

# THE CENTRAL LIMIT THEOREM FOR THE RANGE OF TRANSIENT RANDOM WALK

BY NARESH C. JAIN AND WILLIAM E. PRUITT<sup>1</sup>

Communicated by Henry McKean, November 25, 1969

Let  $\{X_n, n \geq 1\}$  be a sequence of independent identically distributed random variables which take values in the  $d$ -dimensional integer lattice  $E_d$ . The sequence  $S_n$  defined by  $S_0 = 0$ ,  $S_n = \sum_{k=1}^n X_k$  for  $n \geq 1$  is called a random walk. If the distribution of the  $X_n$  assigns mass  $(2d)^{-1}$  to each of the  $2d$  neighbors of the origin, it is called *simple* random walk. Let  $p = P[S_1 \neq 0, S_2 \neq 0, \dots]$ ; the random walk is called transient if  $p$  is positive and recurrent otherwise. The random walk may take place on a proper subgroup of  $E_d$ . In this case, the subgroup is isomorphic to some  $E_k$ ,  $k \leq d$ ; if  $k < d$ , then we should consider the problem in  $k$  dimensions. With this understanding, random walks with summands having mean zero and finite variance are transient if and only if  $d \geq 3$  [3].

Let  $R_n$  denote the cardinality of the set  $\{S_0, S_1, \dots, S_n\}$ ;  $R_n$  is called the range of the random walk (up to time  $n$ ). Dvoretzky and Erdős [1] proved for simple random walk with  $d \geq 2$  that  $R_n/ER_n \rightarrow 1$  with probability one. In the course of their investigation, they obtained the estimate  $\text{Var } R_n = O(n)$  for  $d \geq 5$ . Jain and Orey considered the range of strongly transient random walk in [2]. (Random walks with summands having mean zero and finite variance are strongly transient if and only if  $d \geq 5$ .) They proved that if  $p < 1$  and the random walk is strongly transient, then  $\text{Var } R_n \sim \sigma^2 n$  for some positive constant  $\sigma^2$  and also that  $R_n$  obeys the central limit theorem. The case  $p = 1$  is uninteresting since then  $R_n = n + 1$  almost surely.

We shall consider these problems for general random walk in three and four dimensions. The bounds that Dvoretzky and Erdős obtained here were  $\text{Var } R_n = O(n \log n)$  if  $d = 4$  and  $\text{Var } R_n = O(n^{3/2})$  if  $d = 3$ . We have improved these bounds to  $O(n)$  and  $O(n \log n)$  respectively. Furthermore, we have proved that there is a positive constant  $\sigma^2$  (which may depend on the distribution of the summands) such that  $\text{Var } R_n \sim \sigma^2 n$  if  $d = 4$ , while  $\text{Var } R_n \sim \sigma^2 n \log n$  if  $d = 3$ . (The result for  $d = 3$  is only for the case where the summands have mean zero and finite variance.) With the asymptotic behavior of the variance available, we are then able to prove the central limit theorem.

---

*AMS Subject Classifications.* Primary 6066, 6030.

*Key Words and Phrases.* Random walk, transient, range, central limit theorem.

<sup>1</sup> Research was supported in part by NSF Grant No. GP-7490.

**THEOREM 1.** *Let  $R_n$  be the range of a random walk whose (genuine) dimension is at least four and  $p < 1$ . Then there is a positive constant  $\sigma^2$  such that  $\text{Var } R_n \sim \sigma^2 n$  and  $(R_n - np)/\sqrt{n\sigma}$  converges in distribution to the normal with mean 0 and variance 1.*

**THEOREM 2.** *Let  $R_n$  be the range of a random walk with (genuine) dimension three,  $p < 1$ , and with summands having mean zero and finite variance. Then there is a positive constant  $\sigma^2$  such that  $\text{Var } R_n \sim \sigma^2 n \log n$  and  $(R_n - np)/(n \log n)^{1/2} \sigma$  converges in distribution to the normal with mean 0 and variance 1.*

The situation in three dimensions is particularly interesting. If we write  $R_n = \sum_{k=0}^n Z_k$  where  $Z_k$  is the indicator of the event that a new lattice point is visited at time  $k$ , then we see that  $R_n$  is the sum of  $n+1$  zero-one random variables. They are neither independent nor identically distributed. In fact, there is enough dependence to make the variance grow like  $n \log n$  and yet the central limit theorem applies.

The proofs will appear elsewhere but we will give a brief indication of what is involved. In the four dimensional case, the main problem is to get the asymptotic estimate for the variance. This is a somewhat different problem than in higher dimensions but is fairly straightforward and relies only on the uniform estimate  $P^n(0, x) \leq A n^{-2}$ . Once the variance estimate is obtained, a somewhat simpler version of the blocking technique used in [2] can be applied along with the Lindeberg Theorem for triangular arrays. In three dimensions the estimation of the variance is more delicate and we make use of the asymptotic behavior of the Green function [3]. The Lindeberg Theorem is used once more, but in this case the condition is far more difficult to check due to the faster growth of the variance. In order to get the necessary bound, we have been forced to estimate the fourth central moment of  $R_n$ . The estimate obtained is  $E(R_n - ER_n)^4 = O(n^2 \log^2 n)$ . The same care needed in the variance estimate is also needed here, and since there are so many more terms involved the computation is extremely long.

#### REFERENCES

1. A. Dvoretzky and P. Erdős, *Some problems on random walk in space*, Proc. Second Berkeley Sympos. Statist. and Prob. (Berkeley, Calif., 1950), Univ. of California Press, Berkeley, Calif., 1951, pp. 353–367. MR 13, 852.
2. N. Jain and S. Orey, *On the range of random walk*, Israel J. Math. 6 (1968), 373–380. MR 39 #4944.
3. F. Spitzer, *Principles of random walk*, The University Series in Higher Math., Van Nostrand, Princeton, N. J., 1964. MR 30 #1521.