



POLYNOMIALS THAT ENCODE DEPTH AND NODE INFORMATION FOR THE BINARY TREES CREATED BY QUICKSORT

DAVID M. BRADLEY^{1,*}, JOHANN THIEL²

¹Department of Mathematics & Statistics, University of Maine, Orono, ME 04469, USA

²Department of Mathematics, New York City College of Technology, Brooklyn, NY 11201, USA

To the memory of Jonathan M. Borwein (1951–2016) on the occasion of his 75th birthday

Abstract. We obtain a generating series whose coefficients are polynomials that, for a given positive integer n , encode the depth at which the various list entries appear as labeled nodes in the binary trees obtained by QuickSorting permutations of the list consisting of one copy of each of the first n non-negative integers. Extracting the appropriate coefficients yields information in the form of “node distribution polynomials” that encode the distribution of list entries that appear at different depths, and “depth distribution polynomials” that encode the distribution of depths at which a given list entry appears. These yield formulas for the number of times a given list entry appears at a given depth, the total number of list entries that appear at a given depth, and consequently the average number of list entries that appear at a given depth taken over all $n!$ permutations.

Keywords. Binary search trees; Node distribution; Permutations; QuickSort algorithm.

2020 Mathematics Subject Classification. 05A15, 68R05, 05C30, 05C05.

1. INTRODUCTION

QuickSort is a well-studied algorithm (see Algorithm 1 below) for sorting lists of items drawn from a totally ordered set. The items need not be distinct [18], but for simplicity we assume that they are. Many variations exist, but at its core, the algorithm selects a pivot from the list and creates a left sublist, consisting of the items less than the pivot, and a right sublist, consisting of the items greater than the pivot. It then recurses on each of the sublists. The algorithm terminates with a binary tree whose nodes consist of the entries of the original list, and with the property that for any node, the left child (if there is one) is less than its parent, while the right child (if there is one) is greater than its parent. For an example, see Figure 1. Although not essential to our analysis, for concreteness we also assume that the pivot is the first list entry at each stage, as in the following pseudocode (Algorithm 1).

*Corresponding author.

E-mail address: david.bradley@maine.edu (D. Bradley), JThiel@citytech.cuny.edu (J. Thiel).

Received April 15, 2025; Accepted March 27, 2026.

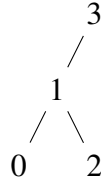


FIGURE 1. Tree produced by QuickSorting the list 3, 1, 0, 2.

Algorithm 1 QuickSort pseudocode

```

1: procedure QUICKSORT( $A = a_0, a_1, \dots, a_{n-1}$  : list of  $n$  real numbers)
2:   if  $n \leq 1$  then
3:     return  $A$ 
4:   else
5:      $pivot := a_0$ 
6:      $left :=$  empty list
7:      $right :=$  empty list
8:     for  $i := 1$  to  $n - 1$  do
9:       if  $a_i \leq pivot$  then
10:         $left := \text{concatenate}(left, a_i)$ 
11:       else
12:         $right := \text{concatenate}(right, a_i)$ 
13:       end if
14:     end for
15:     return  $\text{concatenate}(\text{QUICKSORT}(left), pivot, \text{QUICKSORT}(right))$ 
16:   end if
17: end procedure

```

It is well-known that the height of the tree (and hence the run-time performance of the algorithm) can vary significantly depending on the initial ordering of the list entries and the choice of pivot at each stage. Consequently, it is of interest to collect information concerning the shapes of the trees obtained by QuickSorting every possible list ordering. Recording the number of nodes that appear at various depths would be a step in this direction. We refine that approach here by constructing polynomials that encode the depths at which each of the nodes appears, summed (3.1) over all permutations of the input list, which for convenience and without loss of generality, we assume is an initial segment of the non-negative integers.

Although keeping track of all that information may seem like a gargantuan task, it can nevertheless be organized and encapsulated in the compact form of a single generating function—the exponential generating series of the encoding polynomials—expressed in terms of an integral (Theorem 3.2). As a consequence, the encoding polynomials have a relatively simple representation in terms of the Gegenbauer (ultraspherical) polynomials; see Theorem 3.3 and the remark following its proof. Performing the necessary coefficient extractions on the encoding polynomials yields precise information concerning the distribution of depths at which a given list entry appears (Section 4), and the distribution of list entries that appear at a given depth (Section 5). As result, we obtain an “explicit” formula (not a recurrence) for the exact number of times that

a given list entry appears at a given depth (Theorem 4.1) and a relatively simple formula for the total number of list entries that appear at a given depth (Theorem 6.1). It goes without saying that our apparatus can also be used to obtain the standard formula for the average number of comparisons performed by QuickSort, typically obtained in isolation by quite different methods (Section 6.2).

The study of the average-case behavior of QuickSort originated with Hoare's introduction of the algorithm itself [7], together with his probabilistic analysis that revealed its expected $O(n \log n)$ performance [8]. Knuth (see e.g., [12, §5.2.2, pp. 114–123] and [14, §5.2.2, pp. 113–122]) placed these results within a broader analytic framework, developing systematic methods for determining the run-time complexity and asymptotic behavior of many of the most important algorithms of interest to computer scientists. Knuth's doctoral student Robert Sedgewick (now William O. Baker Professor of Computer Science at Princeton University) deepened this line of work through detailed combinatorial modeling and refined average-case analyses of QuickSort variants, tightly connecting mathematical analysis with practical algorithm engineering [16, 17, 19, 20]. Philippe Flajolet (1948–2011), in collaboration with Sedgewick, later unified and extended these approaches by bringing powerful analytic-combinatorial tools—such as generating functions and singularity analysis—to bear [4], enabling precise distributional results and a general theory for the average-case analysis of QuickSort and related divide-and-conquer algorithms [22].

2. PRELIMINARIES

Notation. For non-negative integer n , let $\mathbf{N}_n := \{0, 1, 2, \dots, n-1\}$ denote the (empty if $n = 0$) set consisting of the first n non-negative integers, and let $\mathfrak{S}_n$ denote the set of permutations (bijections) of $\mathbf{N}_n$. The unsigned Stirling numbers of the first kind are defined as the coefficients of the powers of x in the expansion [13, eq. (2.18), p. 415] [23, eq. (3.5.2), p. 82]

$$\sum_{k=0}^{\infty} \begin{bmatrix} n \\ k \end{bmatrix} x^k = n! \binom{n+x-1}{n} = \prod_{j=0}^{n-1} (x+j). \quad (2.1)$$

They are referred to as Stirling cycle numbers by Knuth [13, p. 420]. We adopt the convention, prevalent in some quarters, for expressing the coefficient a_k of a power series $A(x) = \sum_{k=0}^{\infty} a_k x^k$ by $[x^k]A(x)$. As an example, equation (2.1) is equivalent to the statement that

$$\begin{bmatrix} n \\ k \end{bmatrix} = [x^k] n! \binom{n+x-1}{n} \quad (2.2)$$

for non-negative integers n and k . The notation extends in a straightforward manner to power series in more than one variable [5, Ch. 1, §1.1.2, pp. 1–2]. The harmonic numbers are the partial sums of the harmonic series defined by

$$H_n = \sum_{m=1}^n \frac{1}{m} = 1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n}, \quad (2.3)$$

with $H_0 = 0$ in accordance with the convention that empty sums vanish. Finally, as is now customary, we denote by γ the Euler-Mascheroni constant specified by the limit

$$\gamma := \lim_{n \rightarrow \infty} (H_n - \log n). \quad (2.4)$$

3. POLYNOMIALS THAT ENCODE THE DEPTH OF EACH NODE

Define a sequence of polynomials in the ring $\mathbf{Z}[x, y]$ of polynomials in the indeterminates x and y with integer coefficients by setting $Q_0(x, y) = 0$, and for positive integer n ,

$$Q_n(x, y) = \sum_{\sigma \in \mathfrak{S}_n} \sum_{j \in \mathbf{N}_n} x^{d(\sigma(j))} y^{\sigma(j)}, \quad (3.1)$$

where $d(\sigma(j))$ denotes the depth of the node $\sigma(j)$ in the tree created by QuickSorting the list $\sigma(0), \sigma(1), \dots, \sigma(n-1)$, with the convention that the root $\sigma(0)$ has depth 0, children of the root have depth 1, and so on. Let n be a positive integer. By decomposing the inner sum in (3.1) according to whether the node $\sigma(j)$ appears in the left subtree, the right subtree, or coincides with the root, one readily finds that

$$Q_n(x, y) = \sum_{\sigma \in \mathfrak{S}_n} \left\{ \left(\sum_{\substack{j \in \mathbf{N}_n \\ \sigma(j) < \sigma(0)}} x^{d(\sigma(j))} y^{\sigma(j)} \right) + x^{d(\sigma(0))} y^{\sigma(0)} + \left(\sum_{\substack{j \in \mathbf{N}_n \\ \sigma(j) > \sigma(0)}} x^{d(\sigma(j))} y^{\sigma(j)} \right) \right\}.$$

For each $k \in \{1, 2, \dots, n\}$, there are $(n-1)!$ permutations $\sigma \in \mathfrak{S}_n$ such that $\sigma(0) = k-1$. Of these, there are $(n-1)!/(k-1)!$ with the $k-1$ nodes $\mathbf{N}_{k-1} = \{\sigma(j) : j \in \mathbf{N}_n, \sigma(j) < k-1\}$ of the left subtree that appear in the order specified by σ , but for which the other $n-k$ nodes in $\mathbf{N}_n$ appear in no particular order. Similarly, there are $(n-1)!/(n-k)!$ with the $n-k$ nodes $\{\sigma(j) : j \in \mathbf{N}_n, \sigma(j) > k-1\}$ of the right subtree that appear in the order specified by σ , but for which the other $k-1$ nodes in $\mathbf{N}_n$ appear in no particular order. Since each node $\sigma(j)$ of the left subtree can be regarded as a node $\sigma_L(\ell)$ of the tree generated by QuickSorting some $\sigma_L \in \mathfrak{S}_{k-1}$ with $\ell \in \mathbf{N}_{k-1}$ and $d(\sigma(j)) = 1 + d(\sigma_L(\ell))$, and each node $\sigma(j)$ of the right subtree can be regarded as a node $k + \sigma_R(r)$ of the tree generated by QuickSorting some $\sigma_R \in \mathfrak{S}_{n-k}$ with $r \in \mathbf{N}_{n-k}$ and $d(\sigma(j)) = 1 + d(\sigma_R(r))$, it follows that

$$Q_n(x, y) = \sum_{k=1}^n \left(\frac{(n-1)!}{(k-1)!} \sum_{\sigma_L \in \mathfrak{S}_{k-1}} \sum_{\ell \in \mathbf{N}_{k-1}} x^{1+d(\sigma_L(\ell))} y^{\sigma_L(\ell)} + (n-1)! x^0 y^{k-1} + \frac{(n-1)!}{(n-k)!} \sum_{\sigma_R \in \mathfrak{S}_{n-k}} \sum_{r \in \mathbf{N}_{n-k}} x^{1+d(\sigma_R(r))} y^{k+\sigma_R(r)} \right).$$

Consequently, we have the recurrence

$$\frac{Q_n(x, y)}{n!} = \frac{1}{n} \sum_{k=1}^n \left(\frac{x Q_{k-1}(x, y)}{(k-1)!} + y^{k-1} + \frac{xy^k Q_{n-k}(x, y)}{(n-k)!} \right) \quad (1 \leq n \in \mathbf{Z}). \quad (3.2)$$

Example 3.1. From (3.1) we find that $Q_1(x, y) = 1$, $Q_2(x, y) = (1 + xy) + (x + y)$, and, from Figure 2,

$$Q_3(x, y) = (1 + xy + x^2 y^2) + (1 + xy^2 + x^2 y) + 2(x + y + xy^2) + (x + x^2 y + y^2) + (x^2 + xy + y^2).$$

More simply, the foregoing polynomials can be successively obtained by employing the recurrence (3.2), starting with the initial polynomial $Q_0(x, y) = 0$. $\square$

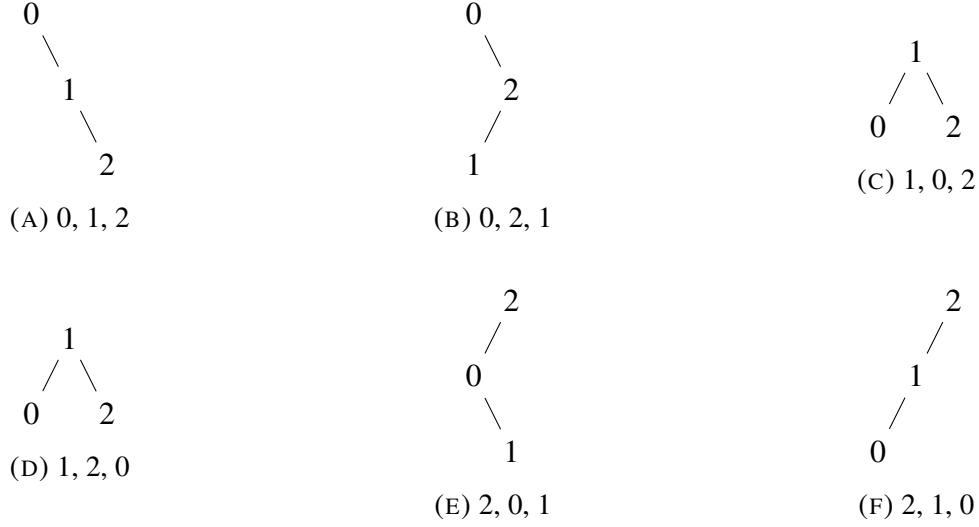


FIGURE 2. Trees produced by QuickSorting the 6 permutations of the list 0, 1, 2.

To solve the recurrence, we begin by forming the exponential generating series

$$Q(x, y, z) := \sum_{n=0}^{\infty} Q_n(x, y) \frac{z^n}{n!} = \sum_{n=1}^{\infty} Q_n(x, y) \frac{z^n}{n!}, \quad (3.3)$$

which as a formal power series in z may start at either $n = 0$ or $n = 1$ since by definition, $Q_0(x, y) = 0$.

Theorem 3.2. *The exponential generating series (3.3) has integral representation*

$$Q(x, y, z) = (1 - z)^{-x} (1 - yz)^{-x} \int_0^z (1 - t)^{x-1} (1 - yt)^{x-1} dt. \quad (3.4)$$

Proof. Let n be a positive integer. By regarding the right hand side of the recurrence (3.2) as three separate summations, adding the terms in the rightmost sum in the opposite order, and then recombining, we obtain the equivalent recurrence

$$\frac{Q_n(x, y)}{n!} = \frac{1}{n} \sum_{k=0}^{n-1} \left(y^k + (1 + y^{n-k}) \frac{x Q_k(x, y)}{k!} \right). \quad (3.5)$$

After multiplying each side of (3.5) by z^n , summing over all positive integers n , and interchanging the order of summation, we find that

$$\begin{aligned} Q(x, y, z) &= \sum_{n=1}^{\infty} \frac{z^n}{n} \sum_{k=0}^{n-1} \left(y^k + (1 + y^{n-k}) \frac{x Q_k(x, y)}{k!} \right) \\ &= \sum_{k=0}^{\infty} \sum_{n=k+1}^{\infty} \frac{z^n}{n} \left(y^k + (1 + y^{n-k}) \frac{x Q_k(x, y)}{k!} \right). \end{aligned}$$

Consequently,

$$\begin{aligned}
\frac{\partial}{\partial z} Q(x, y, z) &= \sum_{k=0}^{\infty} \sum_{n=k+1}^{\infty} z^{n-1} \left(y^k + \left(1 + y^{n-k} \right) \frac{x Q_k(x, y)}{k!} \right) \\
&= (1-z)^{-1} \sum_{k=0}^{\infty} y^k z^k + x \sum_{k=0}^{\infty} \frac{Q_k(x, y)}{k!} \sum_{n=k+1}^{\infty} z^{n-1} \left(1 + y^{n-k} \right) \\
&= (1-z)^{-1} (1-yz)^{-1} + x \sum_{k=0}^{\infty} \frac{Q_k(x, y)}{k!} z^k \sum_{n=k+1}^{\infty} z^{n-k-1} \left(1 + y^{n-k} \right) \\
&= (1-z)^{-1} (1-yz)^{-1} + x \left((1-z)^{-1} + y(1-yz)^{-1} \right) Q(x, y, z).
\end{aligned}$$

The foregoing partial differential equation can be rewritten in the form

$$\begin{aligned}
&\frac{\partial}{\partial z} \left((1-z)^x (1-yz)^x Q(x, y, z) \right) \\
&= (1-z)^x (1-yz)^x \left(\frac{\partial}{\partial z} Q(x, y, z) \right) - x(1-z)^{x-1} (1-yz)^x Q(x, y, z) \\
&\quad - xy(1-z)^x (1-yz)^{x-1} Q(x, y, z) \\
&= (1-z)^x (1-yz)^x \left((1-z)^{-1} (1-yz)^{-1} + x \left((1-z)^{-1} + y(1-yz)^{-1} \right) Q(x, y, z) \right) \\
&\quad - x(1-z)^{x-1} (1-yz)^x Q(x, y, z) - xy(1-z)^x (1-yz)^{x-1} Q(x, y, z) \\
&= (1-z)^{x-1} (1-yz)^{x-1}.
\end{aligned}$$

Since $Q(x, y, 0) = Q_0(x, y) = 0$, integrating the latter partial differential equation yields

$$(1-z)^x (1-yz)^x Q(x, y, z) = \int_0^z (1-t)^{x-1} (1-yt)^{x-1} dt,$$

which is equivalent to the stated integral representation. $\square$

We are now in a position to state and prove the following representation for the coefficient polynomials in the exponential generating series (3.3).

Theorem 3.3. *Let n be a non-negative integer. Then*

$$\frac{Q_{n+1}(x, y)}{(n+1)!} = \sum_{\ell+m=n} \frac{b_\ell(x, y) b_m(1-x, y)}{m+1}, \quad (3.6)$$

where the sum is over all non-negative integers ℓ and m that satisfy the condition $\ell + m = n$, and $b_\ell(x, y)$ is a polynomial of degree ℓ in both x and y which is expressible in terms of the binomial coefficients as

$$b_\ell(x, y) = (-1)^\ell \sum_{j+k=\ell} \binom{-x}{j} \binom{-x}{k} y^k = \sum_{j+k=\ell} \binom{j+x-1}{j} \binom{k+x-1}{k} y^k. \quad (3.7)$$

Proof. Let $g(z) := (1-z)^{-1}$, and let $G(x, y, z) := (g(z)g(yz))^x$. Theorem 3.2 is equivalent to

$$Q(x, y, z) = G(x, y, z) \int_0^z G(1-x, y, t) dt.$$

By Taylor's theorem and Leibniz's rule for higher derivatives of products,

$$\begin{aligned} Q_{n+1}(x, y) &= \left(\frac{\partial}{\partial z} \right)^{n+1} Q(x, y, z) \Big|_{z=0} \\ &= \sum_{\ell+m=n+1} \frac{(n+1)!}{\ell! m!} \left[\left(\frac{\partial}{\partial z} \right)^\ell G(x, y, z) \right] \left(\frac{\partial}{\partial z} \right)^m \int_0^z G(1-x, y, t) dt \Big|_{z=0} \\ &= \sum_{\ell+m+1=n+1} \frac{(n+1)!}{\ell! (m+1)!} \left[\left(\frac{\partial}{\partial z} \right)^\ell G(x, y, z) \right] \left(\frac{\partial}{\partial z} \right)^{m+1} \int_0^z G(1-x, y, t) dt \Big|_{z=0}, \end{aligned}$$

since the zeroth derivative of the integral vanishes at $z = 0$. By the fundamental theorem of calculus,

$$\frac{Q_{n+1}(x, y)}{(n+1)!} = \sum_{\ell+m=n} \left[\frac{1}{\ell!} \left(\frac{\partial}{\partial z} \right)^\ell G(x, y, z) \right] \frac{1}{(m+1)!} \left(\frac{\partial}{\partial z} \right)^m G(1-x, y, z) \Big|_{z=0}.$$

It follows that if

$$b_\ell(x, y) := \frac{1}{\ell!} \left(\frac{\partial}{\partial z} \right)^\ell G(x, y, z) \Big|_{z=0},$$

then

$$\frac{Q_{n+1}(x, y)}{(n+1)!} = \sum_{\ell+m=n} \frac{b_\ell(x, y) b_m(1-x, y)}{m+1},$$

and by the binomial theorem,

$$b_\ell(x, y) = [z^\ell] G(x, y, z) = [z^\ell] (1-z)^{-x} (1-yz)^{-x} = \sum_{j+k=\ell} (-1)^j \binom{-x}{j} (-y)^k \binom{-x}{k},$$

which is equivalent to the two stated representations for $b_\ell(x, y)$. $\square$

Remark 3.4. We can express the polynomial $b_\ell(x, y)$ in terms of the Gegenbauer polynomial $C_\ell^{(x)}(w)$, where $w := (y^{1/2} + y^{-1/2})/2 = \cosh(\frac{1}{2} \log y)$, and [1, eq. (6.4.10), p. 302]

$$\sum_{\ell=0}^{\infty} C_\ell^{(x)}(w) t^\ell = (1 - 2wt + t^2)^{-x}.$$

Indeed, let $t = y^{1/2}z$. Then

$$\begin{aligned} (1-z)(1-yz) &= 1 - (1+y)z + yz^2 = 1 - \left(\frac{z}{t} + \frac{t}{z} \right) t + t^2 = 1 - (y^{-1/2} + y^{1/2})t + t^2 \\ &= 1 - 2wt + t^2. \end{aligned}$$

It follows that

$$\begin{aligned} b_\ell(x, y) &= [z^\ell] ((1-z)(1-yz))^{-x} = [z^\ell] (1 - 2wt + t^2)^{-x} = [(ty^{-1/2})^\ell] (1 - 2wt + t^2)^{-x} \\ &= y^{\ell/2} [t^\ell] (1 - 2wt + t^2)^{-x} = y^{\ell/2} C_\ell^{(x)}(w). \end{aligned}$$

Gegenbauer polynomials are also referred to as ultraspherical polynomials. They have many interesting properties, including orthogonality with respect to the weight function $(1-w^2)^{x-1/2}$. Nevertheless, we found no need to exploit the connection with orthogonal polynomials to obtain our results.

4. NODE DISTRIBUTION POLYNOMIALS

Node distribution polynomials may be defined for non-negative integer n and non-negative integer k by

$$Q_{n+1,k}(x) := [y^k] \frac{Q_{n+1}(x,y)}{(n+1)!}. \quad (4.1)$$

The polynomial (4.1) encodes the number of times the node k appears at various depths after QuickSorting all $(n+1)!$ permutations of the list $0, 1, \dots, n$. Specifically, for any non-negative integer d , the coefficient of x^d in $Q_{n+1,k}(x)$ multiplied by $(n+1)!$ is the number of times the list entry k appears at depth d . We can express the latter in terms of the unsigned Stirling numbers of the first kind (a.k.a. Stirling cycle numbers) (2.1) (2.2).

Theorem 4.1. *Let n , k , and d be not necessarily distinct non-negative integers. The node distribution polynomial (4.1) has representation*

$$Q_{n+1,k}(x) = \sum_{\ell+m=n} \frac{1}{m+1} \sum_{i+j=k} \binom{\ell-i+x-1}{\ell-i} \binom{i+x-1}{i} \binom{m-j-x}{m-j} \binom{j-x}{j}.$$

The number of times the list entry k appears at depth d after QuickSorting all $(n+1)!$ permutations of the list $0, 1, 2, \dots, n$ is given by $(n+1)! Q_{n+1,k,d}$, where

$$Q_{n+1,k,d} = \sum_{\ell+m=n} \sum_{\substack{i+j=k \\ \alpha+\beta+\gamma+\delta=d}} \frac{1}{(\ell-i)!} \begin{bmatrix} \ell-i \\ \alpha \end{bmatrix} \frac{1}{i!} \begin{bmatrix} i \\ \beta \end{bmatrix} \frac{(-1)^\gamma}{(m-j)!} \begin{bmatrix} m-j+1 \\ \gamma+1 \end{bmatrix} \frac{(-1)^\delta}{j!} \begin{bmatrix} j+1 \\ \delta+1 \end{bmatrix}.$$

Proof. From (3.6) and (4.1), we have

$$\begin{aligned} Q_{n+1,k}(x) &= \sum_{\ell+m=n} \frac{1}{m+1} \left([y^k] b_\ell(x,y) b_m(1-x,y) \right) \\ &= \sum_{\ell+m=n} \frac{1}{m+1} \sum_{i+j=k} \left([y^i] b_\ell(x,y) \right) \left([y^j] b_m(1-x,y) \right). \end{aligned}$$

The stated formula for $Q_{n+1,k}(x)$ is now an immediate consequence of (3.7). The stated formula for $Q_{n+1,k,d}$ is obtained by extracting the coefficient of x^d and employing (2.2). $\square$

5. DEPTH DISTRIBUTION POLYNOMIALS

Depth distribution polynomials may be defined for non-negative integer n and non-negative integer d by

$$\mathcal{D}_{n+1,d}(y) := [x^d] \frac{Q_{n+1}(x,y)}{(n+1)!}. \quad (5.1)$$

They encode the number of times various nodes appear at a given depth after QuickSorting all $(n+1)!$ permutations of the list $0, 1, 2, \dots, n$. Specifically, for any non-negative integer k , the coefficient of y^k in $\mathcal{D}_{n+1,d}(y)$ multiplied by $(n+1)!$ is the number of times the list entry k appears at depth d .

Theorem 5.1. *Let n and d be not necessarily distinct non-negative integers. The depth distribution polynomial (5.1) has representation*

$$\begin{aligned} \mathcal{D}_{n+1,d}(y) &= \sum_{\ell+m=n} \frac{1}{m+1} \sum_{\mu+\nu=d} (-1)^\nu \sum_{i=0}^{\ell} \sum_{\alpha+\beta=\mu} \frac{1}{(\ell-i)!} \begin{bmatrix} \ell-i \\ \alpha \end{bmatrix} \frac{y^i}{i!} \begin{bmatrix} i \\ \beta \end{bmatrix} \\ &\quad \times \sum_{j=0}^m \sum_{\gamma+\delta=\nu} \frac{1}{(m-j)!} \begin{bmatrix} m-j+1 \\ \gamma+1 \end{bmatrix} \frac{y^j}{j!} \begin{bmatrix} j+1 \\ \delta+1 \end{bmatrix}. \end{aligned}$$

Proof. From (3.6), (3.7), and (5.1), we have

$$\begin{aligned} \mathcal{D}_{n+1,d}(y) &= \sum_{\ell+m=n} \frac{1}{m+1} \left([x^d] b_\ell(x, y) b_m(1-x, y) \right) \\ &= \sum_{\ell+m=n} \frac{1}{m+1} \sum_{\mu+\nu=d} \left([x^\mu] b_\ell(x, y) \right) \left([x^\nu] b_m(1-x, y) \right) \\ &= \sum_{\ell+m=n} \frac{1}{m+1} \sum_{\mu+\nu=d} \left([x^\mu] \sum_{i=0}^{\ell} \binom{\ell-i+x-1}{\ell-i} \binom{i+x-1}{i} y^i \right) \\ &\quad \times \left([x^\nu] \sum_{j=0}^m \binom{m-j-x}{m-j} \binom{j-x}{j} y^j \right) \\ &= \sum_{\ell+m=n} \frac{1}{m+1} \sum_{\mu+\nu=d} \sum_{i=0}^{\ell} \sum_{\alpha+\beta=\mu} \left([x^\alpha] \binom{\ell-i+x-1}{\ell-i} \right) \left([x^\beta] \binom{i+x-1}{i} y^i \right) \\ &\quad \times \sum_{j=0}^m \sum_{\gamma+\delta=\nu} \left([x^\gamma] \binom{m-j-x}{m-j} \right) \left([x^\delta] \binom{j-x}{j} y^j \right). \end{aligned}$$

To complete the derivation of the stated formula for $\mathcal{D}_{n+1,d}(y)$, apply (2.2). $\square$

6. DISTRIBUTION OF NODE TOTALS

If we're interested in only the total number of list entries that appear at various depths, but not the number of times that specific entries appear at those depths, we can set $y = 1$ in the depth distribution polynomials. However, what we'd prefer is to have a single polynomial that encodes the node totals for the various depths. We can obtain such a polynomial by setting $y = 1$ in the exponential generating series (3.3) and simplifying the resulting expression for $Q(x, 1, z)$ using the formula (3.4) of Theorem 3.2. The polynomial

$$q_n(x) := Q_n(x, 1) = n! [z^n] Q(x, 1, z) = \sum_{d=0}^{n-1} q_{n,d} x^d \quad (6.1)$$

has the desired properties, namely that the coefficient $q_{n,d}$ of x^d is the total number of nodes that appear at depth d in the trees obtained by QuickSorting all $n!$ permutations of the list $0, 1, 2, \dots, n-1$.

Theorem 6.1. *When $y = 1$, the exponential generating series (3.3) reduces to*

$$Q(x, 1, z) = \frac{1}{2x-1} \left((1-z)^{-2x} - (1-z)^{-1} \right). \quad (6.2)$$

If n is any non-negative integer, the polynomial (6.1) is given by

$$q_n(x) = \frac{n!}{2x-1} \left(\binom{n+2x-1}{n} - 1 \right). \quad (6.3)$$

If d is any non-negative integer, the total number of nodes that appear at depth d in the trees obtained by QuickSorting all $n!$ permutations of the list $0, 1, 2, \dots, n-1$ is given by

$$q_{n,d} = [x^d] q_n(x) = 2^d \left(n! - \sum_{m=0}^d \binom{n}{m} \right) = 2^d \sum_{m=d+1}^n \binom{n}{m}. \quad (6.4)$$

Proof. Setting $y = 1$ in (3.4) yields the representation

$$\begin{aligned} Q(x, 1, z) &= (1-z)^{-2x} \int_0^z (1-t)^{2x-2} dt = (1-z)^{-2x} \left[-\frac{(1-t)^{2x-1}}{2x-1} \right]_{t=0}^{t=z} \\ &= (1-z)^{-2x} \left(\frac{1 - (1-z)^{2x-1}}{2x-1} \right), \end{aligned}$$

which is equivalent to (6.2). Extracting the coefficient of z^n in (6.2) via the binomial theorem yields

$$\frac{q_n(x)}{n!} = \frac{Q_n(x, 1)}{n!} = [z^n] Q(x, 1, z) = \frac{1}{2x-1} \left[(-1)^n \binom{-2x}{n} - 1 \right],$$

which is equivalent to (6.3). From (2.2) and (6.3), we have

$$\begin{aligned} q_{n,d} &= n! [x^d] (1-2x)^{-1} \left[1 - \binom{n+2x-1}{n} \right] \\ &= n! [x^d] (1-2x)^{-1} - \sum_{\ell+m=d} \left([x^\ell] (1-2x)^{-1} \right) \left([x^m] n! \binom{n+2x-1}{n} \right) \\ &= n! 2^d - \sum_{\ell+m=d} 2^\ell \left(2^m \binom{n}{m} \right), \end{aligned}$$

which is equivalent to (6.4). □

Corollary 6.2. For fixed positive integer d , the total number $q_{n,d}$ of nodes that appear at depth d in the trees obtained by QuickSorting all $n!$ permutations of the list $0, 1, 2, \dots, n-1$ satisfies the asymptotic formula

$$q_{n,d} = n! 2^d \left(1 - \frac{1}{n} \cdot \frac{(\log n)^{d-1}}{(d-1)!} + o\left(\frac{(\log n)^{d-1}}{n}\right) \right) \quad (n \rightarrow \infty). \quad (6.5)$$

In particular, we have the anticipated limiting behavior $\lim_{n \rightarrow \infty} q_{n,d}/n! = 2^d$.

Proof. The stated formula (6.5) is readily obtained by substituting Jordan's asymptotic formula [10, Ch. IV, §54, p. 161] (namely

$$\binom{n}{m} \sim \frac{(n-1)!}{(m-1)!} (\log n + \gamma)^{m-1} \quad (n \rightarrow \infty)$$

with misprint corrected) into the sum over $1 \leq m \leq d$ in the exact formula (6.4) and noting that the terms corresponding to $0 \leq m < d$ may be absorbed into the error term. □

6.1. Related Sequences. As the referee observed, certain instances of (6.4) have connections to other combinatorial objects of interest. In particular, dividing each side by 2^d and setting $d = 2$ shows that

$$\frac{q_{n,2}}{4} = n! - \sum_{m=0}^2 \begin{bmatrix} n \\ m \end{bmatrix} = (n-1)!(n-1-H_{n-1}),$$

which is equal to the sum of the reflection lengths of all permutations in $\mathfrak{S}_{n-1}$ [3, p. 9] [15, A067318]. When $d = 3$, the quantity $q_{n,d}/2^d$ reduces to $q_{n,3}/8$, which evaluates partial sums of deco polyominoes of height n [2, pp. 31–32] [15, A122105]. Although we have not found similarly compelling connections between $q_{n,d}/2^d$ and other known sequences when $d > 3$, the fact that many of the formulas in this paper involve Stirling cycle numbers suggests that pursuing connections with related counting formulas that arise in other areas of combinatorics may be a fruitful path for future research.

6.2. Comparisons. We note here that our apparatus can also be used to obtain the total (and hence also the average) number of comparisons performed by QuickSorting all $n!$ permutations of the list $0, 1, \dots, n-1$. Observe that two nodes are compared if and only if one is an ancestor of the other. Since the number of ancestors of a node is equal to its depth, the total number of comparisons is given by the sum

$$\sum_{d=0}^{n-1} dq_{n,d} = \frac{\partial}{\partial x} \left[\sum_{d=0}^{n-1} q_{n,d} x^d \right]_{x=1} = q'_n(1).$$

Equivalently, $q'_n(1)/n!$ is the average number of comparisons taken over all $n!$ permutations. In terms of the harmonic numbers (2.3) this is obtainable from (6.3) as

$$\begin{aligned} \frac{q'_n(1)}{n!} &= \frac{1}{n!} \left[\left(\frac{\partial}{\partial s} \right) 2q_n \left(\frac{1+s}{2} \right) \right]_{s=1} = \left[\left(\frac{\partial}{\partial s} \right) 2s^{-1} \left(\binom{n+s}{n} - 1 \right) \right]_{s=1} \\ &= \left[2s^{-1} \sum_{m=1}^n \frac{1}{m+s} \binom{n+s}{n} - 2s^{-2} \left(\binom{n+s}{n} - 1 \right) \right]_{s=1} \\ &= 2(n+1)(H_{n+1} - 1) - 2(n+1) + 2 = 2(n+1)H_n - 4n. \end{aligned}$$

Alternatively, one can first obtain the generating series for the average number of comparisons from (6.2) by partially differentiating with respect to x , and then setting $x = 1$. Thus,

$$\begin{aligned} \left[\frac{\partial}{\partial x} Q(x, 1, z) \right]_{x=1} &= 2 \left[\frac{\partial}{\partial s} Q \left(\frac{s+1}{2}, 1, z \right) \right]_{s=1} = 2 \left[\frac{\partial}{\partial s} s^{-1} \left((1-z)^{-s-1} - (1-z)^{-1} \right) \right]_{s=1} \\ &= 2 \left[-s^{-2} \left((1-z)^{-s-1} - (1-z)^{-1} \right) - s^{-1} (1-z)^{-s-1} \log(1-z) \right]_{s=1} \\ &= 2 \left(\frac{1}{1-z} - \frac{1}{(1-z)^2} - \frac{\log(1-z)}{(1-z)^2} \right). \end{aligned}$$

Extracting the coefficient of z^n from the foregoing yields

$$\frac{q'_n(1)}{n!} = 2 \left(1 - (n+1) + \sum_{m=1}^n \frac{n-m+1}{m} \right) = 2(n+1)H_n - 4n$$

as above. The formula $2(n+1)H_n - 4n$ for the average number of comparisons is in accordance with [11, Ch. 1, p. 8] and [9, Theorem 2.1.3, p. 15], although it is derived there in a quite different manner. Their analysis, as does ours, assumes that no comparisons are made in choosing the pivots, and to create the left and right sublists from a list of positive integer length n , a total of $n - 1$ comparisons are made with the pivot. A minor modification of this assumption yields a slightly different formula for the average number of comparisons in [6, §2.2, eq. (2.14), p. 29] [18, p. 243] [20, Ch. 3, p. 53] [21, Ch. 9, p. 121] [22, Theorem 1.3, p. 21].

Acknowledgements

It is a pleasure to thank the referee for several helpful comments and insights; in particular, for drawing our attention to connections between the instances $d = 2$ and $d = 3$ of (6.4) and the sequences in [15] referenced in Section 6.1.

REFERENCES

- [1] G. E. Andrews, R. Askey, R. Roy, *Special Functions*, Cambridge University Press, 1999.
- [2] E. Barucci, A. Del Lungo and R. Pinzani, “Deco” polyominoes, permutations and random generation, *Theoret. Comput. Sci.* 159 (1996) 29–42.
- [3] M. Bouvel, F. Ferrari, L. B. E. Tenner, Between weak and Bruhat: The middle order on permutations, *Graphs and Combinatorics*, 41 (2025) 1–25.
- [4] P. Flajolet, R. Sedgewick, *Analytic Combinatorics*, Cambridge University Press, 2009.
- [5] I. P. Goulden, D. M. Jackson, *Combinatorial Enumeration*, Dover, Mineola, NY, 2004.
- [6] R. L. Graham, D. E. Knuth, O. Patashnik, *Concrete Mathematics* (2nd ed.) Addison-Wesley, Boston, MA, 1994.
- [7] C. A. R. Hoare, Algorithm 64: Quicksort, *Comm. ACM.* 4 (1961) 321.
- [8] C. A. R. Hoare, Quicksort, *Computer J.* 5 (1962) 10–15.
- [9] V. Iliopoulos, *The Quicksort algorithm and related topics*, PhD dissertation, University of Essex, Colchester, UK, 2013.
- [10] C. Jordan, *Calculus of Finite Differences* (3rd ed. reprinted), Chelsea, New York, 1979.
- [11] M. Kauers, P. Paule, *The Concrete Tetrahedron*, Springer-Verlag, Vienna, 2011.
- [12] D. E. Knuth, *The Art of Computer Programming*, Vol. 3: Sorting and Searching, Addison-Wesley, Reading, MA, 1973.
- [13] D. E. Knuth, Two notes on notation, *Amer. Math. Monthly*, 99 (1992), 403–422.
- [14] D. E. Knuth, *The Art of Computer Programming*, Vol. 3: Sorting and Searching (2nd ed.) Addison-Wesley, Reading, MA, 1998.
- [15] OEIS Foundation Inc. (2026), *The On-Line Encyclopedia of Integer Sequences*, Published electronically at <https://oeis.org>
- [16] R. Sedgewick, *Quicksort*, PhD dissertation, Stanford University, Stanford, CA, 1975.
- [17] R. Sedgewick, The analysis of Quicksort programs, *Acta Informatica* 7 (1977), 327–355.
- [18] R. Sedgewick, Quicksort with equal keys, *SIAM J. Comput.* 6 (1977), 240–267.
- [19] R. Sedgewick, Implementing Quicksort programs, *Comm. ACM*, 21 (1978) 847–857.
- [20] R. Sedgewick, *Quicksort*, Garland Publishing, Inc., New York and London, 1980.
- [21] R. Sedgewick, *Algorithms* (2nd ed.) Addison-Wesley, Reading, MA, 1988.
- [22] R. Sedgewick, P. Flajolet, *An Introduction to the Analysis of Algorithms* (2nd ed.) Addison-Wesley, Upper Saddle River, NJ, 2013.
- [23] H. S. Wilf, *Generatingfunctionology* (2nd ed.) Academic Press, San Diego, CA, 1994.