ELSEVIER

Physica A 269 (1999) 503–510

PHYSICA A

www.elsevier.com/locate/physa

Ramanujan–Fourier series, the Wiener–Khintchine formula and the distribution of prime pairs

H. Gopalkrishna Gadiyar, R. Padma *

Ramanujan Institute for Advanced Study in Mathematics, University of Madras, Chennai 600 005, India

Received 3 February 1999

Abstract

The Wiener–Khintchine formula plays a central role in statistical mechanics. It is shown here that the problem of prime pairs is related to autocorrelation and hence to a Wiener–Khintchine formula. "Experimental" evidence is given for this. © 1999 Elsevier Science B.V. All rights reserved.

PACS: 05.40+j; 02.30.Nw; 02.10.Lh

Keywords: Twin primes; Ramanujan–Fourier series; Wiener–Khintchine formula

1. Introduction

"The Wiener–Khintchine theorem states a relationship between two important characteristics of a random process: the power spectrum of the process and the correlation function of the process" [1]. One of the outstanding problems in number theory is the problem of prime pairs which asks how primes of the form p and $p+h$ (where h is an even integer) are distributed. One immediately notes that this is a problem of finding correlation between primes. We make two key observations. First of all there is an arithmetical function (a function defined on integers) which traps the properties of the primes. This is the von Mangoldt function $\Lambda(n)$ which is defined to be equal to $\log p$ if $n = p^k$, where p is prime and k any positive integer and is equal to 0 otherwise. This is a good approximate "weighted" characteristic function of primes. Recall that a characteristic function $f(x)$ is equal to 1 if $x \in S$ and 0 if $x \notin S$ for some set S. It is standard to use $\Lambda(n)$ in number theory in stead of characteristic function of primes.

* Corresponding author.
E-mail address: padma@imsc.ernet.in (R. Padma)

0378-4371/99/$ - see front matter
PII: S0378-4371(99)00171-5

The second observation is that such arithmetical functions have a Ramanujan–Fourier series. This is due to Ramanujan. The papers of Ramanujan [2] and Hardy [3] are still the best place to read this topic. In this paper we show that the problem of distribution of prime pairs is equivalent to proving a Wiener–Khintchine formula. Further, numerical evidence is provided to justify the plausibility of this approach.

2. Ramanujan–Fourier series

The Ramanujan–Fourier series of an arithmetical function $a(n)$ is an expansion of the form

$$a(n) = \sum_{q=1}^{\infty} a_q c_q(n)\,, \tag{1}$$

where

$$c_q(n) = \sum_{\substack{k=1\\(k,q)=1}}^{q} \mathrm{e}^{2\pi i(k/q)n}$$

and (k,q) denotes the greatest common divisor of k and q. Using extremely simple arguments Ramanujan showed that the commonly known arithmetical functions all have Ramanujan–Fourier series. For example, he showed that

$$d(n) = -\sum_{q=1}^{\infty} \frac{\log q}{q} c_q(n)\,,$$

$$\sigma(n) = \frac{\pi^2 n}{6} \sum_{q=1}^{\infty} \frac{c_q(n)}{q^2}\,,$$

where $d(n)$ is the number of divisors of n and $\sigma(n)$ their sum. He however did not indicate any formula for getting the Ramanujan–Fourier coefficients a_q which are the back bone of Fourier analysis. This was done by Carmichael [4] a little later. He showed that $\mathrm{e}^{2\pi i(k/q)n}$ are almost periodic functions defined on the integers. This led to the orthogonality relations and a method for evaluating the Ramanujan–Fourier coefficient. Denote by $M(g)$ the mean value of an arithmetical function g, that is,

$$M(g) = \lim_{N\to\infty} \frac{1}{N} \sum_{n\leqslant N} g(n)\,.$$

For $1 \leqslant k \leqslant q, (k,q)=1$, let $e_{k/q}(n) = \mathrm{e}^{2\pi i(k/q)n}$, $(n \in \mathcal{N})$. If $a(n)$ is an arithmetical function with expansion (1) then

$$a_q = \frac{1}{\phi(q)} M(ac_q) = \frac{1}{\phi(q)} \lim_{N\to\infty} \frac{1}{N} \sum_{n\leqslant N} a(n)c_q(n)\,.$$

Also,

$$M(e_{k/q}\overline{e_{k'/q'}}) = \begin{cases} 1 & \text{if } k/q = k'/q'\,, \\ 0 & \text{if } k/q \neq k'/q'\,. \end{cases} \tag{2}$$

All this is well known in the field of almost periodic functions.

The arithmetical properties of $c_q(n)$ were elucidated by Hardy in [3] using which he obtained a Ramanujan–Fourier expansion for $(\phi(n)/n)\Lambda(n)$. (Here $\phi(n)$ denotes the number of integers less than n and relatively prime to n and is known as the Euler totient function.) We state these results below.

(i) $c_q(n)$ is multiplicative. That is,

$$c_{qq'}(n) = c_q(n)c_{q'}(n) \quad \text{if } (q,q') = 1 . \tag{3}$$

(ii) If p is a prime, then

$$c_p(n) = \begin{cases} -1 & \text{if } p \nmid n , \\ p-1 & \text{if } p|n , \end{cases} \tag{4}$$

where $a\,|\,b$ means a divides b and $a \nmid b$ means a does not divide b.

(iii)

$$\frac{\phi(n)}{n}\Lambda(n) = \sum_{q=1}^{\infty} \frac{\mu(q)}{\phi(q)} c_q(n) . \tag{5}$$

Here $\mu(q)$ is the Möbius function defined as follows:

$$\mu(q) = \begin{cases} (-1)^k & \text{if } q = p_1 p_2 \cdots p_k,\ p_i^\prime\text{s are distinct primes} , \\ 0 & \text{otherwise} . \end{cases} \tag{6}$$

Also,

(iv) For a given integer h,

$$\begin{aligned} C(h) &\overset{\text{def}}{=} \sum_{q=1}^{\infty} \frac{\mu^2(q)}{\phi^2(q)} c_q(h) \\ &= \begin{cases} 2 \prod_{p>2} \left(1 - \frac{1}{(p-1)^2}\right) \prod_{\substack{p|h \\ p>2}} \left(\frac{p-1}{p-2}\right) & \text{if } h \text{ is even} , \\ 0 & \text{if } h \text{ is odd} , \end{cases} \end{aligned} \tag{7}$$

where $\prod_p$ denotes the product over primes.

The proof of (7) follows from the multiplicative properties of $\mu(q)$, $\phi(q)$ and c_q. For a given h, the series on the left-hand side of (7) is absolutely convergent and hence has the Euler product expansion

$$\prod_p \left(1 + \frac{\mu^2(p)}{\phi^2(p)} c_p(h)\right) = \prod_{p \nmid h} \left(1 - \frac{1}{(p-1)^2}\right) \prod_{p|h} \left(1 + \frac{1}{p-1}\right) , \tag{8}$$

by (4) and (6). When h is odd, the infinite product on the right-hand side of (8) is 0 and when h is even, it is equal to

$$2 \prod_{p>2} \left(1 - \frac{1}{(p-1)^2}\right) \prod_{\substack{p|h \\ p>2}} \left(1 + \frac{1}{p-1}\right) \left(1 - \frac{1}{(p-1)^2}\right)^{-1} ,$$

which on simplification gives the right-hand side of (7).

3. The twin-prime conjecture and the Wiener–Khintchine formula

The Wiener–Khintchine formula basically says that if

$$f(t)=\sum_n f_n e^{i\lambda_n t},$$

then

$$\lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T} f(t+\tau)\overline{f(t)}\,dt=\sum_n |f_n|^2 e^{i\lambda_n\tau}. \tag{9}$$

The left-hand side of (9) is called an autocorrelation function. The right-hand side is nothing but the power spectrum. It is used practically to extract hidden periodicities in seemingly random phenomena.

For an arithmetical function $a(n)$ having the Ramanujan–Fourier series (1), the Wiener–Khintchine formula can be stated as follows:

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n\leqslant N} a(n)a(n+h)=\sum_{q=1}^{\infty} a_q^2 c_q(h). \tag{10}$$

The proof of (10) will follow from (1) and (2) if the Ramanujan–Fourier expansion of $a(n)$ is absolutely and uniformly convergent. But the Ramanujan–Fourier series for $(\phi(n)/n)\Lambda(n)$ does not belong to this class of functions. However, assuming the truth of the theorem in the most general case would give the formula conjectured by Hardy and Littlewood [5] which we now state.

There are infinitely many prime pairs $p, p+h$ for every even integer h and if $\pi_h(N)$ denotes the number of prime pairs less than N, then

$$\pi_h(N)\sim C(h)\frac{N}{\log^2 N}. \tag{11}$$

The Wiener–Khintchine formula for $(\phi(n)/n)\Lambda(n)$ is given by

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n\leqslant N}\frac{\phi(n)}{n}\Lambda(n)\frac{\phi(n+h)}{n+h}\Lambda(n+h)=\sum_{q=1}^{\infty}\frac{\mu^2(q)}{\phi^2(q)}c_q(h)=C(h). \tag{12}$$

Let

$$\Psi(h,N)=\sum_{n\leqslant N}\frac{\phi(n)}{n}\Lambda(n)\frac{\phi(n+h)}{n+h}\Lambda(n+h).$$

Then (12) implies that

$$\frac{\Psi(h,N)}{N}\sim C(h). \tag{13}$$

The terms of the sum on the left-hand side of (13) are non zero if and only if both $\Lambda(n)$ and $\Lambda(n+h)$ are prime powers, say p^k and q^l, respectively, for primes p and q. Then $\phi(n)/n=(p-1/p)$ and $\phi(n+h)/(n+h)=(q-1)/q$. Hence,

$$\sum_{n\leqslant N}\Lambda(n)\Lambda(n+h)-\sum_{n\leqslant N}\frac{\phi(n)}{n}\Lambda(n)\frac{\phi(n+h)}{n+h}\Lambda(n+h)$$

$$= \sum_{\substack{p^k \leqslant N \\ p^k+h=q^l}} \frac{\log p \log q}{p} + \sum_{\substack{p^k \leqslant N \\ p^k+h=q^l}} \frac{\log p \log q}{q} - \sum_{\substack{p^k \leqslant N \\ p^k+h=q^l}} \frac{\log p \log q}{pq}$$

$$= O(\log^2 N \log\log N)\,, \tag{14}$$

where we have used the formula [6]

$$\sum_{p \leqslant x} \frac{\log p}{p} \sim \log x$$

at the last step. Thus if (12) is true and h is even, then from (13) and (14) we get

$$\sum_{n \leqslant N} \Lambda(n)\Lambda(n+h) \sim C(h)N\,. \tag{15}$$

When we pass from this formula to a formula for $\pi_h(N)$ using partial summation, (see for example, Section 3 of [7]), the formula which arises naturally is not (11) but

$$\pi_h(N) \sim C(h) \int_2^N \frac{\mathrm{d}x}{\log^2 x} \tag{16}$$

and this is naturally equivalent to (11).

When h is odd, we know from (7) that $C(h) = 0$. If a term on the left-hand side of (13) is non-zero, then either n or $n+h$ is a power of 2 and hence

$$\sum_{n \leqslant N} \Lambda(n)\Lambda(n+h) = \log 2 \sum_{\substack{2^k \leqslant N \\ |p^l - 2^k| = h}} \log p$$

$$= O(\log^2 N) \tag{17}$$

from which one can deduce that $\pi_h(N) = O(1)$ when h is odd.

It is well-known that one proves the prime number theorem

$$\pi(x) = \sum_{p \leqslant x} 1 \sim \frac{x}{\log x}$$

by proving an equivalent assertion

$$\psi(x) = \sum_{n \leqslant x} \Lambda(n) \sim x\,, \tag{18}$$

for the Chebyshev function $\psi(x)$. Note that passing from (15) to (11) is analogous to this.

4. Numerical evidence

For completeness we give the compelling numerical evidence for the truth of the Wiener–Khintchine formula. For $h = 2, 4$ and 6, the formula (13) gives

$$\Psi(2,N) \sim C(2)N\,,$$

$$\Psi(4,N) \sim C(4)N\,,$$

$$\Psi(6,N) \sim C(6)N\,.$$

Table 1

N	$\Psi(2,N)$	$\frac{\Psi(2,N)}{N}$	Ratio
100000	131522.552204	1.315226	1.003876
200000	264287.347531	1.321437	0.999158
300000	393317.025988	1.311057	1.007068
400000	525523.270611	1.313808	1.004959
500000	654557.716460	1.309115	1.008562
600000	789035.163302	1.315059	1.004004
700000	919941.157912	1.314202	1.004658
800000	1049182.174335	1.311478	1.006745
900000	1180813.946552	1.312015	1.006332
1000000	1312843.985016	1.312844	1.005697

Table 2

N	$\Psi(4,N)$	$\frac{\Psi(4,N)}{N}$	Ratio
100000	130212.335085	1.302123	1.013977
200000	260492.247225	1.302461	1.013714
300000	390320.617781	1.301069	1.014799
400000	527155.226011	1.317888	1.001848
500000	653649.051733	1.307298	1.009964
600000	789177.513123	1.315296	1.003823
700000	923982.224287	1.319975	1.000264
800000	1054670.388142	1.318338	1.001506
900000	1180133.117590	1.311259	1.006913
1000000	1307978.775955	1.307979	1.009438

Note that $C(4)=C(2)$ and $C(6)=2C(2)$. It follows that there should be approximately equal numbers of prime power pairs differing by 2 and by 4, but about twice as many differing by 6. We have tabulated below the actual values of $\Psi(h,N)$ and also the ratio $C(h)/(\Psi(h,N)/N)$ (third column) for $h=2,4$ and 6 and N upto 10^6 (see Tables 1–3). We have used the value of $C(2)=2\prod_{p>2}(1-1/(p-1)^2)\sim 1.320323632$ to compute the ratio.

This shows that there is remarkable agreement between theory and numerical experiment. This is impressive evidence for the truth of the Wiener–Khintchine formula.

5. Conclusion

As Kac [8] remarks "..... Consider the integers divisible by both p and q (q another prime). To be divisible by p and q is equivalent to being divisible by pq and consequently the density of the new set is $1/pq$. Now,

$$\frac{1}{pq}=\frac{1}{p}\frac{1}{q},$$

Table 3

N	$\Psi(6, N)$	$\frac{\Psi(6,N)}{N}$	Ratio
100000	261289.742091	2.612897	1.010620
200000	523391.109218	2.616956	1.009053
300000	787393.641752	2.624645	1.006097
400000	1056087.319082	2.640218	1.000162
500000	1316336.875799	2.632674	1.003029
600000	1579274.310330	2.632124	1.003238
700000	1839327.388416	2.627611	1.004961
800000	2104826.034045	2.631033	1.003654
900000	2368450.398104	2.631612	1.003434
1000000	2631198.406265	2.631198	1.003591

and we can interpret this by saying that the "events" of being divisible by p and q are independent. This holds, of course, for any number of primes, and we can say using a picturesque but not a very precise language, that the primes play a game of chance! This simple, nearly trivial, observation is the beginning of a new development which links in a significant way number theory on the one hand and probability theory on the other". One can see the interesting article [9] for elaborations of this theme. For recent work in this direction see [10–12].

The Wiener–Khintchine formula seems to show that the primes have structure hidden in their seemingly arbitrary behaviour. As Hardy remarks in [3], "... These series have a peculiar interest because they show explicitly the source of the irregularities in the behaviour of their sums. Thus, for example, ...

$$\sigma(n) = \frac{\pi^2 n}{6}\left(1 + \frac{(-1)^n}{2^2} + \frac{2\cos(2/3)n\pi}{3^2} + \frac{2\cos(1/2)n\pi}{4^2} \right.$$
$$\left. + \frac{2(\cos(2/5)n\pi + \cos(4/5)n\pi)}{5^2} + \frac{2\cos(1/3)n\pi}{6^2} + \cdots\right)$$

and we see at once that the most important term in $\sigma(n)$ is $\frac{1}{6}\pi^2 n$, and that irregular variations about this average value are produced by a series of harmonic oscillations of decreasing amplitude". Hence the Ramanujan–Fourier series trap the vagaries in the behaviour of primes and the Wiener–Khintchine formula traps their correlation properties. It is a pleasant surprise that the Wiener–Khintchine formula which normally occurs in practical problems of brownian motion, electrical engineering and other applied areas of technology and statistical physics has a role in the behaviour of prime numbers which are studied by pure mathematicians.

Acknowledgements

The second author wishes to thank the CSIR for financial support.

References

[1] C. Kittel, Elementary Statistical Physics, Wiley, New York, 1958.

[2] S. Ramanujan, On certain trigonometrical sums and their applications in the theory of numbers, Trans. Camb. Phil. Soc. 22 (1918) 259.

[3] G.H. Hardy, Note on Ramanujan's trigonometrical function $c_q(n)$ and certain series of arithmetical functions, Proc. Camb. Phil. Soc. 20 (1921) 263.

[4] R.D. Carmichael, Expansions of arithmetical functions in infinite series, Proc. London Math. Soc. 34 (2) (1932) 1.

[5] G.H. Hardy, J.E. Littlewood, Some problems of Partition Numerorum; III: On the expression of a number as a sum of primes, Acta Math. 44 (1922) 1.

[6] T.M. Apostol, Introduction to Analytic Number Theory, International Springer Student Edition, Springer, Berlin, 1989.

[7] H.G. Diamond, Elementary methods in the study of the distribution of prime numbers, Bull. (New Series) Amer. Math. Soc. 7 (1982) 553.

[8] M. Kac, Statistical Independence in Probability, Analysis and Number Theory, The Carus Mathematical Monograph No. 12, The Mathematical Association of America, 1964.

[9] P. Billingsley, Prime numbers and Brownian motion, Amer. Math. Monthly 80 (1973) 1099.

[10] M. Wolf, $1/f$ noise in the distribution of prime numbers, Physica A 241 (1997) 493.

[11] M. Wolf, Random walk on the prime numbers, Physica A 250 (1998) 335.

[12] N.M. Katz, P. Sarnak, Zeros of zeta functions and symmetry, Bull. (New Series) Amer. Math. Soc. 36 (1999) 1.