ELSEVIER

Physica A 241 (1997) 493–499

PHYSICA A

$1/f$ noise in the distribution of prime numbers

Marek Wolf*

Institute of Theoretical Physics, University of Wrocław, Pl. Maxa Borna 9, PL-50-204 Wrocław, Poland

Received: 13 February 1997

Abstract

The Fourier transform of the "signal" given by the number of primes contained in the successive intervals of equal length $l = 2^{16} = 65\,536$ up to $N = 2^{38} \approx 2.749 \times 10^{11}$ was performed. It turns out that the power spectrum displays the $1/f^{\beta}$ behaviour with the exponent $\beta \approx 1.64$. This slope β does not depend on the length of the sampled intervals, which suggests some kind of self-similarity in the distribution of primes.

PACS: 02.10.L; 02.10.Jf; 42.50.L; 47.27.S

There are a lot of links between number theory and physics, see e.g. [1,2]. As an another example the Fibonnaci numbers can be mentioned: there is an ubiquity of places in the theory of chaos, where they appear (see [3]). On the other hand, it is rather surprising that very often prime numbers provide a toy model for some physical ideas. There are a few papers, where methods used by physicists were applied to the study of primes. For example, in [4] the multifractality of primes was investigated, while in [5] the appropriately defined Lyapunov exponents for the distribution of primes were calculated numerically. In this paper we are going to show that the distribution of prime numbers displays the $1/f$ noise.

It is commonly believed that prime numbers are distributed "randomly" although there exist formulae giving the approximate number of primes $\pi(N)$ less than a given bound N. One of such formula was guessed by 15 years old Gauss and states that

$$\pi(N) \sim \frac{N}{\ln N}\,. \tag{1}$$

Let us consider the interval of integers $1, 2, \ldots, N$ and divide it into $M(l)$ equal subintervals I_t of length l, so $t = 1, 2, \ldots, M(l) = N/l$. To denumerate the intervals the letter t was chosen because it will play the role of time. There is no unique, natural choice

* e-mail: mwolf@ift.uni.wroc.pl.

PII S0378-4371(97)00251-3

of the sampling interval for the case considered here and we are taking it to be simply one: $\Delta t = 1$. Next let us define the signal $X_l(t)$ to be a number of primes inside the tth interval I_t:

$$X_l(t) = \pi((t+1)l) - \pi(tl). \tag{2}$$

It is worth noting that, Gallagher [6] has proved, under assuming the n-tuple conjecture of Hardy–Littlewood, that the fraction of intervals which contain exactly a given number primes follows a Poisson distribution.

It should be stressed that in the definition (2) there is no averaging: in our case the signal $X_l(t)$ is in some sense a "stochastic" variable (see e.g. [7]), but we cannot calculate it over different samples (realizations) and average, as usually.

From (1) it follows that the number of primes contained in each interval decreases for large l like

$$X_l(t) \sim \frac{l}{\ln(tl)}, \tag{3}$$

but the real values display some fluctuations around this expected mean values. In Figs. 1 and 3 values of $X_l(t)$ imposed onto the theoretical curve (3) are shown for $t = 32, \ldots, 512$ with $l = 2^{16}$ and $l = 2^{24}$, respectively.

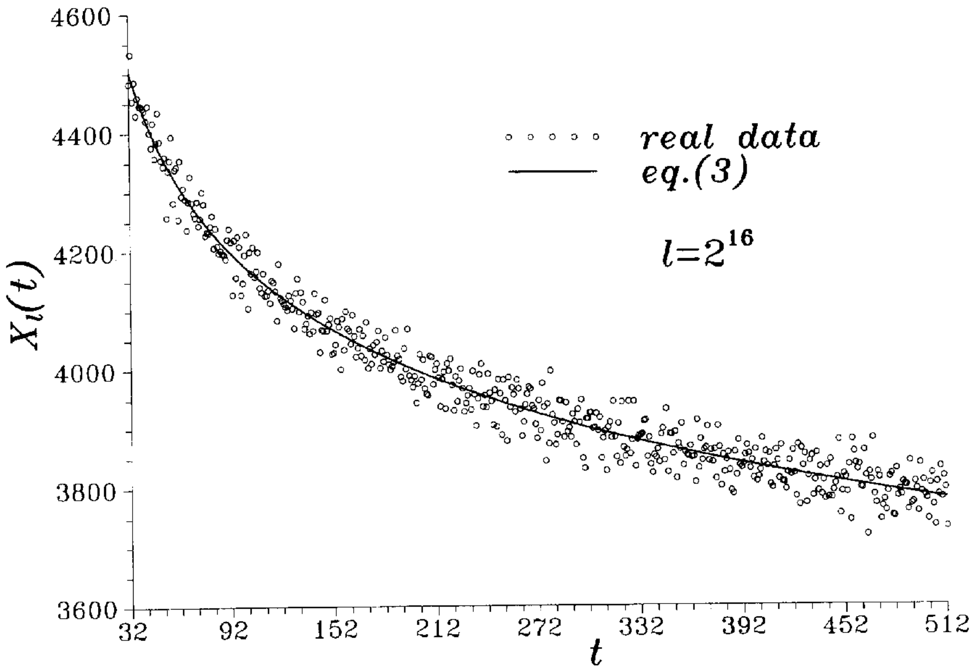

Fig. 1. The plot of the values of numbers of primes contained in the ith interval of the length $l = 2^{16} = 65\,536$ for $i = 32, \ldots, 512$. The solid curve represents Eq. (3). The largest absolute difference between a circle and theoretical curve was 88, and it corresponds to the relative error 2×10^{-2}.

Next we can consider the discrete Fourier transform [8] of the signal $X_l(t)$ defined above by (2):

$$\hat{X}_l(f) = \sum_{t=1}^{M(l)} X_l(t) e^{-2\pi i f t} . \tag{4}$$

Here the frequency f takes k values $f_k = k/M(l)\Delta t$, $k = 1 \ldots, \frac{1}{2}M(l)$ (for $\Delta t = 1$ the Nyquist frequency is just $f_c = \frac{1}{2}$) and the values $f_c < f_k \leqslant 2f_c$ $(k = \frac{1}{2}M(l) + 1, \ldots, M(l))$ are usually considered to correspond to negative frequencies (see [8]). But for real-valued signal $X(t)$ the identity $\hat{X}_l(f) = \hat{X}_l^{\star}(2f_c - f)$ holds and no information is lost by restricting ourselves only to the first half of frequencies. From (4) the power spectrum $S_{X_l}(f)$ can be obtained via the definition

$$S_{X_l}(f) = |\hat{X}_l(f)|^2 . \tag{5}$$

Note that in the literature different normalization conventions in the power spectrum definitions can be found. In the past it was found [9] that a lot natural phenomena lead to the power spectrum depending on frequency as $1/f^{\beta}$. This kind of dependence was named $1/f$ noise, in contrast to the white noise $S(f) = const.$

I have run the computer program to search for primes up to $N = 2^{38} \approx 2.75 \times 10^{11}$ and count the primes in the consecutive intervals of the length $l = 2^{16} = 65\,536$. In other words, there were $M(l) = 2^{22} = 4\,194\,304$ different values of $X_l(t)$. The Fast Fourier transform was used to calculate $\hat{X}_l(f)$ and next to obtain the power spectrum $S_{X_l}(f)$. The results are shown in the Fig. 2. Because of the huge number of points only the first 1000 consecutive values of f_k were plotted, while for $1000 < k < 100\,000$ only every 40th point was plotted and for $k > 100\,000$ only every 250th point was plotted. As it is seen from this figure at low frequencies, $S_X(f)$ displays almost a perfect power-like $1/f$ dependence on the frequencies

$$S_X(f) \sim \frac{1}{f^{\beta}} . \tag{6}$$

A few first points show downward bias and to determine the index β they were discarded. Also data corresponding to larger frequencies were skipped and the determination of the slope was performed via least-squares method only with points following the straight line. (Fitting a line to a set of points always requires a certain amount of human judgment – in fact only two points determine a line !) This procedure results in the estimation of $\beta \approx 1.64$. The scaling (6) seems to be fulfilled over at least 4 orders of the magnitude of f. For larger frequencies there is a systematic bias caused by "leakage" from frequencies larger than the Nyquist frequency and called "spurious aliasing"; see [8]. It manifests as a random scatter of points with tendency to bind upward. I have checked that looking for primes up to N smaller than 2^{38}, i.e. taking only a part of the available data, leads to the plots were the deviations from the power-like dependence start earlier, i.e. departure from straight lines begins for smaller frequencies – the power-like dependence of $S_{X_l}(f)$ spans less than 4 orders of magnitude.

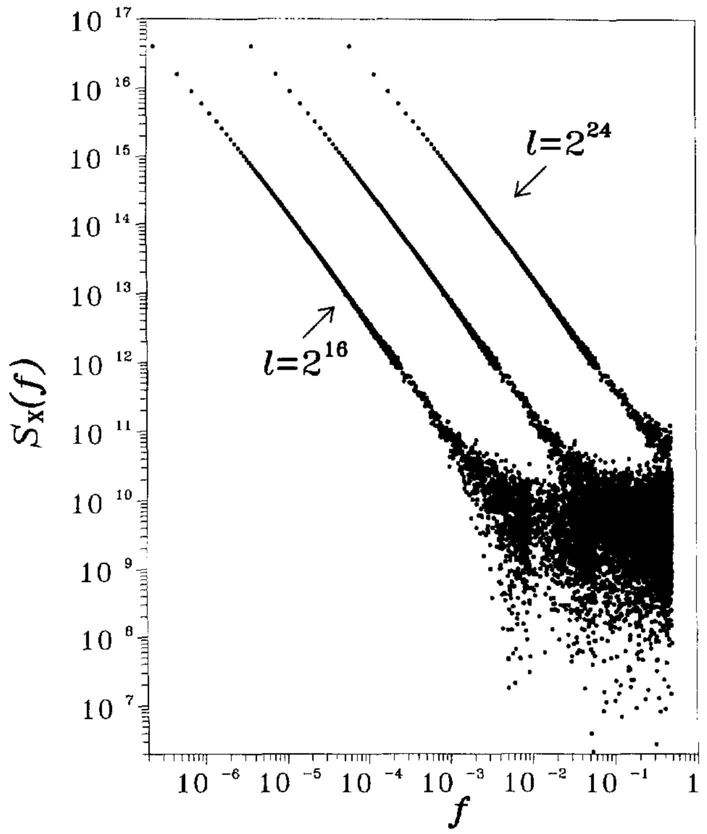

Fig. 2. The plot of the power spectrum $S_X(f)$ on the log–log scale for various lengths l of the interval. The plots for $l = 2^{20}$ and 2^{24} are also shown. There is only a part of data plotted – 2^{22} points is a huge amount of data so they were decimated, as explained in the text. For the largest length $l = 2^{24}$ there were only 8192 points and all of them are consecutively plotted.

The definition (2) depends on the arbitrarily chosen length $l = 2^{16}$. To look for the dependence of power spectrum on the length l of the sampled intervals the appropriate spectra for the lengths $l = 2^{20}$ and 2^{24} are also shown in Fig. 2. For these new lengths I have assumed that again $\Delta t = 1$. If I would choose $\Delta t \rightarrow 2 \times \Delta t$ when $l \rightarrow 2 \times l = l'$, then the initial parts of plots of $S_{X_l}(f)$ should collapse for each l in Fig. 2. As it seen from these plots the slope does not depend on the length of the sampled intervals: $\beta \approx 1.64$ seems to be inherent (generic) to the set of prime numbers. This independence of the β on l suggest some kind of the self-similarity in the distribution of primes, a phenomenon already observed in another context in [4]. It should be noted that the aliasing practically disappears for $l = 2^{24}$ – it can be caused by the fact that fluctuations around theoretical curve (3) are relatively much smaller for $l = 2^{24}$ than for $l = 2^{16}$; see Fig. 3.

From formula (3) it should be (in principle) possible to obtain the power spectrum in the analytical way. But I was not able to find in mathematical tables formulae allowing me to calculate explicitly the sum

$$\hat{X}_l(f) = \sum_{t=1}^{M(l)} l e^{-2\pi i f t} / \ln(tl) . \tag{7}$$

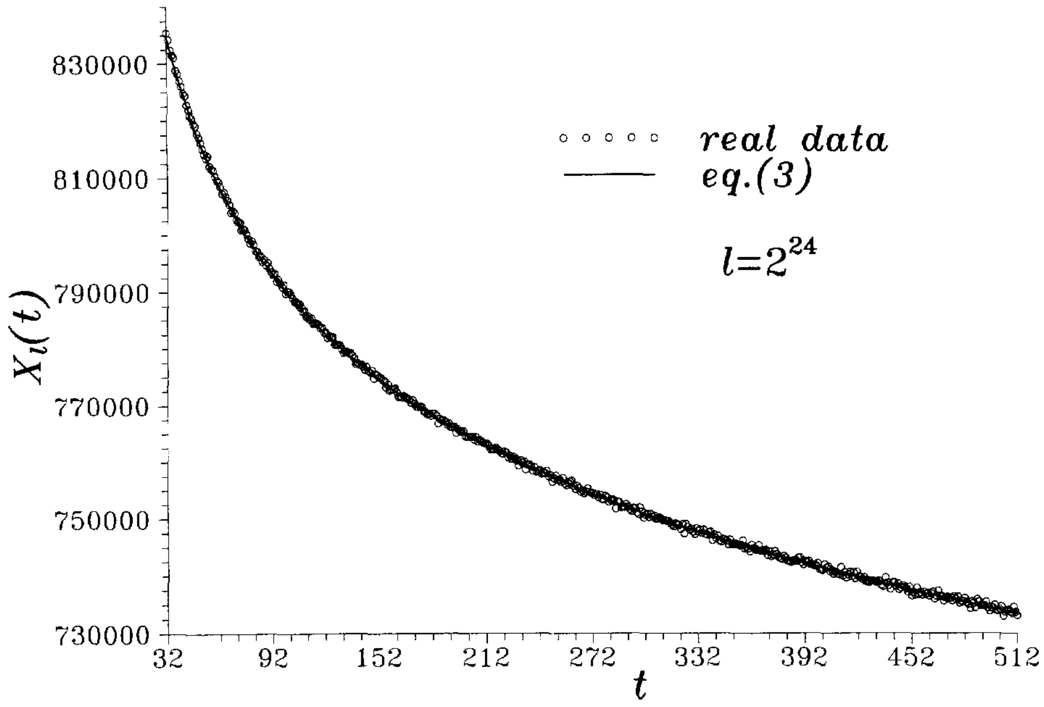

Fig. 3. The plot of the values of numbers of primes contained in the ith interval of the length $l=2^{24}=16\,777\,216$ for $i=32,\ldots,512$. The solid curve represents Eq. (3). The largest absolute difference between a circle and theoretical curve was 1317, but the relative error was 1.6×10^{-3} – over an order smaller than in Fig. 1.

There remains only a possibility to calculate the above sum by the computer. The Fourier transform for $l=2^{16}$ and $N=2^{38}$ (i.e. $M(l)=2^{22}$) is shown in Fig. 4. On the double logarithmic plot points form perfect straight line – the power spectrum behaves like $1/f^{\beta}$ with $\beta\approx1.64$. It is seen that it reproduces the power spectrum observed for the real computer data very well – at small frequencies the values of the power spectrum are practically the same as in Fig. 2. There is very small spurious aliasing visible at f close to the Nyquist frequency $f_c=0.5$. It confirms the observations that "leakage" from frequencies larger than f_c is caused by the rapid (of small periodicities) fluctuations around theoretical, smooth curve. It should be noted that the sum (7) is divergent in the limit $M(l)\rightarrow\infty(N\rightarrow\infty)$. There are two different ways to tackle this problem: First, the normalization in the power spectrum (5) can be fixed by dividing $|\hat{X}(f)|^2$ by $M(l)$, which presumably suffices to make the $S(f)$ convergent for $M(l)\rightarrow\infty$. Second, we can say that we are interested in the asymptotic behaviour of $S(f)$ for large $N\rightarrow\infty$; for example the number of all primes is infinite, but asymptoticaly it tends to infinity like (1).

In the past there were a lot of papers discussing self-organized criticality [10]. The abundance of natural phenomena where the power spectrum displays the $1/f$ noise was attempted to be explained by means of the self-organized critical models. They describe systems without finite characteristic length scale. Primes indeed seems to be

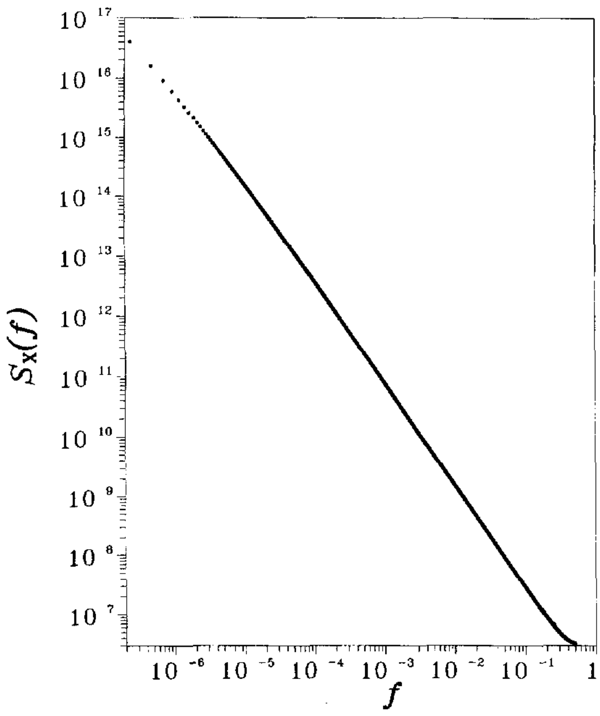

Fig. 4. Power spectrum obtained from the analytical formula (3) for $l = 2^{16}$ and $N = 2^{38}$.

distributed among all natural numbers with all possible gaps between them and there is no natural length scale. We end this short paper by asking the question: Are the prime numbers in a self-organized critical state?

Acknowledgement

The author receives support from the KBN research grant 2 P302 057 07.

Appendix. Details of the used algorithm

I have used a modification of the Eratostenes sieve to search for primes. To avoid the swapping of the Eratostenes sieve of the length $l = 2^{38}$ over all hard disks in our local network some tricks were used. First the primes smaller than $\sqrt{2^{38}} = 2^{19}$ were found and stored. The odd numbers between 1 and 2^{19} were coded as bits, and bits set to 1 represented primes whose values were calculated from the positions in the 2^{19} long array of bits. The small "sieve" of the length $l = 2^{16}$ was "moving" along the successive numbers $n - l, n - l + 1, \ldots, n - 1, n$ and the bits representing numbers divisible by primes smaller than $\sqrt{n}$ were set to zero. The sweep over the array of l bits was done and the number of primes counted giving a new value of the signal $X_l(t)$. Next the procedure was repeated with following set of $\{n, n + l\}$ of numbers.

This algorithm was implemented on the DEC Alpha 300X/175MHz workstation in DEC Fortran which allows some assembly level instructions and it took over 4 days of CPU time to reach the value of $2^{38} \approx 2.75 \times 10^{11}$.

References

[1] J.M. Luck, P. Moussa, M. Waldschmidt, Number Theory and Physics, Springer, Berlin, 1990.
[2] M. Waldschmidt, P. Moussa, J.-M. Luck, C. Itzykson, From Number Theory to Physics, Springer, Berlin, 1992.
[3] Schuster, Deterministic Chaos. An Introduction, Physik Verlag, Weinheim, 1989.
[4] M. Wolf, Physica 160A (1989) 24.
[5] Z. Gamba, J. Hernando, L. Romanelli, Phys. Lett. A 145 (1990) 106.
[6] P.X. Gallagher, Mathematika 23 (1976) 4–9.
[7] P.D.T.A. Elliot, Probabilistic Number Theory, vols. 1–2, Springer, Berlin, 1980.
[8] J.S. Bendat, A.L. Piersol, Random Data, Wiley, New York, 1986.
[9] F.N.H. Robinson, Noise and Fluctuations, Clarendon Press, Oxford, 1974.
[10] P. Bak, Ch. Tang, K. Wiesenfeld, Phys. Rev. Lett. 59 (1987) 381; Phys. Rev. A 38 (1988) 364.