INFORMATION SCIENCES **38**, 121–146 (1986) 121

The Analysis of Simple List Structures

PHILIPPE FLAJOLET

INRIA, 78150 Rocquencourt, France

CLAUDE PUECH

LRI, Université de Paris-Sud, Bât 490, 91405 Orsay Cédex, France

and

JEAN VUILLEMIN

INRIA, 78150 Rocquencourt, France

ABSTRACT

We present an analysis of simple lists, either sorted or unsorted, under the set of all their possible histories (i.e. evolutions considered up to order isomorphism) of length n. Using the theory of continued fractions and orthogonal polynomials, Flajolet, Françon, and Vuillemin have determined average costs of sequences of operations for many data structures of the dictionary or priority queue type. We show here that for the simplest structures variance estimates can also be obtained. The method uses continued fractions and properties of nonclassical q-generalizations of Hermite and Laguerre polynomials.

1. INTRODUCTION

It was shown by Françon [7] and Flajolet, Françon and Vuillemin [3] that several list and tree organizations can be analyzed in a dynamic context. *Integrated costs* for these dynamic structures were defined as averages of costs taken over the set of all possible evolutions of the structure, considered up to order isomorphism.

This paper is concerned with methods for deriving estimates of the *variance*[1] of costs in a similar setting. Since this appears to be a harder problem, we only solve it here for the simplest dictionary and priority queue structures, namely the sorted and unsorted list structures. As in [3], our approach relies on the use

[1] Preliminary results along these lines have appeared in [2].

52 Vanderbilt Ave., New York, NY 10017 0020-0255/86/$03.50

of continued fractions and orthogonal polynomials. It introduces some of the so-called q-generalizations of Laguerre and Hermite polynomials [8], which appear to be nonclassical.

Roughly speaking, a *q-series* is a series of the form

$$\sum_{n,k\geqslant 0} a_{n,k}q^k z^n = \sum_{n\geqslant 0} a_n(q) z^n, \tag{1}$$

where $a_n(q)$ is a polynomial in q with a degree quadratic in n, in practice n^2, $n(n\pm 1)/2,\dots$. Such series regularly occur when analyzing the distribution of costs of algorithms with worst case behavior, on inputs of size n, which is $O(n^2)$. Examples include path length in trees [10, Vol. 1, p. 399], bubble sort, and the distribution of inversions in permutations, for which

$$a_n(q) = 1(1+q)(1+q+q^2)\cdots(1+q+q^2+\cdots+q^{n-1})$$

[10, Vol. 3, p. 15] as well as quicksort or tree sort [10, Vol. 3, p. 142]. We are here in a similar situation, since the total cost of a simple list structure (subjected to a sequence of n operations) ranges between $O(n)$ and $O(n^2)$. The difficulty with *q-series* is that they are not expressible as combinations of elementary functions (exp, log, etc.). Some (but few) of them are related to theta functions in the theory of elliptic functions. Since that does not appear to be the case here, we shall have to resort to indirect methods. Essentially this involves setting up a functional equation of some sort satisfied by the *q-series*. Differentiations, followed by solution of the resulting equations, give access to the moments of the probability distributions associated with the $\{a_{n,k}\}_{k\geqslant 0}$.

We now briefly review some of the definitions of [7, 3] to be used later. We are interested mainly in the analysis of dictionary (DICT) and priority queue (PQ) organizations.

DEFINITION 1. A *schema* is a word

$$\Omega = o_1 o_2 \cdots o_n \in \{I, D, Q^+, Q^-\}^*$$

such that for all j, $1 \leqslant j \leqslant n$,

$$|o_1 o_2 \cdots o_j|_I \geqslant |o_1 o_2 \cdots o_j|_D. \tag{2}$$

A schema is to interpreted as a sequence of requests (the keys operated on not being represented), where I, D, Q^+, Q^- represent respectively an insertion, a deletion, a positive query, and a negative query. The condition (2) is to be interpreted as follows: after the operations $o_1 o_2 \cdots o_j$ have been performed on the structure, the resulting size is

$$\alpha_j(\Omega) = |o_1 o_2 \cdots o_j|_I - |o_1 o_2 \cdots o_j|_D, \tag{3}$$

which should always be nonnegative. Furthermore in the case of priority queues where only insertions and deletion of the minimum are performed, we consider schemas over the restricted alphabet $\{I, D\}$.

DEFINITION 2.

(i) A *dictionary history* is a sequence of the form

$$h = o_1(r_1)\, o_2(r_2) \cdots o_n(r_n)$$

where $\Omega = o_1 o_2 \cdots o_n$ is a schema, and the r_j are integers satisfying

$$0 \leqslant r_j < \alpha_{j-1}(\Omega) \qquad \text{if} \quad o_j = Q^+ \text{ or } D,$$

$$0 \leqslant r_j \leqslant \alpha_{j-1}(\Omega) \qquad \text{if} \quad o_j = Q^- \text{ or } I.$$

(ii) A *priority queue history* is a sequence of the form

$$o_1(r_1)\, o_2(r_2) \cdots o_n(r_n)$$

where $\Omega = o_1 o_2 \cdots o_n$ is a schema over the alphabet $\{I, D\}$ only, and the r_j satisfy

$$r_j = 0 \qquad \text{if} \quad o_j = D,$$

$$0 \leqslant r_j \leqslant \alpha_{j-1}(\Omega) \qquad \text{if} \quad o_j = I.$$

Intuitively speaking, r_j represents the *rank* of the key operated on by o_j. In the case of priority queues, since deletions only operate on minimal elements in the structure, this rank should be 0 for a deletion and can be any integer between 0 and the size of the structure [i.e. $\alpha_j(h)$ at step j] for an insertion. Dictionary structures on the other hand allow the full range of insert, delete, and query operations. For instance, a DICT history is

$$I(0)\, I(0)\, I(2)\, Q^-(1)\, Q^+(1)\, D(1)\, Q^+(1)\, D(1)\, D(0),$$

and, with keys that are real numbers, it could represent (see [3] for details)

$$I(3.1)\,I(2.7)\,I(9.3)\,Q^-(2.9)\,Q^+(2.7)\,D(3.1)\,Q^+(9.3)\,D(9.3)\,D(2.7)$$

as well as

$$I(6.2)\,I(0.5)\,I(7.1)\,Q^-(1.1)\,Q^+(6.5)\,D(6.2)\,Q^+(7.1)\,D(7.1)\,D(0.5).$$

It proves convenient in the sequel, for a history

$$h = o_1(r_1)\,o_2(r_2)\cdots o_n(r_n),$$

to introduce the following notation:

(i) $\alpha_j(h) = \alpha_j(o_1 o_2 \cdots o_n)$ as defined in (3) represents the size of the structure after the first j operations have been performed.

(ii) $\rho_j(h) = r_j$ is the rank of the element operated upon at the jth stage.

(iii) $o_j(h) = o_j$ is the jth operation performed (i.e. an I or D for PQ, a Q^+, Q^-, I, or D for a DICT structure).

To any DICT or PQ data structure which operates by comparison between keys, and to any corresponding history h, is associated a cost cost(h) called the *integrated cost*, which is the sum of the costs of individual operations represented by the history. In the case of the time (τ) or storage (σ) costs for simple list structures, the integrated costs are easy to determine from the histories.

It should be made clear, in the definitions that follow, that there is some arbitrariness in the choices of our cost measures. Nevertheless, all cost measures of interest (number of keys scanned, of pointers used, etc.) bear trivial linear relations to the ones we introduce.

For both the sorted and unsorted list implementation, the (integrated) *storage* cost (σ) for a history h is

$$\sigma(h) = \sum_j \alpha_j(h). \tag{4}$$

This cost measure is the (cumulated) number of keys used in the course of the history, and holds for either a PQ or a DICT history.

The (integrated) *time* cost (τ) for a *sorted list* (SL) used as a priority queue is defined as

$$\tau\mathrm{SL-PQ}(h) = \sum \rho_j(h), \tag{5}$$

where we use the notation

$$\sum_{\omega} \varphi_j(h) = \sum_{o_j(h)=\omega} \varphi_j(h)$$

for any $\omega \in \{I, D, Q^+, Q^-\}$. The above definition is motivated by the fact that, in a sorted list, deletion of the minimum involves no comparison, and insertion in position r involves $r+1$ comparisons.

The (integrated) *time* cost for an *unsorted list* (UL) PQ-structure is similarly

$$\tau\text{UL}-\text{PQ}(h) = \sum_{D} \alpha_j(h), \tag{6.a}$$

since an insertion is now done at no cost (in terms of comparisons), and deletion of the minimum requires scanning of the whole structure. Notice that from this definition it follows that the storage cost for priority queue structures is related to the time cost of the unsorted list implementation of priority queues by

$$\sigma - \text{PQ}(h) = 2\tau\text{UL}-\text{PQ}(h) - |h|, \tag{6.b}$$

where $|h|$ denotes the *length* of the history h (i.e. the length of the associated schema).

For a *dictionary* usage, we have the analogous quantities τSL $-$ DICT and τUL $-$ DICT; τSL $-$ DICT satisfies

$$\tau\text{SL}-\text{DICT}(h) = \sum_{I} \rho_j(h) + \sum_{D} \rho_j(h) + \sum_{Q^+} \rho_j(h) + \sum_{Q^-} \rho_j(h), \tag{7}$$

but τUL $-$ DICT does not appear to have such a symmetrical expression.

Our problem can now be stated precisely: considering all histories of length n and *final altitude* 0 [i.e. such that $\alpha_n(h) = 0$], determine, for each of the cost measures above, their average value and their variance. The results will therefore represent the *analysis of time and storage costs of simple lists structures in a dynamic context*. They cover all the complexity measures defined so far with the sole exception of τUL $-$ DICT.[2]

The plan of the paper is as follows: in Section 2, we recall the continued fraction theorem of [4, 5, 3] and we use it to express, as continued fractions,

[2]We thus have two measures for integrated storage and three for integrated time. Because of the remark (6.b), the number of our analyses is ultimately reduced to four.

generating functions associated to some of the cost measures above. In Section 3, we show how to combine these expressions with properties of q-Laguerre and q-Hermite polynomials in order to attain variance estimates for various cost measures. Finally, we briefly indicate in Section 4 some connections with related combinatorial works.

2. CONTINUED FRACTION AND ORTHOGONAL POLYNOMIAL FORMS FOR GENERATING FUNCTIONS

2.1. *CONTINUED FRACTIONS*

It has been shown [4, 5] that the characteristic series for histories can be expressed under the form of a continued fraction. Let $o_i^{(j)}$ denote the operation o performed on a structure of size i with rank j. In the case of dictionaries, histories are then described by words over the alphabet

$$^{\text{DICT}}X = \left\{ I_i^{(j)}, Q_i^{-(j)} \,|\, 0 \leqslant j \leqslant i,\ i \in N \right\} \cup \left\{ D_i^{(j)}, Q_i^{+(j)} \,|\, 0 \leqslant j < i,\ i \in N \right\},$$

and in the case of priority queues, the alphabet reduces to

$$^{\text{PQ}}X = \left\{ I_i^{(j)} \,|\, 0 \leqslant j \leqslant i,\ i \in N \right\} \cup \left\{ D_i^{(0)} \,|\, i \in N \right\}.$$

One has [4, 5]:

THEOREM 1. *The characteristic series for the set of histories with initial and final altitude* 0,

$$\text{char}(H) = \sum h$$

(*where the sum is taken on the set of all such histories*), *is given by the continued fraction*

$$\text{char}(H) = \cfrac{1}{1 - \Sigma Q_0^{+(j)} - \Sigma Q_0^{-(j)} - \cfrac{\Sigma I_0^{(j)} \Sigma D_1^{(j)}}{1 - \Sigma Q_1^{+(j)} - \Sigma Q_1^{-(j)} - \cfrac{\Sigma I_1^{(j)} \Sigma D_2^{(j)}}{\cdots}}}$$

where the sums $\Sigma o_i^{(j)}$ *are taken for all values of* j *such that* $o_i^{(j)}$ *is in the alphabet* X *of the type of histories considered.*

In particular, replacing each $o_i^{(j)}$ by a single variable leads, on the left, to the series $\Sigma H_n z^n$ (with H_n the number of histories of length n), since z^n appears as many times as there are histories of length n; the right hand side is transformed into a continued fraction of the so-called Jacobi type. In the case of dictionaries and priority queues, we have [4, 5, 3]:

COROLLARY 1. *The generating series relative to dictionary histories and priority queue histories defined by*

$$^{\text{DICT}}H(z) = \sum_{n \geq 0} {}^{\text{DICT}}H_n z^n,$$

$$^{\text{PQ}}H(z) = \sum_{n \geq 0} {}^{\text{PQ}}H_n z^n$$

are expressible as

$$^{\text{DICT}}H(z) = \cfrac{1}{1 - 1z - \cfrac{1^2 z^2}{1 - 3z - \cfrac{2^2 z^2}{1 - 5z - \cfrac{3^2 z^2}{\cdots}}}},$$

$$^{\text{PQ}}H(z) = \cfrac{1}{1 - \cfrac{1z^2}{1 - \cfrac{2z^2}{1 - \cfrac{3z^2}{\cdots}}}}.$$

From this corollary, using identities of Euler and Gauss, it can be proved that

$$^{\text{DICT}}H_n = n!, \tag{8}$$

$$^{\text{PQ}}H_{2n} = 1 \times 3 \times 5 \times \cdots \times (2n-1). \tag{9}$$

Theorem 1 can also be applied to derive information on the distribution of storage cost in simple list structures. With the storage cost defined in (7), we introduce

$$^{\text{DICT}}S_{n,k} = \text{card}\{h \mid |h| = n,\ \sigma(h) = k,\ h \in \text{DICT}-\text{history}\}, \tag{10}$$

$$^{\text{PQ}}S_{n,k} = \text{card}\{h \mid |h| = n,\ \sigma(h) = k,\ h \in \text{PQ}-\text{history}\}. \tag{11}$$

We have:

COROLLARY 2. *The generating functions associated to the distribution of memory costs,*

$$^{\text{DICT}}S(z,q) = \sum_{k,n \geq 0} {}^{\text{DICT}}S_{n,k} q^k z^n,$$

$$^{\text{PQ}}S(z,q) = \sum_{k,n \geq 0} {}^{\text{PQ}}S_{n,k} q^k z^n,$$

satisfy

$$^{\text{DICT}}S(z,q) = \cfrac{1}{1 - 1q^0 z - \cfrac{1^2 q^1 z^2}{1 - 3q^1 z - \cfrac{2^2 q^3 z^2}{1 - 5q^2 z - \cfrac{3^2 q^5 z^2}{\cdots}}}}, \tag{12}$$

$$^{\text{PQ}}S(z,q) = \cfrac{1}{1 - \cfrac{1q^1 z^2}{1 - \cfrac{2q^3 z^2}{1 - \cfrac{3q^5 z^2}{\cdots}}}}. \tag{13}$$

Proof. The same proof applies for both cases. We start from Theorem 1, and substitute for each letter $o_i^{(j)}$ the value zq^j. Substitution on the right hand side leads to the two continued fractions above. Substitution on a term h in char(H) obviously leads to $z^{|h|}q^{\sigma(h)}$. Thus char(H) is changed into

$$\sum_{h \in H} z^{|h|} q^{\sigma(h)} = \sum S_{n,k} q^k z^n.$$

We now introduce, for each integer m, the notation

$$[m] = [m]_q = 1 + q + q + \cdots + q^{m-1} = \frac{1-q^m}{1-q}.$$

The quantity $[m]_q$ is called the q-generalization of integer m (to which it reduces when $q = 1$). For the time cost of sorted lists, in analogy with (10), (11), we define

$$^{\text{DICT}}T_{n,k} = \text{card}\{h \,|\, |h| = n,\ \tau(h) = k,\ h \in \text{DICT}-\text{history}\}, \tag{14}$$

$$^{\text{PQ}}T_{n,k} = \text{card}\{h \,|\, |h| = n,\ \tau(h) = k,\ h \in \text{PQ}-\text{history}\}, \tag{15}$$

and, with the above notation, we can state:

COROLLARY 3. *The generating series for the distribution of comparison cost in sorted lists defined by*

$$^{\mathrm{DICT}}T(z,q)=\sum_{k,n\geqslant 0}{}^{\mathrm{DICT}}T_{n,k}q^k z^n,$$

$$^{\mathrm{PQ}}T(z,q)=\sum_{k,n\geqslant 0}{}^{\mathrm{PQ}}T_{n,k}q^k z^n$$

are given by

$$^{\mathrm{DICT}}T(z,q)=\cfrac{1}{1-([0]+[1])z-\cfrac{[1]^2z^2}{1-([1]+[2])z-\cfrac{[2]^2z^2}{1-([2]+[3])z-\cfrac{[3]^2z^2}{\cdots}}}},\tag{16}$$

$$^{\mathrm{PQ}}T(z,q)=\cfrac{1}{1-\cfrac{[1]z^2}{1-\cfrac{[2]z^2}{1-\cfrac{[3]z^2}{\cdots}}}}.\tag{17}$$

Proof. It follows similarly from Theorem 1, using the substitution

$$o_i^{(j)}\to zq^j.$$

2.2. *ORTHOGONAL POLYNOMIALS*

All the continued fractions considered above are of the form

$$J(z)=\cfrac{1}{1-\kappa_0z-\cfrac{\lambda_1z^2}{1-\kappa_1z-\cfrac{\lambda_2z^2}{1-\kappa_2z-\cfrac{\lambda_3z^2}{\cdots}}}}.$$

The hth convergent $J^{[h]}(z)$ is defined as the finite fraction

$$J^{[h]}(z)=\cfrac{1}{1-\kappa_0 z-\cfrac{\lambda_1 z^2}{1-\kappa_1 z-\cfrac{\lambda_2 z^2}{\cfrac{\cdots}{1-\kappa_h z}}}}$$

and is associated with histories of bounded height. The hth convergent can be put under the form of a quotient of two polynomials in z (with q as a parameter in the case of the fractions of Corollaries 2, 3):

$$J^{[h]}(z)=\frac{P_h(z)}{Q_h(z)}.$$

Finally, as is well known, the P and Q polynomials satisfy linear recurrences; in the case of Q, for instance, we have

$$Q_{-1}(z)=1,\quad Q_0(z)=1-\kappa_0 z,\qquad Q_h(z)=(1-\kappa_h z)Q_{h-1}-\lambda_h z^2 Q_{h-2}. \tag{18}$$

Let J_n be the coefficient of z^n in J. Define a scalar product $\langle\cdot|\cdot\rangle_J$ (relative to J) on the set of polynomials in z, and a corresponding linear form $\langle\cdot\rangle_J$, by

$$\langle z^m|z^n\rangle_J=\langle z^{m+n}\rangle_J=J_{m+n}.$$

The following classical result is crucial to our treatment [14, 11]:

THEOREM 2. *The polynomials* $\overline{Q}_k$ *defined by* $\overline{Q}_{k-1}=z^k Q_{k-1}(1/z)$ *are orthogonal with respect to* $\langle\cdot|\cdot\rangle_J$:

$$\langle\overline{Q}_{k-1}|\overline{Q}_{l-1}\rangle=0\qquad\text{if}\quad 0\leqslant k\neq l,$$

$$\langle\overline{Q}_{k-1}|\overline{Q}_{k-1}\rangle=\lambda_1\lambda_2\cdots\lambda_k\qquad\text{for}\quad 0\leqslant k. \tag{19}$$

Corresponding to the four continued fractions contained in Corollary 2 and Corollary 3, there will thus be associated four families of polynomials and four scalar products. When $q=1$, the fractions associated to the distribution of costs in Corollaries 2, 3 reduce to those of Corollary 1, whose convergents have been expressed [3] in terms of Laguerre and Hermite polynomials respectively. We thus have, corresponding to fractions (12),(13) and (16),(17), q-generalizations of these polynomials which do not appear to be classical.

3. THE DISTRIBUTION OF INTEGRATED COSTS IN SORTED AND UNSORTED LISTS

Returning, for instance, to the storage cost of PQ-structures, we see that

$$\frac{{}^{\text{PQ}}S_{n,k}}{{}^{\text{PQ}}H_n} \tag{20}$$

is the probability, for a history of length n, to have (integrated) storage equal to k. Thus the expected storage utilization of a simple list subjected to n operations (dropping PQ-superscripts) is:

$$\sum_k k\frac{S_{n,k}}{H_n}, \tag{21.1}$$

and the variance is

$$\sum_k k^2\frac{S_{n,k}}{H_n} - \left(\sum_k k\frac{S_{n,k}}{H_n}\right)^2. \tag{21.2}$$

In general, for $\mu = S$ (storage) or T (time), and $\Sigma = \text{DICT}$ or PQ, we define the unnormalized moments of distributions of the type (20) by

$${}^{\Sigma}M_{\mu}^{(j)} = \sum_k k^{j\,\Sigma}\mu_{n,k}, \tag{22}$$

with particular attention to the cases $j = 0,1,2$ related to averages and variances as in (21.1), (21.2).

3.1. *OUTLINE OF THE METHOD*

In order to evaluate the moments (22), we introduce the generating polynomials

$${}^{\Sigma}\mu_n(q) = \sum_k {}^{\Sigma}\mu_{n,k}q^k. \tag{23}$$

(the superscript Σ will be omitted in the sequel for simplicity).

The moments of the distribution can be calculated simply once we know the quantities $\mu_n(1), \mu_n'(1), \mu_n''(1), \ldots$. The method for determining the $\mu_n^{(k)}(1)$ in each case obeys the following general pattern:

(a) We start with the continued fraction expansions previously derived. With the reciprocals of the convergent polynomials given by equation (18), we form a *generating function* of the form

$$K(t,z,q)=\sum_{k\geqslant 0}\overline{Q}_{k-1}(z,q)\,\omega_k(q)\,t^k \tag{24}$$

with an appropriate choice of the coefficients $\omega_k(q)$.

(b) We next let

$$K(t,z,q)=\sum_{n\geqslant 0}R_n(t,q)\,z^n \tag{25}$$

be the expansion of K with respect to z. Applying the linear form $\langle\cdot\rangle$ associated to the continued fraction to both sides, we derive the *orthogonality relation*

$$\sum_{n\geqslant 0}R_n(t,q)\,\mu_n(q)=1, \tag{26}$$

whence, by letting $q=1$ and differentiating with respect to q, for all $k\geqslant 1$,

$$\sum_{n\geqslant 0}R_n(t,1)\,\mu_n(1)=1, \tag{27}$$

$$\sum_{n\geqslant 0}R_n^{(k)}(1)=-\sum_{\substack{1\leqslant j<k\\ n\geqslant 0}}\binom{k}{j}\frac{\partial^{k-j}R_n}{\partial q^{k-j}}(t,1)\cdot\mu_n^{(j)}(1). \tag{28}$$

These relations provide a sort of generating function of the $\mu_n^{(j)}(1)$ once the derivatives of $R_n(t,q)$ with respect to q at $q=1$ have been determined. Although we do not possess closed form expressions for the $R_n(t,q)$ themselves, it turns out that the derivatives at $q=1$ can be determined explicitly, as the continuation of our treatment shows.

(c) The recurrence relation (18) satisfied by the polynomials $\overline{Q}$ yields a *functional relation* verified by $K(t,z,q)$; the choice of $\omega_k(q)$ in (a) is made so as to simplify this relation. This is normally in the form of a *difference-differential equation*.

(d) From this functional relation, we derive a set of *differential equations* satisfied by K and its derivatives with respect to q at $q=1$:

$$L[K(t,z,1)]=0, \tag{29}$$

$$L\left[\frac{\partial^k}{\partial q^k}K(t,z,1)\right]=\Lambda_k\left(K(t,z,1),\frac{\partial K}{\partial q}(t,z,1),\ldots,\frac{\partial^{k-1}K}{\partial q^{k-1}}(t,z,1)\right), \tag{30}$$

where L is some first order linear partial differential operator characteristic of the data type considered, and the Λ_k are linear partial differential operators. Using the initial conditions at $t=0$, we successively determine

$$K(t,z,1),\quad \frac{\partial K}{\partial q}(t,z,1),\ldots,\quad \frac{\partial^k K}{\partial q^k}(t,z,1),$$

from which the expressions for the derivatives of the $R_n(t,q)$ at $q=1$ follow (in the form of generating functions which are actually sufficient for our purposes).

(e) Once in possession of the $R_n(t,1),\ldots,$ we use the orthogonality relations of part (b) to determine generating functions for the $\mu_n^{(j)}(1)$, from which explicit forms for the moments are derived by performing standard *Taylor expansions*.

3.2. *PRIORITIES QUEUES: THE INTEGRATED TIME OF SORTED LIST IMPLEMENTATIONS*

We shall present the method in some detail in the case when Σ is the priority queue data type, and μ is the time cost.

(a) We start with the continued fraction expansion of Corollary 3:

$$\cfrac{1}{1-\cfrac{[1]z^2}{1-\cfrac{[2]z^2}{1-\cfrac{[3]z^2}{\cdots}}}},$$

so that the corresponding Q_h polynomials satisfy for $h \geqslant 1$

$$Q_{-1}=1,\quad Q_0=1,\qquad Q_h=Q_{h-1}-[h]z^2Q_{h-2}.$$

Equivalently, for the reciprocal polynomials

$$\overline{Q}_{h-1}=z^hQ_{h-1}\left(\frac{1}{z}\right),$$

we have for $h \geqslant 1$

$$\overline{Q}_{-1}=1,\quad \overline{Q}_0=1,\qquad \overline{Q}_h=\overline{Q}_{h-1}-[h]z^2\overline{Q}_{h-2}.$$

We now introduce the following generating function of the polynomials $\overline{Q}$:

$$K(t,z,q) = \sum_{h \geqslant 0} \overline{Q}_{h-1}(z,q) \frac{t^h}{[1][2]\cdots[h]}$$

(b) Let us rewrite the generating function of the polynomials $\overline{Q}$:

$$K(t,z,q) = \sum_{n \geqslant 0} R_n(t,q) z^n.$$

Applying the linear form $\langle \cdot \rangle$ to both expressions for K, we obtain

$$\sum_{h \geqslant 0} \langle \overline{Q}_{h-1}(z,q) \rangle \frac{t^h}{[1][2]\cdots[h]} = \sum_{n \geqslant 0} R_n(t,q) \langle z^n \rangle.$$

Since we have

$$\langle z^n \rangle = [z^n]^{PQ}T = \sum_{k \geqslant 0} T_{n,k} q^k = T_n(q)$$

by definition of the linear form,[3] and

$$\langle \overline{Q}_{h-1} \rangle = \langle \overline{Q}_{h-1}, 1 \rangle = 0 \qquad \text{if} \quad h > 0$$

(by virtue of Theorem 2 and $\langle \overline{Q}_{-1} \rangle = 1$), the last equality reduces to

$$\sum_{n \geqslant 0} R_n(t,q) T_n(q) = 1.$$

This implies

$$\sum_{n \geqslant 0} R_n(t,1) T_n(1) = 1, \tag{31}$$

$$\sum_{n \geqslant 0} R_n(t,1) T_n'(1) = -\sum_{n \geqslant 0} R_n'(t,1) T_n(1), \tag{32}$$

$$\sum_{n \geqslant 0} R_n(t,1) T_n''(1) = -2 \sum_{n \geqslant 0} R_n'(t,1) T_n'(1) - \sum_{n \geqslant 0} R_n''(t,1) T_n(1). \tag{33}$$

[3] We denote as usual by $[z^n]f(z)$ the coefficient of z^n in the Taylor expansion of f.

(c) From Equation (23), we deduce the functional relation:

$$K(qt,z,q)-K(t,z,q)=(q-1)t(z-t)K(t,z,q).$$

Denoting by D the *Hahn difference operator*

$$D:f\to\frac{f(tq)-f(t)}{t(q-1)},$$

this relation takes the simpler form

$$DK=(z-t)K. \tag{34}$$

(d) By successive differentiations of Equation (34), we obtain, writing here $K(t)$ for $K(t,z,q)$,

$$t\frac{\partial K}{\partial t}(qt)+\frac{\partial K}{\partial q}(qt)-\frac{\partial K}{\partial q}(t)=t(z-t)K(t)+(q-1)t(z-t)\frac{\partial K}{\partial q}(t),$$

$$t^2\frac{\partial^2 K}{\partial t^2}(qt)+2t\frac{\partial^2 K}{\partial t\,\partial q}(qt)+\frac{\partial^2 K}{\partial q^2}(qt)-\frac{\partial^2 K}{\partial q^2}(t)$$

$$=2t(z-t)\frac{\partial K}{\partial q}(t)+(q-1)t(z-t)\frac{\partial^2 K}{\partial q^2}(t),$$

$$t^3\frac{\partial^2 K}{\partial t^2}(qt)+3t^2\frac{\partial^3 K}{\partial t^2\,\partial q}(qt)+3t\frac{\partial^3 K}{\partial t\,\partial q^2}+\frac{\partial^3 K}{\partial q^3}(qt)-\frac{\partial^3 K}{\partial q^3}(t)$$

$$=3t(z-t)\frac{\partial^2 K}{\partial q^2}(t)+(q-1)t(z-t)\frac{\partial^3 K}{\partial q^3}(t),$$

from which we derive, after letting $q=1$ [K and its derivatives are to be taken here at $(t,z,1)$],

$$\frac{\partial K}{\partial t}-(z-t)K=0, \tag{35}$$

$$\frac{\partial}{\partial t}\left(\frac{\partial K}{\partial q}\right)-(z-t)\frac{\partial K}{\partial q}=-\frac{t}{2}\frac{\partial^2 K}{\partial t^2}, \tag{36}$$

$$\frac{\partial}{\partial t}\left(\frac{\partial^2 K}{\partial q^2}\right)-(z-t)\frac{\partial^2 K}{\partial q^2}=-t\frac{\partial^3 K}{\partial t^2\,\partial q}-\frac{t^2}{3}\frac{\partial^3 K}{\partial t^3}. \tag{37}$$

Let L be the first order linear partial differential operator

$$L: f \to \frac{\partial f}{\partial t} - (z - t) f.$$

Since $K(0, z, 1) = \overline{Q}_{-1}(z, 1) = 1$, the function $K(t, z, 1)$ is the solution of the equation $Lu = 0$ which satisfies the initial condition $u(0) = 0$, i.e.

$$K(t, z, 1) = e^{zt - t^2/2},$$

and thus the polynomials $\overline{Q}_k(z, 1)$ are *Hermite* polynomials. From the differential system (35),(36),(37), we get

$$R_n(t,1), \quad \frac{\partial R_n}{\partial q}(t,1), \quad \frac{\partial^2 R_n}{\partial q^2}(t,1),$$

which are the coefficients of z^n in

$$K(t,z,1), \quad \frac{\partial K}{\partial q}(t,z,1), \quad \frac{\partial^2 K}{\partial q^2}(t,z,1).$$

We find[4]

$$R_n(t,1) = e^{-t^2/2} \frac{t^n}{n!},$$

$$\frac{\partial R_n}{\partial q}(t,1) = e^{-t^2/2}\left[-\frac{t^2}{4} \frac{t^{n-2}}{(n-2)!} + \frac{t^3}{3} \frac{t^{n-1}}{(n-1)!} + \left(-\frac{t^4}{8} + \frac{t^2}{4} \right) \frac{t^n}{n!} \right],$$

$$\frac{\partial^2 R_n}{\partial q^2}(t,1) = e^{-t^2/2}\left[\frac{t^4}{16} \frac{t^{n-4}}{(n-4)!} + \left(-\frac{t^5}{6} + \frac{2t^3}{9} \right) \frac{t^{n-3}}{(n-3)!} \right.$$
$$+ \left(\frac{25t^6}{144} - \frac{5t^4}{8} + \frac{t^2}{4} \right) \frac{t^{n-2}}{(n-2)!} + \left(-\frac{t^7}{12} + \frac{17t^5}{30} - \frac{2t^3}{3} \right)$$
$$\left. \times \frac{t^{n-1}}{(n-1)!} + \left(\frac{t^8}{24} - \frac{25t^6}{144} + \frac{7t^4}{16} - \frac{t^2}{4} \right) \frac{t^n}{n!} \right].$$

[4] The computations that follow have been developed with the help of the MACSYMA system for symbolic calculations.

(e) Equations (31),(32),(33) now enable us to calculate

$$\sum_{n\geqslant 0} T_n(1)\frac{t^n}{n!},\quad \sum_{n\geqslant 0} T_n'(1)\frac{t^n}{n!} \text{ and } \sum_{n\geqslant 0} T_n''(1)\frac{t^n}{n!}.$$

We have

$$\sum_{n\geqslant 0} T_n(1)\frac{t^n}{n!} = e^{t^2/2},$$

$$\sum_{n\geqslant 0} T_n'(1)\frac{t^n}{n!} = \frac{t^4}{24}e^{t^2/2},$$

$$\sum_{n\geqslant 0} T_n''(1)\frac{t^n}{n!} = \left(\frac{t^8}{576}+\frac{t^6}{60}\right)e^{t^2/2}.$$

Performing standard Taylor expansions of these functions, we get

$$T_{2n}(1) = 1.3.5\cdots(2n-1) = n? \qquad [T_{2n+1}(1)=0],$$

$$T_{2n}'(1) = n?\frac{n(n-1)}{6},$$

$$T_{2n}''(1) = n?n(n-1)(n-2)\frac{5n+9}{180}.$$

Combining these expressions of the factorial moments of the distribution of integrated time, we obtain:

THEOREM A. *The distribution of the integrated time of* $2n$ *operations performed on a sorted list implementations of priority queues has mean*

$$M = \frac{n(n-1)}{6}$$

and variance

$$V = \frac{n(n-1)(n+3)}{45}.$$

3.3. *PRIORITY QUEUES: TOTAL STORAGE UTILIZATION OF GENERAL LIST STRUCTURES AND THE INTEGRATED TIME OF UNSORTED LISTS*

The continued fraction relative to the storage utilization of PQ-structures has been given in Corollary 2; it is

$$J(z,q) = \cfrac{1}{1 - \cfrac{1q^1z^2}{1 - \cfrac{2q^3z^2}{1 - \cfrac{3q^5z^2}{\cdots}}}}.$$

We propose to work instead with the modified fraction

$$J_1(z,q) = \cfrac{1}{1 - \cfrac{1q^1z^2}{1 - \cfrac{2q^2z^2}{1 - \cfrac{3q^3z^2}{\cdots}}}},$$

which satisfies

$$J(z,q) \equiv J_1(zq^{-1/2}, q^2).$$

By (6.a), J_1 is nothing but the generating function of the distribution of integrated time costs of unsorted lists.

The convergents of J_1 are such that

$$\overline{Q}_h = z\overline{Q}_{h-1} - hq^h\overline{Q}_{h-2}.$$

As a consequence

$$K(t,z,q) = \sum_{h \geqslant 0} \overline{Q}_{h-1}(z,q)\frac{t^h}{h!}$$

is a solution of

$$\frac{\partial K}{\partial t} = zK - qtK(qt).$$

From that equation we derive that:

(i) $K(t,z,1)$ is the solution of

$$Lu = 0, \qquad u(0) = 1,$$

where L is the differential operator

$$Lu = \frac{\partial u}{\partial t} - (z-t)u.$$

(ii) $\frac{\partial K}{\partial q}(t,z,1)$ is the solution of

$$Lu = (-t^2z + t^3 - t)K; \qquad u(0) = 0$$

(iii) $\frac{\partial^2 K}{\partial q^2}(t,z,1)$ is the solution with $u=0$ of

$$Lu = \left[\left(\tfrac{2}{3}t^5 - t^3\right)z^2 + \left(-\tfrac{7}{6}t^6 + \tfrac{17}{3}t^4 - 2t^2\right)z + \tfrac{1}{2}t^7 - \tfrac{9}{2}t^5 + 6t^3\right]K.$$

Hence, solving these equations, we obtain

$$R_n(t,1) = [z^n]K(t,z,1) = e^{-t^2/2}\frac{t^n}{n!}$$

$$\frac{\partial R_n}{\partial q}(t,1) = [z^n]\frac{\partial K}{\partial q}(t,z,1)$$

$$= e^{-t^2/2}\frac{t^n}{n!}\left[-\frac{t^3}{3}\frac{t^{n-1}}{(n-1)!} + \left(\frac{t^4}{4} - \frac{t^2}{2}\right)\frac{t^n}{n!}\right]$$

$$\frac{\partial^2 R_n}{\partial q^2}(t,1) = [z^n]\frac{\partial^2 K}{\partial q^2}(t,z,1)$$

$$= e^{-t^2/2}\frac{t^n}{n!}\left[\left(\frac{t^6}{9} - \frac{t^4}{4}\right)\frac{t^{n-2}}{(n-2)!}\right.$$

$$\left. + \left(-\frac{t^7}{6} + \frac{17t^5}{15} - \frac{2t^3}{3}\right)\frac{t^{n-1}}{(n-1)!} + \left(\frac{t^8}{16} - \frac{3t^6}{4} + \frac{3t^4}{2}\right)\frac{t^n}{n!}\right],$$

from which we deduce that the exponential generating function of the distribution of total storage utilization, and those of its first two moments, are the following:

$$\sum_{n\geqslant 0} S_n(1)\frac{t^n}{n!}=\frac{e^{t^2}}{2}.$$

$$\sum_{n\geqslant 0} S_n'(1)\frac{t^n}{n!}=\left(\frac{t^4}{12}+\frac{t^2}{2}\right)\frac{e^{t^2}}{2}$$

$$\sum_{n\geqslant 0} S_n''(1)\frac{t^n}{n!}=\left(\frac{1}{144}t^8+\frac{13}{90}t^6+\frac{7}{12}t^4\right)\frac{e^{t^2}}{2}.$$

By identification of left and right hand sides of these equations, we get

$$S_{2n}(1)=1\times 3\times 5\times\cdots\times(2n-1)\equiv n? \qquad [S_{2n+1}(1)=0],$$

$$S_{2n}'(1)=n?\frac{n(n+2)}{3},$$

$$S_{2n}''(1)=n?\frac{n(n-1)(5n^2+27n+31)}{45}.$$

THEOREM B. *The distribution of the integrated time cost of* $2n$ *operations performed on an unsorted list implementation of priority queues has mean and variance given by*

$$M=\frac{n(n+2)}{3},$$

$$V=\frac{n(n-1)(2n+1)}{45}.$$

Using the remarks at the beginning of this section, we get immediately:

THEOREM B′. *The integrated storage* $2n$ *operations of a list implementation of priority queues has mean and variance given by*:

$$M=\frac{n(2n-1)}{3},$$

$$V=\frac{4n(n-1)(2n+1)}{45}.$$

3.4. DISTRIBUTION OF THE NUMBER OF COMPARISONS BETWEEN KEYS UNDER SEQUENCES OF OPERATIONS FOR SORTED LIST IMPLEMENTATIONS OF DICTIONARIES

The continued fraction expansion of $\Sigma T_n(q) z^n$ is

$$\cfrac{1}{1-([0]+[1])z-\cfrac{[1]^2 z^2}{1-([1]+[2])z-\cfrac{[2]^2 z^2}{\cdots}}}.$$

The reciprocal polynomials of the denominators of the convergents are such that

$$Q_h = \{ z-(2[h]+q^h)\} Q_{h-1} - [h]^2 Q_{h-2}.$$

As a consequence,

$$K(t,z,q) = \sum_{h \geqslant 0} \overline{Q}_{h-1}(z,q) \frac{t^h}{h!}$$

is a solution of

$$(1+t)(1+qt) DK = (z-1-t) K$$

(where D is the Hahn operator). From that equation, we derive that:

(i) $K(t,z,1)$ is the solution of

$$L_1 u = 0, \qquad u(0) = 1,$$

where

$$L_1 u = (1+t)^2 \frac{\partial u}{\partial t} - (z-1-t) u.$$

(ii) $\dfrac{\partial K}{\partial q}(t,z,1)$ is the solution of

$$L_1 u = -\frac{t}{2}(1+t)^2 \frac{\partial^2 K}{\partial t^2} - t(1+t)\frac{\partial K}{\partial t}, \qquad u(0) = 0.$$

(iii) $\dfrac{\partial^2 K}{\partial q^2}(t,z,1)$ is the solution, with $u(0)=0$, of

$$L_1 u = -t(1+t)^2 \frac{\partial^3 K}{\partial t^2 \partial q} - 2t(1+t)\frac{\partial^2 K}{\partial t \partial q} - \frac{t^2(1+t)^2}{3}\frac{\partial^3 K}{\partial t^3} - t^2(1+t)\frac{\partial^2 K}{\partial t^2}$$

In order to obtain, in a later step of the calculations, the exponential generating function of the time costs involved, it is convenient to make the change of variable $v = t/(1+t)$. As a function of v, K is such that:

(i) $K(v,z,1)$ is the solution of

$$Lu = 0, \qquad u(0) = 1,$$

where

$$Lu = \frac{\partial u}{\partial v} - \left(z - \frac{1}{1-v}\right)u.$$

(ii) $\frac{\partial K}{\partial q}(v,z,1)$ is the solution of

$$Lu = \left[\left(\frac{v^2}{2} - \frac{v}{2}\right)z^2 + vz\right]K, \qquad u(0) = 0.$$

(iii) $\frac{\partial^2 K}{\partial q^2}(v,z,1)$ is the solution, with $u = 0$, of

$$Lu = \left[\left(\frac{v^5}{6} - \frac{5v^4}{12} + \frac{v^3}{4}\right)z^4 + \left(\frac{3v^4}{2} - \frac{7v^3}{3} + \frac{2v^2}{3}\right)z^3 + \left(3v^3 - \frac{5v^2}{2} + \frac{v}{2}\right)z^2 + (v^2 - v)z\right]K.$$

These relations show that the polynomials $\overline{Q}_k(z,1)$ are *Laguerre* polynomials. From the above equations, we get

$$R_n(v,1) = [z^n]K(v,z,1) = (1-v)\frac{v^n}{n!}$$

$$\frac{\partial R_n}{\partial q}(v,1) = [z^n]\frac{\partial K}{\partial q}(v,z,1) = (1-v)\left[\left(\frac{v^3}{6} - \frac{v^2}{4}\right)\frac{v^{n-2}}{(n-2)!} + \frac{v^2}{2}\frac{v^{n-1}}{(n-1)!}\right]$$

$$\frac{\partial^2 R_n}{\partial q^2}(v,1) = [z^n]\frac{\partial^2 K}{\partial q^2}(v,z,1)$$

$$= (1-v)\left[\left(\frac{v^6}{36} - \frac{v^5}{12} + \frac{v^4}{16}\right)\frac{v^{n-4}}{(n-4)!} + \left(\frac{3v^5}{10} - \frac{7v^4}{12} + \frac{2v^3}{9}\right)\frac{v^{n-3}}{(n-3)!} + \left(\frac{3v^4}{4} - \frac{5v^3}{6} + \frac{v^2}{4}\right)\frac{v^{n-2}}{(n-2)!} + \left(\frac{v^3}{3} - \frac{v^2}{2}\right)\frac{v^{n-1}}{(n-1)!}\right].$$

From these, we can derive the following expansions:

$$\sum_{n\geq 0} T_n(1)\frac{v^n}{n!} = \frac{1}{1-v}$$

$$\sum_{n\geq 0} T_n'(1)\frac{v^n}{n!} = -\frac{1}{6} + \frac{1}{2(1-v)} - \frac{1}{2(1-v)^2} + \frac{1}{6(1-v)^3}$$

$$\sum_{n\geq 0} T_n''(1)\frac{v^n}{n!} = \frac{2}{15} + \frac{v}{30} - \frac{1}{6(1-v)} - \frac{1}{3(1-v)^2} + \frac{5}{6(1-v)^3} - \frac{19}{30(1-v)^4} + \frac{1}{6(1-v)^5}.$$

This leads to the explicit forms

$$T_{2n}(1) = n! \qquad [T_{2n+1}(1) = 0],$$

$$T_{2n}'(1) = n!\frac{(n-1)(n-2)}{12} \quad (n \geq 1), \qquad T_0'(1) = 0,$$

$$T_{2n}''(1) = n!\frac{(n-2)(n-3)(5n^2-n-16)}{720} \quad (n \geq 2), \qquad T_0'(1) = T_2'(1) = 0.$$

We are now in a position to conclude with:

THEOREM C. *The mean and variance of the distribution of the time cost for sorted list implementations of dictionaries subjected to n operations satisfy*:

$$M = \frac{(n-1)(n-2)}{12} \quad \text{if } n \geq 1, \qquad M = 0 \quad \text{if } n = 0;$$
$$V = \frac{(n-2)(2n^2+11n-1)}{360} \quad \text{if } n \geq 2, \qquad V = 0 \quad \text{if } n = 0 \text{ or } n = 1.$$

3.5. DICTIONARIES: TOTAL STORAGE UTILIZATION

The reciprocal polynomials of the denominators of the convergents of the continued fraction

$$\cfrac{1}{1 - 1q^0z - \cfrac{1^2q^1z^2}{1 - 3q^1z - \cfrac{2^2q^3z^2}{1 - 5q^2z - \cfrac{3^2q^5z^2}{\cdots}}}}$$

are such that

$$\overline{Q}_h = \left[z - (2h+1) q^h \right] \overline{Q}_{h-1} - h^2 q^{2h-1} \overline{Q}_{h-2}.$$

As a consequence,

$$K(t, z, q) = \sum_{h \geqslant 0} \overline{Q}_{h-1}(z, q) \frac{t^h}{h!}$$

is a solution of

$$\frac{\partial K}{\partial t} + 2qt \frac{\partial K}{\partial t}(qt) + q^3 t^2 \frac{\partial K}{\partial t}(q^2 t) = zK - K(qt) - qtK(q^2 t)$$

The function $K(t, z, 1)$ is a solution of the partial differential equation

$$L_2 u = 0, \qquad u(0) = 1,$$

where the differential operator is now

$$L_2 u = (1+t)^2 \frac{\partial u}{\partial t} - (z - t - 1) u.$$

From there, we proceed as before (we skip the details here). After the change of variable $v = t/(1+t)$, we get

$$\sum_{n \geqslant 0} S_n(1) \frac{v^n}{n!} = \frac{1}{1-v},$$

$$\sum_{n \geqslant 0} (S_n'(1)) \frac{v^n}{n!} = \frac{1}{6} - \frac{1}{2(1-v)^2} + \frac{1}{3(1-v)^3},$$

$$\sum_{n \geqslant 0} (S_n''(1)) \frac{v^n}{n!} = -407 - 83v + \frac{1200}{1-v} - \frac{1590}{(1-v)^2}$$

$$+ \frac{1225}{(1-v)^3} - \frac{528}{(1-v)^4} + \frac{100}{(1-v)^5}.$$

Finally we find:

THEOREM D. *The total storage utilization for n operations on a sorted list implementation of dictionaries has mean and variance given by*

$$M = \frac{(n-1)(n+1)}{6} \quad \textit{if } n \geqslant 1, \qquad M = 0 \quad \textit{if } n = 0;$$

$$V = \frac{20n^4 - 278n^3 + 1422n^2 - 3073n + 2407}{180} \qquad \textit{if } \; n \geqslant 2,$$

$$V = 0 \qquad \textit{if } \; n = 0 \textit{ or } n = 1.$$

4. CONCLUSIONS

We have developed a method for characterizing (through their moments) the distributions of various cost measures of simple data structures under sequences of operations. The reader should be aware of the fact that our probabilistic model, obtained through a type of "sampling," amounts to considering sequences of operations only up to (local) order isomorphism. Other analyses of dynamic data structures have been performed using different probabilistic models. Most notably, in [9], the authors show how to evaluate integrated costs of binary search trees when the schemas obey a simple repetitive pattern and keys are drawn uniformly over the real interval [0, 1]; the analysis there requires the use of Bessel functions. In [1], Burge has also analyzed integrated costs of stacks using continued fraction techniques.

We thus believe that our results are yet another illustration of the rich mathematical structure behind the performance evaluation of dynamic data structures.

The parameters we have been considering are related to some classical quantities in combinatorial analysis. The continued fraction of Section 3.2 actually appears in a paper of Touchard [13] in connection with chord intersection problems, and a result equivalent to our Theorem B is given without proof in a paper by Riordan [12]. Also, as indicated by A. Odlyzko (private communication), the continued fraction of Section 3.3. appears in connection with the distribution of correlated Gaussian variables.

Let us furthermore mention that it would be of interest to characterize the limiting distribution of integrated costs for large n (numerical calculations clearly show the existence of such limiting distributions). A first step in this direction is a recent result by Louchard, where limiting distributions for the integrated costs of stack structures are obtained. Such problems are related to a more general class of questions, namely how to characterize the distribution of costs for algorithms with a quadratic worst case. We are still in need of general methods to attack them.

REFERENCES

1. W. Burge, Stacks in a two-level store, *J. Algorithms* 2:178–185 (1981).
2. L. Chéno, P. Flajolet, J. Françon, C. Puech, and J. Vuillemin, Dynamic data structures: Finite files, limiting profiles, and variance analysis, in *Proceedings of the 18th Allerton Conference*, Monticello, Ill., 1980, pp. 223–232.
3. P. Flajolet, J. Françon, and J. Vuillemin, Sequence of operations analysis for dynamic data structures, *J. Algorithms* 1:111–141 (1980).
4. P. Flajolet, Combinatorial aspects of continued fractions, *Discrete Math.* 32:125–161 (1980).
5. P. Flajolet, Analyse d'algorithmes de manipulations d'arbres et de fichiers, *Cahiers BURO* 34–35:1–209 (1981).
6. J. Françon, Histoires de fichiers, *RAIRO Inform. Théor.* 12:49–67 (1978).
7. J. Françon, Combinatoire des structures de données, Thèse d'Etat, Univ. Louis Pasteur, Strasbourg, 1979.
8. W. Hahn, Über orthogonal Polynomen die q-Differenzengleichungen genügen, *Math. Nachr.*, 1949, pp. 4–34.
9. A. Jonassen and D. Knuth, A trivial algorithm whose analysis isn't, *J. Comput. System Sci.* 16:301–322 (1978).
10. D. Knuth, *The Art of Computer Programming: Fundamental Algorithms*, Vols. 1,3, Addison-Wesley, Reading, Mass., 1968, 1973.
11. O. Perron, *Die Lehre von den Kettenbrüchen*, Vol. 2, Teubner, Stuttgart, 1954.
12. J. Riordan, The distribution of crossings of chords joining pairs of $2n$ points in a circle, *Math. Comp.* 29:215–222 (1975).
13. Touchard, Sur un problème de configuration et sur les fractions continues. *Canad. J. Math.* 4:2–25 (1952).
14. H. Wall, *Analytic Theory of Continued Fractions*, Chelsea, New York, 1967.

Received 4 November 1983; revised 12 June 1985