

Ballot Paths Avoiding Depth Zero Patterns

Heinrich Niederhausen and Shaun Sullivan
 Florida Atlantic University, Boca Raton, Florida
 niederha@fau.edu, ssulli21@fau.edu

1 Introduction

In a paper by Sapounakis, Tasoulas, and Tsikouras [8], the authors count the number of occurrences of patterns of length four in Dyck paths. In this paper we specify in one direction and generalize in another. We only count ballot paths that avoid a given pattern, where a ballot path stays weakly above the diagonal $y = x$, starts at the origin, and takes steps from the set $\{\uparrow, \rightarrow\} = \{u, r\}$. A pattern is a finite string made from the same step set; it is also a path. Notice that a ballot path ending at a point along the diagonal is a Dyck path. Consider the following enumeration of ballot paths avoiding the pattern rur .

m	1	8	28	62	105	148	178	178	127	0
7	1	7	21	40	59	72	72	51	0	
6	1	6	15	24	30	30	21	0		
5	1	5	10	13	13	9	0			
4	1	4	6	6	4	0				
3	1	3	3	2	0					
2	1	2	1	0						
1	1	1	0							
0	1	0								
	0	1	2	3	4	5	6	7	8	n

The number of ballot paths to (n, m) avoiding rur

$s_n(m)$, the entry at the point (n, m) , is the number of ballot paths from the origin to that point avoiding the pattern rur . We can generate this

table by using the recurrence formula

$$s_n(m) = s_n(m - 1) + s_{n-1}(m) - s_{n-1}(m - 1) + s_{n-1}(m - 2) \quad (1)$$

where $s_0(m) = 1$, and $s_n(n - 1) = 0$ represents the weak boundary $y = x$. The above table is calculated using this recursion and suggests the columns are the values of a polynomial sequence. It will be shown that $s_n(m)$ is a polynomial in m . We shall present a theorem guaranteeing the existence of a polynomial sequence for a given recurrence relation and boundary conditions.

If the number of ballot paths reaching (n, x) is the values of a polynomial sequence $s_n(x)$, the recurrence relation obtained can be transformed into an operator equation. Using the tools from Finite Operator Calculus in Section 2, we can find an explicit formula for the polynomial sequence $s_n(x)$.

In this paper, we will only consider patterns p such that its reverse pattern \bar{p} is a ballot path. For example, the reverse pattern of $p = uururrr$ is $\bar{p} = uuururr$. We call such patterns *depth-zero* patterns. This is a requirement in Theorem 4, although for some patterns that are not depth-zero, the solution may be a polynomial sequence, but not for the entire enumeration. For patterns of nonzero depth see [5]. To develop the recursions, we need to investigate the properties of the pattern we wish to avoid.

Definition 1 *A nonempty string $o \in \{u, r\}^*$ is a bifix of a pattern p if it can be written in the form $p = op'$ and $p = p''o$ for nonempty strings $p', p'' \in \{u, r\}^*$. If a pattern has no bifizes, then we call it bifix-free.*

Definition 2 *If a is the number of r 's in p and c is the number of u 's, then we say p has dimensions $a \times c$, and $d(p) = a - c$.*

If o is a bifix for the pattern p , then we will see in Section 4 that we need that $d(p'') \geq 0$. The following lemma shows that this restriction is not necessary when p is depth zero.

Lemma 3 *If p is a depth zero pattern and o is a bifix of p , then $d(p'') \geq 0$.*

Proof. First we note the additive property of d , that is $d(ab) = d(a) + d(b)$. By definition $p = op' = p''o$, which implies $d(p') = d(p'')$. Since p is depth zero, any suffix of p is also depth zero, in particular p' is depth zero. Therefore, $d(p') \geq 0$, which implies $d(p'') \geq 0$. ■

2 Main Tools

In this section we will present the main tools from Finite Operator Calculus that will be used to solve these enumeration problems. As we noted before, not every pattern we choose to avoid will have a polynomial sequence solution. The following theorem will illustrate why we want a depth-zero pattern.

Theorem 4 *Let x_0, x_1, \dots be a given sequence of integers where $F_n(m) = 0$ for $m < x_n$. For $m > x_n$ define $F_n(m)$ recursively for all $n \geq 0$ by*

$$F_n(m) = F_n(m-1) + \sum_{i=1}^n \sum_j a_{i,j} F_{n-i}(m - b_{i,j}),$$

where $\sum_j a_{1,j} \neq 0$ and with initial values $F_n(x_n)$. If $F_0(m) = c \neq 0$ for all $m \geq x_0$, and if for each j holds $b_{i,j} \leq x_n - x_{n-i} + 1$ for $n \geq 1$ and $i = 1, \dots, n$, then there exists a polynomial sequence $f_n(x)$ where $\deg f_n = n$ such that $F_n(m) = f_n(m)$ for all $m \geq x_n$ and $n \geq 0$ (within the recursive domain).

Before we prove this theorem, we look back at the example of ballot paths avoiding the pattern *rur*.

m	1	8	28	91	105	148	178	178	127	0
7	1	7	21	62	59	72	72	51	0	
6	1	6	15	24	30	30	21	0		
5	1	5	10	13	13	9	0			
4	1	4	6	6	4	0				
3	1	3	3	2	0					
2	1	2	1	0						
1	1	1	0							
0	1	0	0							
-1	1	-1	-1							
	0	1	2	3	4	5	6	7	8	n
	$F_n(m) = F_n(m-1) + F_{n-1}(m) - F_{n-1}(m-1) + F_{n-1}(m-2)$									

Here, $x_n = n - 1$ and $F_n(x_n) = \delta_{n,0}$. The recursion satisfies the conditions $b_{i,j} \leq x_n - x_{n-i} + 1 = i + 1$ and $\sum_j a_{1,j} = 1 \neq 0$. In the table, below the initial values are the polynomial extensions, which are not all 0. Clearly, if the recursion required us to use values below these initial

values, the values generated would not agree with the values generated by the polynomial extensions, thus we need the condition $b_{i,j} \leq x_n - x_{n-i} + 1$, in general. We now prove Theorem 4.

Proof. We first let $f_0(x) = F_0(x_0)$, since $F_0(m) = F_0(m-1)$. Now suppose $F_n(m)$ has been extended to a polynomial of degree n . The recursion tells us that $F_{n+1}(m) - F_{n+1}(m-1)$ can be extended to a polynomial of degree n since $\sum_j a_{1,j} \neq 0$. But if the backwards difference of a function is a polynomial, then the function itself is a polynomial of one degree higher (see [3]). ■

The objects of Finite Operator Calculus are polynomial sequences called Sheffer sequences. A Sheffer sequence is defined by its generating function, which can be written in the following way.

$$p(x, t) = \sum_{n \geq 0} s_n(x) t^n = \sigma(t) e^{x\beta(t)}$$

where $\sigma(t)$ is a power series with a multiplicative inverse and $\beta(t)$ is a delta series, that is, a power series with a compositional inverse. We say (s_n) is a Sheffer sequence for the basic sequence (b_n) where

$$b(x, t) = \sum_{n \geq 0} b_n(x) t^n = e^{x\beta(t)}.$$

Notice that $b_n(0) = \delta_{n,0}$ and $b_0(x) = 1$.

In order to find the solution to the recurrence, we first transform the recursion into an equation with finite operators. For every Sheffer sequence, the linear operator $B = \beta^{-1}(D)$, where $D = \frac{d}{dx}$, maps s_n to s_{n-1} . The operator B also maps its basic sequence b_n to b_{n-1} (Rota, Kahaner, and Odlyzko [7]). Also if B is a *delta* operator, that is

$$B = \sum_{n \geq 1} a_n D^n$$

where $a_1 \neq 0$, then it corresponds to a class of Sheffer sequences associated to a unique basic sequence. We do not have to worry about convergence since these operators are only applied to polynomials, thus only a finite number of the terms will be used for a given polynomial, and thus the name Finite Operator Calculus.

An example of a delta operator that will be frequently used in this paper is $\nabla = 1 - E^{-1}$, where $E^a s_n(x) = s_n(x+a)$ is a shift operator. By Taylor's Theorem, we can write

$$f(x+a) = \sum_{n \geq 0} f^{(n)}(x) \frac{a^n}{n!} \quad \text{or} \quad E^a = \sum_{n \geq 0} \frac{D^n a^n}{n!} = e^{aD}$$

in operators. Clearly ∇ is a delta operator.

If we consider $s_n(x)$ for $x \geq n-1$ in (1), we obtain the operator equation

$$\nabla = B - BE^{-1} + BE^{-2}.$$

We have written one delta operator in terms of an unknown delta operator. The following theorem [6, Theorem 1] lets us find the basic sequence for the unknown operator B in terms of the basic sequence of the known operator.

Theorem 5 (Transfer Formula) *Let A be a delta operator with basic sequence $a_n(x)$. Suppose*

$$A = \tau(B) = \sum_{j \geq 1} T_j B^j,$$

where the T_j are linear operators that commute with shift operators, and T_1 is invertible. Then B is a delta operator with basic sequence

$$b_n(x) = x \sum_{i=1}^n [\tau^i]_n \frac{1}{x} a_i(x).$$

where $[\tau^i]_n$ is the coefficient of t^n in $\tau^i(t)$.

Notice that this Theorem implies that the unknown operator B is a delta operator, and so the underlying polynomial sequence is a Sheffer sequence. The final step is to transform this basic sequence into the correct Sheffer sequence using the initial values. For ballot paths we have that $s_n(n-1) = \delta_{n,0}$, thus the following lemma gives us the solution based on these initial values.

Lemma 6 *If (b_n) is a basic sequence for B , then*

$$t_n(x) = (x - an - c) \frac{b_n(x - c)}{x - c}$$

is the Sheffer sequence for B with initial values $t_n(an + c) = \delta_{n,0}$.

In our case, we would obtain the solution from the basic sequence b_n given in the transfer formula as

$$s_n(x) = (x - n + 1) \frac{b_n(x + 1)}{x + 1} \tag{2}$$

called Abelization [6, (2.5)]. We want to remark that given the generating function $e^{x\beta(t)} = \sum_{n \geq 0} b_n(x) t^n$, we find

$$\sum_{n \geq 0} s_n(x) t^n = e^{(x+1)\beta(t)} - \frac{t}{x+1} \frac{d}{dt} e^{(x+1)\beta(t)} = (1 - t\beta'(t)) e^{(x+1)\beta(t)}.$$

All the examples connected to pattern avoidance are of the form

$$1 - E^{-1} = \nabla = \sum_{k \geq 1} a_k E^{c_k} B^k$$

where $a_k, c_k \in \mathbb{Z}$. If this equation can be solved for E ,

$$E = 1 + \tau(B),$$

where τ is a delta series, then

$$\sum_{n \geq 0} b_n(x) t^n = (1 + \tau(t))^x$$

We first present the results in [5] which covered bifix-free patterns and patterns with exactly one bifix. We then present results for patterns with any number of bifixes, hence for any depth-zero pattern. We also give some examples and special cases.

3 Bifix-free patterns

Let $s_n(x; p)$ be the number of ballot paths avoiding the pattern p ; we occasionally will drop the p in the notation when convenient. If the pattern is bifix-free and depth-zero, we need only to subtract paths that would end in the pattern, thus we have the recurrence

$$s_n(m; p) = s_{n-1}(m; p) + s_n(m - 1; p) - s_{n-a}(m - c; p) \tag{3}$$

where p has dimensions $a \times c$. For example $uurrurrur$ has dimensions 5×4 , and depth zero. If the pattern has depth zero, and $a \geq c \geq 1$, $a \geq 2$ (if $a = 1$ then $p = ur$, a pattern we do not want to avoid), then it is easy

to check that the conditions of the Theorem 4 are satisfied; the solution is polynomial. In operator notation,

$$\nabla = B(1 - B^{a-1}E^{-c}).$$

Since the delta operator ∇ can be written as a delta series in B , the operator B is also a delta operator. The basic sequence can be expressed via the Transfer Formula

$$b_n(x) = x \sum_{i=0}^{\frac{n}{a-1}} \frac{(-1)^i}{x-ci} \binom{n-(a-1)i}{i} \binom{x+n-(a+c-1)i-1}{n-(a-1)i}.$$

Using (2) we obtain

$$s_n(x) = (x-n+1) \sum_{i=0}^{\frac{n}{a-1}} \frac{(-1)^i}{x-ci+1} \binom{n-(a-1)i}{i} \binom{x+n-(a+c-1)i}{n-(a-1)i}.$$

Therefore the number of ballot paths avoiding p and returning to the diagonal (Dyck paths) equals

$$s_n(n) = \sum_{i=0}^{\frac{n}{a-1}} \frac{(-1)^i}{n-ci+1} \binom{n-(a-1)i}{i} \binom{2n-(a+c-1)i}{n-(a-1)i}$$

4 Patterns with exactly one bifix

If the pattern p has exactly one bifix, then there exists a unique nonempty pattern o such that $p = op'o$. If p has depth 0 and dimensions $a \times c$, and p'' has dimensions $b \times d$, we have a recurrence of the form

$$s_n(x) = s_{n-1}(x) + s_n(x-1) - \sum_{i \geq 0} (-1)^i s_{n-a-bi}(x-c-di). \quad (4)$$

For example let $p = urruurr$. This pattern is depth-zero, with dimensions 4×3 and bifix $o = urr$, so $b = 2$ and $d = 2$. From the paths reaching $(n-a, x-c)$, those ending in $p'' = urru$ cannot be included in the recurrence and must be subtracted, and from those again we cannot include paths ending in $urru$, and so on. The p'' piece of the pattern that is responsible for this exclusion-inclusion process may not go below the diagonal; hence it must have $d(p'') \geq 0$ which lemma 3 guarantees. If $d(p'') < 0$, then at some point we would be using numbers below the $y = x$ boundary, which are only the polynomial extensions and do not count paths.

In operators, we have

$$1 = B + E^{-1} - \sum_{i \geq 0} (-1)^i B^{a+bi} E^{-c-di} = B + E^{-1} - \frac{B^a E^{-c}}{1 + B^b E^{-d}}$$

or

$$\begin{aligned} \nabla &= B + B^{b+1} E^{-d} + B^b E^{-d-1} - B^a E^{-c} - B^b E^{-d} \\ &= B - B^a E^{-c} + B^{b+1} E^{-d} - B^b E^{-d} \nabla. \end{aligned} \quad (5)$$

Using the Transfer Formula, we obtain $b_n(x) =$

$$\begin{aligned} x \sum_{j,k,l \geq 0} \binom{n - (a-1)j - bk - (b-1)l}{j, k, l} \frac{(-1)^{j+l}}{n - (a-1)j - bk - (b-1)l} \\ \times \binom{n - (a+c-1)j - (b+d)(k+l) - 1 + x}{n - (a-1)j - b(k+l) - 1}. \end{aligned}$$

Because $s_n(n-1) = \delta_{0,n}$ we get $s_n(x) =$

$$\begin{aligned} \sum_{j,k,l \geq 0} \binom{n - (a-1)j - bk - (b-1)l}{j, k, l} \frac{(-1)^{j+l}(x-n+1)}{n - (a-1)j - bk - (b-1)l} \\ \times \binom{n - (a+c-1)j - (b+d)(k+l) + x}{n - (a-1)j - b(k+l) - 1}. \end{aligned}$$

So for the pattern $urruurr$, we obtain $s_n(x) =$

$$\sum_{j,k,l \geq 0} \binom{n - 3j - 2k - l}{j, k, l} \frac{(-1)^{j+l}(x-n+1)}{n - 3j - 2k - l} \binom{n - 6j - 4(k+l) + x}{n - 3j - 2(k+l) - 1}.$$

m	1	8	35	110	270	544	920	1272	1236	0
7	1	7	27	75	161	279	389	377	0	
6	1	6	20	48	87	122	118	0		
5	1	5	14	28	40	38	0			
4	1	4	9	14	13	0				
3	1	3	5	5	0					
2	1	2	2	0						
1	1	1	0							
0	1	0								
	0	1	2	3	4	5	6	7	8	n

The number of ballot paths avoiding $urruurr$

5 A Special Case

The above operator equation (5) simplifies when $a = b + 1$ and $c = d$. This corresponds to a pattern of the form $rp'r$. For this case we get

$$\nabla = B(1 + B^b E^{-d})^{-1},$$

$$b_n(x) = x \sum_{i \geq 0} \frac{(-1)^i}{x - di} \binom{n - (b-1)i - 1}{i} \binom{x + n - (d+b)i - 1}{n - bi},$$

and

$$s_n(x) = (x - n + 1) \sum_{i \geq 0} \frac{(-1)^i}{x - di + 1} \binom{n - (b-1)i - 1}{i} \binom{x + n - (d+b)i}{n - bi}.$$

The number of Dyck paths equals

$$s_n(n) = \sum_{i \geq 0} \frac{(-1)^i}{n - di + 1} \binom{n - (b-1)i - 1}{i} \binom{2n - (d+b)i}{n - bi}.$$

The first pattern we considered, rur , is of this form where $b = d = 1$. The solution in this case is

$$s_n(x) = (x - n + 1) \sum_{i \geq 0} \frac{(-1)^i}{x - i + 1} \binom{n-1}{i} \binom{x + n - 2i}{n - i}$$

and the number of Dyck paths equals

$$s_n(n) = \sum_{i \geq 0} \frac{(-1)^i}{n - i + 1} \binom{n-1}{i} \binom{2n - 2i}{n - i} = \sum_{i \geq 0} (-1)^i C_{n-i} \binom{n-1}{i}.$$

where C_n is the n th Catalan number.

6 Patterns with any number of bifixes

Now we suppose that the pattern p has m distinct bifixes o_1, \dots, o_m . Let $b_i \times d_i$ be the dimensions of p''_i , the pattern p without the suffix o_i , and $a \times c$ be the dimensions of p . The number of ways to reach the lattice point $(n - a - \sum i_j b_j, x - c - \sum i_j d_j)$ with $k \geq 0$ combinations of the p''_j with

$\sum_{j=0}^m i_j = k$ is $\binom{k}{i_1, \dots, i_m}$ with sign $(-1)^k$. Then the recurrence for the ballot paths avoiding p is

$$s_n(x) = s_{n-1}(x) + s_n(x-1) - \sum_{k \geq 0} (-1)^k \times \sum_{i_1 + \dots + i_m = k} \binom{k}{i_1, \dots, i_m} s_{n-a-\sum i_j b_j}(x-c - \sum i_j d_j). \quad (6)$$

In operators,

$$\begin{aligned} \nabla &= B - \sum_{k \geq 0} (-1)^k \sum_{i_1 + \dots + i_m = k} \binom{k}{i_1, \dots, i_m} B^{a+\sum i_j b_j} E^{-c-\sum i_j d_j} \\ &= B - B^a E^{-c} \sum_{k \geq 0} (-1)^k \left(\sum_{i=0}^m B^{b_i} E^{-d_i} \right)^k \\ &= B - \frac{B^a E^{-c}}{1 + \sum_{i=0}^m B^{b_i} E^{-d_i}} \end{aligned} \quad (7)$$

In general, we can find the basic sequence for any depth zero pattern via the transfer formula, although it would involve many summations. We now consider an example and some special cases.

Example 7 *The pattern rururrur has two bifixes, rur and r. In this case, $a = 5$, $c = 3$, $b_1 = 3$, $d_1 = 2$, $b_2 = 4$ and $d_2 = 3$. Using the above operator equation, we obtain*

$$\begin{aligned} \nabla &= B - \frac{B^5 E^{-3}}{1 + B^3 E^{-2} + B^4 E^{-3}} \\ &= \frac{B + B^4 E^{-3}}{1 + B^3 E^{-2} + B^4 E^{-3}} \end{aligned}$$

Using the transfer formula and (2),

$$\begin{aligned} b_n(x) &= \sum_{i=0}^n x \sum_{j \geq 0} \sum_{k \geq 0} \binom{i}{j} \binom{-i}{k} \binom{k}{n-i-3j-3k} \frac{1}{x-n+i+j+k} \\ &\quad \times \binom{x-n+2i+j+k-1}{i} \end{aligned}$$

and

$$s_n(x) = (x - n + 1) \sum_{i=0}^n \sum_{j \geq 0} \sum_{k \geq 0} \binom{i}{j} \binom{-i}{k} \binom{k}{n - i - 3j - 3k} \\ \times \frac{1}{x - n + i + j + k + 1} \binom{x - n + 2i + j + k}{i}.$$

6.1 Patterns of the form r^a

We had previously determined in [4] by other methods that the recurrence formula for this pattern is

$$s_n(x) = s_{n-1}(x) + s_n(x-1) - s_{n-a}(x-1)$$

giving the operator equation

$$\nabla = B - B^a E^{-1} \quad \text{or} \quad E = \sum_{i=0}^{a-1} B^i.$$

Using the new formula (7), we view r^a as a pattern with $a - 1$ bifixes. Here $c = d_i = 0$ and $b_i = i$ for all $0 \leq i \leq a - 1$, thus the operator equation becomes

$$\nabla = B - \frac{B^a}{1 + \sum_{i=1}^{a-1} B^i} = 1 - \frac{1}{\sum_{i=0}^{a-1} B^i} \quad \text{or} \quad E = \sum_{i=0}^{a-1} B^i$$

as expected.

6.2 Patterns of the form $r(ur)^k$

If the pattern is of the form $r(ur)^k$, then $a = k + 1$, $c = k$, and for each $0 \leq i < k$ there is the bifix $r(ur)^i$. So, $b_j = d_j = j$ for $1 \leq j \leq k$. Let $S = BE^{-1}$. We obtain the following operator equation for this pattern,

$$\nabla = B - \frac{BS^k}{1 + S + S^2 + \dots + S^k} \\ \nabla = B \left(1 - \frac{S^k}{1 + S + S^2 + \dots + S^k} \right).$$

Using the Transfer Formula,

$$b_n(x) = \sum_{i=0}^n x \sum_{j \geq 0} (-1)^j \binom{i}{j} \binom{-j}{n - i - kj}_{k+1} \frac{1}{x - n + i} \binom{x - n + 2i - 1}{i}$$

where $\binom{x}{n}_{k+1} := [t^n](1+t+\dots+t^k)^x$ [4]. The number of Ballot paths is $s_n(x) =$

$$(x-n+1) \sum_{i=0}^n \sum_{j \geq 0} (-1)^j \binom{i}{j} \binom{-j}{n-i-kj}_{k+1} \frac{1}{x-n+i+1} \binom{x-n+2i}{i}.$$

We obtain a very nice formula for the Dyck paths in this case,

$$s_n(n) = \sum_{i=0}^n C_i \sum_{j \geq 0} (-1)^j \binom{i}{j} \binom{-j}{n-i-kj}_{k+1}$$

where C_n is the n th Catalan number.

References

- [1] Deutsch, E., 1999, Dyck path enumeration. *Discrete Math.* **204**, 167 – 202.
- [2] Euler, L., 1801, De evolutione potestatis polynomialis cuiuscunque $(1+x+x^2+x^3+x^4+\dots)^n$. *Nova Acta Academiae Scientiarum Imperialis Petropolitinae* **12**, 47 – 57.
- [3] Jordan, C., 1939, *Calculus of Finite Differences*. Chelsea Publ. Co., New York, 3rd edition 1979.
- [4] Niederhausen, H., Sullivan, S., 2007, Euler Coefficients and Restricted Dyck Paths. *Congr Numer.* **188**, 196 – 210 (arXiv:0705.3065).
- [5] Niederhausen, H. and Sullivan, S., 2007, Pattern Avoiding Ballot Paths and Finite Operator Calculus. To appear in *J. Statistical Inference and Computing* (arXiv:0709.0878).
- [6] Niederhausen, H., 2003, Rota’s umbral calculus and recursions. *Algebra Univers.* **49**, 435 – 457.
- [7] Rota, G.-C., Kahaner, D., Odlyzko, A., 1973, On the Foundations of Combinatorial Theory VIII: Finite operator calculus. *J. Math. Anal. Appl.* **42**, 684 – 760.
- [8] Sapounakis, A., Tasoulas, I., Tsikouras, P., 2007, Counting strings in Dyck paths. *Discrete Math.* **307**, 2909 – 2924.
- [9] Guibas, L. J., Odlyzko, A. M., 1981, String overlaps, pattern matching, and nontransitive games. *J. Comb. Theory* **30**, 183 – 208.