Relations between the random variable $w_x$ and the Dirichlet divisor problem

Dmitry Pyatin/ pyatin@petrsu.ru

sub. 03.07.2020,
rev. 25.09.2021

Abstract

We have developed a heuristic showing that in the Dirichlet divisor problem for almost all $n \in \mathbb{N}^+$:

$$R(n) \leq O(\psi(n)n^{\frac{1}{4}})$$

where

$$R(n) = \left| \sum_{x=1}^{n} \left\lfloor \frac{n}{x} \right\rfloor - n \log n - (2\gamma - 1)n \right|$$

and $\psi(n)$ - any positive function that increases unboundedly as $n \to \infty$. The result is achieved under the hypothesis:

$$\left\{ \frac{n}{x} \right\} \sim w_x$$

where $w_x$ is uniformly distributed over $[0, 1)$ random variable with a values set $\{0, \frac{1}{x}, \ldots, \frac{x-1}{x}\}$ and the value accepting probability $p = \frac{1}{x}$.

The paper concludes with a numerical argument in support of the hypothesis being true. It is shown that the expectation:

$$\mu_1 \left[ \sum_{x=1}^{n} \left( \frac{n}{x} - \frac{x-1}{2x} \right) \right] = (2n + 1)H_{\lfloor \sqrt{n} \rfloor} - \lfloor \sqrt{n} \rfloor^2 - \lfloor \sqrt{n} \rfloor + C$$

has deviation from $D(n)$ is less than $R(n)$ in absolute value for all $n < 10^5$.

Conventions

$\{x\}$ – fractional part of $x$;

$[a,b]$ – least common multiple of $a$ and $b$;

$(a,b)$ – greatest common divisor of $a$ and $b$;

$\mu_k[f(x)]$ – $k$-th central moment of $f(x)$;

$\gamma$ – the Euler-Mascheroni constant;

almost all, almost everywhere – all elements of the set, except for a zero measure subset;

$C$ – some constant.
Introduction

The Dirichlet divisor problem is to determine the lower bound for $\theta$ in the remainder estimate:

$$R(n) = \left| \sum_{x=1}^{n} \left\lfloor \frac{n}{x} \right\rfloor - n \log n - (2\gamma - 1)n \right| = O(n^{\theta+\epsilon})$$

where $D(n) = \sum_{x=1}^{n} \left\lfloor \frac{n}{x} \right\rfloor$ – divisor summatory function.

Using the hyperbola method Dirichlet showed [1], that $\theta \leq \frac{1}{2}$. G. Voronoi (1903) proved [2], that $\theta \leq \frac{1}{3}$. Further the result has improved repeatedly. H. Iwaniec and C. J. Mozzochi (1988) showed [3], that $\theta \leq \frac{7}{22}$. The best known result belongs to M. Huxley (2003), he established [4], that $\theta \leq \frac{131}{416}$.

In 1916, G. H. Hardy and independently E. Landau proved [5], that $\theta \geq \frac{1}{4}$, therefore it has been established that:

$$\frac{1}{4} \leq \theta \leq \frac{131}{416}$$

It is believed that $\theta = \frac{1}{4}$. In this paper, we show that almost everywhere $\theta = \frac{1}{4}$ under some hypothesis.

Content

It can be shown (see appendix) by using the result $\theta < \frac{1}{2}$ ([2]-[4]), the hyperbola method and equality:

$$\sum_{x=1}^{\sqrt{n}} \left\{ \frac{n}{x} \right\} = C[\sqrt{n}] + g(\lfloor \sqrt{n} \rfloor)$$

that for $n \to \infty$ the constant $C = \frac{1}{2}$ and hold:

$$R(n) = \left| \sum_{x=1}^{\sqrt{n}} \left( \left\{ \frac{n}{x} \right\} - \frac{1}{2} \right) \right|$$

From the work of J. Kubilius it is known [4], that as $n \to \infty$:

$$\nu_n \{ |R(n) - \mu_1[R(n)]| \leq \psi(n)\sqrt{\mu_2[R(n)]} \} \to 1$$

, where $\nu_n \{ \ldots \}$ – is the frequency of events with a condition $\{ \ldots \}$. $\nu_n = \frac{1}{n} N_n \{ \ldots \}$, where $N_n$ – number of events with a condition $\{ \ldots \}$, and $\psi(n)$ – any positive function that increases unboundedly as $n \to \infty$.

Thus, for almost all $n$ the following inequation is hold:

$$|R(n) - \mu_1[R(n)]| \leq \psi(n)\sqrt{\mu_2[R(n)]}$$

To find $\mu_1[R(n)]$ and $\mu_2[R(n)]$, we hypothesize:

**Hypothesis 1** For an arbitrary positive integer $n$ and a fixed positive integer $x$:

$$\left\{ \frac{n}{x} \right\} \sim w_x$$
, where \( w_x \) is a random variable uniformly distributed on \([0, 1)\), taking values from \(\{0, \frac{1}{x}, \ldots, \frac{x-1}{x}\}\) with probability \(p = \frac{1}{x}\).

Accepting this hypothesis, we can find \(\mu_1[R(n)]\):

\[
\mu_1[R(n)] = \mu_1\left[\sum_{x=1}^{\sqrt{n}} \left(\left\{\frac{n}{x}\right\} - \frac{1}{2}\right)\right] = \sum_{x=1}^{\sqrt{n}} \mu_1\left[\left\{\frac{n}{x}\right\} - \frac{1}{2}\right] = \sum_{x=1}^{\sqrt{n}} \left(\frac{1}{x} \sum_{k=0}^{x-1} \frac{1}{x} - \frac{1}{2}\right) = -\sum_{x=1}^{\sqrt{n}} \frac{1}{2x} = -\frac{1}{2} H[\sqrt{n}]
\] (6)

and \(\mu_2[R(n)]\):

\[
\mu_2[R(n)] = \mu_2\left[\sum_{x=1}^{\sqrt{n}} \left(w_x - \frac{1}{2}\right)\right] = \sum_{x=1}^{\sqrt{n}} \sum_{y=1}^{\sqrt{n}} \text{Cov}(w_x - \frac{1}{2}, w_y - \frac{1}{2})
\] (7)

where:

\[
\text{Cov}(w_a - \frac{1}{2}, w_b - \frac{1}{2}) = \frac{1}{[a, b]} \sum_{i=1}^{a} \sum_{j=1}^{b} \sum_{k=1}^{((a-1)(a, b) + k)} \left(\frac{a - ((i-1)(a, b) + k)}{a} - \frac{a - 1}{2a}\right) \left(\frac{b - ((j-1)(a, b) + k)}{b} - \frac{b - 1}{2b}\right) = \frac{1}{[a, b]} \frac{(a, b)^2 - 1}{12(a, b)} - \frac{1}{12ab}
\] (8)

The formula (8) comes from the block structure of the covariance matrix \(A(d_1, d_2)\):

\[
A\left(\frac{a}{(a, b)}, \frac{b}{(a, b)}\right) = \begin{bmatrix}
G_{11} & G_{12} & \cdots \\
\vdots & \ddots & \ddots \\
G_{(a,b),1} & G_{(a,b),2} & \cdots & G_{(a,b),a}
\end{bmatrix}
\] (9)

where \(G(d_1, d_2)\) diagonal matrix:

\[
G((a, b), (a, b)) = \begin{bmatrix}
\frac{1}{[a, b]} & 0 & \cdots \\
\vdots & \ddots & \ddots \\
0 & \cdots & \frac{1}{[a, b]}
\end{bmatrix}
\] (10)

using the general formula for finding the covariance of two discrete random variables:

\[
\text{Cov}(w_a, w_b) = \sum_{i=1}^{a} \sum_{j=1}^{b} p_{ij}(w_i - \mu_1(w_a))(w_j - \mu_1(w_b))
\] (11)

in which some terms are equal to zero due to \(p_{ij} = 0\), and the number of nonzero terms is equal to \([a, b]\).

After all we have a second central moment:

\[
\mu_2[R(n)] = \sum_{x=1}^{\sqrt{n}} \sum_{y=1}^{\sqrt{n}} \left(\frac{(a, b)}{12[a, b]} + \frac{1}{12ab}\right) = \sum_{x=1}^{\sqrt{n}} \sum_{y=1}^{\sqrt{n}} \left(\frac{(a, b)}{12[a, b]} + O((\log \sqrt{n})^2)\right)
\] (12)
L. Toth (et al.) gives \cite{7} an explicit formula for the sum with GCD and LCM:

\[
\sum_{a=1}^{n} \sum_{b=1}^{n} \frac{(a,b)}{[a,b]} = 3n + O((\log n)^2)
\]  

(13)

so:

\[
\mu_2[R(n)] = \frac{1}{4} \lfloor \sqrt{n} \rfloor + O((\log n)^2)
\]  

(14)

whence by substitution in (3) and adding \(\frac{1}{2} H_{\lfloor n \rfloor} \) we get that for \(n \to \infty\):

\[
\nu_n \{R(n) > O(\psi(n)n^{\frac{1}{2}})\} \to 0
\]  

(15)

where \(\psi(n)\) – any positive function, increasing unboundedly as \(n \to \infty\), and as a consequence as \(n \to \infty\) almost everywhere performed:

\[
R(n) \leq O(\psi(n)n^{\frac{1}{2}})
\]  

(16)

**Conclusion**

Let’s try to present an argument in support of the truth of hypothesis 1. Knowing that:

\[
D(n) = \sum_{x=1}^{n} \left\lfloor \frac{n}{x} \right\rfloor = \sum_{x=1}^{n} \frac{n}{x} - \sum_{x=1}^{n} \left\{ \frac{n}{x} \right\}
\]  

(17)

Let’s define a random variable:

\[
W(n) = \sum_{x=1}^{n} \frac{n}{x} - \sum_{x=1}^{n} w_x
\]  

(18)

Find \(\mu_1[W(n)]\), using \(\mu_1[w_x] = \frac{x+1}{2x}\) and the Dirichlet hyperbola method:

\[
\mu_1[W(n)] = (2n+1)H_{\lfloor \sqrt{n} \rfloor} - \lfloor \sqrt{n} \rfloor^2 - \lfloor \sqrt{n} \rfloor + C
\]  

(19)

Numerical calculations show that \(\mu_1[W(n)]\) closer to \(D(n)\), that \(n \log n + (2\gamma - 1)n\). Introduce the error functions:

\[
\Delta_R = \sum_{n=1}^{N} |R(n)|, \Delta_W = \sum_{n=1}^{N} \left( |D(n) - \mu_1[W(n)]| \right)
\]

\[
d_R = \sum_{n=1}^{N} R(n), d_W = \sum_{n=1}^{N} \left( D(n) - \mu_1[W(n)] \right)
\]  

(20)

Figure \[\square\] shows that \(\Delta_W < \Delta_R\) and \(d_W \ll d_R\). Figure \[\square\] shows \(d_W\) using the constant \(C = \frac{5}{12}\) in the formula (19). It can be replaced \(\frac{1}{2}\) with \(\frac{\pi - 1}{2\pi}\) in formula (2) for getting \(\mu_1 = 0\).

As a result the proposed heuristic is in good agreement with the numerical data for the parameter up to \(10^5\), however, the proof of the estimates for \(\theta\) in the Dirichlet divisor problem must be carried out using other methods.
Figure 1: Comparison of $\Delta_R$, $\Delta_W$, $d_R$ and $d_W$.

Figure 2: $d_W$ up to $N = 10^5$.

References

[1] G. H. Hardy, E. M. Wright. An Introduction to the Theory of Numbers, 5th ed. Oxford, England: Clarendon Press, 1979.

[2] A. Ivic. The Riemann Zeta-Function. New York: Dover Publications, 2003.

[3] H. Iwaniec, C. J. Mozzochi. On the Divisor and Circle Problem. J. Numb. Th. 29, 60-93, 1988.

[4] M. N. Huxley. Exponential Sums and Lattice Points III. Proc. London Math. Soc. 87, 2003.

[5] G. H. Hardy, Ramanujan: Twelve Lectures on Subjects Suggested by His Life and Work, 3rd ed. New York: Chelsea, 1999.

[6] Jonas Kubilius. Probabilistic Methods in the Theory of Numbers, 1962. p.54 [Russian]

[7] Titus Hilberdink, Florian Luca, Laszlo Toth. On certain sums concerning the gcd’s and lcm’s of k positive integers. International Journal of Number Theory, Vol. 16, No. 1 (2020) 77–90
Appendix

Let us prove that:

\[ \sum_{x=1}^{\sqrt{n}} \left\{ \frac{n}{x} \right\} = \frac{1}{2} [\sqrt{n}] + R(n) \]  

(21)

Using the equality:

\[ \sum_{x=1}^{\sqrt{n}} \left\{ \frac{n}{x} \right\} = C[\sqrt{n}] + g([\sqrt{n}]) \]  

(22)

using the Dirichlet hyperbola method, we obtain:

\[ D(n) = \sum_{x=1}^{n} \left\lfloor \frac{n}{x} \right\rfloor - \sqrt{n} \right\rfloor^{2} = 2 \sum_{x=1}^{\sqrt{n}} \left( \frac{n}{x} - \left\{ \frac{n}{x} \right\} \right) - \sqrt{n}^{2} = \\
= 2n \left( \log \sqrt{n} - 1 + \frac{1}{2 \sqrt{n}} + O\left( \frac{1}{\sqrt{n}} \right) \right) - 2C[\sqrt{n}] + g([\sqrt{n}]) - \sqrt{n}^{2} = \\
= 2n \left( \log \sqrt{n} - 1 + \frac{1}{2 \sqrt{n}} + O\left( \frac{1}{\sqrt{n}} \right) \right) - 2C[\sqrt{n}] + g([\sqrt{n}]) - (\sqrt{n} - 1)^{2} \]  

(23)

because:

\[ \log \sqrt{n} - \left\{ \sqrt{n} \right\} = \log \sqrt{n} - \frac{\left\{ \sqrt{n} \right\}}{\sqrt{n}} + O\left( \frac{1}{n} \right), \text{ and} \]

\[ (\sqrt{n} - \left\{ \sqrt{n} \right\})^{2} = n - 2\sqrt{n}\sqrt{n} + \left\{ \sqrt{n} \right\}^{2} \]

(24)

then, taking into account \( \left\{ \sqrt{n} \right\} < 1 \):

\[ D(n) = 2n \log \sqrt{n} - \frac{2n}{\sqrt{n}} + O(\left\{ \sqrt{n} \right\}) + 2\gamma n + \frac{2n}{2 \sqrt{n}} + O(1) - \\
- 2C[\sqrt{n}] + g([\sqrt{n}]) - n + 2\sqrt{n}\sqrt{n} - \left\{ \sqrt{n} \right\}^{2} \leq \\
\leq n \log n - 2\sqrt{n}\sqrt{n} + 2\sqrt{n}\sqrt{n} + 2\gamma n + \frac{n}{\sqrt{n}} - 2C[\sqrt{n}] - n + O(1) \]  

(25)

because \( \frac{n}{\sqrt{n}} < \left\lfloor \sqrt{n} \right\rfloor + 1 \), so:

\[ D(n) < n \log n + (2\gamma - 1)n + (1 - 2C)[\sqrt{n}] + g([\sqrt{n}]) + O(1) \]  

(26)

whence follows:

\[ R(n) = D(n) - (n \log n + (2\gamma - 1)n) = (1 - 2C)[\sqrt{n}] + g([\sqrt{n}]) \]  

(27)

т. к. \( R(n) \ll \sqrt{n} \), under [2] - [4], we conclude that:

\[ C = \frac{1}{2} \]

As a result:

\[ R(n) = \sum_{x=1}^{\sqrt{n}} \left( \left\{ \frac{n}{x} \right\} - \frac{1}{2} \right) = -\frac{1}{\pi} \sum_{x=1}^{\sqrt{n}} \sum_{k=1}^{\infty} \frac{\sin(2\pi k \frac{n}{x})}{k} \]  

(28)

6