2 > 0$ and $\rho_j = E(|X_j|^3) < \infty$. If*

$$G_n = (X_1 + X_2 + \cdots + X_n) \Big/ \sqrt{\sum_{j=1}^n \sigma_j^2},$$

then

$$\sup_{x\in\mathbb{R}}\left|P(G_n \leq x) - \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{t^2}{2}}\,dt\right| \leq C_1 \frac{\max\limits_{1\leq j\leq n}\frac{\rho_j}{\sigma_j^2}}{\sqrt{\sum_{j=1}^{n}\sigma_j^2}} \tag{4}$$

for some universal constant C_1*.*

Theorem 2.2 *Let* $\mathbf{a} = \{a_n\}_{n\geq 1}$ *be a non-decreasing Erdős-Surányi sequence of positive integers, and let* $\{S_{\mathbf{a}}(n)\}_{n\geq 1}$ *be the sequence counting the number of solutions of its signum equations. If we set* $s_n^2 = a_1^2 + a_2^2 + \cdots + a_n^2$*, then*

$$\left|\frac{1}{2^n}S_{\mathbf{a}}(n) - \sqrt{\frac{2}{\pi}}\frac{1}{s_n}\right| \leq C\frac{a_n}{s_n}, \tag{5}$$

for n *big enough and for some constant* C *(which does not depend on* $\mathbf{a}$*).*

Proof. We are setting $X_j = E_j$ in Theorem 2.1 and observe that $\rho_j = a_j^3$ and so $\max\limits_{1\leq j\leq n}\frac{\rho_j}{\sigma_j^2} = a_n$. Inequality (4) then becomes

$$\sup_{x\in\mathbb{R}}|P(S_n \leq xs_n) - \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{t^2}{2}}\,dt| \leq C_1\frac{a_n}{s_n}.$$

In particular for $y < x$, we have

$$|P(ys_n < S_n \leq xs_n) - \frac{1}{\sqrt{2\pi}}\int_{y}^{x} e^{-\frac{t^2}{2}}\,dt| \leq 2C_1\frac{a_n}{s_n}.$$

For an arbitrary $\epsilon \in (0,1)$ and $x = \frac{1-\epsilon}{s_n}$, $y = -\frac{1}{s_n}$, let us observe that

$$P(ys_n < S_n \leq xs_n) = P(-1 < S_n < 1-\epsilon) = \frac{1}{2^n}S_{\mathbf{a}}(n)$$

and so the inequality above can be written as

$$\left|\frac{1}{2^n}S_{\mathbf{a}}(n) - \frac{1}{\sqrt{2\pi}}\int_{-\frac{1}{s_n}}^{\frac{1-\epsilon}{s_n}} e^{-\frac{t^2}{2}}\,dt\right| \leq 2C_1\frac{a_n}{s_n}.$$

Since

$$\int_{-\frac{1}{s_n}}^{\frac{1-\epsilon}{s_n}} e^{-\frac{t^2}{2}}\,dt \approx \int_{-\frac{1}{s_n}}^{\frac{1-\epsilon}{s_n}} dt = \frac{2-\epsilon}{s_n}$$

letting $\epsilon \to 0$ and increasing the constant $2C_1$ slightly, we get

$$\left|\frac{1}{2^n}S_{\mathbf{a}}(n) - \sqrt{\frac{2}{\pi}}\frac{1}{s_n}\right| \leq C\frac{a_n}{s_n}.$$

□

Remark: In general, the constant C in (5) is the best that one can obtain, but in some particular cases the constant C can be replaced by a sequence convergent to zero. If this sequence is such that $Ca_n = o(1)$ then (5) becomes a genuine asymptotic formula:

$$\frac{S_{\mathbf{a}}(n)}{2^n s_n^{-1}} = \sqrt{\frac{2}{\pi}} + o(1). \tag{6}$$

We next discuss instances when it is very likely that this is the case and an example when this is clearly false.

2.1. The Sequence n^k

If $k \in \mathbb{N}$, and $\mathbf{a} = \{n^k\}_n$, it is easy to show that for $\alpha > 0$ we have

$$\sum_{j=1}^{n} j^\alpha = \frac{n^{\alpha+1}}{\alpha+1} + O(n^\alpha).$$

This shows that in this case $s_n = \frac{1}{\sqrt{2k+1}} n^{k+\frac{1}{2}} \left(1 + O(n^k)\right)$. The asymptotic formula (6) suggests that we can formulate the following:

Conjecture 1. For every positive integer k the following relation holds:

$$\lim_{\substack{n\to\infty \\ n\equiv 0 \ or\ 3\ (mod\ 4)}} \frac{S_k(n)}{2^n n^{-\frac{2k+1}{2}}} = \sqrt{\frac{2(2k+1)}{\pi}}, \tag{7}$$

where $S_k(n)$ stands for $S_{\mathbf{a}}(n)$ in this case.

For $k = 1$ the relation (7) was stated in the paper [2] and it was called the Andrica-Tomescu Conjecture in the recent paper [10]. It was recently proved to be correct by B. Sullivan [10]. For $k \geq 2$, although we do not yet have a proof, we believe that under similar restrictions on n as in the case $k = 1$, (7) is still valid.

2.2. The Sequence of Primes

The sequence $\{\mathbf{a}\} = \{\mathbf{p}\}$ of primes $p_1 = 2$, $p_2 = 3$, ... is an Erdös-Surányi sequence. This was shown in an ingenious way in the paper [4] by using the result in Theorem 1.1. Using the Prime Number Theorem, and a result from ([9]), we derive the following asymptotic identity:

$$s_n^2 = \sum_{j=1}^{n} p_j^2 = (1+o(n))\frac{1}{3}n^3\frac{(\ln n + \ln(\ln n) - 1)^3}{\ln n} = (1+o(n))\frac{1}{3}n^3(\ln n)^2.$$

Since the order of growth of s_n is not greater than n^2, Equation (6) suggests that the following may be true.

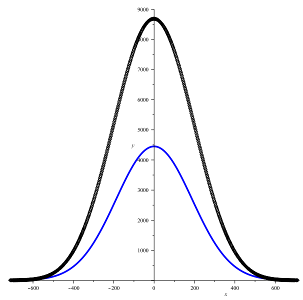

Figure 1: Distribution of $S_{\mathbf{p},\mathbf{21}}$ and $2^{21}\mathcal{N}(0,188)$

Conjecture 2. If $\mathbf{p}$ is the sequence of primes, then

$$\lim_{n\to\infty,\ n\ odd} \frac{S_{\mathbf{p}}(n)}{2^n} n\sqrt{n}\ln n = \sqrt{\frac{6}{\pi}}. \tag{8}$$

In 1830, H. F. Scherk conjectured the following "recursive" formulas generating the m-th prime from the previous $m-1$ primes: for every positive integer n,

$$p_{2n} = 1 \pm p_1 \pm p_2 \pm \cdots \pm p_{2n-2} + p_{2n-1}$$

and

$$p_{2n+1} = 1 \pm p_1 \pm p_2 \pm \cdots \pm p_{2n-1} + 2p_{2n},$$

for some choices of signs + and −. Such representations were shown to exist in 1928 by S. S. Pillai.

According to property (viii) (page 3), the number of representations of p_{2n+1} as

$$p_{2n+1} = \pm p_1 \pm p_2 \pm \cdots \pm p_{2n}$$

for some choices of signs + and − is $N_{\mathbf{p}}(2n+1) = \frac{1}{2}S_{\mathbf{p}}(2n+1)$. It is easy to see that $S_{\mathbf{p}}(n) = 0$ for every even n and for instance $S_{\mathbf{p}}(1) = 0$, $S_{\mathbf{p}}(3) = 2$, $S_{\mathbf{p}}(5) = 2$ $(2-3+5+7-11=0)$,..., which is precisely half the sequence A022894 in the Online Encyclopedia of Integer Sequences. Using the first 1000 terms, we calculated that

$$\frac{S_{\mathbf{p}}(n)}{2^n} n\sqrt{n}\frac{(\ln n + \ln(\ln n) - 1)^{3/2}}{\sqrt{\ln n}} \approx 1.38197 \approx \sqrt{\frac{6}{\pi}} \text{ for } n \text{ odd}, \ n > 199,$$

which supports Conjecture 2.

In Figure 1, we have included the distribution of $X := \pm p_1 \pm p_2 \cdots \pm p_{21}$ which has zero mean and a standard deviation of about 187.98. Here X takes even values from -712 to 712. As we can see, this graph differs clearly from that of the dilated normal distribution $2^{21}\mathcal{N}(0,188)$ but if we account for the half the values, the odd values where the probability is zero, we see an almost perfect matching with $2^{22}\mathcal{N}(0,188)$ (Figure 2) (the factor of 2 which shows up in the proof of Theorem 2.3).

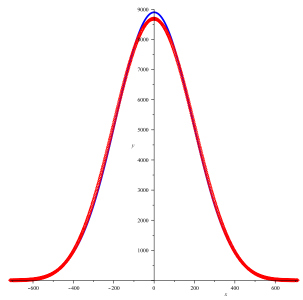

Figure 2: Distribution of $S_{\mathbf{p},\mathbf{21}}$ and $2^{22}\mathcal{N}(0,188)$

2.3. The Fibonacci Sequence

In contrast, let us look at the well-known sequence $F_0 = 0$, $F_1 = F_2 = 1$, $F_3 = 2$, ..., $F_{n+1} = F_n + F_{n-1}$ for $n \geq 2$, and at the related sequence $\{S_{\mathbf{F}}(n)\}_{n\geq 0}$, where $S_{\mathbf{F}}(n)$ is the number of writings of 0 as a sum $\pm F_0 \pm F_1 \pm F_2 \pm \cdots \pm F_n$. For instance, it is easy to check that $S_{\mathbf{F}}(0) = 2$, $S_{\mathbf{F}}(1) = 0$ and $S_{\mathbf{F}}(2) = 4$. In fact, we have the following surprising formula in general.

Theorem 2.3 *For $n = 3(k-1) + r$, $k \in \mathbb{N}$, and $r \in \{0, 1, 2\}$ we have*

$$S_{\mathbf{F}}(n) = \begin{cases} 2^k \ \textit{if} \ \ r = 0, \\ 0 \ \textit{if} \ \ r = 1, \\ 2^{k+1} \ \textit{if} \ \ r = 2. \end{cases} \tag{9}$$

Proof. It is known that for every $n \in \mathbb{N}$, we have

$$F_0 + F_1 + ... + F_n = F_{n+2} - 1.$$

For a fixed $n \in \mathbb{N}$, it is clear that $S_{\mathbf{F}}(n+2)$ is twice the number of writings of F_{n+2} as $\pm F_{n+1} \pm \cdots \pm F_1 \pm F_0$.

Let us assume we have such a writing of F_{n+2}:

$$F_{n+2} = \epsilon_{n+1}F_{n+1} + \epsilon_n F_n + \cdots + \epsilon_0 F_0, \text{ with } \ \epsilon_i \in \{-1, 1\}, i = 0, 1, 2, \cdots, n+1. \tag{10}$$

We claim that that $\epsilon_{n+1} = 1$ and $\epsilon_n = 1$. By way of contradiction, if $\epsilon_{n+1} = -1$ then

$$F_{n+3} = F_{n+2} + F_{n+1} = \epsilon_n F_n + \cdots + \epsilon_0 F_0 \leq F_0 + ... + F_n = F_{n+2} - 1 < F_{n+3}.$$

This contradiction shows that $\epsilon_{n+1} = 1$. Then Equation (10) becomes

$$F_n = F_{n+2} - F_{n+1} = \epsilon_n F_n + \cdots + \epsilon_1 F_1 + \epsilon_0 F_0.$$

Again, if $\epsilon_n = -1$ then we obtain

$$2F_n = \epsilon_{n-1}F_{n-1} + \cdots + \epsilon_1 F_1 + \epsilon_0 F_0 \ \le F_{n+1} - 1 \Leftrightarrow F_n \le F_{n-1} - 1 \Rightarrow F_n < F_{n-1}.$$

So, this contradiction says that we must have $\epsilon_n = 1$ and therefore (10) is equivalent to

$$0 = \epsilon_{n-1}F_{n-1} \cdots + \epsilon_1 F_1 + \epsilon_0 F_0, \text{ with } \ \epsilon_i \in \{-1, 1\}, i = 0, 1, 2, \cdots, n-1.$$

This means that $S_{\mathbf{F}}(n+2) = 2S_{\mathbf{F}}(n-1)$. Taking into account that $S_{\mathbf{F}}(0) = 2$, $S_{\mathbf{F}}(1) = 0$, $S_{\mathbf{F}}(2) = 4$, using an induction argument we see that we get the formula in (9). □

It is known that for every $n \in \mathbb{N}$, $F_0^2 + F_1^2 + \cdots + F_n^2 = F_n F_{n+1}$ and so $s_n = \sqrt{F_n F_{n+1}}$. Hence the left-hand side in (6) can be simplified to

$$\frac{S_{\mathbf{F}}(n)}{2^n s_n^{-1}} = \begin{cases} \frac{8}{2^{2k}}\sqrt{F_{3k-3}F_{3k-2}} \text{ if } \ n = 3k-3, \\ 0 \text{ if } \ n = 3k-2, \\ \frac{4}{2^{2k}}\sqrt{F_{3k-1}F_{3k}} \text{ if } \ n = 3k-1. \end{cases}$$

If we denote golden ratio, as usual, by ϕ, i.e., $\phi = \frac{\sqrt{5}+1}{2}$, then since $F_n = \frac{\phi^n}{\sqrt{5}} + o(1)$, we have $\sqrt{F_{3k-3}F_{3k-2}} = \frac{\phi^{3k-5/2}}{\sqrt{5}} + o(1)$ and $\sqrt{F_{3k-1}F_{3k}} = \frac{\phi^{3k-1/2}}{\sqrt{5}} + o(1)$. Because $\phi^3/4 \approx 1.059$ we see that

$$\lim_{n\to\infty, n\equiv 0 \ or -1 (mod \ 3)} \frac{S_{\mathbf{F}}(n)}{2^n s_n^{-1}} = \infty,$$

which makes (6) very far from giving any asymptotic information about $S_{\mathbf{F}}(n)$.

References

[1] Andrica, D., Ionascu, E. J., *Some Unexpected Connections Between Analysis and Combinatorics*, in "Mathematics without boundaries. Topics in pure Mathematics", Th.M. Rassias and P. Pardalos, Eds., Springer (2014), 1-20

[2] Andrica, D., Tomescu, I., *On an integer sequence related to a product of trigonometric fuctions, and its combinatorial relevance*, J. Integer Seq. **5**(2002), Article 02.2.4.

[3] Berry, A. C., *The Accuracy of the Gaussian Approximation to the Sum of Independent Variates*, Trans. Amer. Math. Soc. **49**(1)(1941), 122-136

[4] Drimbe, M. O., *A problem of representation of integers* (Romanian), G.M.-B, **10-11**(1983), 382-383

[5] Drimbe, M. O., *Generalization of representation theorem of Erdős and Surányi*, Comment. Math. XXVII(1988), No.2, 233-235.

[6] Erdős, P., Surányi, J., *Topics in the Theory of Numbers*, Springer-Verlag, 2003.

[7] Finch, S. R., *Signum equations and extremal coefficients*, people.fas.harvard.edu/ sfinch/

[8] Mitek, J., *Generalization of a theorem of Erdős and Surányi*, Comment. Math. XXI(1979)

[9] T. Salát, and S. Znám, *On the sums of prime powers*, Acta Fac. Rer. Univ. Com. Math. 21 (1968), 2125

[10] Sullivan, B.D., *On a Conjecture of Andrica and Tomescu*, J. Integer Seq. 16(2013), Article 13.3.1.

[11] Zeckendorf, E., *Représentation des nombres naturels par une somme de nombres de Fibonacci ou des nombres de Lucas*, Bull. Soc. R. Sci. Liège **41**(1972), 179-182