

COMPOSITIO MATHEMATICA

WERNER NEUDECKER

DAVID WILLIAMS

The “Riemann hypothesis” for the Hawkins random Sieve

Compositio Mathematica, tome 29, n° 2 (1974), p. 197-200

http://www.numdam.org/item?id=CM_1974__29_2_197_0

© Foundation Compositio Mathematica, 1974, tous droits réservés.

L'accès aux archives de la revue « Compositio Mathematica » (<http://www.compositio.nl/>) implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/legal.php>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme
Numérisation de documents anciens mathématiques

<http://www.numdam.org/>

THE ‘RIEMANN HYPOTHESIS’ FOR THE HAWKINS RANDOM SIEVE

Werner Neudecker and David Williams

1

Hawkins ([1, 2]) and Wunderlich ([6, 7]) have studied the sequence $\{X_n : n \in \mathbb{N}\}$ of ‘random primes’ caught in the Hawkins Sieve which operates inductively as follows.

Let

$$A_1 = \{2, 3, 4, 5, 6, \dots\}.$$

Stage 1. Declare $X_1 = \min A_1$. From the set $A_1 \setminus \{X_1\}$, each number in turn (and each independently of the others) is deleted with probability X_1^{-1} . The set of elements of $A_1 \setminus \{X_1\}$ which remain is denoted by A_2 .

Stage n. Declare $X_n = \min A_n$. From the set $A_n \setminus \{X_n\}$, each number in turn (and each independently of the others) is deleted with probability X_n^{-1} . The set of elements of $A_n \setminus \{X_n\}$ which remain is denoted by A_{n+1} .

Define

$$Y_n = \prod_{k \leq n} (1 - X_k^{-1})^{-1}.$$

Notation and conventions. For $x > 1$, we write $\text{li}(x)$ for the value of the logarithmic integral at x :

$$\text{li}(x) = \lim_{\delta \downarrow 0} \left(\int_0^{1-\delta} + \int_{1+\delta}^x \right) \frac{dz}{\log z} \sim \frac{x}{\log x}.$$

Qualifying ‘with probability one’ phrases will be suppressed. An equation involving the symbol ‘ ε ’ is to be understood as true for every positive ε .

2

Recall that the ‘real’ Riemann Hypothesis about the zeros of the Riemann zeta-function is equivalent to the statement:

$$(1) \quad \text{li}(p_n) = n + O(n^{\frac{1}{2} + \varepsilon})$$

(where p_n denotes the n th prime) and that equation (1) is the result which Riemann really wished to prove. See for example Ingham [3].

THEOREM: *The following ‘Riemann Hypothesis for the Hawkins Sieve’ holds:*

$$(2) \quad L \operatorname{li} (X_n L^{-1}) = n + O(n^{\frac{1}{2} + \epsilon})$$

where L denotes the random limit

$$(3) \quad L = \lim_{n \rightarrow \infty} X_n \exp(-Y_n).$$

This theorem was motivated by the exactly analogous ‘diffusion’ result in Williams [5]. The proof now given mirrors that in [5].

3

The process $\{(X_n, Y_n) : n \in \mathbb{N}\}$ is *Markovian* with

$$(4) \quad \begin{aligned} P[X_{n+1} - X_n = j | \mathcal{F}_n] &= Y_n^{-1} (1 - Y_n^{-1})^{j-1} \quad (j \in \mathbb{N}), \\ Y_{n+1} &= Y_n (1 - X_{n+1}^{-1})^{-1} = Y_n (1 + Z_{n+1}^{-1}) \end{aligned}$$

and $X_1 = Y_1 = 2$. We have written (for $n \in \mathbb{N}$):

$$Z_n = X_n - 1, \quad \mathcal{F}_n = \sigma\{(X_k, Y_k) : k \leq n\},$$

the latter equation signifying that \mathcal{F}_n is the smallest σ -algebra with respect to which X_k and Y_k are measurable for every $k \leq n$. Introduce

$$(5) \quad U_{n+1} = (Z_{n+1} - Z_n) Y_n^{-1} \quad (n \in \mathbb{N}).$$

Elementary properties of geometric distributions now make things particularly neat. For $x > 0$ and $n \in \mathbb{N}$,

$$P[U_{n+1} > x | \mathcal{F}_n] \leq (1 - Y_n^{-1})^{-1} (1 - Y_n^{-1})^{x Y_n} \leq 2e^{-x}.$$

(Recall that $Y_n \geq 2$ for every n .) By the Borel-Cantelli Lemma,

$$(6) \quad U_{n+1} = O(\log n) = O(n^\epsilon).$$

Because

$$E[(U_{n+1} - 1) | \mathcal{F}_n] = 0,$$

$\{(U_{n+1} - 1) : n \in \mathbb{N}\}$ is a family of *orthogonal* random variables. Since also

$$E[(U_{n+1} - 1)^2 | \mathcal{F}_n] = 1 - Y_n^{-1} \leq 1,$$

Theorem 33B(ii) of Loève [4] provides the estimate:

$$(7) \quad \sum_{k \leq n} (U_{k+1} - 1) = O(n^{\frac{1}{2} + \epsilon}).$$

In particular,

$$(8) \quad n^{-1} \sum_{k \leq n} U_{k+1} \rightarrow 1.$$

4

The remainder of the proof is divided into three stages.

PROPOSITION 1: $Y_n \uparrow \infty$ and $Z_n \sim nY_n$.

PROOF: If $Y_n \uparrow Y < \infty$, then we could conclude from (4) that $\sum Z_n^{-1} < \infty$ and from (5) and (8) that (in contradiction) $n^{-1}Z_n \rightarrow Y$.

Thus $Y_n \uparrow \infty$ and

$$(9) \quad Z_{n+1} Y_{n+1}^{-1} - Z_n Y_n^{-1} = U_{n+1} - Y_{n+1}^{-1} = U_{n+1} + o(1)$$

so that (from (8)) $Z_n \sim nY_n$.

PROPOSITION 2: If $H_n = \log Z_n - Y_n$ ($n \in \mathbf{N}$), then

$$(10) \quad H_n = C + O(n^{-\frac{1}{2} + \varepsilon})$$

for some (random) C .

PROOF: Since $x(1+x)^{-1} \leq \log(1+x) \leq x$ for $x \geq 0$,

$$\begin{aligned} H_{n+1} - H_n &= \log(1 + \alpha_n U_{n+1}) - \alpha_n(1 + \alpha_n U_{n+1})^{-1} \\ &= \beta_n(U_{n+1} - 1) + O(\alpha_n^2 U_{n+1}^2) \end{aligned}$$

where

$$\alpha_n = Y_n Z_n^{-1} = O(n^{-1}) \quad \text{and} \quad \beta_n = Y_n Z_{n+1}^{-1} = O(n^{-1}).$$

Proposition 2 now follows by partial summation using (6), (7) and the further estimate:

$$\beta_{n-1} - \beta_n = \beta_{n-1} \beta_n U_{n+1} - \beta_{n-1} Z_{n+1}^{-1} = O(n^{-2+\varepsilon}).$$

Note. From equation (3), in which of course $L = \exp(C)$, and Proposition 1, it follows that

$$(11) \quad Y_n \sim \log n, \quad X_n \sim n \log n.$$

In other words, 'Mertens' Theorem' and the 'Prime Number Theorem' hold. (One could not expect the e^γ factor which is the rather tantalising feature of the *real* Mertens Theorem.) Wunderlich obtained both results at (11) by a more complicated method.

5

On summing equation (9) over n and utilising the exponentiated form:

$$(12) \quad Z_n = L \exp(Y_n)(1 + O(n^{-\frac{1}{2}+\varepsilon}))$$

of equation (10), we obtain

$$LY_n^{-1} \exp(Y_n) = n - \sum_{k \leq n} Y_k^{-1} + O(n^{\frac{1}{2}+\varepsilon}).$$

Extend the random function Y from $\{1, 2, 3, \dots\}$ to $(1, \infty)$ by linear interpolation. Then it is easily checked that

$$LY_t^{-1} \exp(Y_t) = \int_1^t (1 - Y_s^{-1}) ds + f(t)$$

where $f(t) = O(t^{\frac{1}{2}+\varepsilon})$. But now we may compute

$$\begin{aligned} [L \operatorname{li}(\exp(Y_s))]_1^t &= \int_1^t (1 - Y_s^{-1})^{-1} d[LY_s^{-1} \exp(Y_s)] \\ &= t - 1 + \int_1^t f'(s) [1 - Y_s^{-1}]^{-1} ds. \end{aligned}$$

Integration by parts using $Y'_s = O(s^{-1})$ establishes the following strong form of 'Mertens' Theorem':

$$(13) \quad L \operatorname{li}(\exp(Y_t)) = t + O(t^{\frac{1}{2}+\varepsilon}).$$

Equation (2) now follows on combining equations (12) and (13).

REFERENCES

- [1] D. HAWKINS: The random sieve. *Math. Mag.*, 31 (1958) 1-3.
- [2] D. HAWKINS: Random Sieves II. *J. Number Theory* 6 No. 3 (1974) 192-200.
- [3] A. E. INGHAM: *The distribution of prime numbers*. Cambridge University Press, 1932. Reprinted by Stechert-Hafner Service Agency, Inc., New York and London, 1964.
- [4] M. LOËVE: *Probability Theory*, van Nostrand, Princeton, N.J., 1963.
- [5] D. WILLIAMS: A study of a diffusion process motivated by the Sieve of Eratosthenes (to appear in *Bull. London Math. Soc.*).
- [6] M. C. WUNDERLICH: The prime number theorem for random sequences (to appear in *J. Number Theory*).
- [7] M. C. WUNDERLICH: A probabilistic setting for prime number theory (to appear in *Acta Arithmetica* 27, 1).