Negative moments of the gaps between consecutive primes

Marek Wolf

Cardinal Stefan Wyszynski University, Faculty of Mathematics and Natural Sciences. College of Sciences,
ul. Wóycickiego 1/3, PL-01-938 Warsaw, Poland, e-mail: m.wolf@uksw.edu.pl

Abstract

We derive heuristically approximate formulas for the negative k–moments $M_{-k}(x)$ of the gaps between consecutive primes$< x$ represented directly by $\pi(x)$ — the number of primes up to x. In particular we propose an analytical formula for the sum of reciprocals of gaps between consecutive primes $< x$: $M_{-1}(x) \sim \frac{\pi^2(x)}{x-2\pi(x)} \log\left(\frac{x}{2\pi(x)}\right) \sim x \log\log(x)/\log^2(x)$. We illustrate obtained results by the enormous computer data up to $x = 4 \times 10^{18}$.

Let p_n denotes the n-th prime number and $d_n = p_{n+1} - p_n$ denotes the n-th gap between consecutive primes. Let us consider the sum of reciprocals of gaps d_n over primes up to $p_n \leq x$:

$$\sum_{n=2}^{\pi(x)} \frac{1}{p_n - p_{n-1}}, \tag{1}$$

where $\pi(x)$ is, as usual, the number of primes up to x. To our knowledge there is no known formula for the above sum as a function of x. We can consider in general arbitrary negative moments of $p_{n+1} - p_n$:

$$M_{-k}(x) \equiv \sum_{n=2}^{\pi(x)} \frac{1}{(p_n - p_{n-1})^k}. \tag{2}$$

For the positive moments

$$M_k(x) \equiv \sum_{n=2}^{\pi(x)} (p_n - p_{n-1})^k \tag{3}$$

in [5, p.2056] it was conjectured that:

$$M_k(x) \sim k! x \log^{k-1}(x). \tag{4}$$

The symbol $f(x) \sim g(x)$ means here that $\lim_{x\to\infty} f(x)/g(x) = 1$. In [11] we predicted the formula

$$M_k(x) = \frac{\Gamma(k+1)x^k}{\pi^{k-1}(x)} + \mathcal{O}_k(x). \tag{5}$$

Above $\Gamma(x) = \int_0^\infty t^{x-1}e^{-t}dt$ is the Gamma function generalizing the factorial thus the above formula is valid also for non-integer k. By the Prime Number Theorem (PNT) the number of prime numbers below x is very well approximated by the logarithmic integral

$$\pi(x) \sim \text{Li}(x) \equiv \int_2^x \frac{du}{\log(u)}.$$

Integration by parts gives the asymptotic expansion which should be cut at the term $n_0 = \lfloor \log(x) \rfloor$:

$$\text{Li}(x) = \frac{x}{\log(x)} + \frac{x}{\log^2(x)} + \frac{2!x}{\log^3(x)} + \frac{3!x}{\log^4(x)} + \cdots + \frac{n_0!x}{\log^{n_0+1}(x)}. \tag{6}$$

Putting in (5) the approximatation $\pi(x) \sim x/\log(x)$ we recover (4).

Let $\tau_d(x)$ denote the number of pairs of consecutive primes smaller than a given bound x and separated by d:

$$\tau_d(x) = \sharp\{p_n, p_{n+1} < x, \quad \text{with } p_{n+1} - p_n = d\}. \tag{7}$$

In [9] (see also [8], [10]) we proposed the following formula expressing function $\tau_d(x)$ directly by $\pi(x)$:

$$\tau_d(x) \sim C_2 \prod_{p|d, p>2} \frac{p-1}{p-2} \; \frac{\pi^2(x)}{x} \left(1 - \frac{2\pi(x)}{x}\right)^{\frac{d}{2}-1} \quad \text{for } d \geq 6, \tag{8}$$

where the twins prime constant

$$C_2 \equiv 2\prod_{p>2}\left(1 - \frac{1}{(p-1)^2}\right) = 1.320323631693739\ldots$$

The pairs of primes separated by $d = 2$ ("twins") and $d = 4$ ("cousins") are special as they always have to be consecutive primes (with the exception of the pair (3,7) containing 5 in the middle)). For $d = 4$ we adapt the expression obtained from (8) for $d = 2$, which for $\pi(x) \sim x/\log(x)$ goes into the the conjecture B of G. H. Hardy and J.E. Littlewood [2, eqs. (5.311) and (5.312)]:

$$\tau_2(x)\big(\approx \tau_4(x)\big) \sim C_2\frac{\pi^2(x)}{x} \approx C_2\frac{x}{\log^2(x)}. \tag{9}$$

We have

$$M_{-k}(x) = \sum_{n=2}^{\pi(x)} \frac{1}{(p_n - p_{n-1})^k} = \sum_{d=2,4,6,\ldots} \frac{\tau_d(x)}{d^k}. \tag{10}$$

We will assume that for sufficiently regular and decreasing functions $f(n)$ the following formula holds:

$$\sum_{k=1}^{\infty} \prod_{p|k, p>2} \frac{p-1}{p-2}\; f(k) = \frac{1}{\prod_{p>2}(1-\frac{1}{(p-1)^2})} \sum_{k=1}^{\infty} f(k). \tag{11}$$

In other words we will replace the product over $p|d$ in (8) by its mean value as E. Bombieri and H. Davenport [1] have proved that the number $1/\prod_{p>2}(1-\frac{1}{(p-1)^2}) = 2/C_2$ is the arithmetical average of the product $\prod_{p|k} \frac{p-1}{p-2}$:

$$\frac{1}{n}\sum_{k=1}^{n} \prod_{p|k, p>2} \frac{p-1}{p-2} = \frac{1}{\prod_{p>2}(1-\frac{1}{(p-1)^2})} + \mathcal{O}(\log^2(n)). \tag{12}$$

Later H.L. Montgomery [4, eq.(17.11)] has improved the error term to $\mathcal{O}(\log(n))$. Using this trick we get further from (8) and (10)

$$M_{-k}(x) \sim 2\frac{\pi^2(x)}{x-2\pi(x)} \sum_{n=1}^{\infty} \frac{1}{(2n)^k}\Big(1-\frac{2\pi(x)}{x}\Big)^n. \tag{13}$$

To calculate negative moments we need the formula for the series

$$\sum_{n=1}^{\infty} \frac{q^n}{n^k} \equiv \mathrm{Li}_k(q), \qquad |q|<1, \tag{14}$$

where $\mathrm{Li}_k(q)$ is a polylogarithm function of order k, see, for example, [3] or [6, Sect. 25.12]. The $\mathrm{Li}_k(q)$ should not be confused with logarithmic integral in (6), where $\mathrm{Li}(x)$ appears without any subscript. Unfortunately the closed formula for polylogarithm is known only for $k=1$ and is obtained by integrating term by term uniformly convergent geometrical series: $\mathrm{Li}_1(q) = -\log(1-q)$. Hence we obtain

$$M_{-1}(x) = \sum_{n=2}^{\pi(x)} \frac{1}{p_n - p_{n-1}} \sim \widetilde{M}_{-1}^{(2)}(x) \equiv \frac{\pi^2(x)}{x-2\pi(x)} \log\Big(\frac{x}{2\pi(x)}\Big). \tag{15}$$

We use the notation $\widetilde{M}_{-k}^{(2)}(x)$ for the analytical formula for $M_{-k}(x)$ expressed by $\pi(x)$, while $\widetilde{M}_{-k}^{(1)}(x)$ will refer to the formula for $M_{-k}(x)$ expressed by series in $1/\log(x)$, see below. Putting here for $\pi(x)$ a few first terms from the expansion of $\mathrm{Li}(x)$ (6) and expanding in series of $1/\log(x)$ we obtain

$$M_{-1}(x) \sim \widetilde{M}_{-1}^{(1)}(x) \equiv x\left(\frac{\log(\log(x)) - \log(2)}{\log^2(x)} + \frac{8\log(\log(x)) - 1 - 8\log(2)}{\log^3(x)}\right). \tag{16}$$

For large x using the notation $\log_n(x) = \log(\log_{n-1}(x))$ for the iterated logarithm we obtain the pleasant formula:

$$\sum_{n=2}^{\pi(x)} \frac{1}{p_n - p_{n-1}} = 1 + \frac{1}{2} + \frac{1}{2} + \frac{1}{4} + \cdots + \frac{1}{p_{\pi(x)} - p_{\pi(x)-1}} \sim \frac{x\log_2(x)}{\log^2(x)}. \tag{17}$$

During over a seven months long run of the computer program we have collected the values of $\tau_d(x)$ up to $x = 2^{48} \approx 2.8147 \times 10^{14}$. The data representing the function $\tau_d(x)$ were stored at values of x forming the geometrical progression with the ratio 2, i.e. at $x = 2^{15}, 2^{16}, \ldots, 2^{47}, 2^{48}$. Such a choice of the intermediate thresholds as powers of 2 was determined by the employed computer program in which the primes were coded as bits. The data is available for downloading from `http://pracownicy.uksw.edu.pl/mwolf/gaps.zip`. At the Tomás Oliveira e Silva web site `http://sweet.ua.pt/tos/gaps.html` we have found values of $\tau_d(x)$ for $x = 1.61 \times 10^{18}$ and $x = 4 \times 10^{18}$. In Table 1 we give a comparison of formulas (15) and (15) with exact values of $M_{-1}(x)$. We used the values of $\pi(x)$ calculated from the identity $\sum_d \tau_d(x) = \pi(x) - 1$.

TABLE 1

The sum of reciprocals of gaps between consecutive primes< x compared with closed formulas (15) and (16). The ratios initially decrease and next slowly tend towards 1. The convergence in second column is very slow: the ratios are changing only on third places after dot for x spanning over eleven orders.

x	$M_{-1}(x)/\widetilde{M}_{-1}^{(2)}(x)$	$M_{-1}(x)/\widetilde{M}_{-1}^{(1)}(x)$
$2^{24} = 1.6777 \times 10^7$	0.8738	0.7638
$2^{26} = 6.7109 \times 10^7$	0.8731	0.7664
$2^{28} = 2.6844 \times 10^8$	0.8734	0.7699
$2^{30} = 1.0737 \times 10^9$	0.8738	0.7734
$2^{32} = 4.2950 \times 10^9$	0.8741	0.7769
$2^{34} = 1.7180 \times 10^{10}$	0.8744	0.7803
$2^{36} = 6.8719 \times 10^{10}$	0.8748	0.7836
$2^{38} = 2.7488 \times 10^{11}$	0.8751	0.7867
$2^{40} = 1.0995 \times 10^{12}$	0.8755	0.7898
$2^{42} = 4.3980 \times 10^{12}$	0.8759	0.7927
$2^{44} = 1.7592 \times 10^{13}$	0.8762	0.7955
$2^{46} = 7.0369 \times 10^{13}$	0.8766	0.7982
$2^{48} = 2.8147 \times 10^{14}$	0.8770	0.8007
1.61×10^{18}	0.8793	0.8145
4×10^{18}	0.8795	0.8157

For the second negative moment we will use the twice integrated geometrical series

$$\sum_{n=1}^{\infty} \frac{q^n}{n(n+1)} = \frac{q + (1-q)\log(1-q)}{q}, \qquad |q| < 1. \tag{18}$$

as for large n we have $1/(n(n+1)) \approx 1/n^2$. In this way we obtain a crude approximation

$$M_{-2}(x)=\sum_{n=2}^{\pi(x)}\frac{1}{(p_n-p_{n-1})^2}\sim\widetilde{M}_{-2}^{(2)}(x)\equiv\frac{1}{2}\frac{\pi^2(x)}{x-2\pi(x)}\left(1+\frac{2\pi(x)}{x-2\pi(x)}\log\left(\frac{2\pi(x)}{x}\right)\right) \tag{19}$$

and for $\pi(x)\sim\mathrm{Li}(x)$ from (6) we get

$$M_{-2}(x)\sim\widetilde{M}_{-2}^{(1)}(x)\equiv\frac{1}{2}\frac{x}{\log^2(x)}\left(1-\frac{2\log(\log(x))}{\log(x)}+\mathcal{O}\Big(\frac{1}{\log^2(x)}\Big)\right). \tag{20}$$

In Table 2 we give a comparison of these formulas with exact values of $M_{-2}(x)$. The ratios with increasing x tend with decreasing speed to 1.

TABLE 2

The sum of reciprocals of squared gaps between consecutive primes< x compared with closed formulas (19) and (20).

x	$M_{-2}(x)/\widetilde{M}_{-2}^{(2)}(x)$	$M_{-2}(x)/\widetilde{M}_{-2}^{(1)}(x))$
$2^{24}=1.6777\times10^{7}$	1.3318	1.8391
$2^{26}=6.7109\times10^{7}$	1.3224	1.7811
$2^{28}=2.6844\times10^{8}$	1.3167	1.7357
$2^{30}=1.0737\times10^{9}$	1.3122	1.6979
$2^{32}=4.2950\times10^{9}$	1.3081	1.6653
$2^{34}=1.7180\times10^{10}$	1.3044	1.6369
$2^{36}=6.8719\times10^{10}$	1.3009	1.6119
$2^{38}=2.7488\times10^{11}$	1.2978	1.5900
$2^{40}=1.0995\times10^{12}$	1.2951	1.5704
$2^{42}=4.3980\times10^{12}$	1.2926	1.5530
$2^{44}=1.7592\times10^{13}$	1.2903	1.5373
$2^{46}=7.0369\times10^{13}$	1.2882	1.5231
$2^{48}=2.8147\times10^{14}$	1.2862	1.5102
1.61×10^{18}	1.2767	1.4501
4×10^{18}	1.2760	1.4453

Because the closed formula for $\mathrm{Li}_k(q)$ with $k\geq 2$ is unknown we can not obtain the conjecture for negative k similar to (5). The approximate formula given in [7] is not convenient for our purposes. However for large k the negative moments are dominated by smallest gaps $d=2,4,6,\ldots$. Hence in the defining formula for M_{-k} we keep only a few first terms:

$$M_{-k}(x)=C_2\frac{\pi^2(x)}{2^k x}\left(1+\frac{1}{2^k}+\frac{2}{3^k}\left(1-\frac{2\pi(x)}{x}\right)^2+\right.$$
$$\left.+\frac{1}{4^k}\left(1-\frac{2\pi(x)}{x}\right)^3+\frac{4}{3}\frac{1}{5^k}\left(1-\frac{2\pi(x)}{x}\right)^4+\frac{2}{6^k}\left(1-\frac{2\pi(x)}{x}\right)^5+\ldots\right) \tag{21}$$

Above we have used explicit values of the product $\prod_{p|k,p>2} \frac{p-1}{p-2}$ for $d = 2, 4, 6, 8, 10$ and 12. For example, for $k = 4$ keeping gaps up to $d = 10$ and developing powers of $1 - 2\pi(x)/x$ we obtain:

$$M_{-4}(x) \sim \widetilde{M}_{-4}^{(2)} = C_2 \frac{\pi^2(x)}{16x} \Bigg(\frac{34081595473}{31116960000} - \frac{2500235267}{15558480000} \frac{\pi(x)}{x} + \frac{2244748963}{7779240000} \Big(\frac{\pi(x)}{x}\Big)^2 - \frac{1178322017}{3889620000} \Big(\frac{\pi(x)}{x}\Big)^3 + \frac{33735178}{121550625} \Big(\frac{\pi(x)}{x}\Big)^4 \Bigg) \tag{22}$$

Putting above instead of $\pi(x)$ a few terms from the expansion (6) for $\mathrm{Li}(x)$ the series in powers of $1/\log(x)$ follows:

$$M_{-4}(x) \sim \widetilde{M}_{-4}^{(1)} = C_2 \frac{x}{16 \log^2(x)} \Big(\frac{14168273}{12960000} + \frac{14168273}{6480000} \frac{1}{\log(x)} + \frac{4680091}{864000} \Big(\frac{1}{\log(x)}\Big)^2 + \frac{27005921}{6480000} \Big(\frac{1}{\log(x)}\Big)^3 \Big) \tag{23}$$

In Table 3 we show how good the above approximation is. We do not know why ratios in the third columns are closer to one than in second.

TABLE 3

The comparison of the formulas for the sum of reciprocals of fourth powers of gaps between consecutive primes.

x	$M_{-4}(x)/\widetilde{M}_{-4}^{(2)}(x)$	$M_{-4}(x)/\widetilde{M}_{-4}^{(1)}(x))$
$2^{24} = 1.6777 \times 10^7$	1.012003	1.008288
$2^{26} = 6.7109 \times 10^7$	1.007536	1.004022
$2^{28} = 2.6844 \times 10^8$	1.006260	1.002853
$2^{30} = 1.0737 \times 10^9$	1.006038	1.002751
$2^{32} = 4.2950 \times 10^9$	1.005621	1.002445
$2^{34} = 1.7180 \times 10^{10}$	1.005218	1.002190
$2^{36} = 6.8719 \times 10^{10}$	1.004711	1.001826
$2^{38} = 2.7488 \times 10^{11}$	1.004403	1.001666
$2^{40} = 1.0995 \times 10^{12}$	1.004104	1.001513
$2^{42} = 4.3980 \times 10^{12}$	1.003863	1.001417
$2^{44} = 1.7592 \times 10^{13}$	1.003657	1.001349
$2^{46} = 7.0369 \times 10^{13}$	1.003471	1.001297
$2^{48} = 2.8147 \times 10^{14}$	1.003311	1.001265
1.61×10^{18}	1.002630	1.001258
4×10^{18}	1.002580	1.001268

We can obtain another approximate formula. Namely in (21) it is possible to sum over all d and separate terms without $\pi(x)/x$ and with first and second power of $\pi(x)/x$ using the equation (11). In this manner we obtain:

$$M_{-k} \sim \frac{\pi^2(x)}{2^{k-1}(x-2\pi(x))}\Bigg(\zeta(k) - \frac{2\pi(x)}{x}\Big(\zeta(k-1) - \frac{1}{2^k}\Big) + \frac{2\pi^2(x)}{x^2}\Big(\zeta(k-2) - \zeta(k-1) - \frac{1}{2^{k-1}}\Big) + \dots\Bigg). \tag{24}$$

Here $\zeta(k) = \sum_n 1/n^k$ is the Riemann zeta function at integer arguments. Above we have to demand $k \geq 4$ to avoid infinity of $\zeta(1)$.

Acknowledgement: I thank Professor Roger Heath–Brown for e-mail exchange.

References

[1] E. Bombieri and H. Davenport. Small differences between prime numbers. *Proc. Royal Soc.*, A293:1–18, 1966.

[2] G. H. Hardy and J. E. Littlewood. Some problems of 'Partitio Numerorum' III: On the expression of a number as a sum of primes. *Acta Mathematica*, 44:1–70, 1922.

[3] L. Lewin. *Polylogarithms and associated functions.* North Holland, 1981.

[4] H. Montgomery. *Topics in Multiplicative Number Theory.* Springer-Verlag, Heidelberg, New York, Inc, 1971. Springer Lecture Notes 227.

[5] T. Oliveira e Silva, S. Herzog, and S. Pardi. Empirical verification of the even Goldbach conjecture and computation of prime gaps up to $4\cdot 10^{18}$. *Mathematics of Computation*, 83, 2014.

[6] F. W. J. Olver, D. W. Lozier, R. F. Boisvert, and C. W. Clark. *NIST handbook of mathematical functions.* NIST and CUP, 2010.

[7] L. Veptas. An efficient algorithm for accelerating the convergence of oscillatory series, useful for computing the polylogarithm and hurwitz zeta functions. *Numerical Algorithms*, 47(3):211252, Mar 2008.

[8] M. Wolf. Applications of statistical mechanics in number theory. *Physica* **A**, 274(1–2):149 – 157, 1999.

[9] M. Wolf. Some heuristics on the gaps between consecutive primes. *math.NT/1102.0481*, Feb. 2011.

[10] M. Wolf. Nearest-neighbor-spacing distribution of prime numbers and quantum chaos. *Phys. Rev. E*, 89:022922–1–022922–11, Feb 2014.

[11] M. Wolf. On the moments of the gaps between consecutive primes. *arXiv:math.NT/1705.10766*, May 2017.