Reprinted from ILLINOIS JOURNAL OF MATHEMATICS
Vol. 5, No. 2, June 1961
Printed in U.S.A.

A PROBLEM ABOUT PRIME NUMBERS AND THE RANDOM WALK II

BY
P. ERDÖS

I am going to prove $\gamma = 1$. Denote by $u(a, b)$ the probability of the random walk passing through a if it starts at b. It is known and easy to prove that

$$u(a, b) \sim c_1 \mid b - a \mid^{-1} \tag{1}$$

(see K. ITÔ AND H. P. MCKEAN, JR., *Potentials and the random walk*, Illinois J. Math., vol. 4 (1960), pp. 119–132; also a paper of Murdoch cited therein where a sharper result is obtained). In the sequel, the letters p and q denote primes and $u(p, q) = u(a, b)$ in case $a = (p, 0, 0)$ and $b = (q, 0, 0)$.

Consider the number $e(n)$ of points $(p, 0, 0)$ $(p \leqq n)$ that the path hits. We have to prove that for almost all paths $e(n) \uparrow \infty$ as $n \uparrow \infty$.

By (1) and Mertens' estimate $\sum_{p \leqq n} p^{-1} \sim \lg_2 n$ ($\lg_2 = \lg \lg$), we evidently have

$$E[e(n)] = \sum_{p \leqq n} u(0, p) \sim c_1 \sum_{p \leqq n} p^{-1} \sim c_1 \lg_2 n. \tag{2}$$

Next, we prove by a customary argument

$$E[(e(n) - c_1 \lg_2 n)^2] = o(\lg_2 n)^2, \tag{3}$$

which establishes the weak law of large numbers for $e(n)$, i.e., it shows that $e(n) = c_1 \lg_2 n + o(\lg_2 n)$ except for a set of small measure, and this is enough for our purpose.

Clearly by (2)

$$E[(e - c_1 \lg_2 n)^2] = E(e^2) - c_1^2(\lg_2 n)^2 + o(\lg_2 n)^2. \tag{4}$$

Further we evidently have

$$\begin{aligned} E(e^2) &= \sum_{p \leqq n} u(0, p) + \sum_{q<p \leqq n} [u(0, p)u(p, q) + u(0, q)u(q, p)] \\ &= 2c_1^2 \sum_{q<p \leqq n} [1/p(p - q) + 1/q(p - q)] + o(\lg_2 n)^2. \end{aligned} \tag{5}$$

Mertens' estimate cited above gives $\sum_{q<p \leqq n} 1/(qp) = \frac{1}{2}(\lg_2 n)^2 + O(\lg_2 n)$, and so

$$\begin{aligned} \sum_{q<p \leqq n} 1/q(p - q) &= \sum_{q<p<n} 1/(qp) + \sum_{q<p \leqq n} [1/q(p - q) - 1/qp] \\ &= \tfrac{1}{2}(\lg_2 n)^2 + \sum_{q<p \leqq n} 1/p(p - q) + O(\lg_2 n). \end{aligned} \tag{6}$$

Thus we have only to estimate $\sum_{q<p \leqq n} 1/p(p - q)$.

Received May 19, 1960. See the preceding paper for a statement of the problem. This paper is from a letter sent by P. Erdös to the Illinois Journal of Mathematics. It was edited for publication by H. P. McKean, Jr., who then wrote the preceding paper which is a treatment of the same problem.

Put $\varepsilon_k = 0$ if k is not prime and $\varepsilon_p = \sum_{q<p} 1/(p-q)$. We have

$$(7) \quad \sum_{q<p\leq n} 1/p(p-q) = \sum_{k=1}^{n} \varepsilon_k/k = \sum_{k=1}^{n} s_k/k(k+1) + O(1)$$

by partial summation ($s_k = \sum_{i=1}^{k} \varepsilon_i$). A well-known theorem of Schnirelmann states that the number of solutions of $p - q = a$ ($p \leq k$) is less than $c_2\, k(\lg k)^{-2}\prod_{p|a} (1 + p^{-1})$ where c_2 is an absolute constant. Thus

$$(8) \qquad s_k < c_2\, k(\lg k)^{-2} \sum_{a=1}^{k} a^{-1} \prod_{p|a} (1 + p^{-1}) < c_3\, k/\lg k$$

since by interchanging the order of summation we have the well-known

$$\sum_{a=1}^{k} a^{-1} \prod_{p|a} (1 + p^{-1}) = \sum_{d=1}^{k} d^{-1} \sum_{a\equiv 0 (\mathrm{mod}\ d), a\leq k} < c_4 \sum_{d=1}^{\infty} \lg k/d^2 < c_5 \lg k.$$

Thus from (7) and (8)

$$(9) \qquad \sum_{q<p\leq n} 1/p(p-q) < c_6 \lg_2 n.$$

From (9), (6), and (5), we finally obtain $E(e^2) = c_1^2(\lg_2 n)^2 + o(\lg_2 n)^2$ which proves (3), and thus the proof of our theorem is complete.

By using a sharper estimate than (1), it is easy to show that for almost all paths

$$(10) \qquad \lim_{n\to\infty} e(n)/c_1 \lg_2 n = 1.$$

By the same method one can prove that if the integers $q = q_1 < q_2 < \cdots$ satisfy

$$(11) \qquad q_n - q_{n-1} > c_7 \lg n \quad (n \leqq 2), \qquad \sum 1/q = \infty,$$

then almost all paths pass through infinitely many points $(q, 0, 0)$. The primes probably do not satisfy (11) since probably there are an infinite number of prime twins, but one can prove by Brun's method that one can select a subsequence that does satisfy (11).

AUSTRALIAN NATIONAL UNIVERSITY
CANBERRA, AUSTRALIA