Theoretical Informatics and Applications
Theoret. Informatics Appl. **35** (2001) 187–206

ON THE NUMBER OF ITERATIONS REQUIRED BY VON NEUMANN ADDITION

Rudolf Grübel[1] and Anke Reimers[1]

Abstract. We investigate the number of iterations needed by an addition algorithm due to Burks *et al.* if the input is random. Several authors have obtained results on the average case behaviour, mainly using analytic techniques based on generating functions. Here we take a more probabilistic view which leads to a limit theorem for the distribution of the random number of steps required by the algorithm and also helps to explain the limiting logarithmic periodicity as a simple discretization phenomenon.

Mathematics Subject Classification. 68Q25, 65Y20.

1. Introduction and results

Some fifty years ago Burks *et al.* [2] introduced an addition algorithm that since then has become one of standard topics in computer science curricula; see *e.g.* Scott [19] or Wegener [21]. It is also used in practice for multiprecision arithmetic; see *e.g.* Forster [7]. This algorithm proceeds iteratively, with the number of steps and hence the running time depending on the input. In the worst case this number is of the same order as the length n of the input, but on average only about $\log_b n$ steps are needed with a base b representation. The average case behaviour has already been discussed in Burks *et al.* [2], with complements and refinements obtained by Claus [4] and Knuth [12]. Together with Quicksort and a related selection algorithm, also due to Hoare and known as Find or Quickselect, von Neumann addition is one of the classical examples for an algorithm where the worst case behaviour is no better than that of some standard naive procedure. Therefore, the *raison d'être* for these methods, at least

Keywords and phrases: Carry propagation, limit distributions, total variation distance, logarithmic periodicity, Gumbel distributions, discretization, large deviations.

[1] Institut für Mathematische Stochastik, Universität Hannover, Postfach 60 09, 30060 Hannover, Germany; e-mail: rgrubel@stochastik.uni-hannover.de

from a practical point of view, is the improvement obtained "on average" or "for typical input". We mention in passing that there is, similar to the situation with Hoare's algorithms, a deterministic algorithm with worst case behaviour comparable to the average case behaviour of von Neumann addition but requiring more complicated data structures; see Cormen *et al.* [5] or Wegener [21].

In recent years some of the early average case analyses have received renewed interest, with the focus on the complete distribution of the running time rather than the associated expected value. Let Y_n denote the number of comparisons needed by QUICKSORT. Using martingale methods Regnier [15] showed that $(Y_n - n\log n)/n$ converges in distribution as $n \to \infty$, Rösler [17] obtained the same result with a completely different and somewhat more constructive method. For FIND see Grübel and Rösler [10] and Grübel [8,9], where the convergence in distribution of Z_n/n was established for the number Z_n of comparisons needed to find a specific quantile such as the median. It is an immediate consequence of these results that there is a *concentration of mass* phenomenon in the first, but not in the second case: Y_n/EY_n converges to a fixed value as $n \to \infty$, Z_n/EZ_n does not. Also, such results can be used to assess the probability of excessively long running times *via* the quantiles of the limit distribution. The bounds for these quantiles that come out of the results for expected values *via* Chebychev's inequality can be very poor; see *e.g.* the discussion in Grübel [8]. Chassaing *et al.* [3] discuss related optimality concepts; concentration of mass in connection with stochastic algorithms is also discussed in McDiarmid and Hayward [14] and McDiarmid [13].

Let X_n be the number of iterations required by von Neumann addition if the input consists of two independent sequences of length n of independent and uniformly distributed base-b digits (for a description of the algorithm see the beginning of the next section). For $b = 2$ Burks *et al.* [2] obtained the upper bound $EX_n \le \log_2 n + 2$ and Claus [4] showed that $EX_n \ge \lceil \log_2 n \rceil - 1$. The general remarks in the previous paragraph would lead us to expect a result on the convergence in distribution of $(X_n - a_n)/b_n$ to some non-trivial limit law as $n \to \infty$, with suitable sequences $(a_n)_{n\in\mathbb{N}}$ and $(b_n)_{n\in\mathbb{N}}$ of real numbers. In view of the average case results the shift sequence $(a_n)_{n\in\mathbb{N}}$ should be of the order $\log_b n$ and indeed, Claus [4] conjectured on the basis of numerical evidence that $EX_n - \log_2 n$ converges to $1/3$ for $b = 2$ as $n \to \infty$. Knuth [12], however, proved that

$$EX_n \;=\; \log_b n + \frac{\gamma}{\log b} + \frac{1}{2} + \log_b \frac{b-1}{2} - f(n) \;+\; O\left(\frac{(\log n)^4}{n}\right)$$

as $n \to \infty$, where γ is Euler's constant and

$$f(x) \;=\; \frac{2}{\log b} \sum_{k=1}^{\infty} \Re\left(\Gamma\left(\frac{-2\pi i k}{\log b}\right) \exp\left(2\pi i k \log_b \frac{(b-1)x}{2}\right)\right).$$

(The versions of the algorithm considered by Claus [4] and Knuth [12] differ slightly; see Sect. 3.1 below.) We have $f(bx) = f(x)$ for all $x > 0$, but a plot

shows that f is not constant (for small b, f is "almost" constant). This small-fluctuations phenomenon appears in connection with various algorithms such as radix exchange sorting and tries; see Chapter 5.2.2 and Chapter 6.3 in Knuth [11], Chapter 7.8 in Sedgewick and Flajolet [20] and the references given there.

In accordance with the classical approach to such problems both Claus and Knuth based their analysis on a recursion relation for the values $P(X_n \geq i)$, $i = 1, \ldots, n$, $n \in \mathbb{N}$; the use of generating functions then provides the bridge to the impressive array of techniques of asymptotic analysis (Knuth dedicates his paper to de Bruijn). For a review of Mellin transform methods in this context, including an analytic approach to the small-fluctuations phenomenon, see Flajolet *et al.* [6]. Our approach makes more use of probabilistic structures. As a result distributional limits can be obtained (which in turn shows that there is concentration of mass about the mean for large n), the above log-periodicity is seen to arise as a discretization phenomenon and a simple 'non-computational' proof can be given that f is not constant.

We now state our main results, proofs are collected in the next section. Asymptotics for distributions require a suitable notion of distance for probability measures: we work with the total variation distance which for distributions P, Q on the Borel subsets $\mathcal{B}$ of the real line is defined by

$$d_{\mathrm{TV}}(P,Q) \;:=\; \sup_{A\in\mathcal{B}} \big|P(A) - Q(A)\big|.$$

We write $\mathcal{L}(Z)$ for the distribution of the random variable Z and use $d_{\mathrm{TV}}(X,Y)$ as an abbreviation for $d_{\mathrm{TV}}(\mathcal{L}(X),\mathcal{L}(Y))$. If P and Q are concentrated on the set $\mathbb{Z}$ of integers, then

$$d_{\mathrm{TV}}(P,Q) \;=\; \frac{1}{2}\sum_{k\in\mathbb{Z}} \big|P(\{k\}) - Q(\{k\})\big|.$$

Further, if $P, P_1, P_2, \ldots$ all satisfy this condition, then convergence of the atoms, *i.e.*

$$P_n(\{k\}) \;\to\; P(\{k\}) \quad \text{for all } \; k \in \mathbb{Z},$$

is equivalent to convergence in total variation distance (this is a special case of Scheffé's theorem; see *e.g.* Billingsley [1], p. 218). It turns out that no rescaling is required for the distributional asymptotics of X_n (*i.e.* we may take $b_n = 1$ for all $n \in \mathbb{N}$), so that the distributions of interest are indeed concentrated on $\mathbb{Z}$. However, because of the log-periodicity we do not have one single limit distribution but rather a whole family that depends on the behaviour of $\{\log_b n\}$; here $\{x\}$ denotes the fractional part of x, *i.e.* $\{x\} = x - \lfloor x \rfloor$ in the usual floor and ceiling notation. A random variable Z is said to have the Gumbel distribution with scale parameter λ, which we abbreviate to $Z \sim \mathrm{Gu}(\lambda)$, if $P(Z \leq z) = \exp(-\exp(-\lambda z))$ for all $z \in \mathbb{R}$. For typographical convenience we define the constants

$$\zeta_b := \log_b \frac{b-1}{2}\,, \qquad \chi_b := \frac{b-1}{2b}\cdot$$

Note that χ_b is the probability that the sum of two independent random integers, both uniformly distributed on the set $\{0, 1, \ldots, b-1\}$, exceeds the value $b-1$. In our model for the input of the algorithm, this is the probability that a carry is generated at some fixed position of the input sequences.

Theorem 1. *Let X_n denote the random number of iterations required by von Neumann addition with respect to base b if the input consists of independent and uniformly distributed base b digit sequences of length n. Then, with $Z \sim \mathrm{Gu}(\log b)$,*

$$\lim_{n\to\infty} d_{\mathrm{TV}}\Big(X_n - \lfloor \log_b n \rfloor, \lceil Z + \zeta_b + \{\log_b n\} \rceil\Big) \;=\; 0.$$

This shows that the distribution of the integer-valued random variable $X_n - \lfloor \log_b n \rfloor$ can be approximated by a shifted and discretized Gumbel distribution, with the shift depending on n only *via* the fractional part of its base-b logarithm. In particular, we obtain approximations for the probability that X_n is equal to or greater than $\lfloor \log_b n \rfloor + k$ for n large and k fixed. Our next result deals with the quality of the Gumbel approximation in the tails of the distribution, *i.e.* with increasing k, where Theorem 1 would be useless. Of special interest are deviations that are of the same order of magnitude as the mean.

Theorem 2. *Let X_n and Z be as in Theorem 1. Then, for all $t > 0$,*

$$\lim_{n\to\infty} n^t \Big| P\big(X_n - \lfloor \log_b n \rfloor \geq t \log_b n\big) \;-\; P\big(\lceil Z + \zeta_b + \{\log_b n\} \rceil \geq t \log_b n\big)\Big| \;=\; 0.$$

The tail behaviour of Z is easily accessible. From $\lim_{z\to\infty} b^z P(Z \geq z) = 1$ it follows that $n^t P\big(\lceil Z + \zeta_b + \{\log_b n\} \rceil \geq t \log_b n\big)$ oscillates between two positive constants, in Section 3.2 this will be discussed in more detail. There are two points worth noting at this stage: first, the tail probabilities of X_n fluctuate too; second, the distributional approximation by a shifted and discretized Gumbel distribution is close enough to capture these fluctuations.

Closeness in total variation norm of two distributions does not imply that the corresponding moments are close. The next result supplements Theorem 1 in this respect.

Theorem 3. *Let X_n and Z be as in Theorem 1. Then, for all $l \in \mathbb{N}$,*

$$\lim_{n\to\infty} \Big(E\big(X_n - \lfloor \log_b n \rfloor\big)^l \;-\; E\big\lceil Z + \zeta_b + \{\log_b n\}\big\rceil^l\Big) \;=\; 0\,.$$

By Theorem 1, for any fixed η in the interval $[0,1)$ we obtain a limit distribution for the subsequence $(X_{n_k} - \lfloor \log_b n_k \rfloor)_{k\in\mathbb{N}}$ if we take $n_k := \lceil b^{k+\eta} \rceil$ for all $k \in \mathbb{N}$. Theorem 3 shows that, for such sequences,

$$\lim_{k\to\infty} \big(EX_{n_k} - \log_b n_k \big) \;=\; E\big\lceil Z + \zeta_b + \eta \big\rceil \;-\; \eta$$

with $Z \sim \mathrm{Gu}(\log b)$. This representation throws some light on the average case behaviour of the algorithm, supplementing the earlier work by Claus [4] and Knuth [12] cited above. In particular, we have the following result, which also shows that the full sequence $\big(\mathcal{L}(X_n - \lfloor \log_b n \rfloor)\big)_{n\in\mathbb{N}}$ does not converge.

Theorem 4. *With $Z \sim \mathrm{Gu}(\log b)$, the function*

$$\eta \;\mapsto\; E\big\lceil Z + \zeta_b + \eta \big\rceil \;-\; \eta$$

is not constant.

The limiting average case behaviour will be further discussed in Section 3.3 below where we also relate our result to Knuth's formula.

2. Proofs

We first introduce some notation. Throughout, b is a fixed integer greater than 1, b is the basis for the representation of the numbers to be added. Let $\mathbb{Z}_b := \{0, 1, \dots, b-1\}$ be endowed with the operations

$$m \oplus n \;:=\; m + n \pmod b, \quad m \wedge n \;:=\; \begin{cases} 1, & \text{if } n+m \geq b, \\ 0, & \text{otherwise.} \end{cases}$$

Further, $\mathbb{Z}_{b,\infty}$ denotes the set of sequences $u = (u_0, u_1, \dots)$ of elements of $\mathbb{Z}_b$. On sequences, $\oplus$ and $\wedge$ operate componentwise. We define the truncation operators $T_n : \mathbb{Z}_{b,\infty} \to \mathbb{Z}_{b,\infty}$, $n \in \mathbb{N}_0$, and the shift operator $S : \mathbb{Z}_{b,\infty} \to \mathbb{Z}_{b,\infty}$ by

$$T_n(u)_k \;:=\; \begin{cases} u_k, & \text{if } k < n, \\ 0, & \text{otherwise,} \end{cases} \qquad S(u)_k \;:=\; \begin{cases} u_{k-1}, & \text{if } k > 0, \\ 0, & \text{otherwise,} \end{cases}$$

for all $u = (u_k)_{k\in\mathbb{N}_0} \in \mathbb{Z}_{b,\infty}$. Sequences u, v with only a finite number of non-zero elements can be regarded as base-b representations of non-negative integer numbers; in contrast to the usual notation we read the digits from left to right in increasing order. The von Neumann addition algorithm provides the sum $u + v$ of such "terminating" sequences *via* the following iterative procedure: we start with $u^{(0)} := u$, $v^{(0)} := v$. Given $u^{(l)}$, $v^{(l)}$ we first check whether $v^{(l)} = 0$. If this is the

case, then the u-part of the pair contains the required base-b representation of the sum and no (further) iterations are needed. If not, we define the next pair by

$$u^{(l+1)} := u^{(l)} \oplus v^{(l)}, \quad v^{(l+1)} := S\big(u^{(l)} \wedge v^{(l)}\big).$$

From $l = 1$ onwards the $v^{(l)}$-parts contain the respective carry bits. These start in $v^{(1)}$ at those positions k where $u_{k-1} + v_{k-1} \geq b$ and propagate to the right in later v-parts as long as $u_{k-1+j} + v_{k-1+j} = b - 1$, $j = 1, 2, \ldots$ The algorithm terminates as soon as all components of the v-part of the pair are equal to 0, which is the case after at most $n + 1$ steps if n is the maximum of the length of u and the length of v (a slight subtlety arises in this connection, see Sect. 3.1 below). The point of the algorithm is, of course, that "on the average", the number of required iterations is much smaller than the effective length of the input.

In order to make such a statement precise we need a stochastic model. We consider two independent random elements U and V of $\mathbb{Z}_{b,\infty}$, both with independent components that are uniformly distributed on the set $\mathbb{Z}_b$ of availabe digits, and denote by X_n the random number of iterations required by the above algorithm if the input consists of U and V truncated at n, *i.e.* of $T_n(U)$ and $T_n(V)$.

From U and V we obtain two increasing sequences $(\sigma_l)_{l\in\mathbb{N}_0}$ and $(\tau_l)_{l\in\mathbb{N}_0}$ of integer-valued random variables by $\tau_0 := 0$,

$$\sigma_l := \inf\{k \geq \tau_l : U_k + V_k > b - 1\}, \quad \tau_{l+1} := \inf\{k > \sigma_l : U_k + V_k \neq b - 1\}$$

for all $l \in \mathbb{N}_0$. With

$$M_l := \max\{\tau_j - \sigma_{j-1} : 1 \leq j \leq l\}, \quad N_n := \#\{0 \leq k < n : U_k + V_k > b - 1\}$$

we then have the following fundamental relationship on $T_n(V) \neq 0$,

$$M_{N_n-1} + 1 \leq X_n \leq M_{N_n} + 1. \tag{1}$$

On $T_n(V) = 0$ we have $X_n = 0$; as this happens with probability b^{-n} only this will turn out to be without significance for the limiting behaviour. Of course, the point of this construction is that the bounds in (1) are sharp enough to reduce the study of $(X_n)_{n\in\mathbb{N}}$ to that of $(M_l)_{l\in\mathbb{N}}$ and that the latter sequence is easier to analyse. In fact, it is immediately clear from the construction that the distribution of N_n is binomial with parameters n and χ_b; further, it is easily verified that the differences $\tau_{j+1} - \sigma_j$, $j \in \mathbb{N}_0$, are independent and geometrically distributed with parameter $2\chi_b$. Note, however, that $(M_l)_{l\in\mathbb{N}}$ and $(N_n)_{n\in\mathbb{N}}$ are not independent.

We use the well-known fact that geometric distributions arise as discretizations of exponential distributions. Formally, if $\tilde{Y}$ is exponentially distributed with parameter λ, then $Y := \lceil \tilde{Y} \rceil$ has a geometric distribution with parameter $1 - \exp(-\lambda)$. The transition $\tilde{Y} \mapsto Y$ preserves independence and obviously commutes with taking the maximum. We can therefore assume that $M_l = \lceil \tilde{M}_l \rceil$, with

$\tilde{M}_l := \max\{\tilde{Y}_1, \dots, \tilde{Y}_l\}$ and $(\tilde{Y}_j)_{j\in\mathbb{N}}$ a sequence of independent random variables, exponentially distributed with parameter $\lambda := \log b$.

The Gumbel distribution is one of the extreme value distributions, it arises as the limit distribution of the maxima of independent random variables (see *e.g.* Resnick [16]). We require a strong variant of this statement for exponential distributions.

Lemma 5. *Let $(\tilde{Y}_j)_{j\in\mathbb{N}}$ be a sequence of independent random variables, all exponentially distributed with parameter λ, and suppose that $Z \sim \mathrm{Gu}(\lambda)$. Let $\tilde{M}_n := \max\{\tilde{Y}_1, \dots, \tilde{Y}_n\}$. Then, for all $\gamma < 2\lambda$,*

$$\lim_{n\to\infty} \int_{-\infty}^{\infty} \mathrm{e}^{\gamma|z|} \left| f_n(z) - f_\infty(z) \right| \mathrm{d}z \;=\; 0,$$

where f_n and f_∞ denote the densities of $\tilde{M}_n - (\log n)/\lambda$ and Z respectively.

Proof. The distribution of $\lambda\tilde{Y}_j$ is exponential with parameter 1, hence a simple rescaling argument shows that it is enough to consider the case $\lambda = 1$. We split the range of integration for the weighted difference of the densities into the segments $(-\infty, -\log n)$, $[-\log n, 0]$ and $(0, \infty)$. For the first of these we note that $P(\tilde{M}_n < 0) = 0$, hence this part of the integral reduces to

$$\begin{aligned}
\int_{-\infty}^{-\log n} \mathrm{e}^{\gamma|x|} f_\infty(x)\,\mathrm{d}x &= \int_{-\infty}^{-\log n} \exp\bigl(-\gamma x - x - \exp(-x)\bigr)\,\mathrm{d}x \\
&= \int_0^{\infty} \exp\bigl(\gamma\log n + \gamma y + \log n + y - n\exp(y)\bigr)\,\mathrm{d}y \\
&\le\; n^{1+\gamma}\,\mathrm{e}^{-n} \int_0^{\infty} \mathrm{e}^{-y(n-1-\gamma)}\,\mathrm{d}y,
\end{aligned}$$

where we used that $\mathrm{e}^y \ge 1 + y$ for all $y \ge 0$. Evidently, this last integral tends to 0 as $n \to \infty$.

For all $x \ge -\log n$ we have

$$P(\tilde{M}_n \le \log n + x) \;=\; P(\tilde{Y}_1 \le \log n + x, \dots, \tilde{Y}_n \le \log n + x) \;=\; \Bigl(1 - \frac{\mathrm{e}^{-x}}{n}\Bigr)^n,$$

which shows that $\tilde{M}_n - \log n$ has density

$$f_n(x) \;=\; \mathrm{e}^{-x}\Bigl(1 - \frac{\mathrm{e}^{-x}}{n}\Bigr)^{n-1}, \quad x > -\log n.$$

We can therefore rewrite the integral over the middle segment as follows,

$$\begin{aligned}\int_{-\log n}^{0} \mathrm{e}^{\gamma|x|}\bigl|f_n(x)-f_\infty(x)\bigr|\,\mathrm{d}x&\\ &=\int_{-\log n}^{0}\mathrm{e}^{-\gamma x-x}\Bigl|\Bigl(1-\frac{\mathrm{e}^{-x}}{n}\Bigr)^{n-1}-\mathrm{e}^{-\mathrm{e}^{-x}}\Bigr|\,\mathrm{d}x\\ &=\int_1^n y^\gamma\Bigl|\Bigl(1-\frac{y}{n}\Bigr)^{n-1}-\mathrm{e}^{-y}\Bigr|\,\mathrm{d}y\\ &=\int_1^n y^\gamma\,\mathrm{e}^{-y}\Bigl|\exp\Bigl((n-1)\log\Bigl(1-\frac{y}{n}\Bigr)+y\Bigr)-1\Bigr|\,\mathrm{d}y\,.\end{aligned}$$

It is easily seen that the term within the absolute value signs tends to 0 as $n\to\infty$. Using

$$\log(1-z)\le -z\quad\text{for}\quad z<1$$

we see that it is also bounded from above by $\mathrm{e}^{y/n}+1$, which means that the integrand is bounded by the function $y\mapsto y^\gamma\bigl(\mathrm{e}^{-y/2}+\mathrm{e}^{-y}\bigr)$ for $n\ge 2$. Hence Lebesgue's dominated convergence theorem applies and we obtain that the middle term also vanishes in the limit.

The integral over the final segment $(0,\infty)$ can be rewritten as

$$\begin{aligned}\int_0^\infty \mathrm{e}^{\gamma x}\bigl|f_n(x)-f_\infty(x)\bigr|\,\mathrm{d}x&\\ &=\int_0^\infty \mathrm{e}^{\gamma x-x}\Bigl|\Bigl(1-\frac{\mathrm{e}^{-x}}{n}\Bigr)^{n-1}-\mathrm{e}^{-\mathrm{e}^{-x}}\Bigr|\,\mathrm{d}x\\ &=\int_0^\infty \mathrm{e}^{\gamma x-x-\mathrm{e}^{-x}}\Bigl|\exp\Bigl((n-1)\log\Bigl(1-\frac{\mathrm{e}^{-x}}{n}\Bigr)+\mathrm{e}^{-x}\Bigr)-1\Bigr|\,\mathrm{d}x.\end{aligned}$$

We may assume that $n\ge 2$ so that the inequality

$$\bigl|\log(1-y)+y\bigr|\ \le\ 2\,y^2\quad\text{on}\ \ |y|\le 1/2$$

applies with $y=\mathrm{e}^{-x}/n$, resulting in the upper bound

$$\Bigl|(n-1)\log\Bigl(1-\frac{\mathrm{e}^{-x}}{n}\Bigr)+\mathrm{e}^{-x}\Bigr|\ \le\ \frac{\mathrm{e}^{-x}}{n}+2\,\frac{\mathrm{e}^{-2x}}{n}\ \le\ 3.$$

This in turn means that we can use

$$\bigl|\mathrm{e}^y-1\bigr|\ \le\ \mathrm{e}^3\,|y|\quad\text{on}\ \ |y|\le 3$$

to obtain

$$\Bigl|\exp\Bigl((n-1)\log\Bigl(1-\frac{\mathrm{e}^{-x}}{n}\Bigr)+\mathrm{e}^{-x}\Bigr)-1\Bigr|\ \le\ \mathrm{e}^3\Bigl(\frac{\mathrm{e}^{-x}}{n}+2\,\frac{\mathrm{e}^{-2x}}{n}\Bigr).$$

Inserting this upper bound in the above integral and using $\gamma < 2$ we again obtain the limit 0 as $n \to \infty$. □

The proofs of our first three theorems all use the same strategy: X_n is related to M_{N_n} and M_{N_n-1} by (1), these are related to some M-variables with non-random index $l = l(n)$, M_l has a representation *via* a continuous counterpart $\tilde{M}_l$ and Lemma 5 is used to close the gap to a suitably transformed Gumbel variate.

The total variation distance for (the distributions of) random variables X and Y with densities f_X and f_Y is given by

$$\mathrm{d_{TV}}(X,Y) \;=\; \frac{1}{2}\int_{-\infty}^{\infty} \bigl|f_X(u) - f_Y(u)\bigr|\,\mathrm{d}u,$$

it is obviously invariant under shifts, *i.e.*

$$\mathrm{d_{TV}}(X+c, Y+c) \;=\; \mathrm{d_{TV}}(X,Y) \quad \text{for all } c \in \mathbb{R}, \tag{2}$$

and cannot increase under discretization, *i.e.*

$$\mathrm{d_{TV}}\bigl(\lceil X\rceil, \lceil Y\rceil\bigr) \;\le\; \mathrm{d_{TV}}(X,Y). \tag{3}$$

Also, as $\lim_{c\to 0}\int\bigl|f(x+c)-f(x)\bigr|\,\mathrm{d}x = 0$ for integrable functions f (see *e.g.* Th. 9.5 in Rudin [18]), we have the following continuity property for random variables X with an absolutely continuous distribution,

$$\lim_{c\to 0} \mathrm{d_{TV}}(X+c, X) \;=\; 0. \tag{4}$$

With $\gamma = 0$ Lemma 5 leads to a limit result for the total variation distance between the distributions of the shifted maximum and the limit variable, which is enough to carry out the last step of the strategy outlined above in connection with Theorem 1. For the step from X_n to M_l we need the following lemma:

Lemma 6. $\lim_{n\to\infty} \mathrm{d_{TV}}\bigl(X_n, M_{\lceil n\chi_b\rceil}+1\bigr) \;=\; 0.$

Proof. Let

$$A_n \;:=\; \bigl\{|N_n - \lceil n\chi_b\rceil| < n^{3/4}\bigr\}, \quad B_n \;:=\; \bigl\{M_{\lceil n\chi_b\rceil - \lceil n^{3/4}\rceil} = M_{\lceil n\chi_b\rceil + \lceil n^{3/4}\rceil}\bigr\}.$$

We have $X_n = M_{\lceil n\chi_b\rceil}+1$ on $A_n \cap B_n \cap \{X_n \neq 0\}$ so that

$$\bigl|P(X_n \in C) - P(M_{\lceil n\chi_b\rceil}+1 \in C)\bigr| \;\le\; P(A_n^{\mathrm{c}}) + P(B_n^{\mathrm{c}}) + b^{-n}$$

for all $C \subset \mathbb{Z}$ (we write A^{c} for the set-theoretic complement of A). As the right hand side does not depend on C the assertion will follow if we can show that $P(A_n^{\mathrm{c}}) \to 0$ and $P(B_n^{\mathrm{c}}) \to 0$ as $n \to \infty$. Since $EN_n = n\chi_b$ and $\mathrm{var}(N_n) = n\chi_b(1-\chi_b)$ the first of these is an easy consequence of Chebychev's inequality. In connection with B_n we use the above representation by exponentially

distributed random variables $\tilde{Y}_j$, $j \in \mathbb{N}$. Obviously, for $M_{\lceil n\chi_b\rceil - \lceil n^{3/4}\rceil}$ to differ from $M_{\lceil n\chi_b\rceil + \lceil n^{3/4}\rceil}$ we need a record value in the index range $j \in \lceil n\chi_b\rceil \pm \lceil n^{3/4}\rceil$. Let $R_l := \{\tilde{Y}_l > \tilde{Y}_j : j = 1, \dots, l-1\}$ denote the event that the l^{th} value in the $\tilde{Y}$-sequence is a record. It is well known (and easy to see) that $P(R_l) = 1/l$. In particular,

$$P(B_n^c) \;\le\; \sum_{j=\lceil n\chi_b\rceil - \lceil n^{3/4}\rceil}^{\lceil n\chi_b\rceil + \lceil n^{3/4}\rceil} \frac{1}{j} \quad\to\quad 0$$

as $n \to \infty$. □

Proof of Theorem 1. Lemma 5 implies

$$\lim_{n\to\infty} \mathrm{d_{TV}}\big(\tilde{M}_{\lceil n\chi_b\rceil} - \log_b\lceil n\chi_b\rceil\,,\, Z\big) \;=\; 0 \quad \text{with } Z \sim \mathrm{Gu}(\log b).$$

As the relevant distributions are absolutely continuous property (4) yields

$$\lim_{n\to\infty} \mathrm{d_{TV}}\big(\tilde{M}_{\lceil n\chi_b\rceil} + 1 - \log_b n - \zeta_b\,,\, Z\big) \;=\; 0.$$

For $x \in \mathbb{R}$, $k \in \mathbb{Z}$ we have $\lceil x - k\rceil = \lceil x\rceil - k$, hence we obtain on using (2) and (3)

$$\lim_{n\to\infty} \mathrm{d_{TV}}\big(M_{\lceil n\chi_b\rceil} + 1 - \lfloor \log_b n\rfloor\,,\, \lceil Z + \zeta_b + \{\log_b n\}\rceil\big) \;=\; 0.$$

The assertion now follows with Lemma 6, the triangle inequality for the total variation distance and a version of (2) for discrete random variables. □

Proof of Theorem 2. Let $t > 0$ be fixed. As we compare integer-valued random variables we may replace $t\log_b n$ by $a_n := \lceil t\log_b n\rceil$. Let κ be a constant satisfying $1/2 < \kappa < 1$. Chernoff's bound on the tails of binomial distributions gives

$$P\big(|N_n - n\chi_b| \ge \beta n\big) \;\le\; 2\,\exp(-2\beta^2 n) \quad \text{for all } \beta \ge 0,$$

see *e.g.* Section 2 in McDiarmid [13]. Hence, with $A_n := \big\{|N_n - \lceil n\chi_b\rceil| < n^\kappa\big\}$,

$$P(A_n^c) \;\le\; P\big(|N_n - n\chi_b| \ge n^\kappa - 1\big) \;=\; O\Big(\exp\big(-n^{2\kappa-1}\big)\Big) \;=\; o\big(n^{-\gamma}\big) \qquad (5)$$

for all $\gamma > 0$. Let $\phi_\pm : \mathbb{N} \to \mathbb{N}$ be defined by

$$\phi_+(n) := \lceil n\chi_b\rceil + \lfloor n^\kappa\rfloor, \quad \phi_-(n) := \lceil n\chi_b\rceil - \lfloor n^\kappa\rfloor - 1\,.$$

From (1) it follows that

$$\begin{aligned} P\big(X_n - \lfloor\log_b n\rfloor \ge a_n\big) \;&\le\; P\big(M_{\phi_+(n)} + 1 \ge a_n + \lfloor\log_b n\rfloor\big) \;+\; P(A_n^c) \;+\; b^{-n}, \\ P\big(X_n - \lfloor\log_b n\rfloor \ge a_n\big) \;&\ge\; P\big(M_{\phi_-(n)} + 1 \ge a_n + \lfloor\log_b n\rfloor\big) \;-\; P(A_n^c) \;-\; b^{-n}. \end{aligned}$$

In view of (5) it remains to consider the respective first terms on the right hand side. As a_n is integer we may replace $M_{\phi_+(n)}$ by $\tilde{M}_{\phi_+(n)} + 1$. With

$$c_n^{\pm} \;:=\; \lfloor \log_b n \rfloor - \log_b \phi_{\pm}(n) - 2$$

and

$$\Delta_l(z) \;:=\; \Big| P\big(\tilde{M}_l - \log_b l \ge z\big) - P(Z \ge z) \Big|$$

this results in

$$\begin{aligned} P\big(M_{\phi_+(n)} + 1 \ge a_n + \lfloor \log_b n \rfloor\big) \;&=\; P\big(\tilde{M}_{\phi_+(n)} \ge a_n + \lfloor \log_b n \rfloor - 2\big) \\ &\le\; P(Z \ge a_n + c_n^+) \;+\; \Delta_{\phi_+(n)}(a_n + c_n^+)\,, \\ P\big(M_{\phi_-(n)} + 1 \ge a_n + \lfloor \log_b n \rfloor\big) \;&=\; P\big(\tilde{M}_{\phi_-(n)} \ge a_n + \lfloor \log_b n \rfloor - 2\big) \\ &\ge\; P(Z \ge a_n + c_n^-) \;-\; \Delta_{\phi_-(n)}(a_n + c_n^-)\,. \end{aligned}$$

Let f_l and f_∞ be the densities of $\tilde{M}_l - \log_b l$ and Z respectively. Markov's inequality yields

$$\Delta_{\phi_\pm(n)}(a_n + c_n^{\pm}) \;\le\; \exp\big(-\gamma(a_n + c_n^{\pm})\big) \int_{z \ge a_n + c_n^{\pm}} \mathrm{e}^{\gamma z} \big| f_{\phi_\pm(n)}(z) - f_\infty(z) \big| \,\mathrm{d}z,$$

Lemma 5 implies that the integral tends to 0 with $n \to \infty$ for any $\gamma < 2 \log b$. Using $c_n^{\pm} = O(1)$ we see that this yields the rate $o(n^{-\eta t})$ for all $\eta < 2$.

It remains to show that

$$P\big(\lceil Z + \zeta_b + \{\log_b n\} \rceil \ge a_n\big) \;-\; P\big(Z \ge a_n + c_n^{\pm}\big) \;=\; o(n^{-t}).$$

Because of

$$P\big(\lceil Z + \zeta_b + \{\log_b n\} \rceil \ge a_n\big) \;=\; P\big(Z \ge a_n - \zeta_b - \{\log_b n\} - 1\big)$$

this amounts to finding an upper bound for the probability that Z takes its value in a short interval moving to the right. Using $\zeta_b = \log_b \chi_b + 1$ we obtain

$$\big| c_n^+ - (-\zeta_b - \{\log_b n\} - 1) \big| \;=\; \Big| \log_b \frac{\lceil n\chi_b \rceil + \lfloor n^\kappa \rfloor}{n\chi_b} \Big| \;=\; O\big(n^{\kappa - 1}\big),$$

and the same bound holds for the other interval. For any sequence $(\delta_n)_{n \in \mathbb{N}}$ with $\delta_n \downarrow 0$ we have

$$\begin{aligned} P(Z \in a_n \pm \delta_n) \;&\le\; 2\,\delta_n \sup_{z \in a_n \pm \delta_n} \frac{\mathrm{d}}{\mathrm{d}z} P(Z \le z) \\ &\le\; 2\,\delta_n \log b \sup_{z \in a_n \pm \delta_n} \exp\big(-(\log b) z\big) \\ &\le\; 2\,\delta_n \log b \;\; b^{-a_n + \delta_n}\,. \end{aligned}$$

This last term is of the order $O(\delta_n n^{-t})$. With the above bounds on the length of the intervals this leads to the bound $O(n^{-t+\kappa-1})$, which is $o(n^{-t})$ as desired. □

We note that the above proof supplies the rate $o(n^{-\eta})$ for the difference of the tail probabilities, for all $\eta < t + \min\{t, 1/2\}$, which is more than the order $o(n^{-t})$ in the assertion of the theorem.

For convergence in total variation to imply convergence of the associated moments some extra conditions are needed; the following lemma gives the details in the discrete case that is of interest for the proof of Theorem 3. Note that this is a result for distributions, the X- and Y-variables need not be defined on the same probability space.

Lemma 7. *Let X_n, Y_n, $n \in \mathbb{N}$, be random variables with values in $\mathbb{Z}$ and let $l \in \mathbb{N}$. If, for some $\gamma > l$,*

$$\lim_{n\to\infty} \mathrm{d_{TV}}(X_n, Y_n) = 0, \quad \sup_{n\in\mathbb{N}} E|X_n|^\gamma < \infty, \quad \sup_{n\in\mathbb{N}} E|Y_n|^\gamma < \infty,$$

then

$$\lim_{n\to\infty} \big(EX_n^l - EY_n^l\big) \;=\; 0.$$

Proof. A simple decomposition gives, for all $M \in \mathbb{N}$,

$$\begin{aligned}\big|EX_n^l - EY_n^l\big| \;\le\; & \sum_{|k|<M} |k|^l \big|P(X_n = k) - P(Y_n = k)\big| \\ & + \sum_{|k|\ge M} |k|^l\, P(X_n = k) \;+\; \sum_{|k|\ge M} |k|^l\, P(Y_n = k).\end{aligned}$$

For the second term we have

$$\sum_{|k|\ge M} |k|^l\, P(X_n = k) \;\le\; M^{l-\gamma} \sum_{|k|\ge M} |k|^\gamma\, P(X_n = k) \le\; M^{l-\gamma} \sup_{n\in\mathbb{N}} E|X_n|^\gamma,$$

and the analogue of this upper bound obviously also holds for the Y-variables. Hence, for any given $\epsilon > 0$, we can find an M large enough for the second and third term in the decomposition to be less than $\epsilon/3$. For the first term we use

$$\sum_{|k|<M} |k|^l \big|P(X_n = k) - P(Y_n = k)\big| \;\le\; 2\, M^l\, d_{\mathrm{TV}}(X_n, Y_n),$$

i.e. we can further choose n_0 (in dependence on M) large enough for this term to also be less than $\epsilon/3$ for all $n \ge n_0$. Put together this implies the convergence of the moments. □

Proof of Theorem 3. In view of Theorem 1 and Lemma 7 it remains to show that

$$\sup_{n\in\mathbb{N}} E\big|X_n - \lfloor \log_b n\rfloor\big|^\gamma < \infty, \quad \sup_{n\in\mathbb{N}} E\big|\lceil Z + \zeta_b + \{\log_b n\}\rceil\big|^\gamma < \infty$$

for all $\gamma > 0$. For Z this is trivial: The shifts are bounded by $|\zeta_b| + 1$ and Gumbel distributions have exponentially decreasing tails. For the first part we use

$$E\big|X_n - \lfloor \log_b n\rfloor\big|^\gamma \;=\; \left(\int_{A_n\cap\{X_n=0\}} + \int_{A_n\cap\{X_n\neq 0\}} + \int_{A_n^{\mathrm{c}}}\right)\big|X_n - \lfloor \log_b n\rfloor\big|^\gamma \,\mathrm{d}P$$

with A_n as defined in the proof of Theorem 2 and consider the three terms on the right hand side separately. Since $0 \le X_n \le n+1$ we have

$$\int_{A_n^{\mathrm{c}}} \big|X_n - \lfloor \log_b n\rfloor\big|^\gamma \,\mathrm{d}P \;\le\; P(A_n^{\mathrm{c}})\,\big(n + 1 + \lfloor \log_b n\rfloor\big)^\gamma,$$

hence for the third term the desired boundedness follows from (5). On $A_n \cap \{X_n = 0\}$ we can use $P(X_n = 0) = b^{-n}$. On $A_n \cap \{X_n \neq 0\}$ we have, with $\phi_\pm$ as in the proof of Theorem 2,

$$\big|X_n - \lfloor \log_b n\rfloor\big|^\gamma \;\le\; \big|M_{\phi_+(n)} + 1 - \lfloor \log_b n\rfloor\big|^\gamma \;+\; \big|M_{\phi_-(n)} + 1 - \lfloor \log_b n\rfloor\big|^\gamma.$$

Now, with the notation of Lemma 5

$$\big|E|\tilde{M}_n - (\log n)/\lambda|^\gamma - E|Z|^\gamma\big| \;\le\; \int |z|^\gamma \big|f_n(z) - f_\infty(z)\big|\,\mathrm{d}z.$$

As $z \mapsto |z|^\gamma$ increases at a subexponential rate, this lemma implies that the right hand side converges to 0 as $n \to \infty$; in particular, with $\lambda = \log b$,

$$\sup_{n\in\mathbb{N}} E\big|\tilde{M}_n - \log_b n\big|^\gamma \;<\; \infty\,.$$

We have

$$\begin{aligned} M_{\phi_\pm(n)} + 1 - \lfloor \log_b n\rfloor \;=\; &\big(\tilde{M}_{\phi_\pm(n)} - \log_b \phi_\pm(n)\big) \;+\; \big(M_{\phi_\pm(n)} - \tilde{M}_{\phi_\pm(n)}\big) \;+\; 1 \\ &+\; \big(\log_b \phi_\pm(n) - \lfloor \log_b n\rfloor\big), \end{aligned}$$

hence

$$\sup_{n\in\mathbb{N}} E\big|M_{\phi_\pm(n)} - \lfloor \log_b n\rfloor\big|^\gamma \;<\; \infty$$

follows on using Minkowski's inequality. □

Proof of Theorem 4. Suppose the statement is wrong. Then $\eta \mapsto E\{Z + \eta\}$ is constant on $(0, 1)$. From

$$\{Z + \eta\} \;=\; \begin{cases} \{Z\} + \eta, & \text{if } \{Z\} + \eta < 1, \\ \{Z\} + \eta - 1, & \text{if } \{Z\} + \eta \ge 1, \end{cases}$$

for all $\eta \in (0,1)$ it then follows that

$$\begin{aligned} E\{Z+\eta\} &= \int_{(0,1-\eta)} (x+\eta)\, P^{\{Z\}}(\mathrm{d}x) \;+\; \int_{[1-\eta,1]} (x+\eta-1)\, P^{\{Z\}}(\mathrm{d}x) \\ &= \eta \;+\; \int_{(0,1]} x\, P^{\{Z\}}(\mathrm{d}x) \;-\; \int_{[1-\eta,1)} 1\, P^{\{Z\}}(\mathrm{d}x) \\ &= \eta \;+\; E\{Z\} \;-\; P\big(\{Z\} \ge 1-\eta\big). \end{aligned}$$

Hence this function can only be constant if the fractional part of Z is uniformly distributed on the unit interval. This, however, would imply

$$\phi_Z(2\pi) \;=\; E\mathrm{e}^{2\pi i Z} \;=\; E\mathrm{e}^{2\pi i \{Z\}} \;=\; \phi_{\{Z\}}(2\pi) \;=\; 0,$$

where ϕ_Y denotes the characteristic function of the random variable Y. However, the characteristic function of a Gumbel variate Z does not take the value 0. This can either be seen directly from $\phi_Z(t) = \Gamma(1-it/\lambda)$ and the fact that the Gamma function has no zeros, or (more probabilistically...) from the fact that Gumbel distributions are infinitely divisible, which implies that their characteristic functions do not take the value 0. □

3. Miscellaneous comments

As announced above, we briefly comment on the modelling of the algorithm, we discuss some large deviation aspects, and finally we give some more details on the limiting average case behaviour.

3.1. The formal descriptions of von Neumann addition given by Claus [4] and Knuth [12] differ with respect to the way an overflow bit is handled. This is perhaps most easily described with the help of a simple example with $b = 2$: Starting with

$$\begin{array}{rcccccccccc} u &=& (u_0 & \cdots & u_{n-k+1} & 1 & 1 & \cdots & 1) & & \\ v &=& (0 & \cdots & 0 & 1 & 0 & \cdots & 0) & \text{carry: } 0, & \end{array}$$

both of length n, we arrive after k iterations at

$$\begin{array}{rcccccccccc} u^{(k)} &=& (u_0 & \cdots & u_{n-k+1} & 0 & 0 & \cdots & 0) & & \\ v^{(k)} &=& (0 & \cdots & 0 & 0 & 0 & \cdots & 0) & \text{carry: } 1. & \end{array}$$

From a practical point of view no further iterations are required as the binary representation of $u+v$ can be read off from the last display. However, if we pad the original input with zeros we would have

$$\begin{array}{rcccccccccccc} u^{(k)} &=& (u_0 & \cdots & u_{n-k+1} & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots) \\ v^{(k)} &=& (0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 1 & 0 & \cdots), \end{array}$$

and the formal requirement that all components in the v-part are zero would necessitate an additional step – a completely superfluous step as the corresponding u-position will always have the value 0. In our analysis above we have included this step, as in Knuth [12] and the textbooks we are aware of, in contrast to Claus [4]. How does this affect our results? The two formalizations of the algorithm can only differ in the number of iterations required if the maximal run straddles n, *i.e.* if $M_{N_n-1} < M_{N_n}$. As in the proof of Lemma 6 we can construct an event A_n such that

$$0 \;\le\; X_n^{(\mathrm{Knuth})} - X_n^{(\mathrm{Claus})} \;\le\; 1_{A_n}, \quad \text{with } \lim_{n\to\infty} P(A_n) = 0.$$

From this it is obvious that Theorem 1 holds for both versions and so does Theorem 3 in view of Minkowski's inequality, Theorem 4 is not affected. In particular, Claus' conjecture on the asymptotic behaviour of the average number of iterations (which obviously referred to his version) is now disproved.

3.2. The arguments in the proof of Theorem 2 can be used to obtain tail approximations for sequences other than $n \mapsto (1+t)\log_b n$, $t > 0$. In particular, the approximation by a discretized shifted Gumbel distribution turns out to give the correct first order term for the tail probabilities of X_n for sequences of the order $O(n^\gamma)$ with $\gamma < 1$ (some upper bound on the rate of increase of the sequence is obviously needed as $P(X_n > n+2) = 0$ for all $n \in \mathbb{N}$). Using

$$\begin{aligned} P\big(\lceil Z + \zeta_b + \{\log_b n\}\rceil \;&\ge\; (1+t)\log_b n - \lfloor \log_b n \rfloor\big) \\ &=\; P\big(Z + \zeta_b + \{\log_b n\} \ge \lceil (1+t)\log_b n\rceil - \lfloor \log_b n\rfloor - 1\big) \\ &=\; P\big(Z \ge t\log_b n - \zeta_b + \lceil (1+t)\log_b n\rceil \\ &\qquad -(1+t)\log_b n - \;1\big) \end{aligned}$$

and writing $a_n \sim b_n$ if the ratio a_n/b_n tends to 1 with $n \to \infty$, we arrive at

$$P\big(X_n \ge (1+t)\log_b n\big) \;\sim\; n^{-t}\, b^{(1+t)\log_b n + 1 - \lceil (1+t)\log_b n\rceil}\, b\,\chi_b\,.$$

As the exponent of b varies between 0 and 1 this in turn implies

$$\begin{aligned} \limsup_{n\to\infty}\; n^t\, P\big(X_n \ge (1+t)\log_b n\big) \;&=\; b^2\chi_b\,, \\ \liminf_{n\to\infty}\; n^t\, P\big(X_n \ge (1+t)\log_b n\big) \;&=\; b\,\chi_b \qquad \text{for all } t > 0. \end{aligned} \tag{6}$$

On a logarithmic scale, which is much coarser, the fluctuations disappear and we obtain

$$\lim_{n\to\infty} \frac{1}{\log n}\, \log P\big(X_n \ge (1+t)\log_b n\big) \;=\; -t \quad \text{for all } t > 0.$$

The latter could be seen as a large deviation result; see McDiarmid [13] for general comments and McDiarmid and Hayward [14] for a similar result in the QUICKSORT context.

However, if interest is primarily in bounds for the tail probabilities of X_n then, in the present situation, such results can be obtained more directly. Let

$$A_{ik} \;:=\; \{U_i + V_i > b-1\} \cap \bigcap_{l=1}^{k} \{U_{i+l} + V_{i+l} = b-1\}, \quad i \in \mathbb{N},\ k \in \mathbb{N}_0,$$

where $(U_n)_{n\in\mathbb{N}}$ and $(V_n)_{n\in\mathbb{N}}$ refer to the notation introduced at the beginning of Section 2. In words: A_{ik} is the event that a carry is generated at position i and propagates for at least k steps. The independence of the individual bits (or digits) implies $P(A_{ik}) = \chi_b b^{-k}$. For fixed k the events A_{ik}, $i \in \mathbb{N}$, are either disjoint or independent, hence

$$P(A_{ik} \cap A_{jk}) \;=\; \begin{cases} 0, & \text{if } 0 < |i-j| \le k, \\ \chi_b^2 b^{-2k}, & \text{for } |i-j| > k. \end{cases}$$

These events are related to the number X_n of iterations required for input length n by

$$\{X_n \ge k+2\} \;=\; \bigcup_{i=1}^{n-k} A_{ik} \quad \text{for all } n \in \mathbb{N},\ k \in \mathbb{N}_0.$$

Now let $k_n := \lceil (1+t)\log_b n \rceil - 2$. Using Boole's inequality we then obtain

$$\begin{aligned} P\big(X_n \ge (1+t)\log_b n\big) \;&=\; P(X_n \ge k_n + 2) \;\le\; \sum_{i=1}^{n-k_n} P(A_{ik_n}) \;\le\; n\chi_b b^{-k_n} \\ &\le\; n\,\chi_b\, b^{2-(1+t)\log_b n} \;\le\; b^2 \chi_b\, n^{-t}. \end{aligned}$$

Similarly, now with Bonferroni's inequality,

$$\begin{aligned} P\big(X_n \ge (1+t)\log_b n\big) \;&\ge\; \sum_{i=1}^{n-k_n} P(A_{ik_n}) \;-\; \sum_{\substack{i,j=1 \\ i\ne j}}^{n-k_n} P(A_{ik_n} \cap A_{jk_n}) \\ &\ge\; (n-k_n)\chi_b\, b^{-k_n} \;-\; (n-k_n)^2 \chi_b^2\, b^{-2k_n}. \end{aligned}$$

For the second term on the right hand side we obtain the order $O\big(n^2 n^{-2(1+t)}\big) = o\big(n^{-t}\big)$, and $k_n \le (1+t)\log_b n - 1$ yields the lower bound $(n-k_n)\chi_b\, n^{-1-t} b$ for the first term. Put together these elementary estimates result in

$$\begin{aligned} &\limsup_{n\to\infty}\; n^t\, P\big(X_n \ge (1+t)\log_b n\big) \;\le\; b^2\chi_b\,, \\ &\liminf_{n\to\infty}\; n^t\, P\big(X_n \ge (1+t)\log_b n\big) \;\ge\; b\,\chi_b \qquad \text{for all } t > 0. \end{aligned} \tag{7}$$

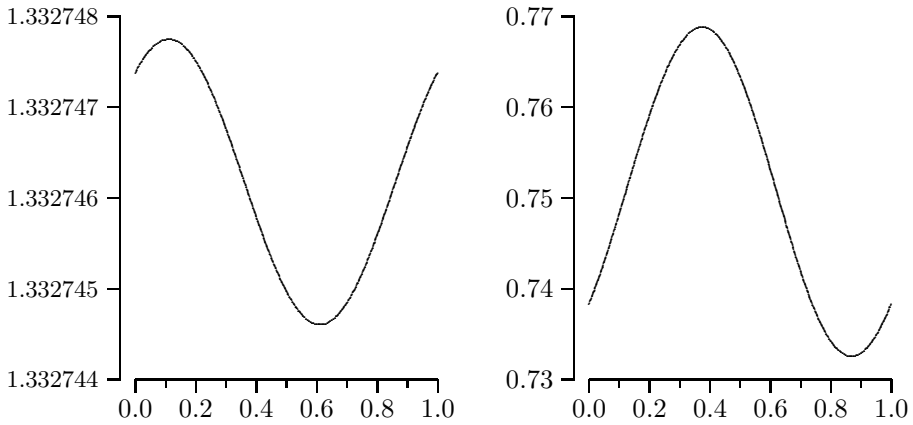
In fact, the above arguments show that the upper bound even holds for the individual n's, a situation not uncommon in the large deviation context. *Via* suitably

chosen subsequences the inequalities in (7) may be strengthened, so we overall obtain a completely elementary proof of (6).

3.3. In connection with the limiting behaviour of the expected number of iterations a central role is played by the function $h : [0,1] \to \mathbb{R},\ \eta \mapsto E\lceil Z+\eta\rceil - \eta$. A straightforward calculation yields

$$E\lceil Z+\eta\rceil \;=\; \sum_{k\in\mathbb{Z}} k\left(\mathrm{e}^{-b^{\eta-k}} - \mathrm{e}^{-b^{\eta+1-k}}\right) \quad \text{for all } \eta\in[0,1].$$

Figure 1 shows the function h for $b=2$ and $b=10$. Apart from being "almost constant", h also "looks very sinusoidal". This we will now investigate; as a corollary we also obtain the connection to Knuth's formula cited in Section 1.



(a) $b=2$ (b) $b=10$

FIGURE 1: The periodic limit function h.

The function h is obviously sufficiently smooth to permit a representation as an absolutely convergent Fourier series,

$$h(\eta) \;=\; \sum_{k\in\mathbb{Z}} a_k\,\mathrm{e}^{2\pi i k\eta}, \quad \text{with } a_k = \int_0^1 h(\eta)\,\mathrm{e}^{-2\pi i k\eta}\,\mathrm{d}\eta\,.$$

Using $EZ=\gamma/\lambda$ for $Z\sim\mathrm{Gu}(\lambda)$, $\lambda=\log b$ and Fubini's theorem we obtain

$$a_0 \;=\; E\left(\int_0^1 \left(\lceil Z+\eta\rceil-\eta\right)\mathrm{d}\eta\right) \;=\; EZ+\frac{1}{2} \;=\; \frac{\gamma}{\log b}+\frac{1}{2}.$$

Suppose that $k\neq 0$. An elementary computation shows that

$$\int_0^1 \left(\lceil z+\eta\rceil-\eta\right)\,\mathrm{e}^{-2\pi i k\eta}\,\mathrm{d}\eta \;=\; \frac{\mathrm{e}^{2\pi i k\{z\}}}{2\pi i k} \quad \text{for all } z\in\mathbb{R},$$

which leads to

$$a_k = E\left(\int_0^1 \left(\lceil Z+\eta\rceil - \eta\right) \mathrm{e}^{-2\pi i k\eta}\,\mathrm{d}\eta\right) = \frac{1}{2\pi i k}\, E\mathrm{e}^{2\pi i k\{Z\}} = \frac{1}{2\pi i k}\,\Gamma\left(1-\frac{2\pi i k}{\log b}\right)$$
$$= -\frac{1}{\log b}\,\Gamma\left(-\frac{2\pi i k}{\log b}\right).$$

In this derivation, Fubini's theorem justifies the first equality, the third uses the arguments for the characteristic functions of $\{Z\}$ and Z given in the proof of Theorem 4 and the last one follows with $\Gamma(z+1) = z\Gamma(z)$. It is well known that $\mathbb{R} \ni t \mapsto \Gamma(it)$ decreases at an exponential rate; Table 1 gives a_k for some values of k and b. We see that indeed the Fourier coefficients decrease rapidly (the near-constancy phenomenon could be seen under this aspect – the decrease sets in at $k = 0$). As h is real-valued, we have $a_{-k} = \overline{a_k}$ for all $k \in \mathbb{Z}$.

TABLE 1: Fourier coefficients of h.

k	$b = 2$	$b = 10$
0	1.332746177	0.750681578
1	$0.603 \times 10^{-6} - 0.506 \times 10^{-6}\, i$	$-0.628 \times 10^{-2} - 0.654 \times 10^{-2}\, i$
2	$0.213 \times 10^{-12} + 0.295 \times 10^{-12}\, i$	$0.873 \times 10^{-4} + 0.122 \times 10^{-4}\, i$
3	$-0.107 \times 10^{-18} - 0.163 \times 10^{-18}\, i$	$0.363 \times 10^{-6} + 0.921 \times 10^{-6}\, i$
10	$0.541 \times 10^{-62} + 0.104 \times 10^{-62}\, i$	$-0.392 \times 10^{-19} - 0.319 \times 10^{-19}\, i$

With the above h and ζ_b as defined at the beginning of this section we can write the special case $l = 1$ of Theorem 3 as

$$EX_n \;=\; \log_b n + \zeta_b + h\big(\zeta_b + \{\log_b n\}\big) \;+\; o(1).$$

Inserting the Fourier series representation with the coefficients as calculated above and using

$$\begin{aligned} a_k \mathrm{e}^{2\pi i k\eta} + a_{-k}\mathrm{e}^{-2\pi i k\eta} &= 2\Re\big(a_k \mathrm{e}^{2\pi i k\eta}\big), \\ \exp\big(2\pi i k(\zeta_b + \{\log_b n\})\big) &= \exp\big(2\pi i k(\zeta_b + \log_b n)\big), \end{aligned}$$

we obtain

$$\begin{aligned} EX_n &= \log_b n + \zeta_b + a_0 + 2\sum_{k=1}^{\infty} \Re\Big(a_k \exp\big(2\pi i k(\zeta_b + \{\log_b n\})\big)\Big) + o(1) \\ &= \log_b n + \frac{\gamma}{\log b} + \frac{1}{2} + \log_b \frac{b-1}{2} \\ &\quad - \frac{2}{\log b} \sum_{k=1}^{\infty} \Re\Big(\Gamma\Big(-\frac{2\pi i k}{\log b}\Big) \exp\Big(2\pi i k \log_b \frac{(b-1)n}{2}\Big)\Big) + o(1). \end{aligned}$$

Hence we finally arrive at a formula for the limiting average case behaviour that agrees with Knuth's result quoted in the first section. Note, however, that Knuth obtained a rate result too. We would expect that a statement on the speed of convergence would also be possible with the methods used here, but we have not carried this out.

Acknowledgements. The first author wishes to thank Professor A. Steger (Technische Universität München) for a most stimulating discussion. Her comments led to a considerable improvement of Theorem 2 and the material in Section 3.2. We also thank both referees for drawing our attention to tries and the related literature.

References

[1] P. Billingsley, *Probability and Measure*, 2nd Ed. Wiley, New York (1986).
[2] A.W. Burks, H.H. Goldstine and J. von Neumann, *Preliminary discussion of the logical design of an electronic computing instrument*. Inst. for Advanced Study Report (1946). Reprinted in *John von Neumann Collected Works*, Vol. 5. Pergamon Press, New York (1961).
[3] P. Chassaing, J.F. Marckert and M. Yor, *A stochastically quasi-optimal algorithm*. Preprint (1999).
[4] V. Claus, Die mittlere Additionsdauer eines Paralleladdierwerks. *Acta Inform.* **2** (1973) 283-291.
[5] Th.H. Cormen, Ch.E. Leiserson and R.L. Rivest, *Introduction to Algorithms*. MIT Press, Cambridge, USA (1997).
[6] Ph. Flajolet, X. Gourdon and Ph. Dumas, Mellin transforms and asymptotics: Harmonic sums. *Theoret. Comput. Sci.* **144** (1995) 3-58.
[7] O. Forster, *Algorithmische Zahlentheorie*. Vieweg, Braunschweig (1996).
[8] R. Grübel, Hoare's selection algorithm: A Markov chain approach. *J. Appl. Probab.* **35** (1998) 36-45.
[9] R. Grübel, On the median-of-k version of Hoare's selection algorithm. *RAIRO: Theoret. Informatics Appl.* **33** (1999) 177-192.
[10] R. Grübel and U. Rösler, Asymptotic distribution theory for Hoare's selection algorithm. *Adv. Appl. Probab.* **28** (1996) 252-269.
[11] D.E. Knuth, *The Art of Computer Programming*, Vol. 3, Sorting and Searching. Addison-Wesley, Reading (1973).
[12] D.E. Knuth, The average time for carry propagation. *Nederl. Akad. Wetensch. Indag. Math.* **40** (1978) 238-242.

[13] C. McDiarmid, Concentration, in *Probabilistic Methods for Algorithmic Discrete Mathematics*, edited by M. Habib, C. McDiarmid, J. Ramirez–Alfonsin and B. Reed. Springer, Berlin (1998).
[14] C. McDiarmid and R.B. Hayward, Large deviations for Quicksort. *J. Algorithms* **21** (1996) 476-507.
[15] M. Régnier, A limiting distribution for quicksort. *RAIRO: Theoret. Informatics Appl.* **23** (1989) 335-343.
[16] S.I. Resnick, *Extreme Values, Regular Variation and Point Processes.* Springer, New York (1987).
[17] U. Rösler, A limit theorem for "Quicksort". *RAIRO: Theoret. Informatics Appl.* **25** (1991) 85-100.
[18] W. Rudin, *Real and Complex Analysis*, 2nd Ed. Tata McGraw-Hill, New Delhi (1974).
[19] N.R. Scott, *Computer Number Systems & Arithmetic.* Prentice-Hall, New Jersey (1985).
[20] R. Sedgewick and Ph. Flajolet, *An Introduction to the Analysis of Algorithms.* Addison-Wesley, Reading (1996).
[21] I. Wegener, *Effiziente Algorithmen für grundlegende Funktionen.* B.G. Teubner, Stuttgart (1996).

Communicated by I. Wegener.
Received November 9, 2000. Accepted March 8, 2001.