STICHTING
MATHEMATISCH CENTRUM
2e Boerhaavestraat 49
AMSTERDAM

AFDELING ZUIVERE WISKUNDE

ZW 1969-004

On the distribution of a specific number-theoretical sequence

by

J. van de Lune

december 1969

Printed at the Mathematical Centre, 49, 2e Boerhaavestraat, Amsterdam.

The Mathematical Centre, founded the 11-th of February 1946, is a non-profit institution aiming at the promotion of pure mathematics and its applications. It is sponsored by the Netherlands Government through the Netherlands Organization for the Advancement of Pure Research (Z.W.O), by the Municipality of Amsterdam, by the University of Amsterdam, by the Free University at Amsterdam, and by industries.

On the distribution of a specific number-theoretical sequence

by

J. van de Lune

Introduction

This note must be considered as a continuation of [1], from which we recall some definitions and theorems.

For a natural number $m \geq 2$ we define $g(m)$ as the largest prime dividing m, whereas $g(1) = 1$. We also write g_m instead of $g(m)$.

Let $G(n,\alpha)$ be the number of natural numbers m with the properties $m \leq n$ and $g(m) \leq m^\alpha$, where α is a fixed real number.

It can be shown that the function

$$G(\alpha) \overset{\text{def}}{=} \lim_{n\to\infty} \frac{1}{n} G(n,\alpha)$$

is continuous and satisfies:

$$\begin{cases} G(\alpha) = 0, & (\alpha \leq 0) \\ G(\alpha) = 1, & (\alpha \geq 1) \\ G'(\alpha) = \frac{1}{\alpha} G(\frac{\alpha}{1-\alpha}), & (0<\alpha<1). \end{cases}$$

Defining

$$\begin{cases} H(x) = 1 & \text{for } 0 \leq x \leq 1 \\ H(x) = G(\frac{1}{x}) & \text{for } x > 1, \end{cases}$$

it is easy to see that $H(x)$ is continuous on $x \geq 0$ and satisfies the equation

$$H'(x) = -\frac{1}{x} H(x-1), \qquad (x>1).$$

From this it follows that

$$xH(x) = \int_{x-1}^{x} H(t)dt, \qquad (x \geq 1),$$

and by means of this formula it is easily shown that H(x) is a positive function which tends to zero very rapidly when x tends to infinity.
We now define the sequence λ_k, $(k=1,2,3,\ldots)$ as follows

$$\begin{cases} \lambda_1 = 1 \\ g_k^{\lambda_k} = k, \qquad (k=2,3,4,\ldots). \end{cases}$$

It is to be expected that λ_k behaves very irregular and while tabulating this sequence one might conjecture for example that the sequence λ_k is uniformly distributed modulo 1.
However, in this note it will be shown that the sequence λ_k is not uniformly distributed modulo a, for any positive a.

1. <u>Lemma 1</u>. If the function f(x) is such that the integral

$$\int_1^A f(x)dH(x), \qquad (A>1)$$

exists as an ordinary Riemann-Stieltjes integral then

$$\lim_{n\to\infty} \frac{1}{n} \sum_{\substack{k\leq n \\ \lambda_k<A}} f(\lambda_k) = - \int_1^A f(x)dH(x).$$

<u>Proof</u>. On the interval [0,A] we construct a subdivision
$1 = a_0 < a_1 < a_2 < \ldots < a_{m-1} < a_m = A$ and we define

$$M_\nu = \sup_{a_{\nu-1} \leq x \leq a_\nu} f(x),$$

$$m_\nu = \inf_{a_{\nu-1} \leq x \leq a_\nu} f(x).$$

Since

$$\int_1^A f(x)dH(x)$$

exists we may choose the subdivision of [1,A] such that

$$\sum_{\nu=1}^{m} M_\nu\{H(a_{\nu-1}) - H(a_\nu)\} < -\int_1^A f(x)dH(x) + \varepsilon$$

and

$$\sum_{\nu=1}^{m} m_\nu\{H(a_{\nu-1}) - H(a_\nu)\} > -\int_1^A f(x)dH(x) - \varepsilon.$$

We now write

$$\frac{1}{n}\sum_{\substack{k\le n\\ \lambda_k<A}} f(\lambda_k) = \frac{1}{n}\sum_{\nu=1}^{m}\sum_{\substack{a_{\nu-1}\le\lambda_k<a_\nu\\ k\le n}} f(\lambda_k)$$

and observe that

$$\frac{1}{n}\sum_{\nu=1}^{m}\sum_{\substack{a_{\nu-1}\le\lambda_k<a_\nu\\ k\le n}} f(\lambda_k) \le \frac{1}{n}\sum_{\nu=1}^{m}\sum_{\substack{a_{\nu-1}\le\lambda_k<a_\nu\\ k\le n}} M_\nu =$$

$$= \frac{1}{n}\sum_{\nu=1}^{m} M_\nu \cdot \{G(n, \frac{1}{a_{\nu-1}}) - G(n, \frac{1}{a_\nu})\},$$

because of the fact that for all ν the number of natural numbers k satisfying the conditions $k \le n$ and $a_{\nu-1} \le \lambda_k < a_\nu$ is equal to

$$G(n, \frac{1}{a_{\nu-1}}) - G(n, \frac{1}{a_\nu}).$$

Since

$$\lim_{n\to\infty}\frac{1}{n}\sum_{\nu=1}^{m} M_\nu\{G(n, \frac{1}{a_{\nu-1}}) - G(n, \frac{1}{a_\nu})\} =$$

$$= \sum_{\nu=1}^{m} M_\nu\{G(\frac{1}{a_{\nu-1}}) - G(\frac{1}{a_\nu})\} =$$

$$= \sum_{\nu=1}^{m} M_\nu\{H(a_{\nu-1}) - H(a_\nu)\} < -\int_1^A f(x)dH(x) + \varepsilon$$

we obtain that

$$\limsup_{n\to\infty} \frac{1}{n} \sum_{\substack{k\le n \\ \lambda_k<A}} f(\lambda_k) < -\int_1^A f(x)\,dH(x) + \varepsilon.$$

In a similar way one also proves that

$$\liminf_{n\to\infty} \frac{1}{n} \sum_{\substack{k\le n \\ \lambda_k<A}} f(\lambda_k) > -\int_1^A f(x)\,dH(x) - \varepsilon.$$

Since this is true for all $\varepsilon > 0$ we may conclude:

$$\lim_{n\to\infty} \frac{1}{n} \sum_{\substack{k\le n \\ \lambda_k<A}} f(\lambda_k) = -\int_1^A f(x)\,dH(x).$$

<u>Theorem</u>. If the function $f(x)$ is such that $|f(x)| \le M$ for all $x \ge 1$ and the integral $\int_1^\infty f(x)\,dH(x)$ exists, then

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} f(\lambda_k) = -\int_1^\infty f(x)\,dH(x).$$

<u>Proof</u>. We write

$$\frac{1}{n} \sum_{k=1}^{n} f(\lambda_k) = \frac{1}{n} \sum_{\substack{k\le n \\ \lambda_k<A}} f(\lambda_k) + \frac{1}{n} \sum_{\substack{k\le n \\ \lambda_k\ge A}} f(\lambda_k)$$

and fix $A > 1$ such that $M.H(A)$ is small.
Then we have

$$\left|\frac{1}{n} \sum_{k=1}^{n} f(\lambda_k) + \int_1^\infty f(x)\,dH(x)\right| \le \left|\frac{1}{n} \sum_{\substack{k\le n \\ \lambda_k<A}} f(\lambda_k) + \int_1^A f(x)\,dH(x)\right| +$$

$$+ \frac{1}{n} \sum_{\substack{k\le n \\ \lambda_k\ge A}} |f(\lambda_k)| + \left|\int_A^\infty f(x)\,dH(x)\right|.$$

According to lemma 1 the first of these terms can be made arbitrarily small by taking n large enough. The second is

$$\leq \frac{M}{n} G(n, \frac{1}{A})$$

which tends to M.H(A) as $n \to \infty$, whereas the third term is

$$\leq - M. \int_A^\infty dH(x) = M\ H(A).$$

From this it is clear that

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} f(\lambda_k) = - \int_1^\infty f(x)dH(x).$$

As an application of this theorem we prove the assertion concerning the distribution of λ_k made in the introduction.
Let a be any fixed positive number and define the set E_t for $0 \leq t < a$ as follows:

$$E_t = \bigcup_{r=0}^{\infty} \{x \in \mathbb{R};\ ra \leq x \leq ra+t\}$$

and let f_t be the characteristic function of E_t.
It is easily seen that this f_t satisfies the conditions of theorem 1.
Thus

$$D(t) \stackrel{\text{def}}{=} \lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} f_t(\lambda_k) = - \int_1^\infty f_t(x)dH(x) = - \int_0^\infty f_t(x)dH(x)$$

$$= - \sum_{r=0}^{\infty} \int_{ra}^{ra+t} dH(x) = - \sum_{r=0}^{\infty} \{H(ra+t) - H(ra)\}.$$

It is rather easy to convince oneself that D(t) is differentiable on the interval

$$1 - a.[\frac{1}{a}] < t < a,$$

such that

$$D'(t) = - \sum_{r=[\frac{1}{a}]}^{\infty} H'(ra+t) = \sum_{r=[\frac{1}{a}]}^{\infty} \frac{1}{ra+1} H(ra+t-1),$$

from which it is obvious that $D'(t)$ is decreasing. Hence $D'(t)$ is not constant.

However, from the definition of $D(t)$ and the assumption that the sequence λ_k is uniformly distributed modulo a it would follow that

$$D(t) = \frac{t}{a}, \qquad (0 \leq t \leq a)$$

and

$$D'(t) = \frac{1}{a} = \text{constant.}$$

<u>Conclusion</u>. The sequence λ_k is not uniformly distributed modulo a, for any $a > 0$.

2. In this section we will make a few remarks on the behaviour of

$$\frac{1}{n} \sum_{k=1}^{n} f(\lambda_k)$$

where f is not bounded.

If we take $f(x) = 2^{2x}$ one has

$$\frac{1}{n} \sum_{k=1}^{n} f(\lambda_k) \geq \frac{1}{n} f(\lambda_n)$$

and for $n = 2^m$

$$\frac{1}{n} f(\lambda_n) = \frac{1}{2^m} f(m) = 2^m$$

and hence it follows that

$$\frac{1}{n} \sum_{k=1}^{n} f(\lambda_k)$$

is divergent.

On the other hand, it is easy to show that

$$\int_1^\infty 2^{2x}\, dH(x)$$

exists.

Thus, it may happen that

$$\int_1^\infty f(x)dH(x)$$

exists whereas

$$\frac{1}{n}\sum_{k=1}^{n} f(\lambda_k)$$

is divergent.

A somewhat more precise result is the next theorem

<u>Theorem</u>. If $f(x) \geq 0$ on $x \geq x_0$ and $f(x)$ is bounded on each finite interval and $\int_1^\infty f(x)dH(x)$ exists then

$$\liminf_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} f(\lambda_k) \geq -\int_1^\infty f(x)dH(x).$$

<u>Proof</u>. For any $A \geq x_0$ one has

$$\frac{1}{n}\sum_{k=1}^{n} f(\lambda_k) \geq \frac{1}{n}\sum_{\substack{k\leq n\\ \lambda_k\leq A}} f(\lambda_k),$$

so that

$$\liminf_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} f(\lambda_k) \geq -\int_1^A f(x)dH(x)$$

for each $A \geq x_0$, and the theorem follows.

A curious consequence of this theorem is, taking $f(x) = x$,

$$\liminf_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \lambda_k \geq - \int_1^\infty x \, dH(x)$$

$$= - \int_0^\infty x \, dH(x) = \int_0^\infty H(x)dx = e^\gamma = 1{,}781\ldots,$$

where γ is Euler's constant (c.f. [1], page 24).

Reference:

[1] J. van de Lune and E. Wattel, On the frequency of natural numbers m whose prime divisors are all smaller than m^α, Mathematical Centre, Amsterdam, Report ZW 1968-007 (1968).