AUSTRALASIAN JOURNAL OF COMBINATORICS
Volume **53** (2012), Pages 31–39

Set partitions as geometric words

Toufik Mansour   Mark Shattuck

Department of Mathematics
University of Haifa, 31905 Haifa
Israel
tmansour@univ.haifa.ac.il   maarkons@excite.com

Abstract

Using an analytic method, we derive an alternative formula for the probability that a geometrically distributed word of length n possesses the restricted growth property. Equating our result with a previously known formula yields an algebraic identity involving alternating sums of binomial coefficients via a probabilistic argument. In addition, we consider refinements of our formula obtained by fixing the number of blocks, levels, rises, or descents.

1 Introduction

If $0 \leq p \leq 1$, then a discrete random variable X is said to be *geometric* if $P(X = i) = pq^{i-1}$ for all integers $i \geq 1$, where $q = 1-p$. We will say that a word $w = w_1w_2\cdots$ over the alphabet of positive integers is *geometrically distributed* if the positions of w are independent and identically distributed geometric random variables. The research in geometrically distributed words has been a recent topic of study in enumerative combinatorics; see, for example, [2, 3, 4] and the references therein.

A nonempty word $w = w_1w_2\cdots$ of finite length over the alphabet of positive integers is said to possess the *restricted growth property* if $w_1 = 1$ and $w_{i+1} \leq \max\{w_1, w_2, \ldots, w_i\} + 1$ for all i. Such words are called *restricted growth functions* and correspond to finite set partitions having a prescribed number of blocks when the range is fixed and to all finite set partitions when it is allowed to vary (see, for example, [7] or [8] for details). If $n \geq 1$, then let P_n denote the probability that a geometrically distributed word of length n possesses the restricted growth property, with $P_0 = 1$. In [6], the following exact formula for P_n was derived using the Cauchy integral formula. Here, $(x; q)_i$ is defined as the product $(1-x)(1-xq)\cdots(1-xq^{i-1})$ if $i \geq 1$, with $(x; q)_0 = 1$.

Theorem 1.1 *If $n \geq 1$, then*

$$P_n = p\sum_{i=0}^{n-1}(-1)^i\binom{n-1}{i}q^i(p;q)_i. \tag{1.1}$$

Here, we provide an alternative formula for P_n by considering solutions to the functional differential equation

$$\frac{d}{dx}h(x) = pe^{px}h(qx),$$

where $h(0) = 1$ and $p + q = 1$. We also consider refinements of this result obtained by restricting the number of blocks or requiring a partition to possess a fixed number of levels, rises, or descents.

2 Another approach to finding P_n

Let P_n denote the probability that a geometrically distributed word of length n with parameter p is a restricted growth function (rgf). Then

$$P_{n+1} = \sum_{j=0}^{n} p^{n+1-j}q^j\binom{n}{j}P_j, \qquad n \geq 0, \tag{2.1}$$

with $P_0 = 1$, upon conditioning on the number, $n - j$, of 1's occurring past the first position. To see this, first note that one may select the positions for these 1's in $\binom{n}{n-j} = \binom{n}{j}$ ways, and the probability that all of these positions are indeed 1's is p^{n-j}. Thus, the probability that there are $n+1-j$ 1's, with a 1 in the first position, is $p^{n+1-j}\binom{n}{j}$. The probability that the remaining j letters form an rgf (on the set $\{2, 3, \ldots\}$) is then q^jP_j, by independence.

Define the exponential generating function $G(x) = \sum_{n\geq 0} P_n\frac{x^n}{n!}$. Multiplying recurrence (2.1) by $\frac{x^n}{n!}$, and summing over $n \geq 0$, implies that $G(x)$ is a solution of the differential equation

$$\frac{d}{dx}h(x) = pe^{px}h(qx), \qquad h(0) = 1. \tag{2.2}$$

We now find a second solution to (2.2) as follows. Let $F(x)$ be given as

$$F(x) = p\sum_{j\geq 0}\frac{(p;q)_j}{j!}(-x)^j.$$

It may be verified that $F(x)$ satisfies

$$F'(x) + F(x) = pF(qx),$$

which implies, upon multiplying by $\frac{1}{p}e^x = \frac{1}{p}e^{px}e^{qx}$, that

$$\frac{d}{dx}\left(\frac{1}{p}e^x F(x)\right) = pe^{px}\left(\frac{1}{p}e^{qx}F(qx)\right).$$

The final equality shows that $\frac{1}{p}e^x F(x)$ is also a solution to (2.2), and thus

$$G(x) = \frac{1}{p}e^x F(x), \tag{2.3}$$

by uniqueness of solutions. From (2.3), we obtain the following explicit formula for P_n.

Theorem 2.1 *If $n \geq 0$, then*

$$P_n = \sum_{i=0}^{n}(-1)^i \binom{n}{i}(p;q)_i. \tag{2.4}$$

Remark: To show that the expressions for P_n in (1.1) and (2.4) are indeed equal, let

$$\tilde{G}(x) = p\sum_{n\geq 1}\left(\sum_{i=0}^{n-1}(-1)^i\binom{n-1}{i}q^i(p;q)_i\right)\frac{x^n}{n!}.$$

Then

$$\begin{aligned}\frac{d}{dx}\tilde{G}(x) &= e^x F(qx) = e^x\left(pe^{-qx}G(qx)\right)\\ &= pe^{px}G(qx) = \frac{d}{dx}G(x),\end{aligned}$$

which implies that the two expressions are the same for all $n \geq 1$. It would be interesting to have a probabilistic or a combinatorial proof in the sense of [1] of either (1.1) or (2.4) or of the identity obtained by equating the expressions in (1.1) and (2.4).

3 Other results

We start with some notation. Suppose q is an indeterminate or a complex number (here, we will assume $q = 1 - p$ lies in the interval $[0, 1]$). Let $0_q := 0$, $n_q := 1 + q + \cdots + q^{n-1} = \frac{1-q^n}{1-q}$ if $n \geq 1$, $0_q! := 1$, $n_q! := 1_q 2_q \cdots n_q$ if $n \geq 1$, and $\binom{n}{k}_q := \frac{n_q!}{k_q!(n-k)_q!}$ if $n \geq 0$ and $0 \leq k \leq n$. The q-binomial coefficient $\binom{n}{k}_q$ is zero if k is negative or if $0 \leq n < k$. If $n \geq 1$, then $[n]$ will denote the set $\{1, 2, \ldots, n\}$, with $[0] = \varnothing$. Let $B(n, k)$ denote the set consisting of all restricted growth functions of length n over the alphabet $[k]$ (which correspond to the partitions of $[n]$ having exactly k blocks).

If $k \geq 1$ is fixed, then let $P_{n,k}$ be the probability that a geometrically distributed word of length n is a member of $B(n, k)$. The following theorem gives an exact formula for $P_{n,k}$.

Theorem 3.1 *If $n \geq k \geq 1$, then*

$$P_{n,k} = \frac{p^n}{k_q!} \sum_{j=0}^{k} (-1)^j q^{\binom{j}{2}} ((k-j)_q)^n \binom{k}{j}_q. \tag{3.1}$$

Proof. First note that $P_{n,k}$ satisfies the recurrence

$$P_{n,k} = pq^{k-1} P_{n-1,k-1} + (1-q^k) P_{n-1,k}, \qquad n, k \geq 1, \tag{3.2}$$

with $P_{0,k} = \delta_{0,k}$ for $k \geq 0$, since if no k occurs amongst the first $n-1$ letters, then the last letter must be a k with probability pq^{k-1}, and if a k does occur amongst the first $n-1$ letters, then the last letter must belong to $[k]$, which has probability

$$p + pq + \cdots + pq^{k-1} = \frac{p(1-q^k)}{1-q} = 1 - q^k.$$

If $k \geq 1$, then let $P_k(z) = \sum_{n \geq k} P_{n,k} z^n$. Multiplying both sides of (3.2) by z^n, summing over $n \geq k$, and solving for $P_k(z)$ implies

$$P_k(z) = \frac{pq^{k-1} z}{1 - (1-q^k) z} P_{k-1}(z), \qquad k \geq 1,$$

with $P_0(z) = 1$, which we iterate to obtain

$$P_k(z) = p^k q^{\binom{k}{2}} z^k \prod_{j=1}^{k} \frac{1}{1-(1-q^j)z}. \tag{3.3}$$

Write

$$\prod_{j=1}^{k} \frac{1}{1-(1-q^j)z} = \sum_{j=1}^{k} \frac{a_{k,j}}{1-(1-q^j)z}.$$

By partial fractions, we have

$$a_{k,j} = \frac{(-1)^{k-j} (j_q)^{k-1}}{q^{\binom{j}{2}+j(k-j)} (k-1)_q!} \binom{k-1}{j-1}_q, \qquad 1 \leq j \leq k.$$

Then

$$\begin{aligned}
[z^n](P_k(z)) &= p^k q^{\binom{k}{2}} \sum_{j=1}^{k} a_{k,j}(1-q^j)^{n-k} \\
&= \frac{p^k q^{\binom{k}{2}}}{(k-1)_q!} \sum_{j=1}^{k} \frac{(-1)^{k-j}(j_q)^{k-1}(1-q^j)^{n-k}}{q^{\binom{j}{2}+j(k-j)}} \binom{k-1}{j-1}_q \\
&= \frac{p^k}{(k-1)_q!} \sum_{j=1}^{k} (-1)^{k-j} q^{\binom{k-j}{2}} (j_q)^{k-1}(1-q^j)^{n-k} \binom{k-1}{j-1}_q \\
&= \frac{p^n}{(k-1)_q!} \sum_{j=1}^{k} (-1)^{k-j} q^{\binom{k-j}{2}} (j_q)^{n-1} \binom{k-1}{j-1}_q \\
&= \frac{p^n}{k_q!} \sum_{j=0}^{k} (-1)^j q^{\binom{j}{2}} ((k-j)_q)^n \binom{k}{j}_q,
\end{aligned}$$

as required, upon changing the indices of summation and noting $\binom{k}{2} = \binom{j}{2} + \binom{k-j}{2} + j(k-j)$, $p+q=1$, and $\binom{k}{j}_q = \frac{k_q}{j_q}\binom{k-1}{j-1}_q$. □

Below, we give a table of values for $P_{n,k}$ for all $1 \le k \le n \le 5$.

$n\backslash k$	1	2	3	4	5
1	p				
2	p^2	p^2q			
3	p^3	$p^3q(q+2)$	p^3q^3		
4	p^4	$p^4q(q^2+3q+3)$	$p^4q^3(q^2+2q+3)$	p^4q^6	
5	p^5	$p^5q(q+2)(q^2+2q+2)$	$p^5q^3(q^2+2q+2)(q^2+q+3)$	$p^5q^6(q^3+2q^2+3q+4)$	p^5q^{10}

The following result provides an asymptotic estimate of $P_{n,k}$ for large n and k fixed.

Corollary 3.2 *If $k \ge 1$ is fixed and $0 < p < 1$, then we have*

$$P_{n,k} \approx \frac{p^k}{(q;q)_k}(1-q^k)^n$$

for n large.

Proof. From (3.3) above, we see that $P_k(z) = \sum_{n \ge k} P_{n,k} z^n$ is a rational function whose smallest positive simple pole is $z_0 = \frac{1}{1-q^k}$. By Theorem IV.9 on p. 256 of [5], we then have $P_{n,k} \approx A_k(1-q^k)^n$, where

$$A_k = \lim_{z \to z_0} \left(\frac{z-z_0}{z_0} P_k(z) \right),$$

which yields the estimate above. □

One may refine Theorem 3.1 as follows. A word $v = v_1v_2\cdots$ is said to have a *level* (at index i) if $v_i = v_{i+1}$. Let $P_{n,k,r}$ denote the probability that a geometrically distributed word of length n belongs to $B(n,k)$ and has r levels. If $k \geq 1$ is fixed, then let $P_k(z,w)$ be given by

$$P_k(z,w) = \sum_{n,r\geq 0} P_{n,k,r} z^n w^r.$$

The generating function $P_k(z,w)$ has the following explicit form.

Theorem 3.3 *If $k \geq 1$, then*

$$P_k(z,w) = \prod_{j=1}^{k} \frac{\frac{pq^{j-1}z}{1-pq^{j-1}z(w-1)}}{1-\sum_{i=1}^{j}\frac{pq^{i-1}z}{1-pq^{i-1}z(w-1)}}. \tag{3.4}$$

Proof. To find $P_k(z,w)$, we first consider a refinement of it as follows. Given $1 \leq i \leq k$, let $P_k(z,w|i)$ be given by

$$P_k(z,w|i) = \sum_{n,r\geq 0} P_{n,k,r}(i) z^n w^r,$$

where $P_{n,k,r}(i)$ denotes the probability that a geometrically distributed word of length n belongs to $B(n,k)$, *ends* in the letter i, and has exactly r levels. Considering whether or not the penultimate letter of a member of $B(n,k)$ ending in i is also an i yields the relations

$$P_k(z,w|i) = pq^{i-1}wzP_k(z,w|i) + pq^{i-1}z(P_k(z,w) - P_k(z,w|i)), \qquad 1 \leq i \leq k-1, \tag{3.5}$$

with

$$P_k(z,w|k) = pq^{k-1}wzP_k(z,w|k) + pq^{k-1}z(P_k(z,w) - P_k(z,w|k)) + pq^{k-1}zP_{k-1}(z,w). \tag{3.6}$$

Solving (3.5) for $P_k(x,w|i)$ and (3.6) for $P_k(z,w|k)$, adding the k equations that result, and noting $\sum_{i=1}^{k} P_k(z,w|i) = P_k(z,w)$ implies

$$P_k(z,w) = \sum_{i=1}^{k} \frac{pq^{i-1}z}{1-pq^{i-1}z(w-1)} P_k(z,w) + \frac{pq^{k-1}z}{1-pq^{k-1}z(w-1)} P_{k-1}(z,w),$$

or

$$P_k(z,w) = \frac{\frac{pq^{k-1}z}{1-pq^{k-1}z(w-1)}}{1-\sum_{i=1}^{k}\frac{pq^{i-1}z}{1-pq^{i-1}z(w-1)}} P_{k-1}(z,w), \qquad k \geq 1,$$

with $P_0(z,w) = 1$, which yields (3.4). □

Letting $w = 1$ in (3.4) gives (3.3). Letting $w = 0$ in (3.4) gives the generating function for the probability that a geometrically distributed word of length n is a member of $B(n,k)$ having no levels, i.e., is a Carlitz set partition with k blocks.

A word $v = v_1v_2\cdots$ is said to have a *rise* (at index i) if $v_i < v_{i+1}$. Let $Q_{n,k,r}$ denote the probability that a geometrically distributed word of length n belongs to $B(n,k)$ and has r rises, and let $Q_k(z,w) = \sum_{n,r\geq 0} Q_{n,k,r}z^nw^r$, where k is fixed. We have the following explicit formula for $Q_k(z,w)$.

Theorem 3.4 *If $k \geq 1$, then*

$$Q_k(z,w) = \frac{p^kq^{\binom{k}{2}}w^{k-1}z^k}{\prod_{\ell=0}^{k}(1-(1-w)pq^\ell z)^{k-\ell} \times \prod_{\ell=0}^{k-1}\left(1-\sum_{j=0}^{\ell}\frac{pq^jwz}{\prod_{i=0}^{j}(1-(1-w)pq^iz)}\right)}. \quad (3.7)$$

Proof. Let $W_{n,k,r}$ denote the probability that a geometrically distributed word of length n has letters only in $[k]$ and has exactly r rises, and let $W_k(z,w) = \sum_{n,r\geq 0} W_{n,k,r}z^nw^r$. Since $\pi \in B(n,k)$ may be decomposed uniquely as $\pi = \pi' k\beta$ for some $\pi' \in B(j,k-1)$, where $k-1 \leq j \leq n-1$, and k-ary word β, we get, by independence,

$$Q_k(z,w) = pq^{k-1}wzQ_{k-1}(z,w)W_k(z,w), \qquad k \geq 2,$$

with $P_1(z,w) = pzW_1(z,w)$, which implies

$$Q_k(z,w) = p^kq^{\binom{k}{2}}w^{k-1}z^k\prod_{j=1}^{k}W_j(z,w). \quad (3.8)$$

To find $W_k(z,w)$, note that it satisfies the recurrence

$$W_k(z,w) = W_{k-1}(z,w) + pq^{k-1}zW_k(z,w) + pq^{k-1}wz(W_{k-1}(z,w)-1)W_k(z,w),\ k \geq 1, \quad (3.9)$$

with $W_0(z,w) = 1$, upon considering whether or not the letter k occurs in a word and, if it does, considering whether or not the first letter of the word is k. To solve recurrence (3.9), we rewrite it as

$$\frac{1}{W_k(z,w)} = \frac{1-(1-w)pq^{k-1}z}{W_{k-1}(z,w)} - pq^{k-1}wz, \qquad k \geq 1,$$

and solve the resulting linear first-order recurrence in $\frac{1}{W_k(z,w)}$ to get

$$\frac{1}{W_k(z,w)} = \left[1 - \sum_{j=0}^{k-1}\frac{pq^jwz}{\prod_{i=0}^{j}(1-(1-w)pq^iz)}\right]\prod_{j=0}^{k-1}(1-(1-w)pq^jz),$$

from which (3.7) follows from (3.8). □

Dividing the right side of (3.7) by w^{k-1}, and letting $w = 0$, gives the generating function for the probability that a geometrically distributed word of length n is a member of $B(n,k)$ having no rises other than those occurring each time that there is an appearance of a new letter.

Finally, a word $v = v_1v_2\cdots$ is said to have a *descent* (at index i) if $v_i > v_{i+1}$. Let $R_{n,k,r}$ denote the probability that a geometrically distributed word of length n belongs to $B(n,k)$ and has r descents, and let $R_k(z,w) = \sum_{n,r\geq 0} R_{n,k,r}z^nw^r$, where k is fixed. Our final result is an explicit formula for $R_k(z,w)$.

Theorem 3.5 *If $k \geq 2$, then*

$$R_k(z,w) = \frac{p^kq^{\binom{k}{2}}z^k}{1-pz}\prod_{j=2}^{k}\frac{wW_j(z,w)+1-w}{1+(w-1)pq^{j-1}z}, \tag{3.10}$$

with $R_1(z,w) = \frac{pz}{1-pz}$, where $W_j(z,w)$ is given by

$$\frac{1}{W_j(z,w)} = \left[1 - \sum_{i=0}^{j-1}\frac{pq^iwz}{\prod_{\ell=0}^{i}(1-(1-w)pq^\ell z)}\right]\prod_{i=0}^{j-1}(1-(1-w)pq^iz).$$

Proof. Given α, a geometrically distributed word of length n, let $W_{n,k,r}$ (respectively, $W'_{n,k,r}$) denote the probability that α (respectively, $k\alpha$) has letters only in $[k]$ and has exactly r descents. Let $W_k(z,w) = \sum_{n,r\geq 0} W_{n,k,r}z^nw^r$ and $W'_k(z,w) = \sum_{n,r\geq 0} W'_{n,k,r}z^nw^r$. Note that $W_k(z,w)$ is as in the proof of Theorem 3.4 above, by symmetry, upon writing words in reverse order. If $k \geq 2$ and $\pi \in B(n,k)$, then $\pi = \pi'k\beta$ for some $\pi' \in B(j,k-1)$, where $k-1 \leq j \leq n-1$, and k-ary word β, which implies $R_k(z,w) = pq^{k-1}zR_{k-1}(z,w)W'_k(z,w)$, by independence, with $R_1(z,w) = \frac{pz}{1-pz}$. Thus, we have

$$R_k(z,w) = \frac{p^kq^{\binom{k}{2}}z^k}{1-pz}\prod_{j=2}^{k}W'_j(z,w), \qquad k \geq 2. \tag{3.11}$$

We now find an expression for $W'_k(z,w)$. Let $W_k(z,w|k)$ be the generating function for the probability that a geometrically distributed word of length n is a k-ary word that starts with k and has exactly r descents. From the definitions, we may write

$$W'_k(z,w) = w(W_k(z,w)-1-W_k(z,w|k)) + W_k(z,w|k) + 1 \tag{3.12}$$

and

$$W_k(z,w|k) = pq^{k-1}z[w(W_k(z,w)-1-W_k(z,w|k)) + W_k(z,w|k)+1]. \tag{3.13}$$

Relations (3.12) and (3.13) together imply

$$W'_k(z,w) = \frac{wW_k(z,w)+1-w}{1+(w-1)pq^{k-1}z},$$

from which (3.10) follows from (3.11). □

References

[1] A. T. Benjamin and J. J. Quinn, *Proofs that Really Count: The Art of Combinatorial Proof*, Mathematical Association of America, Washington, DC 2003.

[2] C. Brennan, Ascents of size less than d in samples of geometric variables: mean, variance and distribution, *Ars Combin.* **102** (2011), 129–138.

[3] C. Brennan and A. Knopfmacher, The first and last ascents of size d or more in samples of geometric random variables, *Quaest. Math.* **28** (2005), 487–500.

[4] C. Brennan and A. Knopfmacher, The distribution of ascents of size d or more in samples of geometric random variables, *Discrete Math. Theor. Comput. Sci.* **11:1** (2009), 1–10.

[5] P. Flajolet and R. Sedgewick, *Analytic Combinatorics*, Cambridge University Press, 2009.

[6] K. Oliver and H. Prodinger, Words coding set partitions, *Appl. Anal. Discrete Math.* **5** (2011), 55–59.

[7] D. Stanton and D. White, *Constructive Combinatorics*, Springer, New York 1986.

[8] C. Wagner, Generalized Stirling and Lah numbers, *Discrete Math.* **160** (1996), 199–218.

(Received 24 Feb 2011; revised 15 Jan 2012)