Can. J. Math. Vol. XLII, No. 5, 1990 pp. 915–932

MORE ON THE WAITING TIME TILL EACH OF SOME GIVEN PATTERNS OCCURS AS A RUN

TAMÁS F. MÓRI

1. Introduction. Let H be a finite set; we can suppose $H = \{1, 2, \ldots, d\}$. Consider H^n, the set of length n words over the alphabet H. For every $A \in H^n$ define the waiting time for A as the number of experiments needed till A appears as a connected sub-sequence of random elements of H. Formally, let $X_1, X_2, \ldots$ be i. i. d. random variables, $P(X_1 = i) = d^{-1}$, $1 \leq i \leq d$; then

$$T(A) = \inf\{m: (X_{m-n+1}X_{m-n+2}\ldots X_m) \equiv A\}\ .$$

A great number of papers have been devoted to problems connected with these waiting times, especially waiting times for pure runs (= elements of the diagonal of H^n).

An interesting problem that involves all the waiting times $T(A)$, $A \in H^n$, is to investigate the limit behaviour of the maximum waiting time

$$W_n = \max\{T(A): A \in H^n\}$$

as the length of the words tends to infinity. In [**1**] two-sided estimations are given for the expected number of coin tossings till each head-and-tail sequence of length n is observed as a run. In [**7**] it is proved that

$$E(W_n) = d^n(\log d^n + O(1)).$$

In a recent paper [**9**] the limit distribution of the maximum waiting time is derived, in a generalized setting. Given an arbitrary subset $H_n \subset H^n$ one can define

$$W(H_n) = \max\{T(A): A \in H_n\}. \tag{1}$$

The main result of [**9**] asserts that for every real y

$$\lim_{n\to\infty} P(d^{-n}W(H_n) - \log|H_n| \leq y) = e^{-e^{-y}}, \tag{2}$$

Received August 1st, 1989.
Research supported by the Hungarian National Foundation for Scientific Research, Grant No. 1808.

provided

$$\lim_{n\to\infty} |H_n| = +\infty.$$

In the same paper a heuristic method called the *independence principle* is introduced which says that in some cases the limit distribution of certain functionals of the waiting times $T(A)$ can be calculated as if they were independent exponentially distributed random variabled with common expectation d^n. The mathematical background of this observation is provided by [**6**], where large deviation type estimations are derived for the joint distribution of the waiting times. However, the exact scope of the independence principle remains to be determined.

The aim of the present paper is to continue the investigation of the maximum waiting times (1). First, the rate of convergence in the limit relation (2) is estimated, which makes it possible to extend (2) to large deviations. These improved results are then used for describing the a. s. behaviour of maximum waiting times and related quantities. Results are formulated in Section 2, while proofs are contained in Sections 3 and 4.

2. Results. As in the Introduction, let $H_n \subset H^n$ and $W(H_n) = \max\{T(A): A \in H_n\}$. Denote by $Y(H_n)$ the normalized maximum waiting time, i.e.

$$Y(H_n) = d^{-n}W(H_n) - \log|H_n|.$$

Finally, let $F(y) = e^{-e^{-y}}$.

THEOREM 1. *Suppose* $\lim_{n\to\infty} |H_n| = +\infty$. *Then*

$$\lim_{n\to\infty} |H_n|^{\frac{1}{6}\frac{d-1}{d+1}} \sup_y |P\big(Y(H_n) \le y\big) - F(y)| = 0, \tag{3}$$

and this holds uniformly in $|H_n|$.

As will be seen from the proof, the exponent of $|H_n|$ could increased in (3); with considerable additional effort our method of proof will yield the best possible exponent. Our result is sufficient for the case of most interest, namely, when $|H_n|$ grows exponentially, for Theorem 1 then implies an exponential rate of convergence.

As an immediate consequence of Theorem 1 we can extend (2) to large deviations.

COROLLARY 1. *Suppose* $|H_n| \to \infty$ *and let* y_n *be such that*

$$F(y_n) \ge |H_n|^{-\frac{1}{6}\frac{d-1}{d+1}} \text{ and}$$
$$1 - F(y_n) \ge |H_n|^{-\frac{1}{6}\frac{d-1}{d+1}}.$$

Then

$$\frac{P\big(Y(H_n) \leq y\big)}{F(y_n)} \longrightarrow 1 \text{ and}$$

$$\frac{P\big(Y(H_n) > y\big)}{1 - F(y_n)} \longrightarrow 1$$

as $n \longrightarrow \infty$.

The following theorem deals with the a. s. behaviour of the sequence $Y(H_n)$, imposing a slight condition on the growth of $|H_n|$.

THEOREM 2. *Suppose*

$$\liminf_{n\to\infty} \log|H_n| / \log n > 6\frac{d+1}{d-1}. \tag{4}$$

Then

$$\liminf_{n\to\infty} Y(H_n)/\log\log n = -1$$

and

$$\limsup_{n\to\infty} Y(H_n)/\log n = 1$$

with probability 1.

Condition (4) is rather weak compared to the exponential bound $|H_n| \leq d^n$; it is, however, far from being necessary. The assertion itself can also be strengthened. For instance, concerning the lower bound we in fact prove a little more than stated, namely

$$\liminf_{n\to\infty} Y(H_n) + \log\log n = 0 \quad \text{w. p. 1}$$

(see also the remark after the proof in Section 4).

In the remainder of this section we confine our attention to the case $W_n = W(H_n)$.

Let $M_k = \max\{n\colon W_n \leq k\}$, then $M_k \geq n$ iff $W_n \leq k$.

COROLLARY 2. *Let* $\varepsilon > 0$ *arbitrary. Then the event that*

$$\left[\frac{1}{\log d}\left(\log k - \log\log k - \varepsilon\frac{\log\log k}{\log k}\right)\right] \leq M_k \leq \left[\frac{1}{\log d}\left(\log k - \log\log k + (1+\varepsilon)\frac{\log\log k}{\log k}\right)\right]$$

holds for every large enough k, is of probability 1.

The two sides of this inequality coincide in most cases; if not, then they are neighbouring integers. Corollary 2 asserts that the sequence M_k is asymptotically quasi-deterministic (AQD) in the sense of [**12**]. Clearly, M_k is not AD; otherwise so would be W_n.

It is well-known that pure runs need the longest time to occur: let $A \in \operatorname{diag} H^n$ and $B \neq A$, then $E(T(B)) \leq E(T(A))$ and also $P(T(B)) < T(A)) \geq 1/2$. This makes it likely that pure runs are more often the last to occur than any other words: $P(W_n = T(A)) \geq P(W_n = T(B))$. An interesting open problem is to determine the probabilities $p_A = P(W_n = T(A))$, $A \in H^n$, i.e. the distribution of the last word to appear.

Comparing the maximum waiting time with the waiting time for a pure run, we can say the former is less random. Let

$$V_n = \min\{T(A): A \in \operatorname{diag} H^n\},$$

the waiting time for a pure run of length n, and let

$$N_k = \max\{n: V_n \leq k\}$$

the length of the longest pure run observed during the first k experiments. Though $\lim M_k/N_k = 1$ a.s., it is well known that the sequence N_k is not AQD (cf. [**2**]):

$$\left[\frac{1}{\log d}(\log k - \log\log\log k) - 2 - \varepsilon\right] \leq N_k \leq \left[\frac{1}{\log d}(\log k + \log\log k + (1+\varepsilon)\log\log\log k)\right]$$

for large enough k, with probability 1. This is because the stopping times W_n are essentially independent and of small dispersion around their mean, while V_n's are not, their joint distribution is approximately of Marshall-Olkin type. (Let us normalize the stopping times $V_n, V_{n+1}, \ldots, V_{n+k}$ by $E(V_n) \sim \frac{d^n}{d-1}$, then the asymptotic distribution of $\{(d-1)d^{-n}V_{n+i},\ 0 \leq i \leq k\}$ is equal to the distribution of $\{\min(\xi_i, \ldots, \xi_k),\ 0 \leq i \leq k$, where $\xi_0, \xi_1, \ldots, \xi_{k-1}, \xi_k$ are independent, exponentially distributed random variables with expectation $\frac{d}{d-1}, \frac{d^2}{d-1}, \ldots, \frac{d^k}{d-1}, d^k$, resp.)

3. Proof of Theorem 1. The proof follows the line of reasoning worked out in [**9**] with suitable modifications and two-sided bounds instead of asymptotic relations. The successive steps of the proof will be formulated in a sequence of lemmas, numbered correspondingly to [**9**]. Lemmas taken over without alteration will not be proved here. In order to reduce duplication, explanations will be kept to the minimum necessary for lucidity.

First we need some preliminary notions and results. Our main tool is the following estimation.

LEMMA 1[**6**]. *Let* $A_1, A_2, \dots, A_r \in H^n$, $Z = \min T(A_i)$ *and* $b = 2nd^{-n}$. *Then for every* $y > 0$

$$\exp\bigl(-(1+rb)y\bigr) \le P\bigl(Z > E(Z)y\bigr) \le \exp(rb - y),$$

provided that $rb < 1/5$.

The expectation $E(Z)$ needed for the application of Lemma 1 can be calculated on the basis of [**5**]. In order to do that let us introduce a certain measure of overlapping between two words. The *leading number* of $A = (a_1 a_2 \dots a_n)$ over $B = (b_1 b_2 \dots b_n)$ is defined as $A * B = \sum_{i=1}^{n} \varepsilon_i d^i$ where $\varepsilon_i = 1$ if the words $(a_{n-i+1} \dots a_n)$ and $(b_1 b_2 \dots b_i)$ are identical, and $\varepsilon_i = 0$ otherwise.

LEMMA 2 [**5**]. *Let* $A_1, A_2, \dots, A_r$ *and* Z *be as above and let* $p_i = P(T(A_i) = Z)$. *Then* $p_1, p_2, \dots, p_r$ *and* $E(Z)$ *are the solution of the following system of linear equations:*

$$\sum_{i=1}^{r} p_i A_i * A_j = E(Z), \quad j = 1, 2, \dots, r,$$

$$\sum_{i=1}^{r} p_i = 1.$$

In particular, $E\bigl(T(A)\bigr) = A * A$.

In the sequel denote $\log |H_n|$ by m and suppose that m is large enough, say $m > 1000 \log d$. Let $q = q(d) = \min\bigl(\frac{1}{6}, \frac{1}{5}\frac{d-1}{d+1}\bigr)$ and let y be bounded by the relations

$$|H_n|^q \le F(y) \text{ and } |H_n|^{-q} \le 1 - F(y).$$

Thus $e^{-y} \le qm$ and $y \le qm$.

Let the words of H_n be numbered as A_i, $1 \le i \le |H_n|$, and let C_i denote the event $\{T(A_i) > d^n(m+y)\}$, $1 \le i \le |H_n|$. Then

$$P\bigl(Y(H_n) \le y\bigr) = P\Bigl(\bigcap_{i \le |H_n|} \overline{C}_i\Bigr).$$

For estimating the right-hand side we shall use the graph-sieve of Rényi (see [**3**], Theorem 1.4.2). A word $A \in H_n$ is said to be *bad* if

$$A * A > (1 + e^{-4qm}) d^n,$$

otherwise *good*; further, the ordered pair (A, B), where $A, B \in H_n, A \ne B$, is said to be *bad* if

$$A * B > 2e^{-2qm} d^n$$

(these bounds slightly differ from those of [**9**]).

Let the exceptional set E_n be defined as

$$E_n = \{(i,j):1 \le i,j \le |H_n|, \text{ not all of } A_i, A_j, (A_i,A_j), (A_j,A_i) \text{ are good}\}.$$

Let $S_0^* = 1$ and for $r \ge 1$ let $S_r^* = \sum_r^* P(C_{i_1} \cap \cdots \cap C_{i_r})$, where $\sum_r^*$ denotes that the summation runs over all r-tuples of indices $1 \le i_1 \le \cdots \le i_r \le |H_n|$, which do not contain any pair from E_n. Further, let $S_r^{**} = \sum_r^{**} P(C_{i_1} \cap \cdots \cap C_{i_r})$, where $\sum_r^{**}$ indicates summation over r-tuples containing *exactly one* pair from E_n. Note that the event $C_{i_1} \cap \cdots \cap C_{i_r}$ means

$$\min\{T(A_{i_1}), \ldots, T(A_{i_r})\} > d^n(m+y),$$

which is just what Lemma 1 is about.

Finally, let $t = [m]$ where [.] stands for integer part. Then the graph-sieve formula implies

$$\left| P\Big(\bigcap_{i \le |H_n|} \bar{C}_i\Big) - \sum_{r=0}^{t-1} (-1)^r S_r^* \right| \le S_t^* + \sum_{r=2}^{t} S_r^{**}. \tag{5}$$

Let us estimate the terms S_r^* and S_r^{**} in (5).

LEMMA 3.

(a) The number of bad words in less than $4e^{4qm}$.

(b) For any given $A \in H_n$ *the number of words* B *for which the pair* (A,B) *(resp.* (B,A)*) is bad, is less than* $2e^{2qm}$.

(c) If both (A,B) *and* (B,A) *are bad, then* A *and* B *are also bad.*

Proof. (a) Suppose the maximum overlapping of A with itself is the length ℓ (apart form the fact that A is identical with itself, which can be interpreted as overlapping of length n), and let k be the minimum of ℓ as A runs over the bad words. Then

$$\begin{aligned} d^n(1+e^{-4qm}) &< A * A \\ &\le d^n + d^k + d^{k-1} + \cdots + d \\ &< d^n + 2d^k, \end{aligned}$$

from which $d^{n-k} < 2e^{4qm}$. The number of words $A \in H^n$ that overlap themselves in length ℓ is $d^{n-\ell}$, thus the number of bad words is not greater than

$$\begin{aligned} d^{n-k} + d^{n-k+1} + \cdots + d &< 2d^{n-k} \\ &< 4e^{4qm}. \end{aligned}$$

(b) Similarly, let k be the minimum of the longest overlapping between A and B (resp. B and A) as B varies in such a way that (A, B) (resp. (B, A)) is bad. Then $2e^{-2qm}d^n < 2d^k$. The number of words that overlap A in length ℓ is $d^{n-\ell}$ again and the proof can be completed in the same way as above.

(c) A has to overlap B and conversely, at least in length k where $2e^{-2qm}d^n < 2d^k$. This implies that A overlaps itself in length not less than $2k - n$, consequently $A * A - d^n \geq d^{2k-n} > e^{-4qm}d^n$.

LEMMA 4. *Let* $|\sum_r^*|$ *and* $|\sum_r^{**}|$ *denote the number of terms of the corresponding sum. Then*

$$(a) \qquad \frac{1}{r!}e^{rm}(1 - 4r^2e^{-m(1-4q)}) \leq |\sum_r^{**}| \leq \frac{1}{r!}e^{rm} \text{ if } r \geq 1,$$

$$(b) \qquad |\sum_r^{**}| \leq \frac{1}{r!}2r^2e^{m(r-1+2q)} \quad \text{if } r > 2$$

Proof. (a) If we disregard the increasing order of indices $i_1, \ldots, i_r$, this will give us a multiplier $r!$. Now the upper bound $|H_n|^r = e^{rm}$ is obvious. For the lower bound choose the words $A_{i_1}, \ldots A_{i_r}$ successively. A_{i_j} should not be chosen (i) from bad words, (ii) from words already chosen, (iii) from words that form a bad pair together with a word already chosen. Thus A_{i_j} is to be chosen from a set of size not less than

$$|H_n| - 4e^{4qm} - (j-1)(1 + 4e^{2qm}) \geq e^m - 4re^{4qm}.$$

Hence

$$\begin{aligned} |\sum_r^*| &\geq \frac{1}{r!}(e^m - 4re^{4qm})^r \\ &\geq \frac{1}{r!}e^{rm}\left(1 - 4r^2e^{-m(1-4q)}\right). \end{aligned}$$

(b) Again, r-tuples in account cannot contain any bad words or else they would contain more than one pairs from E_n. By Lemma 3(b) there are at most $2e^{m(1+2q)}$ bad pairs to choose while for the other $r-2$ indices we have at most $\binom{|H_n|}{n-2} \leq \frac{1}{r!}r^2e^{m(r-2)}$ choices.

The following assertion, which is a simple consequence of Lemma 2, is adopted from [**9**] without modification.

LEMMA 5. *Let* $A_1, \ldots, A_r \in H^n$, $Z = \min T(A_i)$.

(a) Suppose $A_i * A_i \leq (1 + \delta)d^n$ *for* $1 \leq i \leq r$ *and* $A_i * A_j \leq \delta\, d^n$ *for* $i \neq j$. *Then*

$$\frac{1}{r}d^n \leq E(Z) \leq (\frac{1}{r} + \delta)d^n.$$

(b) Suppose the conditions of (a) are met with the only exception $A_k * A_l \le cd^n$, $c > \delta$. *Then*

$$E(Z) \le \frac{1 + r\delta}{r - c} d^n.$$

(c) Suppose the conditions of (a) are met with the only exception $A_k * A_k \le (1 + c)d^n$, $c > \delta$. *Then*

$$E(Z) \le \left[\delta + (r - \frac{c}{1 + c})^{-1}\right] d^n.$$

LEMMA 6. *Suppose* $m > 1000 \log d$. *Then*
(a) $|S_1^* - e^{-y}| \le m^2 e^{-qm}$,
(b) for $2 \le r \le t$, $|S_r^* - \frac{1}{r!} e^{-ry}| \le \frac{1}{r!} e^{-ry} 8m^3 e^{-2qm}$.

Proof. (a) Using Lemma 1 then Lemma 3 we have

$$\begin{aligned}
(5) \qquad S_1^* &= \sum_{A \in H_n} P\big(T(A) > d^n(m + y)\big) \\
&\le \sum_{A \in H_n} \exp\big(b - d^n(m + y)/A * A\big) \\
&\le \sum_{A \text{ good}} \exp\big(b - (m + y)(1 + e^{-4qm})^{-1}\big) + \sum_{A \text{ bad}} \exp\big(b - (m + y)\frac{d - 1}{d}\big) \\
&\le |H_n| \exp\big(b - (m + y)(1 - e^{-4qm})\big) + 4e^{4qm} \exp\big(b - (m + y)\frac{d - 1}{d}\big) \\
&= e^{-y} \exp\big(b + e^{-4qm}(m + y)\big) + 4 \exp\big(-\frac{d - 1}{d} y - m(\frac{d - 1}{d} - 4q)\big)
\end{aligned}$$

In the first term of (5) $m + y \le \frac{7}{6} m$ and $b = 2nd^{-n} < 3me^{-m} < 3me^{-4qm}$. Hence $b + e^{-4qm}(m + y) < 5me^{-4qm} < 1$, from which $\exp\big(b + e^{-4qm}(m + y)\big) \le 1 + 2\big(b + e^{-4qm}(m + y)\big) \le 1 + 10me^{-4qm} < 1 + me^{-qm}$.

In the second term of (5) $e^b < \frac{3}{2}$, $\exp(-\frac{d-1}{d} y) < \frac{1}{6} m$ and $\frac{d-1}{d} - 4q > q$, hence the second term is less than me^{-qm}. Thus

$$\begin{aligned}
S_1^* &\le e^{-y}(1 + me^{-qm}) + me^{-qm} \\
&\le e^{-y} + \frac{1}{6} m^2 e^{-qm} + me^{-qm} \\
&\le e^{-y} + m^3 e^{-qm}.
\end{aligned}$$

On the other hand, by Lemma 1

$$\begin{aligned}
S_1^* &\geq \sum_{A\in H_n} P\big(T(A) > d^n(m+y)\big) \\
&\geq |H_n| \exp\big(-(1+b)d^n(m+y)/A*A\big) \\
&\geq \exp\big(m-(1+b)(m+y)\big) \\
&\geq e^{-y}\exp\big(-b(m+y)\big) \\
&\geq e^{-y}\big(1-b(m+y)\big) \\
&\geq e^{-y}-(\frac{1}{6}m)(3me^{-4qm})(\frac{7}{6}m) \\
&> e^{-y}-m^3e^{-4qm} \\
&> e^{-y}-m^2e^{-qm}.
\end{aligned}$$

(b) Putting $\delta = 2e^{-2qm}$ and combining Lemma 1 with Lemmas 4(a) and 5(a) we have

$$\begin{aligned}
S_r^* &\leq \frac{1}{r!}\exp\big(rm+rb-r(m+y)(1+2re^{-2qm})^{-1}\big) \\
&\leq \frac{1}{r!}\exp\Big(r\big(b+m-(m+y)(1-2re^{-2qm})\big)\Big) \\
&\leq \frac{1}{r!}e^{-ry}\exp\big(rb+2r^2e^{-2qm}(m+y)\big).
\end{aligned}$$

Here $r \leq t \leq m$, $m+y \leq \frac{7}{6}m$, $rb \leq 3m^2e^{-m} < m^3e^{-2qm}$, hence

$$\begin{aligned}
S_r^* &\leq \frac{1}{r!}e^{-ry}\exp(4m^3e^{-2qm}) \\
&\leq \frac{1}{r!}e^{-ry}(1+8m^3e^{-2qm}).
\end{aligned}$$

Further,

$$\begin{aligned}
S_r^* &\geq \frac{1}{r!}e^{-rm}\big(1-4r^2e^{-m(1-4q)}\big)\exp\big(-(1+rb)r(m+y)\big) \\
&= \frac{1}{r!}e^{-ry}\big(1-4r^2e^{-m(1-4q)}\big)\exp(-r^2b(m+y)\big) \\
&\geq \frac{1}{r!}e^{-ry}\big(1-4r^2e^{-m(1-4q)}\big)(1-r^2b(m+y)) \\
&\geq \frac{1}{r!}e^{-ry}\big(1-4r^2e^{-m(1-4q)}-r^2b(m+y)\big) \\
&\geq \frac{1}{r!}e^{-ry}(1-4m^2e^{-2qm}-3m^3e^{-m}\cdot\frac{7}{6}m) \\
&\geq \frac{1}{r!}e^{-ry}(1-8m^3e^{-2qm}).
\end{aligned}$$

LEMMA 7. *Suppose* $m > 1000 \log d$. *Then*
(a) $S_2^{**} \le \frac{1}{2!} e^{-2y} 6m^2 e^{-qm} F(y) + 3m^2 e^{-qm}$,
(b) for $2 < r \le t$, $S_r^{**} \le \frac{1}{r!} e^{-ry} 6m^2 e^{-qm} F(y)$.

Proof. (b) Every r-tuple of indices, corresponding to the summands of S_r^{**}, contains exactly one pair from E_n, say (i,j). Then both A_i, A_j have to be good; furthermore, by Lemma 3(c) (A_i, A_j) and (A_j, A_i) cannot be bad at the same time. Now let us apply Lemmas 1, 4(b) and 5(b) with $\delta = 2e^{-2qm}$. We shall separately treat the summands whose bad pair (A_i, A_j) is of maximal leading number $d^{n-1} + d^{n-2} + \cdots + d$. There are at most d^3 of them, because in these pairs the same letter has to stand at every position but the first in A_i and the last in A_j. For these pairs thus $c = \frac{1}{d-1}$ in Lemma 5(b). All the other pairs have leading number

$$\begin{aligned} A_i * A_j &\le d^{n-1} + d^{n-3} + d^{n-5} + \cdots \\ &< \frac{d}{d^2 - 1} d^n \\ &\le \frac{2}{d+1} d^n, \end{aligned}$$

which gives $c = \frac{2}{d+1}$ in Lemma 5(b).

Hence

$$\begin{aligned} S_r^{**} \le d^3 \frac{1}{(r-2)!} |H_n|^{r-2} \exp\Big(& rb - (m+y)(r - \frac{1}{d-1})(1 + 2re^{-2qm})^{-1} \Big) \\ & + \frac{1}{r!} 2r^2 \exp\Big(m(r - 1 + 2q) + rb - (m+y) \\ & \qquad (r - \frac{2}{d+1})(1 + 2re^{-2qm})^{-1} \Big). \end{aligned}$$

Let us examine the exponents. Let c denote either $\frac{1}{d-1}$ or $\frac{2}{d+1}$. Then

$$\begin{aligned} -(m+y)(r-c)&(1 + 2re^{-2qm})^{-1} \\ &< -(m+y)(r-c)(1 - 2re^{-2qm}) \\ &< (m+y)r + mc - e^{-y} + (cy + e^{-y}) + \frac{7}{6} mre^{-2qm}. \end{aligned}$$

Since the function $y \longmapsto cy + e^{-y}$ is decreasing in the given domain of y,

$$\begin{aligned} cy + e^{-y} &\le qm - c \log(qm) \\ &< qm \end{aligned}$$

holds there. Now $\exp(rb + \frac{7}{6}mre^{-2qm})$ is majorized by 2, hence

$$S_r^{**} \le \frac{1}{r!}e^{-ry}2d^3m^2F(y)\exp\Big(-m(2 - \frac{1}{d-1} - q)\Big)$$
$$+ \frac{1}{r!}e^{-ry}4m^2F(y)\exp\Big(-m(\frac{d-1}{d+1} - 3q)\Big)$$
$$\le \frac{1}{r!}e^{-ry}6m^2e^{-qm}F(y).$$

(a) The sum Σ_2^{**} will be divided into four parts:

(i) summands, where both words are good,

(ii) summands, where one of the words is good, one is bad and the pair they form is good whatever be their order,

(iii) summands, where one of the words is good, one is bad and the pair is also bad in one order,

(iv) summands, where both words are bad.

Case (i) can be treated in the same way as (b).

In case (ii) Lemma 5(c) with $c = \frac{1}{d-1}$, $\delta = 2e^{-2qm}$ gives

$$E(Z) \le d^n\Big(\frac{d}{2d-1} + 2e^{-2qm}\Big).$$

The number of such summands is less than $4e^{m(1+4q)}$. By Lemma 1 this part of Σ_2^{**} is bounded by

$$4\exp\Big(m(1+4q) + 2b - (m+y)\Big(\frac{d}{2d-1} + 2e^{-2qm}\Big)^{-1}\Big)$$
$$\le 4\exp\Big(m(1+4q) + 2b - (m+y)\frac{2d-1}{d}\Big(1 - \frac{2d-1}{d}2e^{-2qm}\Big)\Big)$$
$$\le 4\exp\Big(2b - m\Big(\frac{2d-1}{d} - 1 - 4q\Big) - y\frac{2d-1}{d} + (m+y)8e^{-2qm}\Big).$$

Here $\exp\Big(-\frac{2d-1}{d}y\Big) \le (qm)^2$ and $\exp(2b + (m+y)8e^{-2qm}) < 2$, hence for part (ii) we obtain that it is $\le m^2\exp\Big(-m(\frac{d-1}{d} - 4q)\Big) \le m^2e^{-qm}$.

In cases (iii) and (iv) Lemma 5(a) with $\delta = \frac{1}{d-1}$ gives

$$E(Z) \le \frac{1}{2}\frac{d+1}{d-1}d^n.$$

The number of such terms is bounded by $(4e^{4qm})^2$ and $2 \cdot 4e^{4qm}2e^{2qm}$, resp., thus their constribution is less than

$$3(4e^{4qm})^2\exp\Big(2b - 2(m+y)\frac{d-1}{d+1}\Big)$$
$$= 3\Big[4\exp\Big(b - y\frac{d-1}{d+1} - m\Big(\frac{d-1}{d+1} - 4q\Big)\Big)\Big]^2$$
$$< 2m^2e^{-qm}.$$

Our last lemma completes the proof of Theorem 1.

LEMMA 8. *Suppose* $m > 1000 \log d$. *Then*

$$\sup_y |P\big(Y(H_n) \le y\big) - F(y)| \le 10m^3 e^{-qm}.$$

Proof. First, let y belong to the domain, where the estimations of Lemma 6 and 7 are valid: $y_1 \le y \le y_2$, where $F(y_1) = e^{-qm} = 1 - F(y_2)$. Then

$$\begin{aligned}(6)\qquad |P\big(Y(H_n) \le y\big) - F(y)| &\le |P\big(Y(H_n) \le y\big) - \sum_{r=0}^{t-1} (-1)^r S_r^*| \\ &+ |\sum_{r=0}^{t-1} (-1)^r S_r^* - \sum_{r=0}^{t-1} (-1)^r \frac{1}{r!} e^{-ry}| \\ &+ |\sum_{r=0}^{t-1} (-1)^r \frac{1}{r!} e^{-ry} - F(y)|.\end{aligned}$$

By (5), the first term is bounded by

$$(7)\qquad S_t^* + \sum_{r=2}^{t} S_r^{**} \le |S_t^* - \frac{1}{t!} e^{-ty}| + \sum_{r=2}^{t} S_r^{**} + \frac{1}{t!} e^{-ty}.$$

Applying Lemma 7 to the last sum we obtain

$$\begin{aligned}\sum_{r=2}^{t} S_r^{**} &\le 6m^2 e^{-qm} F(y) \sum_{r=2}^{t} \frac{1}{r!} e^{-ry} + 3m^2 e^{-qm} \\ &\le 6m^2 e^{-qm} F(y) e^{-e^{-y}} + 3m^2 e^{-qm} \\ &\le 9m^2 e^{-qm}.\end{aligned}$$

The second term in the right-hand side of (6) and the first term of (7) are estimated by Lemma 6: They together are not greater than

$$\begin{aligned}\sum_{r=1}^{t} |S_r^* - \frac{1}{r!} e^{-ry}| &\le m^2 e^{-qm} + 8m^3 e^{-2qm} \sum_{r=2}^{t} \frac{1}{r!} e^{-ry} \\ &\le m^2 e^{-qm} + 8m^3 e^{-2qm} e^{e^{-y}} \\ &\le (m^2 + 8m^3) e^{-qm}.\end{aligned}$$

The last term of (6) is well-known to be less than $\frac{1}{t!} e^{-ty}$. Hence

$$|P\big(Y(H_n) \le y\big) - F(y)| \le \frac{2}{t!} e^{-ty} + (10m^2 + 8m^3) e^{-qm}.$$

Here

$$\begin{aligned}\frac{2}{t!}e^{-ty} &< 2\frac{e^m}{m^m}e^{-my}\\ &< 2(qe)^m\\ &< m^2e^{-qm},\end{aligned}$$

hence

$$\sup_{y_1\le y\le y_2}\left|P\big(Y(H_n)\le y\big)-F(y)\right| \le 9m^3e^{-qm}.$$

For $y < y_1$,

$$\begin{aligned}\left|P\big(Y(H_n)\le y\big)-F(y)\right| &\le \max\{P\big(Y(H_n)\le y\big),F(y)\}\\ &\le \max\{P\big(Y(H_n)\le y_1\big),F(y_1)\}\\ &\le \left|P\big(Y(H_n)\le y_1\big)-F(y_1)\right|+F(y_1)\\ &\le 10m^3e^{-qm},\end{aligned}$$

and similarly, for $y > y_2$

$$\begin{aligned}\left|P\big(Y(H_n)\le y\big)-F(y)\right| &= \left|P\big(Y(H_n)> y\big)-(1-F(y))\right|\\ &\le \max\{P(Y\big(H_n)> y\big),1-F(y)\}\\ &\le \max\{P\big(Y(H_n)> y_2\big),1-F(y_2)\}\\ &\le \left|P\big(Y(H_n)> y_2\big)-\big(1-F(y_2)\big)\right|+1-F(y_2)\\ &\le 10m^3e^{-qm}.\end{aligned}$$

4. Proof of Theorem 2 and Corollary 2. By Corollary 1, for sufficiently small positive ε

$$\begin{aligned}P\big(Y(H_n)\le -\log\log n-\varepsilon\big) &\sim F(-\log\log n-\varepsilon)\\ &< \frac{1}{n^{1+\varepsilon}}.\end{aligned}$$

(Condition (4) assures the applicability of Corollary 1.) The sum of these probabilities is convergent, hence the Borel-Cantelli lemma is applicable. Since ε can be arbitrarily small,

$$\liminf_{n\to\infty} Y(H_n)/\log\log n \ge -1.$$

Similarly,

$$P\big(Y(H_n) > (1+\varepsilon)\log n\big) \sim 1 - f\big((1+\varepsilon)\log n\big) \sim \frac{1}{n^{1+\varepsilon}}.$$

Again, the Borel-Cantelli lemma gives

$$\limsup_{n\to\infty} Y(H_n)/\log n \le 1.$$

Now let $a(1) = 1$ and for $n \ge 2$ $a(n) = [3n\log n]$. With the notation $b(n) = d^{a(n)}\big(\log|H_{a(n)}| + \log a(n)\big)$, let us define the set E_n as

$$E_n = \{A \in H_{a(n)} : T(A) > b(n-1)\}, \quad n \ge 2.$$

First we show that

$$|E_n| = |H_{a(n)}|\Big(1 - o\Big(\frac{1}{\log a(n)}\Big)\Big) \quad \text{a. s.} \tag{8}$$

In order to do that let us compute the expectation $E(|E_n|)$. By Lemma 1

$$\begin{aligned} E(|E_n|) &= E\sum_{A\in H_{a(n)}} I\big(T(A) > b(n-1)\big) \\ &= \sum_{A\in H_{a(n)}} P\big(T(A) > b(n-1)\big) \\ &\ge \sum_{A\in H_{a(n)}} \exp\big(-2b(n-1)/E(T(A))\big) \\ &\ge |H_{a(n)}|\exp\big(-2b(n-1)d^{-a(n)}\big). \end{aligned}$$

Now we apply the Markov inequality to the random variable $|H_{a(n)}| - |E_n|$ to obtain

$$\begin{aligned} P\big(|H_{a(n)}| - |E_n| > |H_{a(n)}|/\log^2 a(n)\big) &\\ &\le \log^2 a(n)\Big(1 - \exp\big(-2b(n-1)d^{-a(n)}\big)\Big) \\ &= 2\log^2 a(n) b(n-1) d^{-a(n)}\big(1 + o(1)\big) \\ &\le (6\log d) n(\log^3 n) d^{a(n-1)-a(n)}(1 + o(1)). \end{aligned}$$

Here $a(n) - a(n-1) = 3\log n + 3 + o(1)$; thus the above estimation is asymptotically equal to

$$(6\log d) d^{-3} n^{1-3\log d}\log^3 n.$$

Since $1 - 3\log d < -1$, the sum of these probabilities is finite and the Borel-Cantelli lemma leads us to (8).

Let $\mathcal{F}_n$ denote the σ-field generated by the first $b(n)$ experiments, further, let the events C_n and D_n be defined as

$$\begin{aligned} C_n &= \{ Y(H_{a(n)}) \leq -\log\log a(n) \} \\ &= \{ W(H_{a(n)}) \leq d^{a(n)}\big(\log|H_{a(n)}| - \log\log a(n)\big) \}, \\ D_n &= \{ Y(H_{a(n)}) > \log a(n) \} \\ &= \{ W(H_{a(n)}) > b(n) \}. \end{aligned}$$

Then C_n and D_n belong to $\mathcal{F}_n$.

In order to estimate $P(C_n|\,\mathcal{F}_{n-1})$ from below let us replace the observer after $b(n-1)$ experiments, then the new observer has only to take care of words belonging to E_n. In that way we might make an error by not observing the occurence of runs from E_n during the first $a(n) - 1$ experiments after $b(n-1)$. In this case we wait longer than needed. Waiting times concerning the second observer are distinguished by $'$. Then

$$\begin{aligned} &P(C_n|\,\mathcal{F}_{n-1}) \\ &\quad \geq P\Big(W'(E_n) \leq d^{a(n)}\big(\log|H_{a(n)}| - \log\log a(n)\big) - b(n-1)|\,E_n \Big) \\ &\quad = P\Big(Y'(E_n) \leq \log|H_{a(n)}| - \log|E_n| \\ &\qquad\qquad - \log\log a(n) - d^{-a(n)}b(n-1)|\,E_n \Big). \end{aligned}$$

By Corollary 1 this latter is asymptotically

$$\sim F\Big(-\log\log a(n) + o\Big(\frac{1}{\log a(n)}\Big)\Big) = e^{\log a(n)+o(1)} \sim \frac{1}{a(n)}$$

with probability 1. The sum of these conditional probabilities diverges a. s., hence by the Lévy generalization of the Borel-Cantelli lemma (see [**11**], Corollary VII. 2. 6), infinitely many of the events C_n occur a. s., thus

$$\liminf_{n\to\infty} Y(H_n)/\log\log n \leq -1.$$

On the other hand, in order to estimate $P(D_n|\,\mathcal{F}_{n-1})$ from below let us estimate the conditional probability of the error we make when replacing the observer after $b(n-1)$ experiments. Obviously,

$$\begin{aligned} P(\text{error}\,|\,\mathcal{F}_{n-1}) = P\big(&|T(A)| < b(n-1) + a(n) \text{ and} \\ &T'(A) > b(n) - b(n-1) \text{ for some } A \in E_n|\,\mathcal{F}_{n-1}\big). \end{aligned}$$

This will be majorized if the new observer gets active just after $b(n-1)+a(n)$ experiments. Denote this new waiting time by $T''(A)$; it is independent of $\mathcal{F}_{n-1}$ and of the event $\{T(A) < b(n-1)+a(n)\}$. Hence

$$\begin{aligned} P(\text{error} \mid \mathcal{F}_{n-1}) &\le \sum_{A\in E_n} P\big(T(A) < b(n-1)+a(n), \\ &\qquad T''(A) > b(n) - b(n-1) - a(n) \mid \mathcal{F}_{n-1}\big) \\ &= \sum_{A\in E_n} \Big(P\big(T''(A) > b(n) - b(n-1) - a(n)\big) \\ &\qquad P\big(T(A)| < b(n-1)+a(n) \mid \mathcal{F}_{n-1}\big)\Big). \end{aligned}$$

Since $E\big(T''(A)\big) < 2d^{a(n)}$, Lemma 1 gives for the first term

$$\begin{aligned} &P\big(T''(A) > b(n) - b(n-1) - a(n)\big) \\ &\qquad \le 2\exp\Big(-\big(b(n) - b(n-1) - a(n)\big)/E\big(T''(A)\big)\Big) \\ &\qquad \le 2\exp\Big(-\frac{1}{2}d^{-a(n)}\big(b(n) - b(n-1) - a(n)\big)\Big) \\ &\qquad = 2\exp\Big(-\frac{1}{2}\big(\log|H_{a(n)}| + \log a(n)\big) + o(1)\Big) \\ &\qquad \sim 2\Big(a(n)|H_{a(n)}|\Big)^{-1/2}, \end{aligned}$$

while

$$\begin{aligned} &\sum_{A\in E_n} P\big(T(A) < b(n-1)+a(n) \mid \mathcal{F}_{n-1}\big) \\ &\qquad \le \sum_{A\in E_n} \sum_{k=1}^{a(n)-1} P\big(T(A) = b(n-1)+k \mid \mathcal{F}_{n-1}\big) \\ &\qquad = \sum_{k=1}^{a(n)-1} \sum_{A\in E_n} d^{-k} I(A \text{ is compatible with } \mathcal{F}_{n-1}) \\ &\qquad \le \sum_{k=1}^{a(n)-1} d^{-k} \min\{d^k, |E_n|\} \\ &\qquad = O(\log|E_n|) \\ &\qquad = O(\log|H_{a(n)}|). \end{aligned}$$

Under "A is compatible with $\mathcal{F}_{n-1}$" we mean that the last $a(n)-k$ experiments before $b(n-1)$ should result just in the first $a(n)-k$ letters of A. The number of such $A \in E_n$ is at most $\min\{d^k, |E_n|\}$. After all,

$$P(\text{error} \mid \mathcal{F}_{n-1}) = o\Big(\frac{1}{a(n)}\Big).$$

Hence

$$P(D_n|\mathcal{F}_{n-1}) \geq P\Big(W'(E_n) > b(n) - b(n-1)|E_n\Big) - o\Big(\frac{1}{a(n)}\Big)$$
$$= P\Big(Y'(E_n) > \log a(n) + o(1)|E_n\Big) - o\Big(\frac{1}{a(n)}\Big).$$

The principal term is $\sim 1 - F\big(\log a(n) + o(1)\big) \sim \frac{1}{a(n)}$ by Corollary 1. Thus $\sum P\big(D_n|\mathcal{F}_{n-1}\big) = +\infty$ a. s., which implies that infinitely many D_n occur simultaneously, with probability 1. Hence

$$\lim\sup_{n\to\infty} Y(H_n)/\log n \geq 1.$$

This completes the proof of Theorem 2.

Remark. Actually, the proof utilized something equivalent to the approximate independence of the variables $Y(H_{a(n)})$. As a matter of fact, the variables $Y(H_n)$ themselves are asymptotically independent, which makes it possible to determine the LL, LU, UL, UU classes for them.

We now give the proof of Corollary 2.

For the sake of brevity let us denote the stated bounds of M_k by $[a]$ and $[b]$, resp. Then $[a] \leq M_k \leq [b]$ is equivalent to $W_{[a]} \leq k < W_{[b]+1}$.

By Theorem 2, for sufficiently large a, i. e., for sufficiently large k

$$W_{[a]} < d^{[a]}\big(\log d^{[a]} + (1+\varepsilon/2)\log[a]\big)$$
$$\leq d^{a}\big(\log d^{a} + (1+\varepsilon/2)\log a\big)$$
$$= \frac{k}{\log k}\exp(-\varepsilon\frac{\log\log k}{\log k})$$
$$\big(\log k - \log\log k + (1+\varepsilon/2)\log\log k + O(1)\big)$$
$$< k\exp\Big(-\varepsilon\frac{\log\log k}{\log k}\Big)\Big(1+\varepsilon\frac{\log\log k}{\log k}\Big)$$
$$< k.$$

Further,

$$W_{[b]+1} > d^{[b]+1}\big(\log d^{[b]+1} - \log\log([b]+1) - \varepsilon\big)$$
$$\geq d^{b}\big(\log d^{b} - \log\log b - \varepsilon\big)$$
$$\geq \frac{k}{\log k}\exp\Big((1+\varepsilon)\frac{\log\log k}{\log k}\Big)\Big(\log k - (1+\frac{\varepsilon}{2})\log\log k\Big)$$
$$> k.$$

The proof is complete.

REFERENCES

1. A. Benczúr, *On the expected time to the first occurence of every k bit long pattern in the symmetric Bernoulli process*, Acta Math., Hungar. **47** (1986) 233–238.

2. P. Erdős, and P. Révész, *On the length of the longest head run.* In: *Topics in Information Theory*, (Keszthely, Hungary, 1975), Colloq. Math. Soc. J. Bolyai, **16** (1977) 219–228.

3. J. Galambos, *The Asymptotic Theory of Extreme Order Statistics*, (1978) Wiley, New York.

4. L. J. Guibas, and A. M. Odlyzko, *Long repetitive patterns in random sequences*, Z. Wahrsch. verw. Geb., **53** (1980) 241–262.

5. S. R. Li, *A martingale approach to the study of occurence of sequence patterns in repeated experiments*, Ann. Prob. **8** (1980) 1171–1176.

6. T. F. Móri, *Large deviation results for waiting times in repeated experiments*, Acta Math. Hungar., **45** (1985) 213–221.

7. T. F. Móri, *On the expectation of the maximum waiting time*, Ann. Univ. Sci. Bud. R. Eötvös Nom., Sect. Comput., **7** (1987) 111–115.

8. T. F. Móri, *Asymptotic joint distribution of waiting times in repeated experiments*, Lecture read at the 4th EYSM, Varna, Bulgaria, (1985).

9. T. F. Móri, *On the waiting time till each of some given patterns occurs as a run*, Prob. Th. Rel. Fields, to appear.

10. T. F. Móri, and G. J. Székely, *Asymptotic independence of 'pure head' stopping times*, Statist. Prob. Letters, **2** (1984) 5–8.

11. J. Neveu, *Discrete Parameter Martingales*, North-Holland, Amsterdam, (1975).

12. P. Révész, *Strong theorems on coin tossing*, In: Proc. 1978 Internat. Congress of Mathematicians, Helsinki, (1980) 749–754.

Eötvös University
Budapest, Múzeum krt. 6–8
H-1088 Hungary