AN ANALYSIS OF (h,k,l) -SHELLSORT

by

Andrew Chi-Chih Yao

STAN-CS-79-726
March 1979

COMPUTER SCIENCE DEPARTMENT
School of Humanities and Sciences
STANFORD UNIVERSITY

An Analysis of (h,k,l) -Shellsort */

Andrew Chi-Chih Yao

Computer Science Department
Stanford University
Stanford, California 94305

Abstract.

One classical sorting algorithm, whose performance in many cases remains unanalyzed, is Shellsort. Let $\vec{h}$ be a t-component vector of positive integers. An $\vec{h}$-Shellsort will sort any given n elements in t passes, by means of comparisons and exchanges of elements. Let $S_j(\vec{h};n)$ denote the average number of element exchanges in the j-th pass, assuming that all the n! initial orderings are equally likely. In this paper we derive asymptotic formulas of $S_j(\vec{h};n)$ for any fixed $\vec{h} = (h,k,l)$, making use of a new combinatorial interpretation of S_3. For the special case $\vec{h} = (3,2,1)$, the analysis is further sharpened to yield exact expressions.

*/ This work was supported in part by National Science Foundation under grant MCS77-05313. Part of this work was done while the author was on leave at Bell Laboratories, 600 Mountain Avenue, Murray Hill, New Jersey 07974.

1. Introduction.

The analysis of sorting algorithms has been a prototype for the mathematical analysis of algorithms (Knuth [2][3], Sedgewick [7]). One classical sorting algorithm, whose performance remains unanalyzed in most cases, is the Shellsort proposed by D. L. Shell [8] in 1959. All the known analytic results about this algorithm can be found in Knuth [2, Sec. 5.2.1] and Pratt [4]. In this paper, we will present some new results concerning the average-case performance of Shellsort,

Let $\vec{h} = (h_t, h_{t-1}, \ldots, h_2, h_1)$ be a vector of positive integers with $h_1 = 1$. An $\vec{h}$-shellsort on a list (or, an array) $L[0: n-1]$ of n elements performs an in-place sort in t passes, using comparisons and exchanges of the elements. In the j-th pass, $(1 \le j \le t)$, a straight-insertion sort*/ is done to each of the h_{t-j+1} sublists, where the i-th sublist $(0 \le i < h_{t-j+1})$ consists of $L[i]$, $L[i+h_{t-j+1}]$, $L[i+2h_{t-j+1}]$, ... Assuming that all $n!$ initial orderings of the elements are equally likely, let $S_j(\vec{h};n)$ be the average number of element exchanges in the j-th pass. The determination of $S_j(\vec{h};n)$, a standard performance measure for Shellsort, poses challenging mathematical questions. So far, the only completely analyzed case is when h_i divides h_{i+1} for each $1 \le i < t$ (see Knuth [2]). In the present paper, we derive asymptotic formulas for $S_j(\vec{h};n)$ when $\vec{h} = (h,k,1)$ is fixed and $n \to \infty$. In the derivation an interesting combinatorial interpretation of $S_3((h,k,1);n)$ will be introduced. For the special case $\vec{h} = (3,2,1)$, we further refine the analysis to give exact expressions for $S_j(\vec{h};n)$.

---

*/ See Knuth [2] for a description of the straight-insertion sort.

2. Preliminaries.

Let $L = (a_0, a_1, a_2, \ldots, a_{n-1})$ be a list of distinct real numbers. An inversion in L is a pair $\{i,j\}$ such that $i < j$ and $a_i > a_j$. The total number of inversions in L is denoted by $I(L)$. Clearly, the concept of inversion depends only on the ordering of the a_i. It is known (Knuth [2, equation 5.1.1-(12)]) that the expected value of $I(L)$ is $n(n-1)/4$ for a random list L (i.e., all n! permutations of a_i are equally likely).

For any sublist L' of L, the number of inversions $I(L')$ can be defined in an obvious way. An important property of inversions is that, when we perform a straight-insertion sort into ascending order a sublist L', the number of element exchanges is exactly equal to $I(L')$. Thus, $S_j(\vec{h};n)$ is the sum of the average number of inversions in all the h_{t-j+1} sublists that are to be sorted in the j-th pass.

For a list L of n elements, let $L^{(h,j)}$ $(0 \le j < h)$ denote the sublist $(L[j], L[j+h], L[j+2h], \ldots)$ of length $\lfloor (n+h-1-j)/h \rfloor$. We will call L h-ordered if, for each $0 < j < h$, the elements in $L^{(h,j)}$ are in ascending order. We say that we h-sort L, if we sort each $L^{(h,j)}$ $(0 < j < h)$ separately into ascending order.

Instead of drawing a list L as a single array, it is often convenient to show L in an h-row representation (Figure 1). The list is arranged in h rows, so that the j-th row $(1 \le j \le h)$ contains the sublist $L^{(h,j-1)}$. Thus, to h-sort L is to sort the elements in each row separately.

An h-ordered list $L[0:n-1]$ is a random h-ordered list if any ordering of its elements consistent with h-ordering is equally likely,

$L \rightarrow$	a_0	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8	a_9	a_{10}	a_{11}	a_{12}	a_{13}	a_{14}	a_{15}

$L^{(3,0)} \rightarrow$	a_0	a_3	a_6	a_9	a_{12}	a_{15}
$L^{(3,1)} \rightarrow$	a_1	a_4	a_7	a_{10}	a_{13}	
$L^{(3,2)} \rightarrow$	a_2	a_5	a_8	a_{11}	a_{14}	

Figure 1. A list L and its 3-row representation.

It is not difficult to see that, if we h-sort a random list L , the resulting array is a random h-ordered list. We remark that, if L is a random h-ordered list and $0 < i < j < h$, then the union of row $i+1$ and row $j+1$ forms a random 2-ordered list, ie., the sublist L[i] , L[j] , L[i+h] , L[j+h] , L[i+2h], L[j+3h], . . . is a random 2-ordered list. Note that A_n , the average number of inversions in a random 2-ordered list of n elements, is given by (see Knuth [2, Sec. 5.2.1])

$$A_n = \lfloor n/2 \rfloor 2^{n-2} \Big/ \binom{n}{\lfloor n/2 \rfloor} \quad . \tag{1}$$

Asymptotically,

$$A_n = \sqrt{\pi/128}\, n^{3/2} + O(\sqrt{n}) \quad . \tag{2}$$

<u>Remark on the O-notation.</u> In Section 4, Appendix, and in the statement of Lemma 2, the constants in the O-notation are dependent on h and k. Everywhere else, the constants are absolute constants.

3. Asymptotics for the (3,2,1) Case.

In this section we present our basic approach, by analyzing the asymptotic performance of the (3,2,1)-Shellsort on a random list L of n elements. It will be shown that, for $\vec{h} = (3,2,1)$,

$$S_1(\vec{h};n) = n^2/12 + O(n), \tag{3}$$

$$S_2(\vec{h};n) = \sqrt{\pi/192}\, n^{3/2} + O(n), \tag{4}$$

$$S_3(\vec{h};n) = n/4 + O(n^{2/3}). \tag{5}$$

Several facts for use in later sections will also be given.

Analysis of Pass 1.

Consider L in a 3-row representation (see Figure 1). In the first pass of the (3,2,1)-Shellsort, each row is sorted by a straight-insertion sort. Thus, $S_1(\vec{h};n)$ is equal to the expected value of the sum $\sum_j I(L^{(3,j)})$. As each row is initially a random list, we have */

$$S_1(h;n) = \sum_{0 \le i \le 2} n_i(n_i-1)/4, \tag{6}$$

with $n_i = \lfloor (n+2-i)/3 \rfloor$. Asymptotically,

$$S_1(h;n) = n^2/12 + O(n),$$

which is (3).

---

*/ Here and hereafter, we will often use the fact $E\left(\sum_i X_i\right) = \sum_i E(X_i)$ for any variables X_i, without explicit reference to it. ($E(X_i)$ is the expected value of X_i.)

Analysis of Pass 2.

For the moment, assume that $n = 3m$ for some integer $m > 0$. At the end of Pass 1, we have a random 3-ordered list L. Pass 2 will perform a straight-insertion sort for the sublists $L^{(2,0)}$ (the "shaded" list in Figure 2) and $L^{(2,1)}$ (the "blank" list in Figure 2), separately. Let $S_2^{(j)}(n)$ denote the average number of inversions in $L^{(2,j)}$, $j \in \{0,1\}$. We have

$$S_2(\vec{h};n) = S_2^{(0)}(n) + S_2^{(1)}(n) \quad . \tag{7}$$

Now, consider a 2-ordered list $L' = (a_0, b_0, a_1, b_1, a_2, b_2, \ldots, a_{m-1}, b_{m-1})$ in a 2-row representation (Figure 3). Define four sublists as follows:

$$L_{\text{even,even}} = (a_0, b_0, a_2, b_2, a_4, b_4, \ldots) \quad ,$$

$$L_{\text{even,odd}} = (a_0, b_1, a_2, b_3, a_4, b_5, \ldots) \quad ,$$

$$L_{\text{odd,even}} = (b_0, a_1, b_2, a_3, b_4, a_5, \ldots) \quad ,$$

$$L_{\text{odd,odd}} = (a_1, b_1, a_3, b_3, a_5, b_5, \ldots) \quad .$$

For each $\alpha, \beta \in \{\text{even}, \text{odd}\}$, let $X_{\alpha,\beta}$ denote the random variable, defined on the set of L', corresponding to the number of inversions in $L_{\alpha,\beta}$, and let $B_{\alpha,\beta}(2m) = E(X_{\alpha,\beta})$ for a random 2-ordered L'. Clearly,

$$\sum_{\alpha,\beta} B_{\alpha,\beta}(2m) = \text{average \# of inversions in } L'$$

$$= A_{2m} \quad . \tag{8}$$

Returning to the evaluation of S_2, we observe that the union of any two rows in Figure 2 is a random 2-ordered list at the start of Pass 2. It follows that

$$S_2^{(0)}(n) = B_{\text{even,odd}}(2m) + B_{\text{odd,even}}(2m) + B_{\text{even,even}}(2m) \quad ,$$

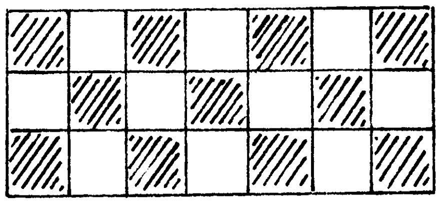

Figure 2. The sublists $L^{(2,0)}$ (the "shaded" list) and $L^{(2,1)}$ (the "blank" list).

a_0	a_1	a_2	a_3	a_4	a_5	a_6
b_0	b_1	b_2	b_3	b_4	b_5	b_6

Figure 3. A 2-ordered list L' .

and

$$S_2^{(1)}(n) = B_{odd,even}(2m) + B_{even,odd}(2m) + B_{odd,odd}(2m) .$$

Together with (7) and (8), these lead to

$$S_2(\vec{h};n) = A_{2m} + (B_{even,odd}(2m) + B_{odd,even}(2m)) . \tag{9}$$

It remains to evaluate $B_{even,odd}(2m) + B_{odd,even}(2m)$. A 'precise calculation is possible (see Section 5), but here we will determine it only asymptotically.

We assert that

$$B_{\alpha,\beta}(2m) = \frac{1}{4} A_{2m} + O(m) \quad \text{for } \alpha,\beta \in \{\text{even},\text{odd}\} . \tag{10}$$

Suppose (10) is true. It then follows from (9), (10) and (2) that

$$S_2(h;n) = \sqrt{\pi/192}\, n^{3/2} + O(n) ,$$

which is (4).

It remains to prove (10), which we will show in a more general form. Let $k > 1$ be an integer and $L' = (a_0,b_0,a_1,b_1,a_2,b_2,\ldots)$ a 2-ordered list of n elements. For each $0 < i,j < k$, let $L_{i,j}$ denote the list $(a_i,b_j,a_{i+k},b_{j+k},a_{i+2k},b_{j+2k},\ldots)$ if $i \le j$, and the list $(b_j,a_i,b_{j+k},a_{i+k},b_{j+2k},a_{i+2k},\ldots)$ if $i > j$. Define $Y_{i,j}$ to be the random variable whose value for L' is the number of inversions in $L_{i,j}$, and let $B_{i,j}^{(k)}(n) = E(Y_{i,j})$ be the expected value of $Y_{i,j}$ for a random 2 -ordered list L' , It is clear that $B_{i,j}^{(2)}(2m)$ are the $B_{i,j}(2m)$ defined earlier, provided we identify " 0 " with "even" and " 1 " with "odd" in the subscripts. Thus, formula (10) is a special case of the following result.

Lemma 1. For any fixed $k > 1$, $B^{(k)}_{i,j}(n) = A_n/k^2 + O(n/k)$ for any $i,j \in \{0,1,2,\ldots,k-1\}$.

Corollary. $B^{(k)}_{i,j}(n) = \sqrt{\pi/128}\, n^{3/2}/k^2 + O(n/k)$.

Proof. Define the following random variables on the set of 2-ordered lists L' :

$$Y_{s,t} = \begin{cases} \begin{cases} 1 & \text{if } a_s < b_t \\ 0 & \text{otherwise} \end{cases} & \text{for } 0 \le t < s < \lceil n/2 \rceil, \\ \begin{cases} 1 & \text{if } a_s > b_t \\ 0 & \text{otherwise} \end{cases} & \text{for } 0 \le s \le t < \lfloor n/2 \rfloor, \\ 0 & \text{otherwise}. \end{cases}$$

Then,

$$Y_{i,j} = \sum_{\substack{s \bmod k = i \\ t \bmod k = j}} y_{s,t}, \qquad (11)$$

We wish to prove that,*/ for $0 \le i \ne j < k$,

$$Y_{i,i} = Y_{i,j} + O(n/k), \qquad (12)$$

and

$$Y_{j,j} = Y_{i,j} + O(n/k). \qquad (13)$$

This would imply the lemma, since all $Y_{i,j}$ (and hence $B^{(k)}_{i,j}(n)$) would be equal up to an additive $O(n/k)$ term, whereas $\sum_{i,j} B^{(k)}_{i,j}(n) = A_n$.

*/ An equality (or inequality) involving random variables is valid if and only if, for every event in the sample space, the values of the random variables satisfy the given formula. For example, (12) is equivalent to the following statement: there exists a constant c such that, for any L', $|Y_{i,i} - Y_{i,j}| \le c\, n/k$.

We will only prove (12); the proof of (13) is similar. For any t satisfying $t \bmod k = i$, let $t_+ = t + (j-i)$ and $t_- = t - (k+i-j)$ if $i < j$, and $t_+ = t + (k+j-i)$ and $t_- = t - (i-j)$ if $i > j$. Then t_+ and t_- are, respectively, the smallest $t' > t$ and the largest $t' < t$ that satisfy $t' \bmod k = j$. For the rest of the proof of Lemma 1, we use variables s, t, t' exclusively for integers satisfying $s \bmod k = t \bmod k = i$ and $t' \bmod k = j$, and when they appear in a summation, it is understood that they only range over such values.

From the definition of $y_{u,v}$ and the fact that L' is 2-ordered, we deduce

$$y_{s,t_-} \leq y_{s,t} \leq y_{s,t_+} \qquad \text{if } s > t+k, \tag{14}$$

and

$$y_{s,t_+} \leq y_{s,t} \leq y_{s,t_-} \qquad \text{if } t > s+k. \tag{15}$$

Now, noting that $0 \leq y_{u,v} \leq 1$, we have

$$\sum_{s>t+k} y_{s,t} = \sum_{s>t} y_{s,t} - \sum_{t+k>s>t} y_{s,t}$$

$$= \sum_{s>t} y_{s,t} - O(n/k),$$

$$\sum_{s>t+k} y_{s,t_+} = \sum_{s>t} y_{s,t_+} - \sum_{t+k>s>t} y_{s,t_+}$$

$$= \sum_{s>t'} y_{s,t'} - O(n/k),$$

and

$$\sum_{s>t+k} y_{s,t_-} = \sum_{s>t} y_{s,t_-} - \sum_{t+k\geq s\geq t} y_{s,t_-}$$

$$= \sum_{s>t'} y_{s,t'} - O(n/k).$$

Together with (14), this implies

$$\sum_{s>t'} y_{s,t'} - O(n/k) \le \sum_{s>t} y_{s,t} \le \sum_{s>t'} y_{s,t'} + O(n/k) . \tag{16}$$

A similar argument using (15) gives

$$\sum_{s<t'} y_{s,t'} - O(n/k) \le \sum_{s\le t} y_{s,t} \le \sum_{s\le t'} y_{s,t'} + O(n/k) . \tag{17}$$

Adding up (16) and (17), we obtain (12).

This completes the proof of Lemma 1. The corollary follows by using the asymptotic expression (2) of A_n . □

We have derived (4) for the case n mod 3 = 0 . The other cases can be handled in the same fashion. In fact, one obtains the following generalization of (9): For $\vec{h} = (3,2,1)$,

$$S_2(\vec{h};n) = \begin{cases} A_{2m} + (B_{\text{even},\text{odd}}(2m) + B_{\text{odd},\text{even}}(2m)) & \text{if } n = 3m \\ A_{2m+1} + (B_{\text{even},\text{odd}}(2m) + B_{\text{odd},\text{even}}(2m)) & \text{if } n = 3m+1 \\ A_{2m+1} + (B_{\text{even},\text{odd}}(2m+2) + B_{\text{odd},\text{even}}(2m+2)) & \text{if } n = 3m+2 \end{cases} , \tag{18}$$

The asymptotic formula (4), for general n , can be proved using 2), (10) and (18).

Analysis of Pass 3.

We now come to the analysis of Pass 3, which is the most interesting part combinatorially. The question is "What is the average number of inversions in a list, obtained by first performing a 3-sort and then a 2-sort on a random list?" It will be convenient to work with the equivalent form "Given a random 3-ordered list L of n elements, what is the expected number of inversions in the new list L' obtained from 2-sorting L ?".

Consider the following random variables $z_{i,j}$ (i, j integers) defined on $\mathcal{L}$, the set of 3-ordered lists L of n elements $\{1,2,\ldots,n\}$:

(a) For $1 \leq j < i \leq \lceil n/2 \rceil$, $z_{i,j} = 1$ if the i-th smallest element in $L^{(2,0)}$ is less than the j-th smallest element in $L^{(2,1)}$, and $z_{i,j} = 0$ otherwise;

(b) For $1 \leq i \leq j \leq \lfloor n/2 \rfloor$, $z_{i,j} = 1$ if the i-th smallest element in $L^{(2,0)}$ is greater than the j-th smallest element in $L^{(2,1)}$, and $z_{i,j} = 0$ otherwise;

(c) $z_{i,j} = 0$ otherwise.

These random variables have the following interpretation. Let $L'[0:n-1]$ be the 2-ordered list resulting from 2-sorting L . (We remark that L' also remains 3-ordered. See Knuth [2, Sec. 5.2.1 Theorem K].) Then $z_{i,j} = 1$ if and only if the pair $\{2i-2,2j-1\}$ is an inversion in L' , i.e., the elements in $L'^{(2,0)}[i-1]$ and $L'^{(2,1)}[j-1]$ are out of order in L' . It follows that

$$S_3(\vec{h};n) = \sum_{i,j} E(z_{i,j}) , \qquad (19)$$

where the expected values are for a random 3-ordered list L .

Formula (19) can be simplified, if we observe that $z_{i,j}$ is 0 unless $i = j$ or $i = j+1$. This is due to the fact that only adjacent elements may be out of order for L' , a list both 3-ordered and 2-ordered (see Knuth [2, Ex. 5.2.1-25]). Thus,

$$S_3(\vec{h};n) = \sum_{1 \le i \le \lfloor n/2 \rfloor} E(z_{i,i}) + \sum_{1 \le i < \lceil n/2 \rceil} E(z_{i+1,i}) . \qquad (20)$$

We shall derive (5) from (20). Assume that $n = 3m$ is a multiple of 3 for the present. A 3-ordered list L of n elements $\{1\ 2,\ldots,n\}$ can be represented as a ternary sequence of n symbols in $\{1,2,3\}$, with $n/3$ j's for each $j \in \{1,2,3\}$. The i-th symbol in the sequence is j if and only if the integer i in the list appears in the j-th row. Note that this representation is a 1-1 mapping from the set $\mathcal{L}$ <u>onto</u> the set of ternary sequences with exactly $n/3$ j's for each $j \in \{1,2,3\}$. This shows, incidentally, $|\mathcal{L}| = \binom{3m}{m,m,m}$. Figure 4 shows a 3-ordered list (in its 3-row representation) whose associated sequence is $(1,1,1,2,3,1,3,2,1,2,1,2,2,3,3,3,3,2,1,3,2)$. To evaluate $E(z_{i,i})$, we need to count the number of 3-ordered lists in $\mathcal{L}$ for which $z_{i,i} = 1$. Consider the 3-row representation of L, with positions of $L^{(2,0)}$ "shaded" as in Figure 2. It is easy to see that $z_{i,i} = 1$ if and only if there are more "blank" cells than "shaded" cells in the positions occupied by the smallest $2i-1$ elements $\{1,2,\ldots,2i-1\}$. For example, $z_{4,4} = 1$ in the example shown in Figure 4, as there are 4 blank cells but only 3 shaded cells in positions occupied by $\{1,2,\ldots,7\}$ (see Figure 5). This condition can easily be tested from the ternary sequence representation of L discussed above. Suppose there are k_j j's $(j \in \{1,2,3\})$ in the first $2i-1$ components of the sequence. Then, using Table 1, one immediately sees that

1	2	3	6	9	11	19
4	8	10	12	13	18	21
5	7	14	15	16	17	20

Figure 4. A sample 3-ordered list L in $\mathcal{L}$.

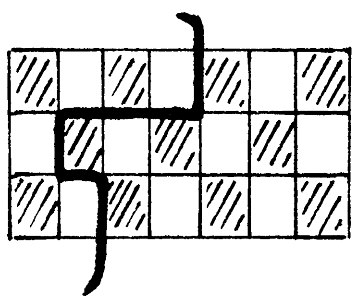

Figure 5. The positions occupied by elements $\{1,2,\ldots,7\}$ in the list of Figure 4.

	row 1	row 2	row 3
$k_{\cdot j}$ = even	0	0	0
$k_{\cdot j}$ = odd	-1	+1	-1

Table 1. The contribution to (# of blank cells - # of shaded cells) by positions in row j as a function of $k_{\cdot j}$.

# of blank cells - # of shaded cells

$$= k_2 \bmod 2 - k_1 \bmod 2 - k_3 \bmod 2 .$$

It follows that $z_{i,i} = 1$ if and only if k_1 and k_3 are even and k_2 is odd. Thus,*/ for $1 \le i \le \lfloor n/2 \rfloor$,

$$E(z_{i,i}) = \frac{1}{\binom{3m}{m,m,m}} \sum_{\substack{k_1+k_2+k_3 = 2i-1 \\ k_1,k_3 = \text{even} \\ k_2 = \text{odd}}} \binom{2i-1}{k_1,k_2,k_3} \binom{3m-(2i-1)}{m-k_1,m-k_2,m-k_3} .$$

Or, equivalently,

$$E(z_{i,i}) = \frac{1}{\binom{3m}{2i-1}} \sum_{\substack{k_1+k_2+k_3 = 2i-1 \\ k_1,k_3 = \text{even} \\ k_2 = \text{odd}}} \binom{m}{k_1} \binom{m}{k_2} \binom{m}{k_3}$$

$$\text{for } 1 \le i \le \lfloor n/2 \rfloor. \tag{21}$$

A similar argument for $z_{i+1,i}$ leads to

$$E(z_{i+1,i}) = \frac{1}{\binom{3m}{2i}} \sum_{\substack{k_1+k_2+k_3 = 2i \\ k_1,k_3 = \text{odd} \\ k_2 = \text{even}}} \binom{m}{k_1} \binom{m}{k_2} \binom{m}{k_3}$$

$$\text{for } 1 \le i < \lceil n/2 \rceil . \tag{22}$$

Formulas (20), (21) and (22) give an exact formula for $S_3(\vec{h};n)$.

---

*/ We use the convention that a multinomial coefficient is zero whenever any of its lower indices is negative.

We now assert that, as $n \to \infty$, (21) and (22) lead to

$$\sum_{1 \le i \le \lfloor n/2 \rfloor} E(z_{i,i}) = n/8 + O(n^{2/3}) ,$$

(23)

$$\sum_{1 \le i < \lceil n/2 \rceil} E(z_{i+1,i}) = n/8 + O(n^{2/3} .$$

Intuitively, for each w , the summation

$$1 = \frac{1}{\binom{3m}{w}} \sum_{k_1+k_2+k_3=w} \binom{m}{k_1}\binom{m}{k_2}\binom{m}{k_3}$$

can be partitioned into four approximately equal parts (each $\approx 1/4$), according to the parities of k_1 , k_2 and k_3 . Hence each sum in (23) is roughly equal to $1/4$ times the number of terms. This argument can be made precise to prove (23), and in fact the next lemma.

<u>Definition.</u> Let $h,k > 1$ be positive integers, and $C_{h,k}$ the set of vectors $\vec{c} = (c_0,c_1,\ldots,c_{k-1})$ with integer components $0 \le c_i < k$. Suppose $m > 0$ is an integer, and $\vec{n} = (n_0,n_1,\ldots,n_{k\ 1})$ is a vector of integer components satisfying $|n_i-m| < 2$ for all i . For each $\vec{c} \in C_{h,k}$, we define

$$G_{h,\ k}(\vec{c};\vec{n}) = \sum_{0 \le w \le N} \left[\frac{1}{\binom{N}{w}} \sum_{\substack{\sum_i j_i = w \\ j_i \bmod k = c_i \\ \text{for all } i}} \binom{n_0}{j_0}\binom{n_1}{j_1}\cdots\binom{n_{h-1}}{j_{h-1}} \right] .$$

Lemma 2. Let $h,k > 1$ be fixed integers. As $N \to \infty$,

$$G_{h,k}(\vec{c};\vec{n}) = \frac{1}{k^h}(N+1) + O(N^{2/3}) \qquad \text{for each } \vec{c} \in C_{h,k} .$$

We emphasize that the constants in the O-notation in Lemma 2 are dependent on h and k. The proof of Lemma 2 will be given in the appendix.

Clearly, (23) is a special case of Lemma 2 with $h = 3$, $k = 2$. It now follows from (20) and (23) that, for $n \bmod 3 = 0$, (5) is true, i.e.,

$$S_3(\vec{h};n) = n/4 + O(n^{2/3}) .$$

One can prove (5) for $n \bmod 3 \neq 0$ in a similar way. In fact, for any $n > 3$, formula (20) and the analogue to (21), (22) read: For $\vec{h} = (3,2,1)$

$$S_3(\vec{h};n) = \sum_{1 \le i \le \lfloor n/2 \rfloor} E(z_{i,i}) + \sum_{1 \le i < \lceil n/2 \rceil} E(z_{i+1,i}) , \tag{24}$$

$$E(z_{i,i}) = \frac{1}{\binom{n}{2i-1}} \sum_{\substack{k_1+k_2+k_3 = 2i-1 \\ k_1,k_3 \text{ are even} \\ k_2 \text{ is odd}}} \binom{\lfloor (n+2)/3 \rfloor}{k_1} \binom{\lfloor (n+1)/3 \rfloor}{k_2} \binom{\lfloor n/3 \rfloor}{k_3} ,$$

$$1 \le i \le \lfloor n/2 \rfloor . \tag{25}$$

and

$$E(z_{i+1,i}) = \frac{1}{\binom{n}{2i}} \sum_{\substack{k_1+k_2+k_3=2i \\ k_1,k_3 \text{ are odd} \\ k_2 \text{ is even}}} \binom{\lfloor (n+2)/3 \rfloor}{k_1} \binom{\lfloor (n+1)/3 \rfloor}{k_2} \binom{\lfloor n/3 \rfloor}{k_3},$$

$$1 \leq i < \lceil n/2 \rceil \quad . \tag{26}$$

Formula (5) then follows from (24)-(26) and Lemma 2, We shall see in Section 5 that $O(n^{2/3})$ is an overestimate of the error term for S_3.

We have finished the asymptotic analysis for the (3,2,1) case. Generalizations and refinements will be made in the next two sections.

4. Generalization to the $(h,k,1)$ Case.

Let h, $k > 1$ be fixed distinct positive integers and $\vec{h} = (h,k,1)$. In this section, we will derive asymptotic formulas for the $(h,k,1)$-Shellsort on n elements. Let $\psi(h,k)$ denote a function to be defined in a moment. We will establish the following results,

Theorem 1. Suppse $\gcd(h,k) = 1$. Then, as $n \to \infty$,

$$S_1(\vec{h};n) = \frac{n^2}{4h} + O(n),$$

$$S_2(\vec{h};n) = \frac{\sqrt{\pi}}{8} \frac{\sqrt{h} - (\sqrt{h})^{-1}}{k} n^{3/2} + O(n),$$

and

$$S_3(\vec{h};n) = \psi(h,k)n + O(n^{2/3}).$$

Theorem 2. Suppose $\gcd(h,k) = d > 1$. Then, as $n \to \infty$,

$$S_1(\vec{h};n) = \frac{n^2}{4h} + O(n),$$

$$S_2(\vec{h};n) = \frac{\sqrt{\pi}}{8} \frac{\sqrt{h} - d(\sqrt{h})^{-1}}{k} n^{3/2} + O(n),$$

and

$$S_3(\vec{h};n) = \frac{\sqrt{\pi}}{8} \frac{(d-1)}{\sqrt{d}} n^{3/2} + O(n).$$

We will now define $\psi(h,k)$ and some other terms. Assume that $\gcd(h,k) = 1$. Consider the $h \times k$ matrix $V[1{:}h, 1{:}k]$, where $V[i,j] = ((j-1)h + (i-1)) \bmod k$. Each row of V is then a permutation of the integers in $\{0,1,2,\ldots,k-1\}$. For each $\alpha \in \{0,1,2,\ldots,k-1\}$ and each $0 < i < h$, let $u_i(\alpha)$ denote the position in row $i+1$ where α appears, i.e., $1 \le u_i(\alpha) \le k$ and $V[i+1, u_i(\alpha)] = \alpha$. For each

$\alpha,\beta \in \{0,1,2,\ldots,k-1\}$, there are some rows $i+1$ in which α appears before β in the permutation, i.e., $u_i(\alpha) < u_i(\beta)$; denote the set of such i as $K_{\alpha,\beta}$, and define $K'_{\alpha,\beta} = \{0,1,\ldots,h-1\} - K_{\alpha,\beta}$. Let $m_{\alpha,\beta} = \mathrm{I\$,,\backslash}$, and $\ell_{\alpha,\beta} = |K'_{\alpha,\beta}| = h - m_{\alpha,\beta}$. It is easy to see that, for any $i \in K_{\alpha,\beta}$, the number of positions between the appearances of α and β in row $i+1$ is $u_i(\beta) - u_i(\alpha)$ is independent of i , which we denote by $\Delta_{\alpha,\beta}$. Also, for any $i \in K'_{\alpha,\beta}$, the number of positions between the appearances of β and α is $u_i(\alpha) - u_i(\beta) = h - \Delta_{\alpha,\beta}$. An illustration of these definitions is given in Figure 6. Note that the matrix V can be obtained by filling in the sequence $0,1,2,\ldots,k-1,0,1,2,\ldots,k-1,0,1,2,\ldots,k-1,0,1,\ldots$, in a column by column manner; this is in general true.

For any integers p , ℓ satisfying $\ell+p \geq 0$ and any real number $0 < q < 1$, let

$$f(p,\ell,q) = (1-q)^p q^\ell \sum_j \binom{\ell+p}{\ell+j} \left(\frac{q}{1-q}\right)^j |j| \quad . \tag{27}$$

Define ψ by

$$\psi(h,k) = \frac{1}{k} \sum_{0 \leq \alpha < \beta < k} f(m_{\alpha,\beta}-1,\ h-m_{\alpha,\beta},\ \Delta_{\alpha,\beta}/k) \quad . \tag{28}$$

Finally, throughout this section, we use the symbol n_i for $\lfloor (n+h-i-1)/h \rfloor$ $(0 \leq i < h)$.

0	3	1	4	2
1	4	2	0	3
2	0	3	1	4
3	1	4	2	0
4	2	0	3	1
0	3	1	4	2
1	4	2	0	3
2	0	3	1	4

$h = 8$, $k = 5$;

$m_{32} = 3$ (3 appears before 2 in rows 1, 4 and 6);

$\ell_{32} = 8-3 = 5$;

$4_{2} = u_0(2) - u_0(3) = 5-2 = 3$.

Figure 6. An illustration of definitions for the matrix V and related terms.

4.1 Proof of Theorem 1.

Consider the first pass on a random list L . As each $L^{(h,i)}$ is initially a random list of n_i elements, its average number of inversions is $n_i(n_i-1)/4$. Thus,

$$S_1(\vec{h};n) = \sum_{0 \le i < h} n_i(n_i-1)/4 = \frac{1}{4h} n^2 + O(n) .$$

This proves the formula for S_1 in the theorem.

Consider a random h-ordered list L of n elements in the h-row representation. For each $0 < r < k$, $0 \le s < h$, $0 \le t < h$, let $L_{r;s,t}$ denote the sublist of elements in $L^{(h,r)}$ that are in the (s+1) -st and (t+1) -st row. Then

$$S_2(\vec{h};n) = \sum_{0 \le r < k} T_r(n)$$

$$= \sum_{0<r<k} \sum_{0<s<t<h} T_{r;s,t} , \qquad (30)$$

where $T_r(n)$ is the average number of inversions in $L^{(h,r)}$, and $T_{r;s,t}$ is the average number of inversions in $L_{r;s,t}$.

Let $P_{ij} = ((j-1)h+(i-1)) \bmod k$ for $1 \le i \le h$, $1 \le j \le n_{i-1}$. Then the j-th element of the i-th row (in the h-row representation) of L is in $L^{(k,r)}$ where $r = P_{ij}$. Clearly, the first k columns of the matrix (P_{ij}) form a matrix identical to the matrix V defined earlier. As each row of (P_{ij}) is periodic with period k , the sublist $L^{(k,r)}$ occupies positions $v(s,r), v(s,r)+k, v(s,r)+2k, \ldots$ in row s+1 , where $v(s,r)$ is the position of the integer r in the (s+1) -st row of matrix V . It follows that, for $0 < r < k$, $0 \le s < h$, $0 \le t < h$,

$$T_{r;s,t} = B_{i,j}^{(k)}(n_{s,t}) , \tag{31}$$

where $i = v(s,r)-1$, $j = v(t,r)-1$ and $n_{s,t} = n_s + n_t$.

According to the corollary of Lemma 1,

$$B_{i,j}^{(k)}(n_{s,t}) = \sqrt{\frac{\pi}{128}} \; \frac{1}{k^2} \left(\frac{2n}{h} \right)^{3/2} + O(n) . \tag{32}$$

Substituting (32) into (30), we obtain

$$\begin{aligned} S_2(\vec{h};n) &= k\binom{h}{2} \sqrt{\frac{\pi}{128}} \; \frac{1}{k^2} \left(\frac{2n}{h} \right)^{3/2} + O(n) \\ &= \frac{\sqrt{\pi}}{8} \; \frac{\sqrt{h} - (\sqrt{h})^{-1}}{k} \; n^{3/2} + O(n) . \end{aligned}$$

This proves the formula for S_2 in Theorem 1.

We will now analyze Pass 3. Let $\mathcal{L}$ be the set of all h-ordered lists of n elements $\{1,2,3,\ldots,n\}$. For each $0 \leq \alpha < \beta < k$, let $I_{\alpha,\beta}(L)$ denote the number of inversions between elements in $L'^{(k,\alpha)}$ and $L'^{(k,\beta)}$, and let

$$\mathcal{I}_{\alpha,\beta} = \sum_{L \in \mathcal{L}} I_{\alpha,\beta}(L) . \tag{33}$$

Clearly

$$\begin{aligned} S_3(\vec{h};n) &= \frac{1}{|\mathcal{L}|} \sum_{L \in \mathcal{L}} \; \sum_{0 \leq \alpha < \beta < k} I_{\alpha,\beta}(L) \\ &= \sum_{0 \leq \alpha < \beta < k} \bar{I}_{\alpha,\beta} , \end{aligned} \tag{34}$$

where

$$\bar{I}_{\alpha,\beta} = \mathcal{I}_{\alpha,\beta} \Big/ \binom{n}{n_0, n_1, \ldots, n_{k-1}} . \tag{35}$$

Consider any list $L \in \mathcal{L}$ in the h-row representation (Figure 7). A position is called <u>of type α</u>, or an <u>α-cell</u> $(0 < \alpha < k)$ if it is in $L^{(k,\alpha)}$. It is easy to see that, for each $1 \leq i \leq h$, $1 \leq j \leq n_i$, the j-th position of row i is a $P_{i,j}$-cell. For each $0 \leq m < n$ and each $L \in \mathcal{L}$, let $Q_m(L)$ be the set of positions occupied by elements $\{1,2,\ldots,m\}$, and $D_{\alpha,\beta}(m;L) = |\#$ of a-cells - # of p-cells$|$ in $Q_m(L)$ for each $0 \leq \alpha < \beta < k$. We shall say that $Q_m(L)$ has <u>shape</u> (relative to the h-row representation) $(j_0, j_1, \ldots, j_{h-1})$ if it consists of the leftmost j_i positions of row $i+1$ for $0 < i < h$. In the example shown in Figure 7, $Q_m(L)$ has shape (4, 6, 3, 3, 8) and $D_{0,1}(m,L) = |9-8| = 1$ when $m = 24$ (see Figure 8). Let $s(m+1;L)$ denote the type of the -position that the element $m+1$ occupies, We have, for $0 < \alpha < \beta < k$,

$$I_{\alpha,\beta}(L) = \sum_{0<m<n} \delta_{\alpha,s(m+1;L)} D_{\alpha,\beta}(m;L) \quad . \tag{36}$$

Formula (36) can be proved as follows. For any given $0 \leq \alpha < \beta < k$, an inversion between $L'^{(k,\alpha)}$ and $L'^{(k,\beta)}$ can be uniquely labeled as $(m+1,i,j)$, meaning that the i-th smallest element in $L'^{(k,\alpha)}$ is the element $m+1$, and is less (or greater) than the j-th smallest element in $L'^{(k,\beta)}$ where $j < i$ (or $j \geq i$). Foranygiven m and L, there are $\delta_{\alpha,s(m+1;L)} D_{\alpha,\beta}(m;L)$ such triplets $(m+1,i,j)$. Formula (36) follows.

From (33) and. (36), we have

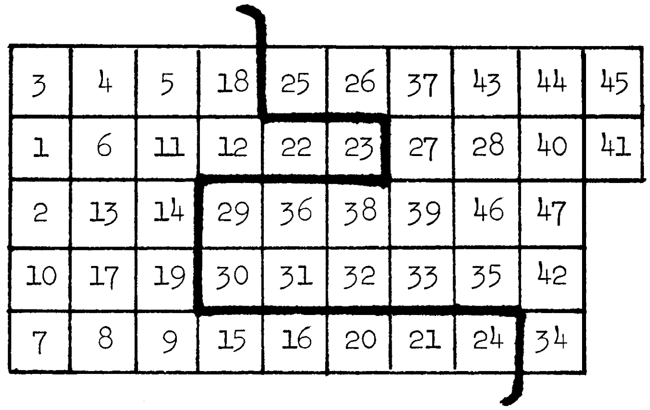

Figure 7. A 5-ordered list L . $Q_m(L)$ consists of all the cells to the left of the heavy line, where m = 24 .

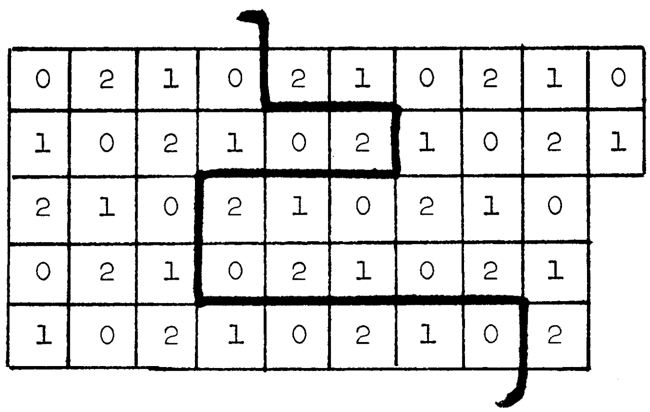

Figure 8. In the example in Figure 6, $Q_m(L)$ contains 9 0-cells, 8 1-cells, and 7 2-cells. Thus $D_{0,1}(m,L) = |9-8| = 1$. (Each cell is marked with its type.)

$$\vartheta_{\alpha,\beta} = \sum_{0<m<n} \sum_{L \in \mathfrak{L}} \delta_{\alpha, s(m+1;L)} D_{\alpha,\beta}(m;L)$$

$$= \sum_{0 \le m < n} \sum_{\substack{0 \le j_i \le n_i, \forall i \\ \sum_i j_i = m}} \sum_{\substack{0 \le i < h \\ P_{i+1,\ j_i+1} = \alpha}} \left(\begin{array}{l} \text{\#of } L \in \mathfrak{L} \text{ with} \\ \text{element } m+1 \text{ in} \\ \text{row } i+1\text{, and shape} \\ (j_0, j_1, \ldots \text{[illegible]} \\ \text{for } Q_m(L) \end{array} \right)$$

$$\times g_{\alpha,\beta}(j_0, j_1, \ldots, j_{h-1}) \quad ,$$

where $g_{\alpha,\beta}(j_0, j_1, \ldots, j_{h-1}) = D_{\alpha,\beta}(m;L)$ for all L whose $Q_m(L)$ have shape $(j_0, j_1, \ldots, j_{h-1})$. Clearly, whether the (j_i+1)-st position in row $i+1$ is an a-cell or not depends only on i , α , and $c = j_i \bmod k$. (Remember that $P_{i,j}$ is periodic in j with period k .) Define $x(i,\alpha, c) = 1$ if it is an a-cell, and 0 otherwise. Then

$$\vartheta_{\alpha,\beta} = \sum_{0 \le m < n} \sum_{\substack{j_0, \ldots, j_{h-1} \\ \sum_t j_t = m}} \sum_{\substack{0 \le i < h \\ x(i,\alpha, j_i \bmod k) = 1}} g_{\alpha,\beta}(j_0, j_1, \ldots, j_{h-1})$$

$$\times \binom{m}{j_0, j_1, \ldots, j_{h-1}} \binom{n-m-1}{n_0-j_0, \ldots, n_i-j_i-1, \ldots, n_{h-1}-j_{h-1}}$$

$$= \sum_{0<i<h} \sum_{0<m<n} \sum_{\substack{j_0, \ldots, j_{h-1} \\ \sum_t j_t = m \\ x(i,\alpha, j_i \bmod k) = 1}} g_{\alpha,\beta}(j_0, j_1, \ldots, j_{h-1})$$

$$\times \binom{m}{j_0, j_1, \ldots, j_{h-1}} \binom{n-m-1}{n_0-j_0, \ldots, n_i-j_i-1, \ldots, n_{h-1}-j_{h-1}} . \qquad (37)$$

Now, note that the value of $D_{\alpha,\beta}(m;L)$ depends only on $j_0 \bmod k, \ldots, j_{h-1} \bmod k$ where $(j_0, j_1, \ldots, j_{h-1})$ is the shape of $Q_m(L)$, because of the fact that $P_{i,j}$ is periodic on j with period k. Thus, we have $g_{\alpha,\beta}(j_0, j_1, \ldots, j_{h-1}) = g_{\alpha,\beta}(j_0 \bmod k, j_1 \bmod k, \ldots, j_{h-1} \bmod k)$. Thus, (37) leads to

$$\vartheta_{\alpha,\beta} = \sum_{0 \le i < h} \sum_{\substack{0 \le c_0, c_1, \ldots, c_{h-1} < k \\ x(i,\alpha,c_i) = 1}} g_{\alpha,\beta}(c_0, c_1, \ldots, c_{h-1})$$

$$\sum_{0 < m < n} \sum_{\substack{j_0, j_1, \ldots, j_{h-1} \\ \sum_t j_t = m \\ j_t \bmod k = c_t, \ \forall t}} \left(\begin{matrix} m \\ j_0, j_1, \ldots, j_{h-1} \end{matrix} \right)$$

$$\cdot \left(\begin{matrix} n-m-1 \\ n_0 - j_0, \ldots, n_i - j_i - 1, \ldots \text{[illegible]} \end{matrix} \right) . \qquad (38)$$

From (35) and (38), we obtain

$$\bar{I}_{\alpha,\beta} = \sum_{0 < i < h} \sum_{\substack{0 \le c_0, c_1, \ldots, c_{h-1} < k \\ x(i,\alpha,c_i) = 1}} g_{\alpha,\beta}(c_0, c_1, \ldots, c_{h-1})$$

$$\sum_{0 < m < n} \frac{1}{\binom{n}{m}} \sum_{\substack{j_0, \ldots, j_{h-1} \\ \sum_t j_t = m \\ j_t \bmod k = c_t, \ \forall t}} \binom{n_0}{j_0} \binom{n_1}{j_1} \cdots \binom{n_{h-1}}{j_{h-1}} \frac{n_i - j_i}{n-m} . \qquad (39)$$

We can use Lemma 2 to estimate the inner sum as follows:

$$\sum_{0 \leq m < n} \frac{1}{\binom{n}{m}} \sum_{\substack{j_0,\ldots,j_{h-1} \\ \sum_t j_t = m \\ j_t \bmod k = c_t,\ \forall t}} \binom{n_0}{j_0} \cdots \binom{n_{h-1}}{j_{h-1}} \frac{n_i - j_i}{n-m}$$

$$= \frac{n_i}{n} \sum_{0 < m < n} \frac{1}{\binom{n-1}{m}} \sum_{\substack{j_0,\ldots,j_{h-1} \\ \sum_t j_t = m \\ j_t \bmod k = c_t,\ \forall t}} \binom{n_0}{j_0} \cdots \binom{n_i - 1}{j_i} \cdots \binom{n_{h-1}}{j_{h-1}}$$

$$= \frac{n_i}{n} \left(\frac{n}{k^h} + O(n^{2/3}) \right)$$

$$= \frac{1}{hk^h} n + O(n^{2/3}) \qquad (40)$$

Thus, (39) implies

$$\bar{I}_{\alpha,\beta} = \frac{1}{hk^h} \left[\sum_{0 \leq c_0, c_1, \ldots, c_{h-1} < k} g_{\alpha,\beta}(c_0,\ldots,c_{h-1}) \cdot (\#\text{ of } i \text{ with } x(i,\alpha,c_i) = 1) \right] n$$

$$+ O(n^{2/3}) \ . \qquad (41)$$

Let

$$\xi_{\alpha,\beta} = \frac{1}{k^{h-1}} \sum_{0 \leq i < h} \ \sum_{0 \leq c_0, c_1, \ldots, c_{h-1} \leq k} g_{\alpha,\beta}(c_0, c_1, \ldots, c_{h-1}) x(i,\alpha,c_i) \ , \qquad (42)$$

and

$$\xi = \frac{1}{hk} \sum_{0 \leq \alpha < \beta < k} \xi_{\alpha,\beta} \ . \qquad (43)$$

It follows from (34 and (41) that,

$$S_3(\vec{h};n) = \xi n + O(n^{2/3}) .$$

To prove the formula for S_3 in Theorem 1, it remains to show

$$\xi = \psi(h,k) . \tag{44}$$

<u>The Evaluation of ξ .</u>

Write (42) as

$$\xi_{\alpha,\beta} = \sum_{0<i<h} \xi^{(i)}_{\alpha,\beta} , \tag{45}$$

where

$$\xi_{\alpha,\beta} = \frac{1}{k^{h-1}} \sum_{0 \leq c_0,\ldots,c_{h-1}<k} g_{\alpha,\beta}(c_0,c_1,\ldots,c_{h-1})x(i,\alpha,c_i) . \tag{46}$$

In (46), there are at most k^{h-1} non-vanishing terms for each i , since $r(i,\alpha,c_i) = 0$ except for one value of c_i . In fact, we can write

$$\xi^{(i)}_{\alpha,\beta} = \begin{cases} \dfrac{1}{k^{h-1}} \displaystyle\sum_{0 \leq c_0,\ldots,c_{i-1},c_{i+1},\ldots,c_{h-1}<k} \left(\sum_{j \neq i} \eta^{(j)}_{\alpha,\beta}(c_j) \right) \Big| & \text{if } i \in K_{\alpha,\beta} , \\[2ex] \dfrac{1}{k^{h-1}} \displaystyle\sum_{0 \leq c_0,\ldots,c_{i-1},c_{i+1},\ldots,c_{h-1}<k} \left(-1 + \sum_{j \neq i} \eta^{(j)}_{\alpha,\beta}(c_j) \right) \Big| & \text{if } i \in K'_{\alpha,\beta} . \end{cases} \tag{47}$$

The functions $\eta^{(j)}_{\alpha,\beta}$ are defined as follows:

$$\text{If } j \in K_{\alpha,\beta}\text{, then } \eta^{(j)'}_{\alpha,\beta}(c) = \begin{cases} 1 & \text{if } u_j(\alpha) \le c < u_j(\beta)\,, \\ 0 & \text{otherwise,} \end{cases}$$

and (48)

$$\text{if } j \in K'_{\alpha,\beta}\text{, then } \eta^{(j)'}_{\alpha,\beta}(c) = \begin{cases} -1 & \text{if } u_j(\beta) \le c < u_j(\alpha)\,, \\ 0 & \text{otherwise.} \end{cases}$$

Formula (47) follows from (46), by writing $g_{\alpha,\beta}(c_0,c_1,\ldots,c_{h-1})$ as a sum of $\eta^{(j)}_{\alpha,\beta}(c_j)$, which are the contributions to $|\#$ of a-cells - $\#$ of β-cells$|$ from rows j , with the row i contribution explicitly taken care of.

To simplify (47) further, consider the following game using a biased coin with probability q to be a "Head". Suppose we first flip p times, collecting $1 for each occurrence of "Head", and then flip it ℓ times, losing $1 for each "Tail". What is the expected absolute value of payoff? The probability for making $j is,

$$\begin{aligned}
\Pr(\text{payoff} = j) &= \sum_{\substack{a-b=j \\ a,b}} \binom{p}{a} q^a (1-q)^{p-a} \binom{\ell}{b} (1-q)^b q^{\ell-b} \\
&= (1-q)^p q^\ell \sum_{\substack{a-b=j \\ a,b}} \binom{p}{a}\binom{\ell}{b}\left(\frac{q}{1-q}\right)^{a-b} \\
&= (1-q)^p q^\ell \left(\sum_b \binom{\ell}{b}\binom{p}{b+j}\right)\left(\frac{q}{1-q}\right)^j \\
&= (1-q)^p q^\ell \binom{\ell+p}{\ell+j}\left(\frac{q}{1-q}\right)^j \quad .
\end{aligned} \tag{49}$$

The expected absolute value of payoff is then

$$f(p,\ell,q) = \sum_j \Pr(\text{payoff} = j)\cdot |j|$$

$$= (1-q)^p q^\ell \sum_j \binom{\ell+p}{\ell+j}\left(\frac{q}{1-q}\right)^j \cdot |j| \quad , \tag{50}$$

as defined earlier. Let us also define a related function f by

$$f_-(p,\ell,q) = \sum_j \Pr(\text{payoff} = j)\cdot |j-1| \quad . \tag{51}$$

It is easy to verify from (50) and (51) that

$$f_-(p,\ell-1,q) = f(p-1,\ell,q) \quad . \tag{52}$$

Returning to the evaluation of $\xi^{(i)}_{\alpha,\beta}$ from (47)' we note that $\xi^{(i)}_{\alpha,\beta}$ can be regarded as the expected absolute payoffs in the coin game with $\eta^{(\ell)}_{\alpha,\beta}$ the payoff from the ℓ-th coin toss $(\ell \neq i)$. For $i \in K_{\alpha,\beta}$, the parameters of the game are $p = |K_{\alpha,\beta}|-1 = m_{\alpha,\beta}-1$, $\ell = |K'_{\alpha,\beta}| = h-m_{\alpha,\beta}$, and $q = \Delta_{\alpha,\beta}/k$. Thus,

$$\xi^{(i)}_{\alpha,\beta} = f(m_{\alpha,\beta}-1,\ h-m_{\alpha,\beta},\ \Delta_{\alpha,\beta}/k) \text{ , for } i \in K_{\alpha,\beta} \quad . \tag{53}$$

Similarly,

$$\xi^{(i)}_{\alpha,\beta} = f_-(m_{\alpha,\beta},\ h-m_{\alpha,\beta}-1,\ \Delta_{\alpha,\beta}/k) \quad \text{for } i \in K'_{\alpha,\beta} \quad . \tag{54}$$

From (45), (52), (53), (54), we obtain

$$\xi_{\alpha,\beta} = \sum_{0<i<h} \xi^{(i)}_{\alpha,\beta}$$

$$= |K_{\alpha,\beta}|f(m_{\alpha,\beta}-1,\ h-m_{\alpha,\beta},\ \Delta_{\alpha,\beta}/k)$$

$$+ |K'_{\alpha,\beta}|f_-(m_{\alpha,\beta},\ h-m_{\alpha,\beta}-1,\ \Delta_{\alpha,\beta}/k)$$

$$= h\cdot f(m_{\alpha,\beta}-1,\ h-m_{\alpha,\beta},\ \Delta_{\alpha,\beta}/k) \quad . \tag{55}$$

Substituting (55) into (43), we obtain

$$\xi = \frac{1}{k} \sum_{0<\alpha<\beta<k} f(m_{\alpha,\beta}-1,\ h-m_{\alpha,\beta},\ \Delta_{\alpha,\beta}/k$$

$$= \psi(h,k) \quad .$$

This proves (44). The proof of Theorem 1 is complete. $\square$

As an illustration of the formula for $\psi(h,k)$, consider the example $h = 5$, $k = 3$, whose V matrix consists of the first three columns in Figure 8,. It is easy to see that $m_{0,1} = 2$, $\Delta_{0,1} = 2$, $m_{0,2} = 4$, $\Delta_{0,2} = 1$, $m_{1,2} = 2$, $\Delta_{1,2} = 2$. Thus,

$$\psi(5,3) = \frac{1}{3}(f(1,4,2/5) + f(3,3,1/5) + f(1,4,2/5))$$

$$= \frac{2}{3} f(1,4,2/5) + \frac{1}{3} f(3,3,1/5) \quad .$$

4.2 Proof of Theorem 2.

The derivation of the expression for $S_1(\vec{h};n)$ is exactly the same as in the -proof of Theorem 1.

To prove the formula for $S_2(\vec{h};n)$, note that, at the end of Pass 1, we have h independent sorted sublists $L^{(h,\, i)}$, $0 < i < h$. These sublists can be grouped into d lists, with the $(s+1)$-st list M_s $(0 < s < d)$ containing the sublists $L^{(h,\, s+jd)}$ for $0 < j < h'$. The action of Pass 2 is equivalent to performing a k'-sort (using straight-insertions) on each M_s . Therefore,

$$S_2((h,k,1);n) = \sum_{0\le s<d} S_2((h',k',1);\lfloor (n+d-s-1)/d \rfloor) \quad . \tag{56}$$

As $\gcd(h',k') = 1$, we obtain from (56) and Theorem 1 that,

$$S_2((h,k,1);n) = d \times \frac{\sqrt{\pi}}{8} \frac{\sqrt{h'} - (\sqrt{h'})^{-1}}{k'} \left(\frac{n}{d}\right)^{3/2} + O(n)$$

$$= \frac{\sqrt{\pi}}{8} \frac{\sqrt{h} - d(\sqrt{h})^{-1}}{k} n^{3/2} + O(n) \quad .$$

We turn to the evaluation of S_3. Let $\mathcal{L}$ be the set of h-ordered lists L of n elements $\{1,2,\ldots,n\}$. For each $0 \leq m < n$, let Z_m denote the random variable on $\mathcal{L}$ defined as follows. For any $L \in \mathcal{L}$, let L' be the list obtained by k-sorting L, then $Z_m(L)$ is equal to the number of inversions in L' that involve the $(m+1)$-st smallest element, i.e., $m+1$. Clearly,

$$S_3(\vec{h};n) = \frac{1}{2} \sum_{0<m<n} E(Z_m) \quad , \tag{57}$$

where the expectation values are taken for a random $L \in \mathcal{L}$.

Let $\Lambda_s = \{s + \lambda d \mid \lambda = 0,1,2,\ldots,h'-1\}$ and $\Lambda'_s = \{s + \lambda d \mid \lambda = 0,1,2,\ldots,k'-1\}$ for $0 < s < d$. Consider any list $L \in \mathcal{L}$ in its h-row representation. Suppose $Q_m(L)$ has shape $(j_0, j_1, \ldots, j_{h-1})$. Define $\bar{j}_s = \sum_{r \in \Lambda_s} j_r$.

Lemma 3. Suppose that the element $m+1$ is in $L^{(i)}$. Let $t = i \bmod d$, then

$$Z_m(L) = \sum_{0<s<d} |\bar{j}_s - \bar{j}_t| + O(1) \quad .$$

Proof. We first prove the following fact.

Fact 1. Let L' be the list obtained from L by k-sorting. Then in its k-row representation, $Q_m(L')$ has shape $(\ell_0, \ell_1, \ldots, \ell_{k-1})$ with

$$\ell_p = \frac{1}{k'} \bar{j}_{p \bmod d} + O(1) \quad , \quad \text{for} \quad 0 \leq p < k \; .$$

Proof of Fact 1. Let M'_s $(0 \leq s < d)$ be the sublist of L' that consists of $L'^{(k,r)}$, $r \in \Lambda'_s$. Then M'_s can be viewed as obtained from the h'-sorted list M_s by k'-sorting. As $\gcd(h',k') = 1$, the number of r-cells in each row of M_s are the same for all $r \in \Lambda'_s$, up to an additive $O(1)$ term. Thus, for each $p \in \Lambda'_s$,

$$\ell_p = \frac{1}{|\Lambda'_s|} \sum_{r \in \Lambda_s} j_r + O(1) \ . \ \Box$$

To prove Lemma 3, suppose that, in L', element $m+1$ appears in $L'^{(u)}$. Note that $u \bmod d = t$. The number of inversions in L' involving $m+1$ is then

$$Z_m(L) = \sum_{\substack{v \neq u \\ 0 < v < k}} |\ell_v - \ell_u| + O(1) \ .$$

Dividing the range of v into groups Λ'_s, and making use of Fact 1, we obtain

$$Z_m(L) = \sum_{0 \leq s < d} |\bar{j}_s - \bar{j}_{u \bmod d}| + O(1)$$

$$= \sum_{0 < s < d} |\bar{j}_s - \bar{j}_t| + O(1) \ .$$

This proves Lemma 3. $\Box$

Let $n_i = \lfloor (n+h-1-i)/h \rfloor$ for $0 \leq i < h$, and $\bar{n}_s = \sum_{r \in \Lambda_s} n_r$, $0 < s < d$. It is easy to see that $\bar{n}_s = \lfloor (n+d-1-s)/d \rfloor$. Clearly n_i is the length of list $L^{(h,i)}$. It follows from Lemma 3 that,

$$E(Z_m) = \frac{1}{\binom{n}{n_0, n_1, \ldots, n_{h-1}}} \sum_{\substack{j_0, j_1, \ldots, j_{h-1} \\ \sum j_\ell = m}} \sum_{0 \le i < h} \binom{m}{j_0, j_1, \ldots, j_{h-1}}$$

$$\times \binom{n-m-1}{n_0 - j_0, \ldots, n_{i-1} - j_{i-1}, n_i - j_i - 1, n_{i+1} - j_{i+1}, \ldots, n_{h-1} - j_{h-1}}$$

$$\times \left(\sum_{0 \le s < d} |\bar{j}_s - \bar{j}_{i \bmod d}| + O(1) \right) \quad . \tag{58}$$

Formula (58) leads to

$$E(Z_m) = \frac{}{\binom{n}{\bar{n}_0, \bar{n}_1, \ldots, \bar{n}}} \sum_{\substack{\bar{j}_0, \bar{j}_1, \ldots, \bar{j}_{d-1} \\ \sum \bar{j}_r = m}} \sum_{0 \le t < h} \binom{m}{\bar{j}_0, \bar{j}_1, \ldots, \bar{j}_{d-1}}$$

$$\times \binom{n-m-1}{\bar{n}_0 - \bar{j}_0, \ldots, \bar{n}_{t-1} - \bar{j}_{t-1}, \bar{n}_t - \bar{j}_t - 1, \bar{n}_{t+1} - \bar{j}_{t+1}, \ldots, \bar{n}_{d-1} - \bar{j}_{d-1}}$$

$$\times \left(\sum_{0 \le s < d} |\bar{j}_s - \bar{j}_t| \quad O(1) \right) \quad . \tag{59}$$

The derivation of (59) from (58) is elementary but tedious, and will be sketched later. We now observe that (59) can be regarded as, up to an additive $O(1)$ term, the expected number of inversions involving the element $m+1$ in a random d-ordered list of n elements $\{1,2,\ldots,n\}$. Thus, from (57), we have as desired

$$S_3(\vec{h};n) = \frac{1}{2} \sum_{0 \le m < n} E(Z_m)$$

$$= S_2((d,1);n) + O(n)$$

$$= \sum_{0 \le s < t < d} A_{\bar{n}_s + \bar{n}_t} + O(n)$$

$$= \binom{d}{2} \sqrt{\frac{\pi}{128}} \left(\frac{2n}{d}\right)^{3/2} + O(n)$$

$$= \frac{\sqrt{\pi}}{8} \frac{d-1}{\sqrt{d}} n^{3/2} + O(n) \quad .$$

It remains to derive (59) from (58). We write (58) as

$$E(Z_m) = \sum_{\substack{j_0, j_1, \ldots, j_{h-1} \\ \sum j_\ell = m}} \sum_{0 \le i < h} \frac{\binom{n_0}{j_0} \cdots \binom{n_{i-1}}{j_{i-1}} \binom{n_{i+1}}{j_{i+1}} \cdots \binom{n_{h-1}}{j_{h-1}} n_i \binom{n_i - 1}{j_i}}{n \binom{n-1}{m}}$$

$$\times \left(\sum_{0 \le s < d} \bar{j}_s - \bar{j}_{i \bmod d} | \right) + O(1) \quad . \qquad (60$$

For any $0 \le s < d$ and integer J_s , define

$$\Theta_s(J_s) = \sum_{\substack{j_r \ (r \in \Lambda_s) \\ \sum_r j_r = J_s}} \left[\prod_{r \in \Lambda_s} \binom{n_r}{j_r} \right] \quad ,$$

$$\Theta_s'(J_s) = \sum_{i \in \Lambda_s} \sum_{\substack{j_r \ (r \in \Lambda_s \\ \sum_r j_r = J_s}} \left| n_i \binom{n_i - 1}{j_i} \prod_{r \in \Lambda_s - \{i\}} \binom{n_r}{j_r} \right| \quad .$$

Fact 2. $\Theta_s(J_s) = \binom{\bar{n}_s}{J_s}$, and $\Theta'_s(J_s = \bar{n}_s\binom{\bar{n}_s-1}{J_s}$.

Proof of Fact 2. Observe that

$$\prod_{r\in\Lambda_s}(1+x)^{n_r} = (1+x)^{\bar{n}_s} \quad . \tag{61}$$

This gives

$$\sum_{J_s}\Theta_s(J_s)x^{J_s} = \sum_{J_s}\binom{\bar{n}_s}{J_s}x^{J_s} \quad ,$$

and hence the first equality.

To obtain the other equality, we differentiate both sides in (61).

This gives

$$\sum_{i\in\Lambda_s}\left[n_i(1+x)^{n_i-1}\prod_{r\in\Lambda_s-\{i\}}(1+x)^{n_r}\right] = \bar{n}_s(1+x)^{\bar{n}_s-1} \quad .$$

The formula follows immediately by equating terms. $\square$

Now write (60) as

$$E(Z_m) = \sum_{\substack{J_0,J_1,\ldots,J_{d-1}\\ \sum_v J_v = m}} \frac{1}{n\binom{n-1}{m}} \sum_{\substack{j_0,j_1,\ldots,j_{h-1}\\ \sum_{r\in\Lambda_v} j_r = J_v,\ \forall v}} \sum_{0\le t<d}\ \sum_{i\in\Lambda_t}$$

$$\forall v$$

$$\times\left[n_i\binom{n_i-1}{j_i}\prod_{\substack{\ell\ne i\\ 0\le\ell<h}}\binom{n_\ell}{j_\ell}\right]\left(\sum_{0\le s<d}|J_s-J_t|\right) + O(1)$$

$$= \sum_{\substack{J_0, J_1, \ldots, J_{d-1} \\ \sum_v J_v = m}} \frac{1}{n\binom{n-1}{m}} \sum_{0 \le t < d}$$

$$\times \left[\left(\sum_{0 \le s < d} |J_s - J_t| \right) \Theta_t'(J_t) \prod_{\substack{0 < s < d \\ s \ne t}} \Theta_s(J_s) \right] + o(1) \; .$$

Using Fact 2, we obtain

$$E(Z_m) = \sum_{\substack{J_0, J_1, \ldots, J_{d-1} \\ \sum_v J_v = m}} \frac{1}{n\binom{n-1}{m}} \sum_{0 \le t < d} \binom{\bar{n}_0}{J_0} \cdots \binom{\bar{n}_{t-1}}{J_{t-1}} \binom{\bar{n}_{t+1}}{J_{t+1}} \cdots$$

$$\binom{\bar{n}_{d-1}}{J_{d-1}} \bar{n}_t \binom{\bar{n}_t - 1}{J_t} \left(\sum_{0 \le s < d} |J_s - J_t| \right) + o(1) \; .$$

This is just an alternative way of writing (59).

This completes the proof of Theorem 2. $\square$

5. An Exact Analysis of (3,2,1) -Shellsort.

In this section we prove the following theorem. Recall that A_n, the expected number of inversions in a random 2-ordered list of n elements, is $\lfloor n/2 \rfloor 2^{n-2} / \binom{n}{\lfloor n/2 \rfloor}$.

Theorem 3. For $n > 3$,

$$S_1((3,2,1);n) = \frac{1}{4} \sum_{0 \le i \le 2} \lfloor (n+i)/3 \rfloor^2 - \frac{1}{4} n ,$$

$$S_2((3,2,1);n) = \begin{cases} \frac{3}{2} A_{2m} - \frac{1}{8} \dfrac{4^m}{\binom{2m}{m}} & \text{if } n = 3m , \\ A_{2m+1} + \frac{1}{2} A_{2m} - \frac{1}{8} \dfrac{4^m}{\binom{2m}{m}} & \text{if } n = 3m+1 , \\ A_{2m+1} + \frac{1}{2} A_{2m+2} - \frac{1}{8} \dfrac{4^{m+1}}{\binom{2m+2}{m+1}} & \text{if } n = 3m+2 , \end{cases}$$

and

$$S_3((3,2,1);n) = \frac{1}{4} n - \frac{1}{8} + R(n) ,$$

$$\text{where} \quad R(n) = \begin{cases} \dfrac{1}{8(2m+1)} & \text{if } n = 3m , \\ -\dfrac{1}{8(2m+1)} & \text{if } n = 3m+1 , \\ \dfrac{3}{8(2m+3)} & \text{if } n = 3m+2 , \end{cases}$$

The expression for S_1 follows directly from (6). To derive the formula for S_2, we start with formula (18). Write $B_m = B_{odd,even}(2m) + B_{even,odd}(2m)$. We have

$$S_2((3,2,1);n) = \begin{cases} A_{2m} + B_m & \text{if } n = 3m , \\ A_{2m+1} + B_m & \text{if } n = 3m+1 , \\ A_{2m+1} + B_{m+1} & \text{if } n = 3m+2 . \end{cases} \tag{62}$$

Let

$$\mu(i,j,m) = \binom{i+j}{i}\binom{2m-(i+j+1)}{m-j} .$$

It was known (Knuth [1, Exercise 5.2.1-14]) that

$$\binom{2m}{m} A_{2m} = \sum_{0 \le i,j \le m} |i-j|\,\mu(i, j,m) . \tag{63}$$

We extend it to show the following lemma. For $\alpha,\beta \in \{\text{even},\text{odd}\}$, we agree that $\overline{\text{even}} = \text{odd}$, $\overline{\text{odd}} = \text{even}$, and $(-1)^{\alpha+\beta} = 1$ if α , β are both odd or both even, and -1 otherwise.

Lemma 4. For each $\alpha,\beta \in \{\text{even},\text{odd}\}$,

$$\binom{2m}{m} B_{\alpha,\beta}(2m) = \sum_{\substack{0 \le i,j \le m \\ i = a \\ j = \alpha}} \frac{|i-j|}{2} \mu(i,j,m) + \sum_{\substack{0 < i,j < m \\ i = \alpha \\ j = \bar{\alpha}}} \frac{|i-j-(-1)^{\alpha+\beta}|}{2} \mu(i,j,m) .$$

Proof. We will prove the lemma for $\alpha = \text{even}$, $\beta = \text{odd}$; the other three cases can be proved in a similar way.

Let $\mathcal{L}$ be the set of 2 ordered lists $(a_0,b_0,a_1,b_1,\ldots,a_{m-1},b_{m\ 1})$ of 2m elements $\{1,2,\ldots,2m\}$. For each $L \in \mathcal{L}$, let $\Omega_i(L)$ be the set of inversions of the form (a_i,b_ℓ) with odd ℓ . Then

$$\binom{2m}{m} B_{\text{even},\text{odd}}(2m) = \sum_{L \in \mathcal{L}} \sum_{i = \text{even}} |\Omega_i(L)| . \tag{64}$$

For any even i , if a_i is the (i+j+1)-st smallest element in L with $j < i$, then

$$\Omega_i(L) = \{(a_i,b_\ell) \mid \ell \text{ is odd}, j \le \ell < i\} ,$$

implying

$$|\Omega_i(L)| = \begin{cases} \frac{i-j}{2} & \text{if } j \text{ is even,} \\ \frac{i-j+1}{2} & \text{if } j \text{ is odd .} \end{cases}$$

Similarly, if $j > i$ then

$$\Omega_i(L) = \{(a_i,b_\ell) \mid \ell \text{ is odd}, j > \ell \ge i\} ,$$

and

$$|\Omega_i(L)| = \begin{cases} \frac{j-i}{2} & \text{if } j \text{ is even,} \\ \frac{j-i-1}{2} & \text{if } j \text{ is odd ,} \end{cases}$$

Thus, for any even i , if a_i is the (i+j+1)-st smallest element in L , then

$$|\Omega_i(L)| = \begin{cases} \frac{|i-j|}{2} & \text{if } j \text{ is even,} \\ \frac{|i-j+1|}{2} & \text{if } j \text{ is odd .} \end{cases} \tag{65}$$

Observing that $\mu(i,j,m)$ is the number of $L \in \mathcal{L}$ such that a_i is the (i+j+1)-st smallest element in L , we have from (64) and (65)

$$\binom{2m}{m} B_{\text{even,odd}}(2m) = \sum_{i=\text{even}} \sum_{L \in \mathcal{L}} |\Omega_i(L)|$$

$$= \sum_{i=\text{even}} \left(\sum_{j=\text{even}} \frac{|i-j|}{2} \mu(i,j,m) + \sum_{j=\text{odd}} \frac{|i-j+1|}{2} \mu(i,j,m) \right) .$$

This proves Lemma 4 when α = even , β = odd . □

Define

$$W_m = \sum_{\substack{j<i \\ i=\text{odd} \\ j=\text{even}}} \mu(i,j,m) + \sum_{\substack{j<i \\ i=\text{even} \\ j=\text{odd}}} \mu(i,j,m) - \sum_{\substack{j>i \\ i=\text{even} \\ j=\text{odd}}} \mu(i,j,m) - \sum_{\substack{j>i \\ i=\text{odd} \\ j=\text{even}}} \mu(i,j,m) .$$

From Lemma 4 and (63), we obtain after some manipulations,

$$\binom{2m}{m} B_m = \frac{1}{2} \sum_{0 \le i,j \le m} |i-j| \mu(i,j,m) + \frac{1}{2} W_m$$

$$= \frac{1}{2} \binom{2m}{m} A_{2m} + \frac{1}{2} W_m \quad . \tag{66}$$

We shall now show that, for $m > 1$,

$$W_m = -4^{m-1} \quad . \tag{67}$$

Formulas (62), (66) and (67) imply immediately the expression of S_2 given in the theorem.

To prove (67), we use a result due to R. Sedgewick [6].

<u>Lemma 5</u> (Sedgewick [6, Theorem 2]). Let $f(i,j)$ be a function defined for integers $0 < i,j < m$, and satisfying $f(i,j) = f(i-j,0)$ for $i > j$, $f(i,j) = f(0,j-i)$ for $j \ge i$, and $f(0,j+1) = f(j,0)+1$. Then

$$\sum_{0 \le i,j < m} f(i,j) \mu(i,j,m) = \sum_{j \ge 1} \binom{2m}{m-j} (2F(j)+j) ,$$

where $F(j) = \sum_{0 \le \ell < j} f(\ell, 0)$.

In our problem, we write

$$W_m = - \sum_{0 \le i, j \le m} f(i,j)\mu(i,j,m) ,$$

where

$$f(i,j) = \begin{cases} 0 & \text{for } i, j \text{ both even or both odd,} \\ -1 & 0 \le j < i \le m \\ 1 & 0 \le i < j \le m \end{cases} \quad \text{otherwise.}$$

It is easy to verify that all conditions in Lemma 5 on f are satisfied. Note that

$$F(k) = \sum_{0 < j < k} f(j,0) = -\lfloor k/2 \rfloor .$$

We have, using Lemma 5

$$\begin{aligned} W_m &= - \sum_{k \ge 1} \binom{2m}{m-k} (-2\lfloor k/2 \rfloor + k) \\ &= - \sum_{k \ge 1} \binom{2m}{m-2k+1} \\ &= - \sum_{k > 1} \left[\binom{2m-1}{m-2k+1} + \binom{2m-1}{m-2k} \right] \\ &= - \sum_{k \ge 1} \binom{2m-1}{m-k} \\ &= -2^{2m-2} . \end{aligned}$$

This proves (67), and hence the expression for S_2 in Theorem 3.

We now derive the expression for S_3 . The derivation will be given only for the case $n = 3m$ (integer $m \ge 1$). The other two cases $n = 3m+1$, $n = 3m+2$ can be similarly treated.

Recall the formulas (24), (25), (26) for the case $n = 3m$:

$$S_3((3,2,1);n) = \sum_{1 \le i \le \lfloor n/2 \rfloor} E(z_{i,i}) + \sum_{1 \le i \le \lceil n/2 \rceil - 1} E(z_{i+1,i}) , \qquad (68)$$

$$E(z_{i,i}) = \frac{1}{\binom{3m}{2i-1}} \sum_{\substack{k_1+k_2+k_3 = 2i-1 \\ k_1,k_3 \text{ are even} \\ k_2 \text{ is odd}}} \binom{m}{k_1}\binom{m}{k_2}\binom{m}{k_3} , \quad 1 \le i \le \lfloor n/2 \rfloor , \qquad (69)$$

and

$$E(z_{i+1,i}) = \frac{1}{\binom{3m}{2i}} \sum_{\substack{k_1+k_2+k_3 = 2i \\ k_1,k_3 \text{ are odd} \\ k_2 \text{ is even}}} \binom{m}{k_1}\binom{m}{k_2}\binom{m}{k_3} , \quad 1 \le i < \lceil n/2 \rceil . \qquad (70)$$

Consider the expansions

$$\frac{1}{2} [(1+x)^n + (1-x)^n] = \sum_{k = \text{even}} \binom{n}{k} x^k ,$$

and

$$\frac{1}{2} [(1+x)^n - (1-x)^n] = \sum_{k = \text{odd}} \binom{n}{k} x^k .$$

It is easy to see that the quantity $E(z_{i,i})$, as given by (69), is $\binom{3m}{2i-1}^{-1}$ times the coefficient of x^{2i-1} in the function

$$\frac{1}{8} ((1+x)^m + (1-x)^m)^2((1+x)^m - (1-x)^m)$$

$$= \frac{1}{8} ((1+x)^m + (1-x)^m)((1+x)^{2m} - (1-x)^{2m})$$

$$= \frac{1}{8} ((1+x)^{3m} - (1-x)^{3m}) + \frac{1}{8} ((1-x)^m(1+x)^{2m} - (1-x)^{2m}(1+x)^m)$$

$$= \frac{1}{4} \sum_i \binom{3m}{2i-1} x^{2i-1} + \frac{1}{4} \times (\text{terms of odd -powers in } (1-x)^m(1+x)^{2m}) .$$

Let $a_{m,k}$ be defined by

$$(1-x)^m(1+x)^{2m} = \sum_{k\geq 0} a_{m,k}x^k \quad . \tag{71}$$

Then

$$E(z_{i,i}) = \frac{1}{4} + \frac{1}{4}\frac{a_{m,2i-1}}{\binom{3m}{2i-1}} \qquad \text{for } 1 \leq i \leq \lfloor 3m/2 \rfloor \quad . \tag{72}$$

Similarly, we can show that

$$E(z_{i+1,i}) = \frac{1}{4} - \frac{1}{4}\frac{a_{m,2i}}{\binom{3m}{2i}} \qquad \text{for } 1 \leq i < \lceil 3m/2 \rceil \quad . \tag{73}$$

From (68), (72) and (73), we obtain

$$S_3((3,2,1);n) = \frac{1}{4}(n-1) - \frac{1}{4}\sum_{0<k<3m}(-1)^k \frac{a_{m,k}}{\binom{3m}{k}} \quad .$$

Noting that $a_{m,0} = 1$ and $a_{m,3m} = (-1)^m$, we can write this as

$$S_3((3,2,1);n) = \frac{1}{4}(n+1) - \frac{1}{4}\sum_{0\leq k\leq 3m}(-1)^k \frac{a_{m,k}}{\binom{3m}{k}} \tag{74}$$

Let

$$a_m = \sum_{0\leq k\leq 3m}(-1)^k \frac{a_{m,k}}{\binom{3m}{k}}$$

$$= \sum_{0\leq k\leq 3m}(-1)^k \frac{\Gamma(k+1)\Gamma(3m-k+1)}{\Gamma(3m+1)} a_{m,k} \quad , \tag{75}$$

where $\Gamma(x)$ is the <u>Gamma function</u> (see, e.g. [1]). The <u>Beta function</u> (see, e.g. [1]) defined by

$$B(x,y) = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}$$

has an integral representation (when $\mathrm{Re}\ x > 0$, $\mathrm{Re}\ y > 0$),

$$B(x,y) = \int_0^1 dt\ t^{x-1}(1-t)^{y-1}\ dt\ .$$

We can write (75) as

$$\begin{aligned} a_m &= \sum_{0\le k\le 3m} (-1)^k a_{m,k}(3m+1)B(k+1,3m-k+1) \\ &= (3m+1) \sum_{0\le k\le 3m} (-1)^k a_{m,k} \int_0^1 t^k(1-t)^{3m-k}\ dt \\ &= (3m+1) \int_0^1 \left\{ \sum_{0\le k\le 3m} a_{m,k}\left(\frac{-t}{1-t}\right)^k \right\} (1-t)^{3m}\ dt\ . \end{aligned}$$

Using (71), we have

$$\begin{aligned} a_m &= (3m+1) \int_0^1 \left(1 + \frac{t}{1-t}\right)^m \left(1 - \frac{t}{1-t}\right)^{2m} (1-t)^{3m}\ dt \\ &= (3m+1) \int_0^1 (1-2t)^{2m}\ dt \\ &= \frac{3m+1}{2m+1}\ . \end{aligned}$$

Therefore, from (74), we obtain

$$\begin{aligned} S_3((3,2,1);n) &= \frac{1}{4}(n+1) - \frac{1}{4}\alpha_m \\ &= \frac{1}{4}(n+1) - \frac{1}{4}\frac{3m+1}{2m+1} \\ &= \frac{1}{4}n - \frac{1}{8} + \frac{1}{8(2m+1)}\ , \end{aligned}$$

for $n = 3m$. This proves the formula for S_3 when $n = 3m$.

We have completed the proof of Theorem 3. $\square$

6. Concluding Remarks.

In this paper we have analyzed the asymptotic behavior of $(h,k,1)$-Shellsort for fixed h, k. This procedure can be generalized to analyze $\vec{h}$-Shellsort with more than 3 increments. We shall report the results in a future paper, where we shall also study the situation when $\vec{h}$ varies with n.

Acknowledgements. I wish to thank Bob Sedgewick for helpful suggestions that led to the simplification of several derivations.

Appendix: Proof of Lemma 2.

In this appendix we will show that, for each $\vec{c} \in C_{h,k}$,

$$G_{h,k}(\vec{c};\vec{n}) = (N+1)/k^h + O(N^{2/3}) , \tag{A.1}$$

For notations and definitions, see Section 3 of the paper. In this appendix, the constants in the O-notation can depend on h and k , which are fixed integers.

Write $c_i' = (n_i - c_i) \bmod k$ for $0 \le i < h$, and

$$G^{(1)}_{h,k}(\vec{c};\vec{n}) = \sum_{0 \le w \le \lfloor N/2 \rfloor} X_{h,k}(\vec{c};\vec{n};w) , \tag{A.2}$$

where

$$X_{h,k}(\vec{c};\vec{n};w) = \frac{1}{\binom{N}{w}} \sum_{\substack{\sum_i j_i = w \\ j_i \bmod k = c_i \\ \forall i}} \binom{n_0}{j_0}\binom{n_1}{j_1}\cdots\binom{n_{h-1}}{j_{h-1}} .$$

It is easy to see that

$$G_{h,k}(\vec{c};\vec{n}) = G^{(1)}_{h,k}(\vec{c};\vec{n}) + G^{(1)}_{h,k}(\vec{c}\,';\vec{n}) + O(1) .$$

To prove (A.1), it clearly suffices to prove the following result:

$$G^{(1)}_{h,k}(\vec{c};\vec{n}) = (\lfloor N/2 \rfloor + 1)/k^h + O(N^{2/3}) \quad \text{for each } \vec{c} . \tag{A.3}$$

We shall prove (A.3) by establishing the following claim.

<u>Claim</u>. If $\vec{c} = (c_0, c_1, c_2, \ldots, c_{h-1}) \in C$ and $\vec{d} = ((c_0+1) \bmod k, c_1, c_2, \ldots$ [illegible]$)$ then

$$G^{(1)}_{h,k}(\vec{c};\vec{n}) \le G^{(1)}_{h,k}(\vec{d};\vec{n}) + O(N^{2/3}) .$$

The Claim would imply that, if $\vec{c},\vec{d} \in C$ differ only in the first component, then their values of $G^{(1)}_{h,k}$ differ at most by $O(N^{2/3})$. By symmetry, this conclusion is also true for $\vec{c}$, $\vec{d}$ differing only in any one component. It then follows that, for any $\vec{c},\vec{d} \in C$,

$$G^{(1)}_{h,k}(\vec{c};\vec{n}) = G^{(1)}_{h,k}(\vec{d};\vec{n}) + O(N^{2/3}) .$$

Formula (A.3) follows as

$$\sum_{\vec{c} \in C} G^{(1)}_{h,k}(\vec{c};\vec{n}) = \sum_{0 \le w \le \lfloor N/2 \rfloor} \sum_{\vec{c} \in C} x_{h,k}(\vec{c};\vec{n})$$

$$= \lfloor N/2 \rfloor + 1 .$$

It remains to prove the Claim. For $i \in \{1,3\}$, let $R_i(w)$ denote the set of integers in the interval $(w/h - iw^{2/3}, w/h + iw^{2/3})$.

Lemma A1. If $1 \le w \le \lfloor N/2 \rfloor$, then

$$\sum_{j_0 \notin R_i(w)} \frac{\binom{n_0}{j_0}\binom{N-n_0}{w-j_0}}{\binom{N}{w}} = O(w^{-1/3}) \quad , \text{ for each } i \in \{1,3\} .$$

Proof. The hypergeometric distribution $p_k = \binom{n_0}{k}\binom{N-n_0}{w-k}\Big/\binom{N}{w}$, $(k = 0,1,2,\dots)$, has expected value $wn_0/N = w/h + O(1)$ and variance

$$w \frac{n_0}{N}\left(1 - \frac{n_0}{N}\right)\left(1 - \frac{w-1}{N-1}\right) = O(w) \quad ,$$

(see, e.g. Rényi [5, p. 1051). The lemma then follows from Chebychev's Inequality (see, e.g. Rény [5, p. 373]). □

Lemma A2. Let $1 \leq w \leq \lfloor N/2 \rfloor$, $0 \leq j_0 < n_0$, and $j_0 - \frac{w}{h} = O(w^{2/3})$.

Then

$$\frac{\binom{n_0}{j_0}}{\binom{N}{w}} = \frac{\binom{n_0}{j_0+1}}{\binom{N}{w+1}} (1 + O(w^{-1/3})) \quad .$$

Proof.

$$\frac{\binom{n_0}{j_0}}{\binom{N}{w}} \times \frac{\binom{N}{w+1}}{\binom{n_0}{j_0+1}} = \frac{(j_0+1)(N-w)}{(n_0-j_0)(w+1)}$$

$$= \frac{w/h + O(w^{2/3})}{(N-w)/h - O(w^{2/3})} \; \frac{N-w}{w+1}$$

$$= 1 + O(w^{-1/3}) \quad . \quad \square$$

We shall now prove the Claim. Without loss of generality, assume N is-large enough so that $n_0 \notin R_3(w)$. Let $\mathcal{J}(\vec{c};w)$ denote the set of h-tuples $(j_0, j_1, \ldots, j_{h-1})$ of non-negative integers such that $\sum_i j_i = w$, and $j_i \bmod k = c_i$ for all i. Then, for each $1 \leq w \leq \lfloor N/2 \rfloor$,

$$\chi_{h,k}(c;n;w) = \sum_{\vec{j} \in \mathcal{J}(\vec{c};w)} \frac{1}{\binom{N}{w}} \binom{n_0}{j_0}\binom{n_1}{j_1}\cdots\binom{n_{h-1}}{j_{h-1}}$$

$$= \sum_{\substack{j_0 \in R_3(w) \\ \vec{j} \in \mathcal{J}(\vec{c},w)}} \frac{1}{\binom{N}{w}} \binom{n_0}{j_0}\binom{n_1}{j_1}\cdots\binom{n_{h-1}}{j_{h-1}}$$

$$+ \sum_{\substack{j_0 \notin R_3(w) \\ \vec{j} \in \mathcal{J}(\vec{c},w)}} \frac{1}{\binom{N}{w}} \binom{n_0}{j_0}\binom{n_1}{j_1}\cdots\binom{n_{h-1}}{j_{h-1}}$$

$$= \sum_{\substack{j_0 \in R_3(w) \\ \vec{j} \in \mathcal{J}(\vec{c}, w)}} \frac{1}{\binom{N}{w}} \binom{n_0}{j_0}\binom{n_1}{j_1}\cdots\binom{n_{h-1}}{j_{h-1}}$$

$$+ \; O\left(\sum_{j_0 \notin R_3(w)} \frac{\binom{n_0}{j_0}}{\binom{N}{w}} \sum_{\substack{j_1,\ldots,j_{h-1} \\ j_1+\cdots+j_{h-1} = w-j_0}} \binom{n_1}{j_1}\cdots\binom{n_{h-1}}{j_{h-1}} \right).$$

Using*/ Lemmas A1 and A2, we have then

$$\chi_{h,k}(\vec{c};\vec{n};w) = \sum_{\substack{j_0 \in R_3(w) \\ \vec{j} \in \mathcal{J}(\vec{c}, w)}} (1+ O(w^{-1/3}))\frac{1}{\binom{N}{w+1}} \binom{n_0}{j_0+1}\binom{n_1}{j_1}\cdots\binom{n_{h-1}}{j_{h-1}}$$

$$+ \; O\left(\sum_{j_0 \notin R_3(w)} \frac{\binom{n_0}{j_0}\binom{N-n_0}{w-j_0}}{\binom{N}{w}} \right)$$

$$= \sum_{\substack{j_0 \in R_3(w) \\ \vec{j} \in \mathcal{J}(\vec{c}, w)}} \binom{n_0}{j_0+1}\binom{n_1}{j_1}\cdots\binom{n_{h-1}}{j_{h-1}} + O(w^{-1/3}) \quad . \qquad \text{(A.4)}$$

It is straightforward to check that $R_1(w+1) \subseteq \{j_0+1 \mid j_0 \in R_3(w)\}$, Thus, from (A.4),

*/ Note that $j_0 \in R_3(w)$ implies $j_0 \neq n_0$, because we have assumed $n_0 \notin R_3(w)$. This enables us to apply Lemma A2 in the ensuing derivation.

$$\chi_{h,k}(\vec{c};\vec{n};w) \geq \sum_{\substack{j_0 \in R_1(w+1 \\ \vec{j} \in \mathcal{J}(\vec{d},w+1)}} \frac{1}{\binom{N}{w+1}} \binom{n_0}{j_0}\binom{n_1}{j_1}\cdots\binom{n_{h-1}}{j_{h-1}} + O(w^{-1/3})$$

$$= \chi_{h,k}(\vec{d};\vec{n};w+1) + O(w^{-1/3}) \quad . \tag{A.5}$$

In the last line of the above derivation, we used a formula similar to (A.4) for $\chi_{h,k}(\vec{d};\vec{n};w+1$. On the other hand, it follows directly from (A.4) that

$$\chi_{h,k}(\vec{c};\vec{n};w) \leq \chi_{h,k}(\vec{d};\vec{n};w+1) + O(w^{-1/3}) \quad . \tag{A.6}$$

Therefore, we obtain from (A.5) and (A.6) that

$$\chi_{h,k}(\vec{c};\vec{n};w) = \chi_{h,k}(\vec{d};\vec{n};w+1 + O(w^{-1/3}) \quad . \tag{A.7}$$

From (A.2) and (A.7), we obtain

$$G^{(1)}_{h,k}(\vec{c};\vec{n}) = \sum_{1 \leq w \leq \lfloor N/2 \rfloor} \chi_{h,k}(\vec{c};\vec{n};w) + O(1)$$

$$= \sum_{1 \leq w \leq \lfloor N/2 \rfloor} \chi_{h,k}(\vec{d};n;w+1) + \sum_{1 \leq w \leq \lfloor N/2 \rfloor} O(w^{-1/3}) + O(1)$$

$$= G^{(1)}_{h,k}(\vec{d};\vec{n}) + O(N^{2/3})$$

This proves the Claim, and completes the proof of Lemma 2. □

References

[1] G. F. Carrier, M. Krook, and C. E. Pearson, Functions of A Complex Variable, McGraw-Hill, New York, 1966.

[2] D. E. Knuth, The Art of Computer Programming Vol. 3, Addison-Wesley, Reading, Mass., 2nd printing, 1975.

[3] D. E. Knuth, "Mathematical Analysis of Algorithms," Information Processing (IFIP) 71 (1972), North Holland, 19-27.

[4] V. R. Pratt, "Shellsort and Sorting Networks," Ph.D. Thesis, Computer Science Department, Stanford University, 1972.

[5] A. Rényi, Probability Theory, North-Holland/ American Elsevier, New York, 1970.

[6] R. Sedgewick, "Data Movement in Odd-Even Merging," SIAM J. on Computing 7 (1978), 239-272.

[7] R. Sedgewick, "The Analysis of Quicksort Programs," Acta Informatica 7 (1977), 327-355.

[8] D. L. Shell, "A High-Speed Sorting Procedure," Comm. ACM 2 (1959), 30-32.