JOURNAL OF ALGORITHMS 8, 236–249 (1987)

Linear Probing: The Probable Largest Search Time Grows Logarithmically with the Number of Records

B. PITTEL

Department of Mathematics, The Ohio State University, Columbus, Ohio 43210

Received July 3, 1985

The worst-case performance of a linear probing algorithm is studied under assumption that each of m locations can contain at most one record. It is shown that the length of the longest sequence of probes needed to insert (locate) a record grows, in probability, as $c^{-1}\log m$ if m and n, the total number of records, approach ∞ in such a way that the load factor $a = n/m$ is bounded away from 0 and 1, here $c = a - 1 - \log a$. The argument is based on the study of the size distribution of the longest block of occupied locations. © 1987 Academic Press, Inc.

1. INTRODUCTION: RESULTS

Linear probing is a hashing technique in which a record x hashed to an already occupied cell $h(x)$ is sent to the next empty cell modulo m. ($h(\cdot)$ is a hashing function, and m is the size of the table.) This empty location is found via the sequential search beginning at the cell $h(x)$, and continuing at the cell 1 if no cell with index $\geq h(x)$ is empty. (It is convenient to think that the cells $1, \ldots, m$ form an oriented cycle.) As usual, we assume that the sequence of records $x_1, \ldots, x_n$ ($n \leq m$) and the hash function $h(\cdot)$ are such that $h(x_1), \ldots, h(x_m)$ are independent, uniformly distributed over $\{1, \ldots, m\}$. Denote T_j the random number of probes needed to insert (locate) the record x_j, $1 \leq j \leq n$. The distribution and the moments of T_j were studied in [4, 5, 6]. (A more general scheme with each location capable to accommodate more than one record was analyzed in [1].)

Our goal is to study $L_n = \max_{1 \leq j \leq n} T_j$, which is interpreted as the worst-case search time for the first n records. To this end, we introduce another random variable B_n, the length of the longest block of cells

0196-6774/87 $3.00

occupied by the records $x_1, \dots, x_n$. It should be clear that

$$L_n \le B_{n-1} + 1. \tag{1.1}$$

Suppose that $m, n \to \infty$ and that the load factor a $(= n/m)$ remains bounded away from 0 and 1. Introduce a parameter

$$c = a - 1 - \log a,$$

it is clear that $c \in (0, \infty)$ and is bounded away from 0 and ∞. We shall prove

THEOREM. *Define a random variable B_n^* by the relation*

$$B_n = c^{-1}\left[\log m - \tfrac{3}{2}\log\log m + B_n^*\right] - 1.$$

Then, B_n^* is double exponentially distributed in the limit. More precisely, if $y = O(1)$ is such that

$$b = c^{-1}\left[\log m - \tfrac{3}{2}\log\log m + y\right] - 1$$

is a positive integer then

$$P(B_n^* < y) - \exp(-\lambda e^{-y}) \to 0.$$

Here

$$\lambda = c^{3/2}(a^{-1} - 1)\left[(1 - e^{-c})(2\pi)^{1/2}\right]^{-1}. \tag{1.2}$$

Thus, in probability,

$$B_n = c^{-1}\left[\log m - \tfrac{3}{2}\log\log m\right] + O(1). \tag{1.3}$$

COROLLARY. *For every $\delta \in (\frac{1}{2}, 1)$,*

$$c^{-1}\log m + O_p(\log^\delta m) \le L_n \le c^{-1}\left[\log m - \tfrac{3}{2}\log\log m\right] + O_p(1). \tag{1.4}$$

Here $O_p(\log^\delta m), O_p(1)$ stand for random terms Δ_1, Δ_2 such that $\Delta_1/\omega(n)\log^\delta m$, $\Delta_2/\omega(n) \to 0$ in probability, for every $\omega(n) \to \infty$.

Proof. According to (1.1), we only have to prove the lower estimate. Introduce $n_1 = n - [n \log^{\delta-1} m]$, then

$$a_1 = n_1/m = a + O(\log^{\delta-1} m), \qquad c_1 = c + O(\log^{\delta-1} m),$$

and

$$c_1^{-1}\left[\log m - \tfrac{3}{2}\log\log m\right] = c^{-1}\log m + O(\log^{\delta} m).$$

By (1.3),

$$P\left(B_{n_1} < c^{-1}\log m - \gamma \log^{\delta} m\right) \to 0, \tag{1.5}$$

if γ is sufficiently large. If, on the other hand, the hash sequence $h(x_1), \dots, h(x_{n_1})$ is such that $B_{n_1} \ge c^{-1}\log m - \gamma \log^{\delta} m$, then the conditional probability that $L_n < c^{-1}\log m - 2\gamma \log^{\delta} m$ is at most

$$\left(1 - [\gamma \log^{\delta} m]/m\right)^{n-n_1} \le \exp(-2^{-1}\gamma a \log^{2\delta-1} m) = o(1). \tag{1.6}$$

(Indeed, if at least one of the records $x_{n_1+1}, \dots, x_n$ is hashed into one of $[\gamma \log^{\delta} m]$ *first* cells of a maximal block corresponding to the records $x_1, \dots, x_{n_1}$ then the number of probes for this record is at least $c^{-1}\log m - 2\gamma \log^{\delta} m$.) The relations (1.5), (1.6) together imply that, with probability approaching 1,

$$L_n \ge c^{-1}\log m - 2\gamma \log^{\delta} m.$$

How sharp are the estimates (1.4)? Using the explicit formulae for $P(T_j > k)$, [4, 5], we shall also prove that

$$L_n \le c^{-1}[\log m - (5/2)\log\log m] + O_p(1). \tag{1.7}$$

We conjecture that in this relation $\le$ can be replaced by $=$.

The reader interested in the largest search time analysis of other hashing methods is referred to [3] (uniform probing, random probing, direct chaining), [2] (direct chaining for a nonuniform distribution of records), and [8, 9] (coalesced hashing).

2. Proofs

We begin with

LEMMA 1. *For every* $b \ge 1$,

$$P(B_n < b) = [n!/m^{n-1}(m-n)]\operatorname{coeff}_{x^m}[t_b(x)]^{m-n}, \tag{2.1}$$

where

$$t_b(x) = \sum_{k=1}^{b} x^k k^{k-1}/k!.$$

Notes. (1) $t_b(x)$ is the bth partial sum of the exponential generating function of rooted trees. (2) Setting $m = \alpha$, $m - n = \beta$, and $b = \infty$ we get from (2.1) that

$$\operatorname{coeff}_{x^\alpha}[t(x)]^\beta = \beta\alpha^{\alpha-\beta-1}/(\alpha-\beta)!,$$

a relation crucial for enumeration of (labelled) trees and forests, [7].

Proof of the Lemma. Denote $S = \{1, \dots, m\}$. Let $A(s^n)$ be the set of n occupied cells if the algorithm is applied to s^n, $s^n = (x_1, \dots, s_n) \in S^n$. Since

$$P(h(x_1) = s_1, \dots, h(x_n) = s_n) = m^{-n}, \qquad \forall s^n \in S^n,$$

the problem is to evaluate the number of s^n such that $A(s^n)$ does not contain a block of size $\geq b$.

Fix $\nu < m$ and $i \leq m$. According to [4, 5], the number of s^ν such that

$$A(s^\nu) = \{i, i+1, \dots, i+\nu-1\}$$

(addition modulo m) equals

$$f(\nu) = (\nu+1)^{\nu-1}. \tag{2.2}$$

Now, given nonnegative integers $k_1, \dots, k_n$ satisfying

$$\sum_{j=1}^{n} jk_j = n, \qquad k_j = 0 \quad \text{for } j \geq b, \tag{2.3}$$

denote $H_{mn}(k)$ the number of all s^n such that $A(s^n)$ contains k_j blocks of size j, $i \leq j \leq n$. It turns out that

$$H_{mn}(k) = mn!(m-n-1)_{(|k|-1)} \prod_{j=1}^{n} [f(j)/j!]^{k_j}/k_j!, \tag{2.4}$$

where

$$|k| =_{\text{def}} \sum_{j=1}^{n} k_j, \qquad (\alpha)_\beta =_{\text{def}} \alpha(\alpha-1)\cdots(\alpha-\beta+1).$$

To prove this formula, we describe a procedure which generates all such sequences s^n. First, we partition the set $\{1, \dots, n\}$ into k_1 subsets of size 1,

k_2 subsets of size $2, \ldots,$ and k_n subsets of size n, which can be done in

$$N_1 = n! \Big/ \prod_{j=1}^{n} (j!)^{k_j} k_j!$$

ways. Given a partition $\mathscr{P}$, each of its $|k|$ subsets is assigned to a block of cells from the oriented cycle of m cells, so that the cardinality of a subset equals the size of the corresponding block, and all the blocks are disjoint. To define such an assignment we (1) select one of $N_{21} = (|k| - 1)!$ possible cyclic orders of the blocks, (2) select one of $N_{22} = m$ possible positions for the block corresponding to the subset which contains, say, the element 1, and finally (3) select the sizes of all $|k|$ gaps between the neighboring blocks. Since the total length of the gaps is $m - n$, the selection (3) can be done in the number of ways N_{23} equal to the number of integer solutions of

$$\sum_{t=1}^{|k|} \Delta_t = m - n, \qquad \Delta_t \geq 1, \quad 1 \leq t \leq |k|,$$

whence

$$N_{23} = \binom{m - n - 1}{|k| - 1}.$$

After this, it remains to select for every subset $\{\alpha_1, \ldots, \alpha_\nu\}$ of the partition $\mathscr{P}$ a sequence $s^\nu \in S^\nu$ such that $A(s^\nu)$ coincides with the block of cells that subset $\{\alpha_1, \ldots, \alpha_\nu\}$ is assigned to. Since it can be done in $f(\nu)$ ways, the total selection can be done in

$$N_3 = \prod_{j=1}^{n} [f(j)]^{k_j}$$

ways. Clearly,

$$H_{mn}(k) = N_1 \cdot (N_{21} N_{22} N_{23}) \cdot N_3,$$

which is equivalent to (2.4).

Denote D_n the random number of blocks of all sizes. By (2.4),

$$P(D_n = d, B_n < b) = m^{-n}\left[mn!(m - n - 1)_{(d-1)}\right] P(n, d, b), \quad (2.5)$$

where

$$P(n, d, b) = \sum_k \prod_{j=1}^{n} [f(j)/j!]^{k_j} \Big/ k_j!,$$

k satisfies (2.3) and such that $|k| = d$. But

$$\begin{aligned}1 + \sum_{n \geq 1,\, d \geq 1} & P(n,d,b) x^n y^d \\ &= 1 + \sum_{n \geq 1,\, d \geq 1} \sum_k \prod_{j=1}^{n} \left[y x^j f(j)/j! \right]^{k_j} \Big/ k_j! \\ &= \sum_{k_1,\dots,k_{b-1} \geq 0} \prod_{j=1}^{b-1} \left[y x^j f(j)/j! \right]^{k_j} \Big/ k_j! \\ &= \prod_{j=1}^{b-1} \left\{ \sum_{k_j \geq 0} \left[y x^j f(j)/j! \right]^{k_j} \Big/ k_j! \right\} \\ &= \exp\left[y \sum_{j=1}^{b-1} x^j f(j)/j! \right].\end{aligned}$$

Hence

$$\begin{aligned}P(n,d,b) &= \operatorname{coeff}_{x^n} \operatorname{coeff}_{y^d} \exp\left[y \sum_{j=1}^{b-1} x^j f(j)/j! \right] \\ &= \operatorname{coeff}_{x^n} \left[\sum_{j=1}^{b-1} x^j f(j)/j! \right]^d \Big/ d!. \qquad (2.6)\end{aligned}$$

So, (see (2.5)),

$$\begin{aligned}P(B_n < b) = {} & \left[n!/m^{n-1}(m-n) \right] \\ & \times \operatorname{coeff}_{x^n} \left[\sum_{d=0}^{m-n} \binom{m-n}{d} \left(\sum_{j=1}^{b-1} x^j f(j)/j! \right)^d \right],\end{aligned}$$

where, according to (2.2),

$$\begin{aligned}\sum_{d=0}^{m-n} \binom{m-n}{d} \left(\sum_{j=1}^{b-1} x^j f(j)/j! \right)^d &= \left[1 + \sum_{j=1}^{b-1} x^j (j+1)^{j-1}/j! \right]^{m-n} \\ &= x^{n-m} \left(\sum_{j=1}^{b} x^j j^{j-1}/j! \right)^{m-n}.\end{aligned}$$

The last two relations are equivalent to (2.1).

This lemma and contour integration enables us to prove

LEMMA 2. *Let $n, m \to \infty$. Suppose that an integer $b \to \infty$ in such a way that*

$$mb^{-5/2}e^{-cb} = o(1), \tag{2.7}$$

(recall that $c = a - 1 - \log a$, $a = n/m$). Then

$$P(B_n < b) \sim \exp(-\delta m\beta^{-3/2}e^{-c\beta}), \qquad \beta = b + 1,$$

where

$$\delta = (a^{-1} - 1)\left[(1 - e^{-c})(2\pi)^{1/2}\right]^{-1}. \tag{2.8}$$

Applying Lemma 2 to

$$b = c^{-1}[\log m - (3/2)\log\log m + y] - 1, \qquad y = O(1),$$

we immediately obtain the proof of theorem.

Proof of Lemma 2. From (2.1) and Cauchy's formula, it follows that

$$P(B_n < b) = \left[n!/m^{n-1}(m-n)\right](2\pi i)^{-1}I, \tag{2.9}$$

$$I = \int_C z^{-(m+1)}\left[t_b(z)\right]^{m-n}dz, \tag{2.10}$$

where C is a circle with center at $z = 0$. To estimate I, we use the saddle-point method. To this end, we select the radius of C equal to r, which is a minimum point of the function $F(x) = x^{-m}[t_b(x)]^{m-n}$, $x \in (0, \infty)$. (Since n/m is bounded away from 0 and 1, the minimum is achieved if b is sufficiently large.) So, r is a root of an equation

$$xt_b'(x)/t_b(x) = (1-a)^{-1}, \qquad a = n/m. \tag{2.11}$$

This equation has precisely one root, since by the Cauchy–Schwartz inequality

$$\frac{d}{dx}[xt_b'(x)/t_b(x)] = \left[xt_b^2(x)\right]^{-1}\left[\left(\sum_{k=1}^{b} x^k k^2 c_k\right)\right.$$
$$\left.\times\left(\sum_{k=1}^{b} x^k c_k\right) - \left(\sum_{k=1}^{b} x^k k c_k\right)^2\right] > 0,$$
$$\left(c_k = k^{k-1}/k!\right).$$

It is possible to estimate r rather sharply for large bs. Namely, it is known

[7] that the series $\Sigma_{j=1}^{\infty} x^j j^{j-1}/j!$ converges for $|x| \le e^{-1}$, and its sum $t(x)$ is the solution of an equation $t = xe^t$, so that, in particular,

$$xt'(x)/t(x) = [1 - t(x)]^{-1}, \qquad |x| \le e^{-1}. \tag{2.12}$$

In view of these relations, $\bar{r} = ae^{-a}$ is the solution of an equation

$$xt'(x)/t(x) = (1 - a)^{-1}, \tag{2.13}$$

and $t(\bar{r}) = a$. Since $t_b(x)$, $t_b'(x)$ converge (resp. to $t(x)$, $t'(x)$) uniformly in every interval $|x| \le \xi$, $\xi < e^{-1}$, and

$$[xt'(x)/t(x)]'|_{x=\bar{r}} = e^a(1 - a)^{-3} > 0,$$

we have then

$$r = \bar{r} + \Delta, \qquad \Delta = o(1), \quad (b \to \infty). \tag{2.14}$$

In the Appendix, we show how this estimate leads, via the estimations of $t(r) - t_b(r)$, $t'(r) - t'b(r)$, to

$$\Delta = \theta_1 \beta^{-1/2} \rho^\beta + \theta_2 \beta^{-3/2} \rho^\beta + O(\beta^{-5/2}\rho^\beta), \tag{2.15}$$

$$t_b(r) = a + e^a(1 - a)^{-1}\Delta + \theta_3 \beta^{-3/2}\rho^\beta + O(\beta^{-5/2}\rho^\beta), \tag{2.16}$$

here

$$\theta_1 = (1 - a)^3\left[ae^a(2\pi)^{1/2}(1 - \rho)\right]^{-1}, \tag{2.17}$$

$$\theta_2 = -\theta_1\left[12^{-1}(1 + 5\rho)(1 - \rho)^{-1} + (1 - a)^{-1}\right], \tag{2.18}$$

$$\theta_3 = -(2\pi)^{-1/2}(1 - \rho)^{-1}, \tag{2.19}$$

$\beta = b + 1$ and $\rho = er$. (Observe that $e\bar{r} = ae^{1-a} < 1$.) According to (2.14)–(2.19), we have then

$$\begin{aligned} F(r) &= r^{-m}[t_b(r)]^{m-n} = (a^{m-n}/\bar{r}^m)\exp\Big\{-m\log(1 + \Delta/r) \\ &\quad + (m - n)\log\left[1 + \Delta(1 - a)^{-1}/\bar{r} + a^{-1}\theta_3\beta^{-3/2}\rho^\beta + O(\beta^{-5/2}\rho^\beta)\right]\Big\} \\ &= (a^{m-n}/\bar{r}^m)\exp\left[a^{-1}\theta_3(m - n)\beta^{-3/2}\rho^\beta + O(m\beta^{-5/2}\rho^\beta)\right] \\ &\sim (a^{m-n}/\bar{r}^m)\exp(-\delta m\beta^{-3/2}e^{-c\beta}), \qquad c = -\log(e\bar{r}); \end{aligned} \tag{2.20}$$

(by (2.15), $\rho^\beta = (e\bar{r})^\beta + O(\beta^{1/2}(e\bar{r})^{2\beta})$).

Now, set in (2.10) $z = re^{i\phi}$, $\phi \in (-\pi, \pi]$ and write

$$I = \int_C F(z) z^{-1} dz = i \int_{-\pi}^{\pi} F(re^{i\phi}) d\phi$$

$$= i \int_{|\phi| \le \phi_0} + i \int_{|\phi| > \phi_0} = iI_1 + iI_2,$$

where $\phi_0 = m^{-5/12}$. Consider I_1. Since $F'(r) = 0$, we have

$$F(re^{i\phi}) = \exp[\log F(re^{i\phi})]$$

$$= F(r) \exp[-\alpha(r)\phi^2/2 + O(m\phi_0^3)],$$

where (see (2.10), (2.11)),

$$\alpha(r) = r^2 [\log F(z)]''|_{z=r} = m + (m-n)\left[r^2 t_b''(r)/t_b(r) - (1-a)^{-2}\right]$$

$$\sim m\left[1 + (1-a)\bar{r}^2 t''(\bar{r})/t(\bar{r}) - (1-a)^{-1}\right] = ma(1-a)^{-2}.$$

($t''(r)$ is obtained by differentiating both sides of (2.12) at $x = \bar{r}$.) Since $m\phi_0^3 \to 0$, $m\phi_0^2 \to \infty$, we conclude that

$$I_1 \sim F(r) \int_{|\phi| \le \phi_0} \exp[-\alpha(r)\phi^2/2] d\phi$$

$$\sim F(r)\left[2\pi(1-a)^2 m^{-1} a^{-1}\right]^{1/2}. \tag{2.21}$$

It remains to estimate I_2. We shall prove in the Appendix that, for all ϕ,

$$|t_b(re^{i\phi})| \le a \exp\left[r(\cos\phi - 1)/2 + O(\beta^{-3/2} e^{-c\beta})\right]. \tag{2.22}$$

Invoking (2.7) and an inequality $\cos\phi - 1 \le -\omega\phi^2, (\phi \in (-\pi, \pi]), \omega > 0$, we obtain then for $|\phi| > \phi_0$, and large enough m,

$$|t_b(re^{i\phi})| \le a \exp[-ae^{-a}\omega\phi^2 + o(m^{-1}\log m)]$$

$$\le a \exp[-ae^{-a}\omega m^{-5/6} + o(m^{-5/6})]$$

$$\le a \exp(-\omega_0 m^{-5/6}), \qquad \omega_0 > 0. \tag{2.23}$$

Using (2.23), and also (2.15), we arrive at

$$\begin{aligned}|I_2| &\le 2\pi(a^{m-n}/r^m)\exp\left[-\omega_0(1-a)m^{1/6}\right]\\ &= (a^{m-n}/\bar{r}^m)\exp\left[-\omega_0(1-a)m^{1/6} + O(m\beta^{-1/2}e^{-c\beta})\right],\end{aligned}$$

which, together with (2.7), (2.20), and (2.21), implies that

$$|I_2| = o(I_1). \tag{2.24}$$

Combination of (2.9), (2.10), (2.20), (2.21), (2.24) yields, after numerous cancellations,

$$\begin{aligned}P(B_n < b) &\sim \left[(2\pi n)^{1/2}(n/e)^n/(m^{n-1}(m-n))\right](2\pi)^{-1}\\ &\quad\times(a^{m-n}/\bar{r}^m)\left[2\pi(1-a)^2 m^{-1}a^{-1}\right]^{1/2}\\ &\quad\times\exp(-\delta m\beta^{-3/2}e^{-c\beta})\\ &= \exp(-\delta m\beta^{-3/2}e^{-c\beta}). \qquad (!)\end{aligned}$$

The lemma is proven.

Let us finally prove that

$$L_n \le c^{-1}\left[\log m - \tfrac{5}{2}\log\log m\right] + O_p(1). \tag{2.25}$$

Suppose that an integer b is such that

$$b = c^{-1}\left[\log m - \tfrac{5}{2}\log\log m + \omega(m)\right], \tag{2.26}$$

where $\omega(m) \to \infty$. Denote $n_1 = [n/2]$. By the theorem,

$$\begin{aligned}P(L_n > b) &= o(1) + P\Big(\max_{n_1 \le \nu \le n} T_\nu > b\Big)\\ &\le o(1) + \sum_{\nu=n_1}^{n} P(T_\nu > b) = o(1) + \sum_{\nu=n_1}^{n}\sum_{k=b+1}^{\nu} P(T_\nu = k).\end{aligned} \tag{2.27}$$

Here [4, 5],

$$\begin{aligned}P(T_\nu = k) &= m^{-\nu+1}\sum_{j=k}^{\nu}\binom{\nu-1}{j-1} j^{j-2}(m-j)^{\nu-1-j}(m-\nu)\\ &=_{\text{def}} \sum_{j=k}^{\nu} P(\nu, j).\end{aligned} \tag{2.28}$$

Furthermore, for $k \le j < \nu$,

$$P(\nu, j+1)/P(\nu, j) = \left(\frac{\nu - j}{m - j - 1}\right)(1 + j^{-1})^{j-1}\left(1 - (m-j)^{-1}\right)^{\nu - 1 - j}$$
$$\le \left(\frac{\nu - j}{m - j - 1}\right)\exp\left(1 - \frac{\nu - 1 - j}{m - j}\right)$$
$$= \left[\eta + O(m^{-1})\right]e^{1-\eta}, \qquad (2.29)$$

where

$$\eta = \frac{\nu - 1 - j}{m - j} \le n/m = a.$$

Since $ae^{1-a} < 1$ and bounded away from 1, (2.28) implies that

$$P(T_\nu = k) = O(P(\nu, k)),$$

and, (cf. (2.28) and the innermost sum in (2.27)),

$$P(L_n > b) = O\left(\sum_{\nu = n_1}^{n} P(\nu, b+1)\right) + o(1). \qquad (2.30)$$

Here, by Stirling's formula,

$$P(\nu, b+1) \le \text{const.}\, m^{-b}\binom{\nu - 1}{b} b^{b-1}\exp(-b\nu/m)$$
$$\le \text{const.}\, b^{-3/2} m^{-b}\exp(-b\nu/m)\frac{(\nu-1)^{\nu-1}}{(\nu - 1 - b)^{\nu - 1 - b}}$$
$$\le \text{const.}\, b^{-3/2} m^{-b}\exp[-b\nu/m + b(1 + \log \nu)]$$
$$= \text{const.}\, b^{-3/2}\exp[-bc(\nu/m)],$$

where

$$c(x) = x - 1 - \log x, \qquad x \in (0,1].$$

(Note that $c(a) = c$.) Since $c'(x) < 0$, $c''(x) > 0$ for $x \in (0,1)$,

$$\sum_{\nu = n}^{n} P(\nu, b+1) \le \text{const.}\, b^{-3/2}e^{-cb}\sum_{\nu = n_1}^{n}\exp[m^{-1}bc'(a)(n - \nu)]$$
$$\le \text{const.}\, m^{-1}b^{-5/2}e^{-cb}.$$

Hence (see (2.26), (2.30)),

$$P(L_n > b) = O(\exp(-\omega(m))) = o(1),$$

and the relation (2.25) is proven.

Appendix

Show first that uniformly in any disk $|z| \leq R$, $R < e^{-1}$,

$$\sum_{k \geq \beta} k^{k-1} z^k / k! = (2\pi)^{-1/2} \beta^{-3/2} \xi^\beta (1 - \xi)^{-1} + O(\beta^{-5/2} |\xi|^\beta), \quad (1)$$

$$\sum_{k \geq \beta} k^k z^k / k! = (2\pi)^{-1/2} \beta^{-1/2} \xi^\beta (1 - \xi)^{-1} - 12^{-1} (2\pi)^{-1/2} \beta^{-3/2} \xi^\beta (1 + 5\xi)(1 - \xi)^{-2} + O(\beta^{-5/2} |\xi|^\beta), \quad (2)$$

$\xi = ze$. Really, using the Stirling series for $\log k!$, we have

$$k^{k-1} z^k / k! = (2\pi)^{-1/2} k^{-3/2} \xi^k + O(k^{-5/2} |\xi|^k), \quad (3)$$

$$k^k z^k / k! = (2\pi)^{-1/2} k^{-3/2} \xi^k - 12^{-1} (2\pi)^{-1/2} k^{-3/2} \xi^k + O(k^{-5/2} |\xi|^k). \quad (4)$$

Together with

$$k^{-1/2} = \beta^{-1/2} - 2^{-1} \beta^{-3/2} (k - \beta) + O\left(\beta^{-5/2} (k - \beta)^2\right), \qquad k \geq \beta,$$

$$k^{-3/2} = \beta^{-3/2} + O(\beta^{-5/2} (k - \beta)), \qquad k \geq \beta,$$

the relations (3)–(4) imply (1)–(2).

Proof of (2.15)–(2.16). According to (1)–(2),

$$t_b(r) = t(r) + \omega_1 \beta^{-3/2} \rho^\beta + O(\beta^{-5/2} \rho^\beta), \quad (5)$$

$$r t_b'(r) = r t'(r) + \omega_1 \beta^{-1/2} \rho^\beta + \omega_2 \beta^{-3/2} \rho^\beta + O(\beta^{-5/2} \rho^\beta), \quad (6)$$

where

$$\omega_1 = -(2\pi)^{-1/2} (1 - \rho)^{-1}, \qquad \omega_2 = 12^{-1} (2\pi)^{-1/2} (1 + 5\rho)(1 - \rho)^{-2},$$

$\beta = b + 1$, and $\rho = er$. Further, since $r = \bar{r} + \Delta$, $\Delta = o(1)$, $\bar{r} = ae^{-a}$ and

$t(\bar{r}) = a$, we have (see (2.12–(2.13)),

$$t(r) = t(\bar{r}) + t'(\bar{r})\Delta + O(\Delta^2) = a + e^a(1-a)^{-1}\Delta + O(\Delta^2), \qquad (7)$$

$$rt'(r) = t(r)[1 - t(r)]^{-1} = a(1-a)^{-1} + e^a(1-a)^{-3}\Delta + O(\Delta^2). \quad (8)$$

Therefore (see (2.11), (5–8)),

$$\begin{aligned} a(1-a)^{-1} &+ e^a(1-a)^{-3}\Delta + O(\Delta^2) + \omega_1\beta^{-1/2}\rho^\beta + \omega_2\beta^{-3/2}\rho^\beta \\ &= a(1-a)^{-1} + e^a(1-a)^{-2}\Delta + \omega_1(1-a)^{-1}\beta^{-3/2}\rho^\beta \\ &\quad + O(\beta^{-5/2}\rho^\beta), \end{aligned}$$

which yields (see (2.17)–(2.19)),

$$\Delta = \theta_1\beta^{-1/2}\rho^\beta + \theta_2\beta^{-3/2}\rho^\beta + O(\beta^{-5/2}\rho^\beta).$$

Also, the relation (2.16) follows immediately from (5), (7), so that $\theta_3 = \omega_1$.

Proof of (2.22). Since by (1)

$$|t_b(re^{i\phi})| \le |t(re^{i\phi})| + O(\beta^{-3/2}e^{-c\beta}),$$

it suffices to show that, for $|z| \le e^{-1}$,

$$|t(z)| \le t(|z|)\exp[(\mathrm{Rl}\, z - |z|)/2],$$

(cf. [10], an inequality (2.16)). To this end, fix a number $c > 0$, and write

$$t(z) = e^{cz}[e^{-cz}t(z)] = e^{cz}\sum_{n=1}^{\infty} a_n z^n, \qquad (9)$$

where

$$a_n = \sum_{j=0}^{n-1} a_{nj}, \qquad a_{nj} = (-1)^j c^j (n-j)^{n-j-1}/j!(n-j)!.$$

Further, for $1 \le j \le n-2$,

$$\begin{aligned} |a_{n,j+1}|/|a_{nj}| &= c(j+1)^{-1}\left[1 - (n-j)^{-1}\right]^{n-j-2} \\ &\le ce^{-1}(j+1)^{-1}\left[1 - (n-j)^{-1}\right]^{-2} \le 2ce^{-1}. \end{aligned}$$

If $\chi = 2ce^{-1} < 1$ then

$$\sum_{j=1}^{n-1} |a_{nj}| \le a_{no}\chi(1-\chi)^{-1}$$

and

$$a_n \geq a_{no} - \sum_{j=1}^{n-1} |a_{nj}| \geq a_{no}\left[1 - \chi(1-\chi)^{-1}\right].$$

Since $a_{no} = n^{n-1}/n! > 0$, all the coefficients are therefore positive provided that $\chi < \frac{1}{2}$, or that $c < e/4$. Take $c = \frac{1}{2}$; as $a_n > 0$, $n = 1, 2, \ldots$, it follows from (9) that

$$\begin{aligned} |t(z)| &= |e^{z/2}| \cdot \left| \sum_{n=1}^{\infty} a_n z^n \right| \leq \exp(\text{Rl}\, z/2) \cdot \sum_{n=1}^{\infty} a_n |z|^n \\ &= \exp[(\text{Rl}\, z - |z|)/2] e^{|z|/2} \cdot \sum_{n=1}^{\infty} a_n |z|^n \\ &= \exp[(\text{Rl}\, z - |z|)/2] t(|z|). \end{aligned}$$

Acknowledgment

I am indebted to the Editor for valuable suggestions which helped me to revise the paper.

References

1. I. F. Blake and A. G. Konheim, Big buckets are (are not) better!, *J. Assoc. Comput. Mach.* **24** (1977) 591–606.
2. L. Devroye, The expected length of the longest probe sequence for bucket searching when the distribution is not uniform, *J. Algorithms* **6** (1985), 1–9.
3. G. H. Gonnet, Expected length of the longest probe sequence in hash code searching, *J. Assoc. Comput. Mach.* **28** (1981), 289–304.
4. D. E. Knuth, "The Art of Computer Programming. 3. Sorting and Searching," Addison–Wesley, Reading, Mass. 1973.
5. A. G. Konheim and B. Weiss, An occupancy discipline and applications, *SIAM J. Appl. Math.* **14** (1966), 1266–1274.
6. H. Mendelson and U. Yechiali, A new approach to the analysis of linear probing schemes, *J. Assoc. Comput. Mach.* **27** (1980), 474–483.
7. J. W. Moon, "Counting Labelled Trees," Canadian Mathematical Congress, 1970.
8. B. Pittel, Probabilistic analysis of a coalesced hashing algorithm, *Ann. Probab.*, in press.
9. B. Pittle and J. H. Yu, On search times for an early-insertion coalesced hashing, *SIAM J. Computing*, submitted.
10. B. Pittel, Paths in a random digital tree: limiting distributions, *Advan. Appl. Probab.*, **18** (1986), 139–155.