August 1996, IFTUWr 909/96

On the Twin and Cousin Primes

Marek Wolf

Institute of Theoretical Physics, University of Wrocław
Pl.Maxa Borna 9, PL-50-204 Wrocław, Poland
e-mail: mwolf@ift.uni.wroc.pl

Abstract

The computer results of the investigation of the number of pairs of primes separated by gap $d = 2$ ("twins") and gap $d = 4$ ("cousins") are reported. The number of twins and cousins turn out to be is almost the same. The plot of the function $W(x)$ giving the difference of the number of twins and cousins for $x \in (1, 10^{12})$ is presented . This function has fractal properties and the fractal dimension is approximately 1.48 — what is very close to the fractal dimension of the usual Brownian motion. The set of primes, up to which the numbers of twins and cousins are *exactly the same* seems to have the fractal structure. The box-counting method gives the fractal dimension of this set approximately 0.51. The statistics of distances between primes being the zeros of $W(x)$ display the cross-over from the exponential decrease to the power like dependence with the exponent equal 1.48. Arguments that $W(x)$ has the same properties as a *typical* sample path of the random walk are given. The analog of the Brun's constant is defined for cousins.

1. In the paper [1] Hardy and Littlewood have proposed about 15 conjectures. The conjecture B of their paper states[1]:

There are infinitely many primes pairs

$$p, p' = p + d, \tag{1}$$

for every even d. If $\pi_d(x)$ is the number of pairs less than x, then

$$\pi_d(x) \sim c_2 \frac{x}{\ln^2(x)} \prod_{p|d} \frac{p-1}{p-2}. \tag{2}$$

Here the constant c_2 (sometimes called "twin–prime " constant, see [3]) is defined in the following way:

$$c_2 \equiv 2 \prod_{p>2} \left(1 - \frac{1}{(p-1)^2}\right) = 1.32032\ldots \tag{3}$$

Nobody has proved as yet (2), even there is no proof that there is infinity of twin ($d = 2$) primes. The largest twins known officially are:

$$697053813 \times 2^{16352} \pm 1 \tag{4}$$

found recently by Indlekofer and Jarai, [2]. The pairs of primes separated by $d = 2$ and $d = 4$ are special as they always have to be consecutive primes (with exception of the pair (3,7) containing in the middle 5) and for $d \geq 6$ the function $\pi_d(x)$ counts *all* pairs $p, p' = p + d$, not necessarily successive. On the page 43 of [1] Hardy and Littlewood wrote: *Thus there should be approximately equal numbers prime-pairs differing by 2 and 4, but twice as many differing by 6.* In this paper I am going to present the results of the computer study of the behavior of $\pi_d(x)$ for $d = 2$ and $d = 4$ and, in particular, of the set of such x for which the equality $\pi_2(x) = \pi_4(x)$ holds. By analogy with twins I will call pairs $p, p + 4$ cousins.

2. Hardy and Littlewood have said nothing about the error terms in (2). Under the assumption of the Riemann Hypothesis, the Prime Number Theorem (PNT) was proved in the form:

$$\pi(x) = Li(x) + \mathcal{O}(\sqrt{x}\ln(x)), \tag{5}$$

where the logarithmic integral is defined by;

$$Li(x) = \int_2^x \frac{du}{\ln(u)}, \tag{6}$$

[1] I have changed notations to modern ones.

and one can expect similar error terms in (2). But it turns out that the computer search up to $N = 2^{44}$ shows that the relation

$$\pi_2(x) \sim \pi_4(x) \left(\approx \frac{c_2 x}{\ln^2(x)} \right) \tag{7}$$

holds with much higher accuracy than can be expected from the analogy with PNT. The relation (7) is well satisfied already for small values of x — it holds not necessarily assymptotically for large x. The table I shows the values of the numbers of twins and cousins captured during the computer scan at the values of N forming the geometrical progression $N = 2^{18}, 2^{20}, \ldots, 2^{44}$. The differences between the numbers of twins and cousins are much smaller than $\sqrt{N}$.

TABLE 1

The numbers of twins ($d = 2$) and cousins ($d = 4$).

N	$\pi_2(N)$	$\pi_4(N)$	$\pi_2(N)/\pi_4(N)$
2^{18}	2679	2678	1.00037
2^{20}	8535	8500	1.00412
2^{22}	27995	27764	1.00832
2^{24}	92246	91995	1.00273
2^{26}	309561	309293	1.00087
2^{28}	1056281	1057146	0.99918
2^{30}	3650557	3650515	1.00001
2^{32}	12739574	12740283	0.99994
2^{34}	44849427	44842399	1.00016
2^{36}	159082253	159089620	0.99995
2^{38}	568237005	568225073	1.00002
2^{40}	2042054332	2042077653	0.99999
2^{42}	7378928530	7378989766	0.99999
2^{44}	26795709320	26795628686	1.00000

3. To measure the discrepancy between the number of twins and cousins I will use the function

$$W(x) = \pi_2(x) - \pi_4(x). \tag{8}$$

The character W was introduced here because the plot of (8) resembles the fractal Weierstrass-Mandelbrot $W(x)$ functions [4], [5] or the Brownian motion, that is also often denoted by $W(x)$. The function $W(x)$ is piecewise constant, because $\pi_d(N)$ can change values only for N being the prime. The Fig.1 shows the plot of $W(x)$ in the range $x \in (1, 10^{12})$. The argument is shown on the logarithmical x-axis. The comparison of three parts (a), (b) and (c) of Fig.1 reveals a self–affinity of $W(x)$: left parts starts with oscillations of relatively small amplitudes and they increase at right parts (figures (a), (b) and (c) have increasing scales on the y-axis). Let me remark that, if instead of using the logarithmic scale, the x-axis would be drawn linearly with the same yardstick as the vertical axis on the original of Fig. 1(c), where the interval (0, 25000) had the length 11 cm, then the total plot for $1 \leq x \leq 10^{12}$ should be 4400 km long! It took over 9 days of CPU time on the DEC 3000/800 200MHz workstation to produce the data for Fig.1.

4. As it is seen from the Table 1 the ratio $\pi_2(N)/\pi_4(N)$ is sometimes larger than 1, and for other N is smaller than 1. It means that there is a set of such x that the numbers of Twins and Cousins are equal, so it is reasonably to look for zeroset of the function $W(x)$. Because $\pi_d(x)$ changes value only at x being prime, I have looked for such primes $p^{(z)}$ at which the numbers of twins or cousins are the same:

$$\mathcal{T}(x) = \{p^{(z)} < x : W(p^{(z)}) = 0\}. \tag{9}$$

and let $\pi_z(x)$ = number of $p^{(z)} < x$ such that $W(p^{(z)}) = 0$. The direct computer search shows that up to $N = 2^{43} \approx 8.8 \times 10^{12}$ there are 2823290 such primes $p^{(z)}$ that $W(p^{(z)}) = 0$ holds. First the same number of twins and cousins appears between 101 and 103 (besides the trivial zeros 2 and 3, when $\pi_2(x) = 0$ and $\pi_4(x) = 0$). The largest captured zero below 2^{43} was 8205034088567 $\approx 2^{42.899646}$. On the Fig.1(a) there are 2334 zeros, on (b) $W(x)$ touches x-axis 13019 times and in (c) there were 1035496 prime zeros of $W(x)$. In the Table 2 the numbers of primes $p^{(z)}$ up to 10^{13} every one order of magnitude are given. The values of $\pi_z(x)$ in this table display rather large fluctuations however using the analogy with random walk (see next paragraphs) one can expect that the $\pi_z(x)$ is of the same order as the number of visits of random walk to the origin. As it is well known, see e.g. [6], the average number of returns of the random walk to the origin during n steps is $\sqrt{n/\pi}$, I guess that:

$$\pi_z(x) \sim \sqrt{x/\pi} \tag{10}$$

The comparison of this formula with the actual data is provided on the figure Fig.2. Of course it is not obvious that there is infinity of $p^{(z)}$.

TABLE 2

The values of the function $\pi_z(x)$ giving the number of primes $p^{(z)} < x$ fulfilling $W(p^{(z)}) = 0$.

x	$\pi_z(x)$
1000	31
10000	60
100000	592
1000000	2332
10000000	2332
100000000	4718
1000000000	15351
10000000000	68440
100000000000	278503
1000000000000	1787793
10000000000000	2823290

4. The interesting information about the structure of the set $\mathcal{T}(x)$ can be obtained from the distribution $\mathcal{D}_N(\Delta z)$ of spacings Δz between consecutive $p^{(z)} < N$

$$\begin{aligned}\mathcal{D}_N(\Delta z) = \text{number of consecutive } & p^{(z)}, p^{(z')} < N \\ \text{such that } & p^{(z')} - p^{(z)} = \Delta z \end{aligned} \tag{11}$$

The plot of $\mathcal{D}_N(\Delta z)$ is shown on the Fig.3. It should be stressed that on the x-axis up to $\Delta z = 360$ there is a linear scale, while for larger Δz the scale is logarithmic. There are 2790362 spacings shown on this figure — only 32927 distances between consecutive $p^{(z)}$ were larger than 10^4. In fact these larger distances have very scattered values (the largest gap between two clusters of the same numbers of twins and cousins was 314267840234) and they appeared only once — it resembles the intermittent behaviour in some dynamical systems. There is a cross-over from the exponential dependence of $\mathcal{D}_N(\Delta z)$ to the power-like decrease starting at around $\Delta z = 360$. In the power like-regime all Δz were multiplicities of 6: $\Delta z = 6n$ (see below), while on the left part there appeared arbitrary even gaps $\Delta z = 2, 4, 6, \ldots 358$ and local jumps are for $\Delta z = 6n$.

The two-type functional dependence of $\mathcal{D}_N(\Delta z)$ means that the set of $p^{(z)}$ is formed by "clusters" separated by distances obeying the power law. The first elements of these clusters are characterized by the equation

$$W(p^{(z)}) = 0 \wedge W(p^{(z)} - 1) \neq 0 \tag{12}$$

and the ends of clusters satisfy:

$$W(p^{(z)}) \neq 0 \wedge W(p^{(z)} - 1) = 0. \tag{13}$$

Inside "clusters" the values of $\pi_2(x)$ and $\pi_4(x)$ do not change their values and are equal each other. So for $p^{(z)}$ inside clusters the dependence of $\mathcal{D}_N(\Delta z)$ is the same as the dependence of the number of gaps between consecutive primes. Because the first part (up to $\Delta z < 360$) of the plot of $\mathcal{D}_N(\Delta z)$ in the Fig.3 has the x-axis linear and y-axis logarithmic, one expects exponential decrease. It is really the case, as the function

$$h_N(d) = \text{number of pairs } p_n, p_{n+1} < N \text{ with } d = p_{n+1} - p_n. \tag{14}$$

decreases exponentially with d:

$$h_N(d) \sim \frac{c_2 N}{\ln^2(N)} \prod_{p|d, p>2} \frac{p-1}{p-2} e^{-d/\ln(N)}, \tag{15}$$

see [7]. Exactly the product in (15) is responsible for jigsaw pattern overimposed onto the exponential decreasing in the quite left part of Fig.3. On the contrary, the number of spacings $\Delta z > 360$ (this threshold 360 is for $N = 2^{43}$) decrease in the power–like manner:

$$\mathcal{D}_N(\Delta z) \sim (\Delta z)^{-\gamma}. \tag{16}$$

It means that the "clusters" (or "islands") of such x that $\pi_2(x) = \pi_4(x)$ are organized in a hierarchical, selfsimilar set. The power–like dependences are characteristic for fractal sets [4]. The exponent γ in the power–like part $\mathcal{D}_N(\Delta z)$ has the value of $\gamma \approx 1.48$ and it appears that this "critical exponent" *does not* depend on N, see the plot for $N = 2^{37}$ on the Fig.3 (the plots of $\mathcal{D}_N(\Delta z)$ for other values of N display the same slope, but I did not plot them because they overlap). Since this index γ does not depend on N, it can be regarded as the fractal dimension of the curve represented by $W(x)$ [8].

However, in the region $\Delta z \in (z_0, 360)$, where z_0 is roughly 120 for $N = 2^{43}$, the two behaviors overlap and it causes big 'jumps' of the values of $\mathcal{D}_N(\Delta z)$ at $\Delta z = 6n < 360$; for distances larger than 360 only multiplicities of 6 appears (all twins are of the form $6n \pm 1$, where n are not necessary consecutive integers). But it is a little bit difficult to split the values of $\mathcal{D}_N(\Delta z)$ into the part arising from the distribution of gaps between consecutive primes and distances between "clusters" of primes at which $W(p^{(z)}) = 0$ holds.

There is an open question what is the functional form of the prefactor, depending on N, in (16). This prefactor is connected with the form of $\pi_z(x)$ by the following selfconsistency equation:

$$\sum_{\Delta z} \mathcal{D}_N(\Delta z) = \pi_z(N). \tag{17}$$

5. Another argument for the fractal structure of the set $\{p^{(z)}\}$ is supplied by calculation of the Haussdorf dimension [4]. I have used the direct box-counting method. Namely the whole interval $(1, 2^{43})$ was covered by consecutive intervals of the length $l = 16$ and the number $N(l)$ of "boxes" containing primes $p^{(z)}$ was calculated. This procedure was successively repeated for lengths two times larger, up to $l = 2^{39} \approx 5.5 \times 10^{11}$. The obtained values of $N(l)$ are plotted in the Fig.4 on the double logarithmic axes. The large linear part in the middle tells us that:

$$N(l) \sim l^{-D_{fr}}, \qquad D_{fr} \approx 0.509 \tag{18}$$

where D_{fr} is the fractal dimension of $\mathcal{T}(x)$, see e.g. [4]. I have calculated D_{fr} by fitting the straight line to points for $l = 2^{12}, \ldots 2^{33}$. There is a surplus of small boxes caused by the cluster–like organization of the set $\mathcal{T}(x)$: short boxes grasp separate primes at which $W(x)$ is not changing value (equal 0). In other words, there is a minimal length (depending on x), below which $\mathcal{T}(x)$ is not a fractal: the small boxes intersect with zeros of $W(x)$ which are inside "clusters". The fractal, selfsimilar hierarchy is formed only by primes $p^{(z)}$ marking the beginning or the end of clusters and those are distinguished by conditions (12) and (13), respectively. Because the largest cluster encountered during the computer search had the length 358, boxes with $l > 1024$ follow perfectly the power–like dependence: they contain the clusters totally inside. Let us notice, that an approximate relation $\gamma = D_{fr} + 1$ holds, see [4].

6. In 1919 Brun [9] has shown that the sum of the reciprocals of all twin primes is finite:

$$B_2 = \left(\frac{1}{3} + \frac{1}{5}\right) + \left(\frac{1}{5} + \frac{1}{7}\right) + \left(\frac{1}{11} + \frac{1}{13}\right) + \ldots < \infty. \tag{19}$$

The numerical estimations give [11] $B = 1.9021605\ldots$. The sum of the finite number of terms in (19)

$$B_2(x) = \sum_{p\ twin, p<x} \frac{1}{p} \tag{20}$$

gives the finite approximations to the constant B_2. Probabilistic arguments show that [10]:

$$B_2(x) = B_2 - \frac{2c_2}{\ln(x)}. \tag{21}$$

By analogy the appropriate constant for cousins can be defined:

$$B_4 = \left(\frac{1}{7} + \frac{1}{11}\right) + \left(\frac{1}{13} + \frac{1}{17}\right) + \left(\frac{1}{19} + \frac{1}{23}\right) + \ldots. \tag{22}$$

The reasoning of Brent applies *mutatis mutandis* to the cousins and:

$$B_4(x) = \sum_{p\ cousin, p<x} \frac{1}{p} = B_4 - \frac{2c_2}{\ln(x)} \tag{23}$$

I have calculated on the computer the finite approximations $B_2(N)$ and $B_4(N)$ up to $N = 2^{42}$. The intermediate values were stored at $N = 2^{22}, 2^{23}, 2^{24}, \ldots, 2^{42}$ and obtained values are plotted versus $1/\ln(N)$ on the Fig.5. On this plot the points really are lying on the straight lines. As predicted by (21) and (23), the two lines on the Fig.5 are practically parallel with slope 2.6399 for twins and 2.6401 for cousins ($2c_2 \approx 2.64064$). The value of the "cousin"– Brun constant estimated from (23), where $B_4(2^{42}) = 1.1063389965796880$, is $B_4 = 1.1970449\ldots$

7. One can ask a lot of questions on the structure of the set $\mathcal{T}(x)$. Is this set infinite? What is the measure of this set up to a given x? How grows with x the largest distance between two consecutive clusters of $p^{(z)}$. Are there arbitrarily large values of the function $W(x)$? More precisely, does there exist for every $M > 0$ such x_M that

$$\mid \pi_2(x_M) - \pi_4(x_M) \mid > M? \tag{24}$$

Maybe answers to the above questions can be obtained by exploiting the analogy between primes and Brownian motion [12]. Namely, the plot of $W(x)$, see Fig.1, resembles the sample path of random walk. The graph of usual random walk has the fractal dimension 1.5 and it is very close to the dimensions of $W(x)$ estimated via box – counting method and from the histogram of spacings Δz. Also the graph 2 confirms, that $W(x)$ behaves as the *typical* realization of the random walk. However, for usual random walk the steps are made at every instant of time, while $W(x)$ is piecewise constant. Maybe $W(x)$ is a projection of a higher dimensional random walk, which performs steps at every instant of time, onto the one of coordinates?

Acknowledgment: The author receives support from the KBN research grant No 2 P302 057 07

References

[1] 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** (1922), p.1-70

[2] K.H. Indlekofer and A.Jarai, *Largest known twin primes*, Math. Comp. **65** (1996), p. 427-428

[3] J.W.Wrench, "Evaluation of Artin's constant and the Twin Prime constant", *Math.Comp.***15** (1961), p.396-398

[4] B.B.Mandelbrot, *The Fractal Geometry of Nature*, (W.H.Freeman and Co., San Francisco, 1982)

[5] M.V.Berry, Z.V.Lewis, *On the Weierstrass–Mandelbrot fractal function* Proc.R. Soc. Lond. **A370** (1980), pp.459-484;

[6] W.Feller, *An Introduction to the Probability Theory and its Applications*, (John Wiley and sons, New York, 1970), vol.1

[7] M.Wolf *On the gaps between consecutive primes*, preprint IFT UWr IFTUWr 894/95, submitted to J.Number Theory; available via anonymous ftp from: orchid.ift.uni.wroc.pl (156.17.88.118), directory pub/primes

[8] M.T.Barlow and S.J. Taylor, *Defining Fractal Subsets of* Z^d, Proc. London Math. Soc. **64** (1992), pp. 125–152

[9] V.Brun, "La serie 1/5 + 1/7 +... est convergente ou finie", *Bull. Sci.Math.* **43** (1919), pp.124-128

[10] R.P.Brent, *Irregularities in the Distribution of Primes and Twin Primes*, Math. Comp. **29** (1975), pp.43-56

[11] T. R. Nicely, "Enumeration to 10^{14} of the Twin Primes and Brun's constant", *Virginia Journal Of Science* **46** (1995), pp.195-204

[12] P.Billingsley, "Prime Numbers and Brownian Motion", *Amer. Math. Monthly* **80** (1973), pp.1099-1115

Figure captions:

The postscript files of the parts (b) and (c) are 2.3 MB long and are available separately under the names fig1b.ps, fig1c.ps

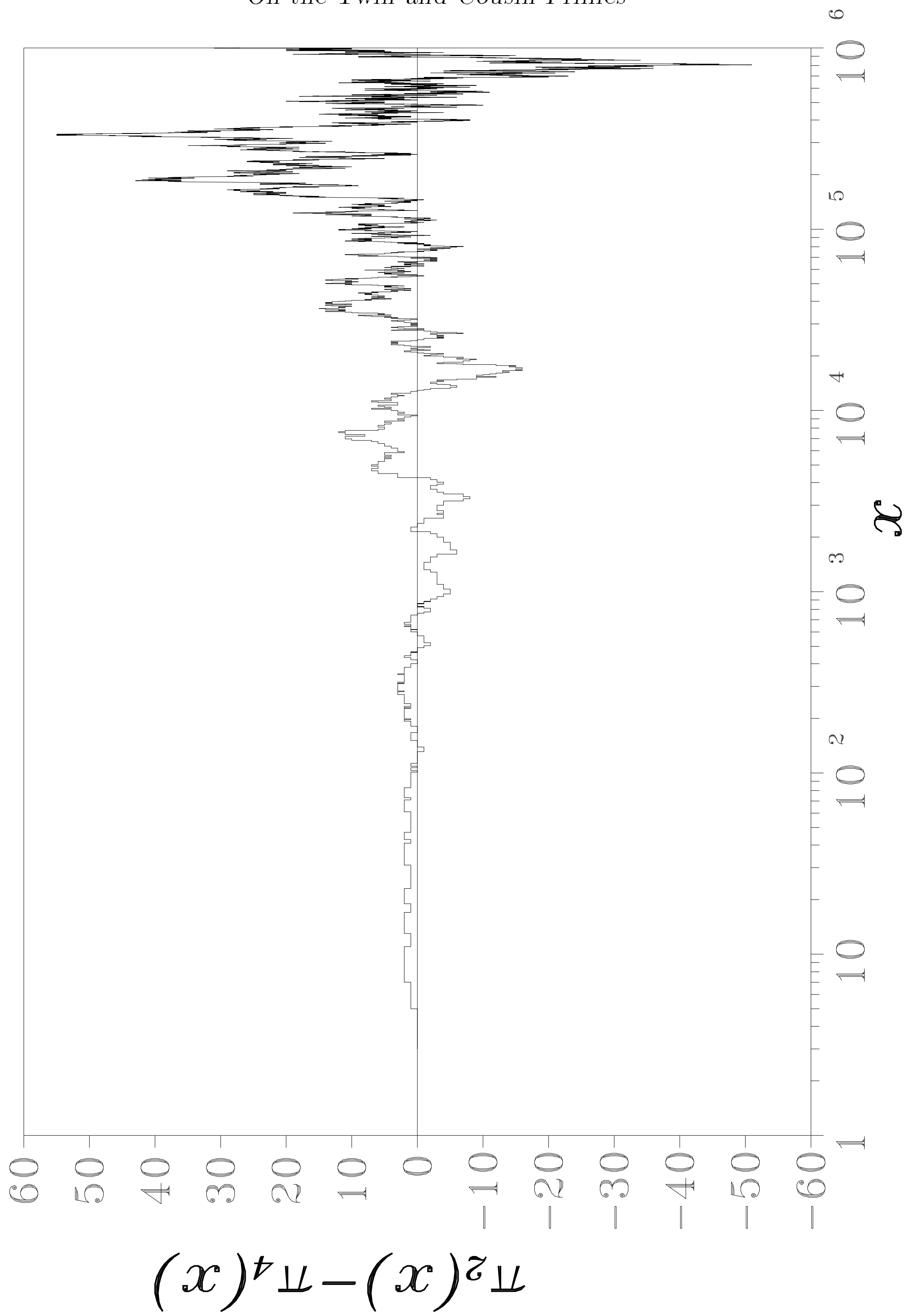

Fig.1 The function $W(x)$ is plotted in the range $x \in (1, 10^6)$ in the part (a), for $x \in (10^6, 10^9)$ in (b) and for $x \in (10^9, 10^{12})$ in (c). The range of values on the y-axis changes in each case. Up to $x = 5 \times 10^6$ all arguments x are plotted — for arguments larger some decimation procedure was employed. Namely for $W(x) > 100$ only changes of values larger than 8% were recorded, while smaller values of the function $W(x)$ were updated only after changes larger than 30%. There are 16314 points in (a) plotted.

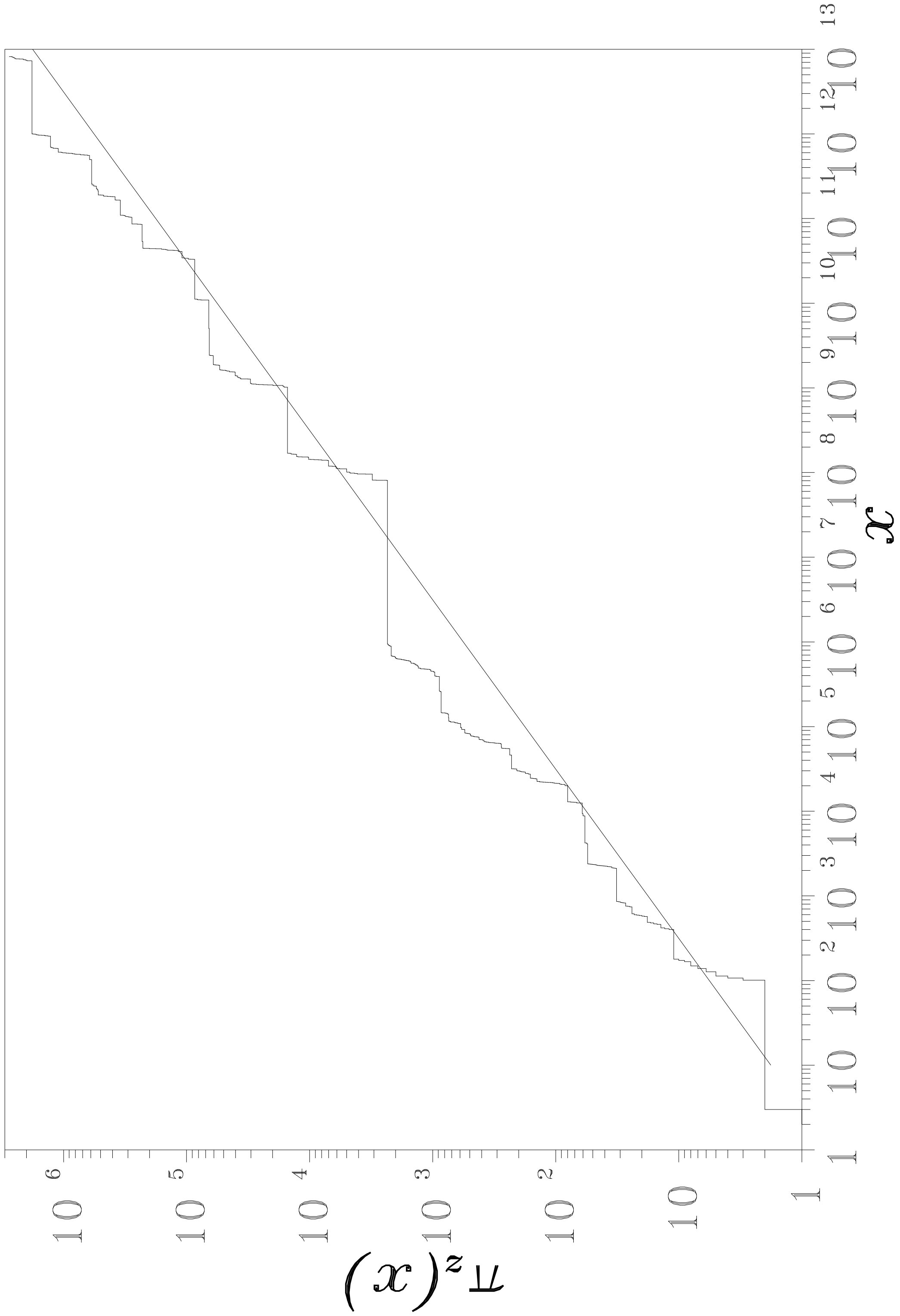

Fig.2 The plot showing the actual dependence of $\pi_z(x)$ on x. The straight line represents the plot of $\sqrt{x/\pi}$ — the conjectured dependence of the $\pi_z(x)$.

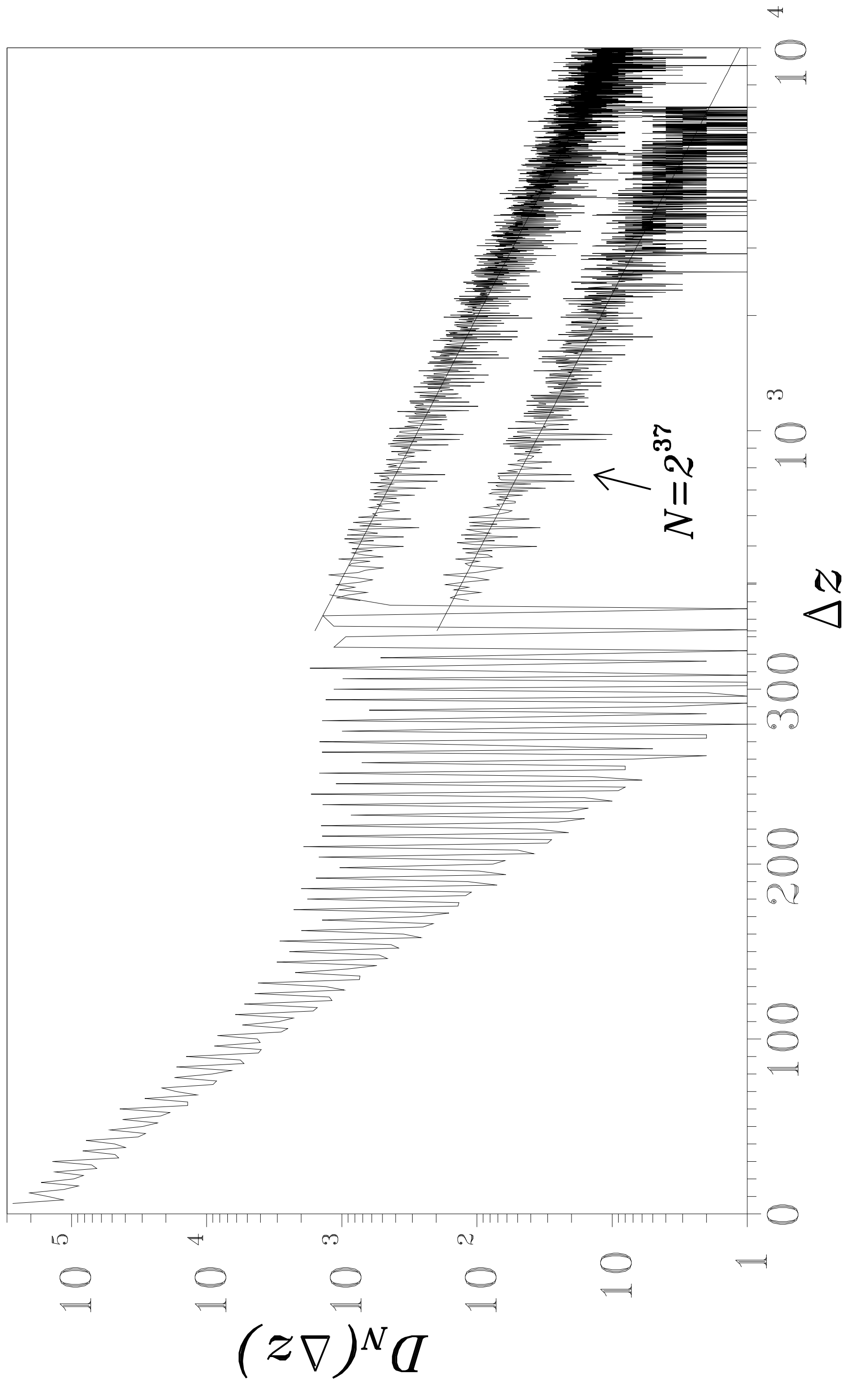

Fig.3 The plot showing the dependence of the distribution $\mathcal{D}_N(\Delta z)$ of spacings Δz between consecutive $p^{(z)} < 2^{43}$. There is a logarithmic scale on the y-axis , while on the x-axis there is a linear scale up to $\Delta z = 360$ while for larger Δz the scale is logarithmic. There is also the power-like part of $\mathcal{D}_N(\Delta z)$ for $N = 2^{37}$ shown and for this case the slope is $\gamma \approx 1.47$.

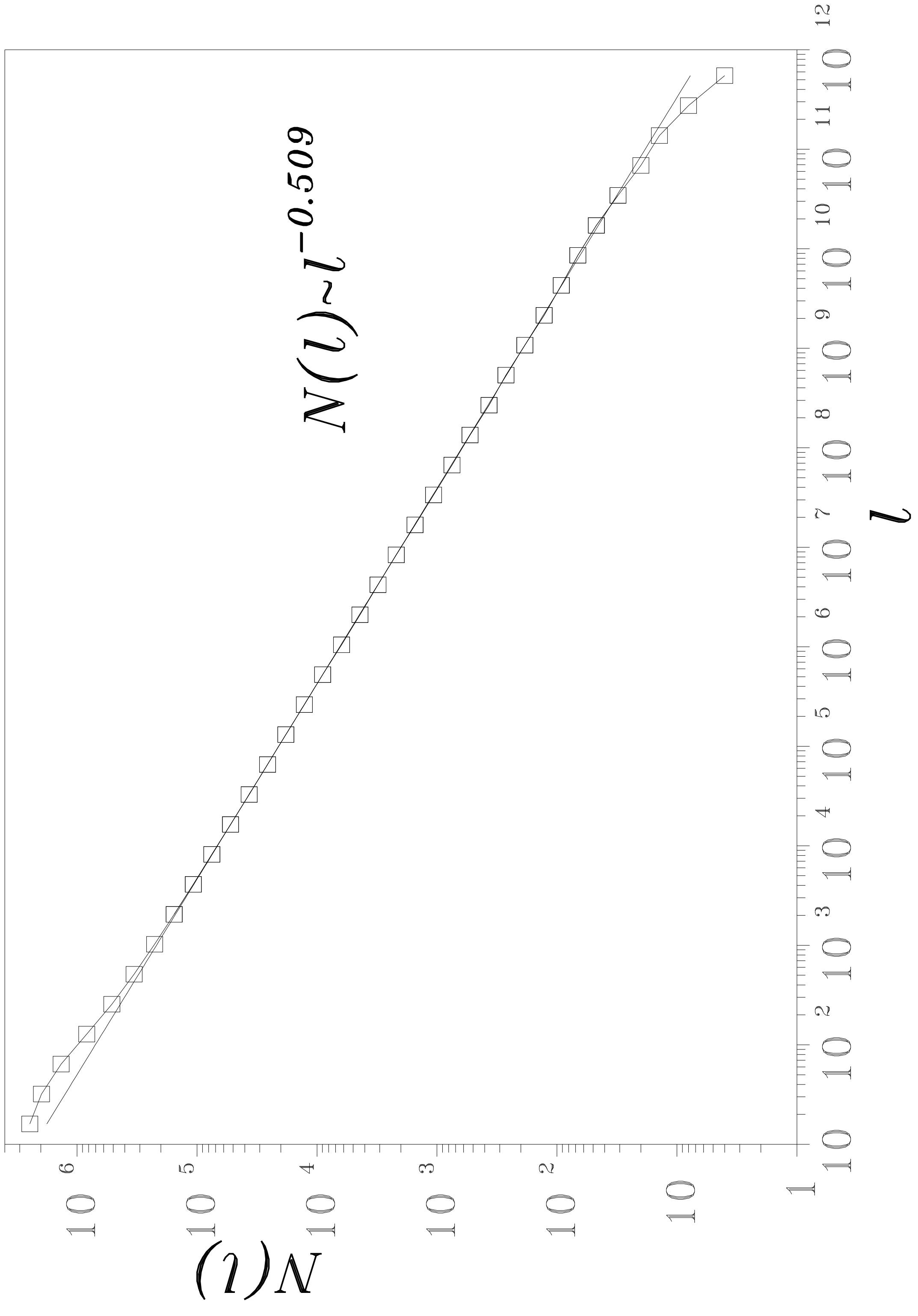

Fig.4 The plot showing the dependence of the distribution $N(l)$ vs l. There is a logarithmic scale on the x and y-axis.

Fig.5 The plot showing the dependence of the finite approximations to the Brun constant for twins (circles) and cousins (squares). On the horizontal axis values of $1/\ln(N)$ are used.