Faculty of Sciences and Mathematics, University of Niš, Serbia
Available at: http://www.pmf.ni.ac.rs/filomat

Filomat **25:3** (2011), 191–201 DOI: 10.2298/FIL1103191I

ON THE NUMBER OF RESTRICTED DYCK PATHS

Aleksandar Ilić and Andreja Ilić

Abstract

In this note we examine the number of integer lattice paths consisting of up-steps $(1,1)$ and down-steps $(1,-1)$ that do not touch the lines $y = m$ and $y = -k$, and in particular Theorem 3.2 in [P. Mladenović, *Combinatorics*, Mathematical Society of Serbia, Belgrade, 2001]. The theorem is shown to be incorrect for $n \geqslant m + k + \min(m, k)$, and using similar combinatorial technique we proved the upper and lower bound for the number of such restricted Dyck paths. In conclusion, we present some relations between the Chebyshev polynomials of the second kind and generating function for the number of restricted Dyck paths, and connections with the spectral moments of graphs and the Estrada index.

1 Introduction

A Dyck path is a lattice path in the plane integer lattice $\mathbb{Z}^2$ consisting of up-steps $(1,1)$ and down-steps $(1,-1)$, which never passes below the x-axis (but may touch). The number of Dyck paths from $(0,0)$ to $(2n,0)$ is given by the Catalan number

$$C_n = \frac{1}{n+1}\binom{2n}{n}.$$

The Catalan numbers arise in many combinatorial problems - see Stanley [21] for several combinatorial interpretations of these numbers, and see [20] for matrix approach. These numbers also satisfy the recurrence relation $C_0 = 1$ and

$$C_{n+1} = \sum_{k=0}^{n} C_k \cdot C_{n-k}.$$

2010 *Mathematics Subject Classifications.* 05A50, 05C50.
Key words and Phrases. Dyck paths; walks in graphs; dynamic programming; Catalan number; spectral moments.
Received: 12.10.2010.
Communicated by Dragan Stevanović
This work was supported by Research Grants 174010 and 174033 of Serbian Ministry of Education and Science. The authors are grateful to Dragan Stevanović and Slobodan Simić for several useful remarks on the earlier draft of the paper.

Let $\pi = \pi_1\pi_2 \dots \pi_n$ be a permutation of numbers $\{1, 2, \dots, n\}$ and $\sigma = \sigma_1\sigma_2 \dots \sigma_k$ be a permutation of $\{1, 2, \dots, k\}$, $k \leqslant n$. We say that the permutation π contains the pattern σ, if there are indices $1 \leqslant i_1 < i_2 < \cdots < i_k \leqslant n$ such that $\pi_{i_1}\pi_{i_2} \dots \pi_{i_k}$ is in the same relative order as $\sigma_1\sigma_2 \dots \sigma_k$. Otherwise, π avoids the pattern σ and we say that π is σ-avoiding. For example, the permutation 598376412 avoids 123 and contains 1432 exactly twice. There is a bijection between 132-avoiding permutations of $\{1, 2, \dots, n\}$ and Dyck paths from $(0, 0)$ to $(2n, 0)$ (see [11, 12, 15] for various results concerning permutations with restricted patterns and Dyck paths).

There are many modifications and generalizations of Dyck paths. In production quality control, we are often faced with problems to count the number of variations of elements 0 and 1 of length n in which before every zero, there are at least k times more ones than zeros. In this paper, we will consider *restricted Dyck paths* that do not touch the lines $y = m$ and $y = -k$ for some $m, k > 0$. Mladenović in [13] counted the number of such paths and proved the following

Theorem 1. *Let $m, n, k \in \mathbb{N}$. The number of integer lattice paths from $(0, 0)$ to $(2n, 0)$ that do not touch lines $y = m$ and $y = -k$ is equal to*

$$\binom{2n}{n} - \binom{2n}{n+m} - \binom{2n}{n+k} + 2\binom{2n}{n+m+k}.$$

We will show in Section 2 using combinatorial methods that this theorem is not true. Furthermore, the restriction of Theorem 1 when $k = 0$ is itself difficult problem and the number of such paths does not have closed formula. We implemented a dynamic programming algorithm for counting such paths and compare correct and proposed values for the number of restricted Dyck paths that do not touch the lines $y = m$ and $y = -k$ for different values of n. In Section 3 we show that the proposed number of paths is actually an upper bound, while in Section 4 we prove a similar lower bound with equality if and only if $n < 2(m + k)$. In Section 4 we present some relations between the Chebyshev polynomials of the second kind and generating function of the number of restricted Dyck paths, and connections with spectral moments of graphs and the Estrada index.

2 Dynamic programming approach

There is an obvious bijection between the restricted Dyck paths from $(0, 0)$ to $(2n, 0)$ that do not touch the lines $y = m$ and $y = -k$, and the closed walks of length $2n$ starting from the fixed vertex v_k in the path $P = v_1v_2v_3 \dots v_{k+m-1}$.

Denote by $d[i][j]$ the number of paths of length j starting from the vertex v_k and ending in v_i. The initial values are $d[k][0] = 1$ and $d[i][0] = 0$ for $i \neq k$. The following recurrent formula for calculating the matrix d holds

$$d[i][j] = d[i-1][j-1] + d[i+1][j-1]$$

and taking zeros $d[i][j] = 0$ for $i \leq 0$ or $i > k + m - 1$. Time complexity of this algorithm is $O(n(m + k))$. Since we have to know only the values of elements in the $(i-1)$-th row for calculating the element (i, j), we can reduce the memory complexity to $O(m+k)$. We also implemented operations involving big integers, since the number of walks can be out of limits for all standard integer types in computer architecture.

Finally, the number of restricted Dyck paths of length $2n$ is equal to $d[k][2n]$. The computational results for the case $k = 2$ and $m = 5$ are presented in Table 1 and Figure 1.

Path length	**Correct number**	**Proposed number**	**Difference**
2	2	2	0
4	5	5	0
6	14	14	0
8	42	42	0
10	131	131	0
12	417	417	0
14	1341	1341	0
16	4334	4334	0
18	14041	14042	1
20	45542	45562	20
22	147798	148029	231
24	479779	481804	2025
26	1557649	1572625	14976
28	5057369	5156025	98656
30	16420730	17018505	597775
32	53317085	56717910	3400825
34	173118414	191540940	18422526
36	562110290	658084182	95973892
38	1825158051	2309516616	484358565
40	5926246929	8307367724	2381120795
42	19242396629	30693300198	11450903569
44	62479659622	116529049786	54049390164
46	202870165265	453938745602	251068580337
48	658715265222	1808941473996	1150226208774
50	2138834994142	7345198595742	5206363601600

Table 1: Comparison of proposed and correct number of paths for $k = 2$ and $m = 5$.

3 An upper bound on restricted Dyck paths

We will show that the original proof of Theorem 1 from [13] is incorrect.

Let S be the set of integer lattice paths in $\mathbb{Z}^2$ consisting of up-steps $(1, 1)$ and down-steps $(1, -1)$, from $(0, 0)$ to $(2n, 0)$. It is well known that the number of such lattice paths equals

$$|S| = \binom{2n}{n}.$$

Let S_1 be the set of integer lattice paths from S that touch the line $y = m$, and S_2 be the set of integer lattice paths from S that touch the line $y = -k$.

The number of integer lattice paths that touch the line $y = -k$ can be calculated as follows: consider the first time when a lattice path t from the origin $(0, 0)$ to $(2n, 0)$ touches the line $y = -k$. Denote this point as P. The part of path t from the origin to the point P is reflected with respect to the line $y = -k$. This way we established a bijection from the set of all paths that cross the line $y = -k$ and to the set of all lattice paths from the new origin $(0, -2k)$ to $(2n, 0)$, or equivalently

$$|S_1| = \binom{2n}{n+k}.$$

Analogously, we get

$$|S_2| = \binom{2n}{n+m}.$$

The bad lattice paths are those that belong to both S_1 and S_2. For every such lattice path t, denote by $x_1(t)$ the lattice point that belongs to the line $y = m$ with the minimal x coordinate – the first time when the path t touches the line $y = m$. In the same way, denote by $x_2(t)$ the lattice point that belongs to the line $y = -k$ with the minimal x coordinate – the first time when the path t touches the line $y = -k$. Let T_1 be the set of bad lattice paths for which holds $x_1(t) < x_2(t)$ and T_2 the set of bad paths with $x_1(t) > x_2(t)$.

We will use a slightly modified counting argument than the one described in [13]. Namely, for a path $t \in T_1$, let t_1 be the first part with the origin $(0, 0)$ and the end $(x_1(t), m)$, t_2 the second part with the origin $(x_1(t), m)$ and the end $(x_2(t), -k)$, and t_3 the third part with the origin $(x_2(t), -k)$ and the end $(2n, 0)$. We will construct the corresponding path t' as follows. Let t'_3 be the reflected third part with respect to the line $y = -k$, and $t'_2 \cup t''_3$ be the reflected parts $t_2 \cup t'_3$ with the respect to the line $y = m$. Finally, we get $t' = t_1 \cup t'_2 \cup t''_3$ and the origin of the corresponding path t' is $(0, 0)$ and the end $(2n, 2m + 2k)$.

This is not a bijection, as the author in [13] pointed out. Namely, for

$$2n \geqslant 2(m + k) + 2k,$$

there are paths from $(0, 0)$ to $(2n, 2m+2k)$ that after reconstruction do not touch the line $y = m$ before touching $y = -k$. One counterexample is presented in Figure 2.

Therefore, we have

$$|T_1| \leqslant \binom{2n}{n+m+k},$$

with equality if and only if $n < m + 2k$. Analogously, we can establish an injection between paths from T_2 and paths from $(0, 0)$ to $(2n, -2k - 2m)$.

Using the inclusion–exclusion argument, it follows

$$\begin{aligned}|S \setminus (S_1 \cup S_2)| &= |S| - |S_1| - |S_2| + |S_1 \cup S_2| \leqslant |S| - |S_1| - |S_2| + |T_1| + |T_2| \\ &= \binom{2n}{n} - \binom{2n}{n+k} - \binom{2n}{n+m} + 2\binom{2n}{n+k+m}.\end{aligned}$$

Therefore, we proved the following upper bound

Theorem 2. *Let $m, n, k \in \mathbb{N}$. The number of integer lattice paths from $(0,0)$ to $(2n,0)$ that do not touch the lines $y = m$ and $y = -k$ is less than or equal to*

$$\binom{2n}{n} - \binom{2n}{n+m} - \binom{2n}{n+k} + 2\binom{2n}{n+m+k},$$

with equality if and only if $n < m + k + \min(m,k)$.

4 A lower bound on restricted Dyck paths

By using a similar method as in Section 3, we can derive a lower bound for the number of restricted Dyck paths. According to the inclusion–exclusion formula, we have to subtract the number of integer lattice paths from $(0,0)$ to $(2n, 2m+2k)$ that touch the line $y = -k$ before the line $y = m$. After reflection with respect to the line $y = -k$, this is equivalent as counting the number of lattice paths from $(0,0)$ to $(2n, 2m+4k)$, which is equal to $\binom{2n}{n+m+2k}$. Similarly, we have $\binom{2n}{n+2m+k}$ lattice paths that first touch the line $y = m$ before the line $y = -k$.

This way we subtracted all paths that after intersection $(x_k, -k)$ with $y = -k$ have at least one intersection with both lines. The bad paths in this case are those that touch the line $y = m$ at x coordinate that is less than x_k, or in other words $x_m < x_k$. The first time this can happen is exactly when a path touches the lines $y = m$ and $y = -k$ twice, or equivalently when $n = 2(m+k)$.

Theorem 3. *Let $m, n, k \in \mathbb{N}$. The number of integer lattice paths from $(0,0)$ to $(2n,0)$ that do not touch the lines $y = m$ and $y = -k$ is greater than or equal to*

$$\binom{2n}{n} - \binom{2n}{n+m} - \binom{2n}{n+k} + 2\binom{2n}{n+m+k} - \binom{2n}{n+m+2k} - \binom{2n}{n+2m+k},$$

with equality if and only if $n < 2(m+k)$.

5 Spectral moments and generating functions

Let G be a simple connected graph on n vertices. A walk of length k in G is any sequence of vertices and edges of G,

$$w = w_0, e_1, w_1, e_2, \ldots, w_{k-1}, e_k, w_k,$$

such that e_i is the edge joining w_{i-1} and w_i for every $i = 1, 2, \ldots, k$. The walk is closed if $w_0 = w_k$. The spectrum of G consists of the eigenvalues $\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_n$ of its adjacency matrix $A(G)$. Let M_k denotes the k-th spectral moment of the graph G,

$$M_k(G) = \sum_{i=1}^{n} \lambda_i^k.$$

For every $k \geqslant 0$, the number $M_k(G)$ represents the number of closed walks of length k (see [1]). It is well-known that for every graph holds $M_1 = 0$ and $M_2 = 2m$. For the third moment we have $M_3 = 6t$, where t is the number of triangles in the graph G.

Let $M_k(n,i)$ be the number of closed walks of length k starting from the vertex v_i in the path $P_n = v_0 v_1 v_2 \dots v_n$. By simple parity argument, we have identity $M_{2k+1}(n,i) = 0$ for $0 \leqslant i \leqslant n$; and by symmetry, we have $M_k(n,i) = M_k(n, n-i)$. Recall that the sequence of numbers $c_1, c_2, \dots, c_k$ is *unimodal* if there exists no indices $1 \leqslant p < q < r \leqslant k$ such that $c_p > c_q < c_r$. The authors in [10] proved the following

Theorem 4. *For every even $k > 2$ holds*

$$M_k(n,0) \leqslant M_k(n,1) \leqslant \cdots \leqslant M_k(n, \left\lceil \frac{n}{2} \right\rceil - 1) \leqslant M_k(n, \left\lceil \frac{n}{2} \right\rceil).$$

For sufficiently large k, strict inequalities hold.

These are main tools for analyzing the novel graph invariant, called *Estrada index*, defined as

$$EE(G) = \sum_{i=1}^{n} e^{\lambda_i}.$$

In the last ten years, the Estrada index found applications in measuring the degree of protein folding [3], the centrality of complex networks (such as neural, social, metabolic, protein–protein interaction networks, and the World Wide Web) [6], and it was also proposed as a measure of molecular branching, accounting the effects of all atoms in the molecule, giving higher weight to the nearest neighbors [4]. Deng [2] and Ilić and Stevanović [10] proved the following conjecture from [16]:

Theorem 5. *For any tree T on n vertices it holds*

$$EE(S_n) \geq EE(T) \geq EE(P_n),$$

with the left equality if and only if $T \cong S_n$ and with the right equality if and only if $T \cong P_n$, where S_n and P_n denote the star and the path on n vertices, respectively.

From the Taylor expansion of e^x, the Estrada index and the spectral moments of G are related by

$$EE(G) = \sum_{k=0}^{\infty} \frac{M_k}{k!}.$$

The Chebyshev polynomials of first and second kinds $T_k(x)$ and $U_k(x)$ are defined implicitly as $\cos k\theta = T_k(cos\theta)$ and $\sin(k+1)\theta / \sin\theta = U_k(cos\theta)$. These polynomials occur in diverse areas (see Rivlin [17]), but their application in combinatorics and lattice path counting is less well known. For this purpose, it is convenient to define modified Chebyshev polynomials by

$$T_k^*(x) = 2x^k \; T_k\left(\frac{1}{2x}\right) \qquad\qquad U_k^*(x) = x^k \; U_k\left(\frac{1}{2x}\right).$$

In [9] Gutman and Graovac estimated the Estrada index of a path P_n and proved that

$$EE(P_n) \approx (n+1)I_0 - \cosh 2,$$

where $I_0 = \sum_{k=0}^{\infty} \frac{1}{(k!)^2} = 2.27958530$. In [8] the authors point out certain classes of graphs whose characteristic polynomials are closely connected to the Chebyshev polynomials of the second kind. The precision of these approximations is remarkably good.

For the restricted Dyck paths with $k = 0$, we get the sequence A080934 from the On-Line Encyclopedia of Integer Sequences [19]. This simpler sequence can be represented as m-th coefficient in expansion of the rational function $R(n)$, where $R(1) = 1$ and

$$R(n+1) = \frac{1}{1 - x \cdot R(n)}.$$

The generating function for the number of restricted Dyck paths from $(0, 0)$ to $(2n, 0)$ that do not touch the line $y = m$ and do not pass below the line $y = 0$ is given by [7, 14, 18]

$$R_m(x) = \frac{U^*_m(x)}{U^*_{m+1}(x)}.$$

This can be proven by a recurrent formula for the Chebyshev polynomials of the second kind $U_{k+1}(x) = U_k(x) - x^2 U_{k-1}$. Using the first return decompositions of the paths from Section 3, we obtain that the generating function for the number of Dyck paths consisting of up-steps and down-steps that do not touch the lines $y = m$ and $y = -k$ is expressed in terms of Chebyshev polynomials by

$$\frac{1}{1 - xR_{m-1}(x) - xR_{k-1}(x)}. \tag{1}$$

This generating function can be used for establishing further properties of restricted Dyck paths. However, we focused our research on combinatorial approach and it is hard to derive upper and lower bounds from Theorem 2 and Theorem 3 using analytic methods.

References

[1] D. Cvetković, M. Doob, H. Sachs, *Spectra of Graphs – Theory and Application*, Academic Press, New York, 1980.

[2] H. Deng, *A Proof of a Conjecture on the Estrada Index*, MATCH Commun. Math. Comput. Chem. **62** (2009) 599–606.

[3] E. Estrada, *Characterization of the folding degree of proteins*, Bioinformatics **18** (2002) 697–704.

[4] E. Estrada, J. A. Rodríguez-Valázquez, M. Randić, *Atomic branching in molecules*, Int. J. Quantum Chem. **106** (2006) 823–832.

[5] E. Estrada, N. Hatano, *Returnability in complex directed networks (digraphs)*, Linear Algebra Appl. **430** (2009) 1886–1896.

[6] E. Estrada, *Topological structural classes of complex networks*, Phys. Rev. E **75** (2007) 016103-1-12.

[7] E. Georgiadis, D. Callan, Q. H. Hou, *Circular Digraph Walks, k-Balanced Strings, Lattice Paths and Chebyshev Polynomials*, Electronic J. Comb. **15** (2008) #R108.

[8] Y. Ginosar, I. Gutman, T. Mansour, M. Schork, *Estrada index and Chebyshev polynomials*, Chem. Phys. Lett. **454** (2008) 145–147.

[9] I. Gutman, A. Graovac, *Estrada index of cycles and paths*, Chem. Phys. Lett. **436** (2007) 294–296.

[10] A. Ilić, D. Stevanović, *The Estrada index of chemical trees*, J. Math. Chem. **47** (2010) 305–314.

[11] C. Krattenthaler, *Permutations with restricted patterns and Dyck paths*, Adv. Appl. Math. **27** (2001) 510–530.

[12] T. Mansour, E. Deng, R. Du, *Dyck paths and restricted permutations*, Discrete Appl. Math. **154** (2006) 1593–1605.

[13] P. Mladenović, *Combinatorics* (in Serbian), Mathematical Society of Serbia, Third edition, Belgrade, 2001.

[14] H. Niederhausen, S. Sullivan, *Euler Coefficients and Restricted Dyck Paths*, Congr Numer. **188** (2007) 196–210.

[15] P. Peart, W. J. Woan, *Dyck Paths with no peaks at height k*, J. Integer Sequences **4** (2001) 01.1.3

[16] J. A. de la Peña, I. Gutman, J. Rada, *Estimating the Estrada index*, Linear Algebra Appl. **427** (2007) 70–76.

[17] T. J. Rivlin, *Chebyshev polynomials, From Approximation Theory to Algebra and Number Theory*, Second edition, Wiley, New York, 1990.

[18] M. Sato, T. Sado, *Lattice Paths Restricted by Two Parallel Hyperplanes*, Bull. Inf. Cybernetics **21** (1985) 97–105.

[19] N. J. A. Sloane, The On-Line Encyclopedia of Integer Sequences, `http://www.research.att.com/~njas/sequences/A080934`

[20] S. Stanimirović, P. Stanimirović, M. Miladinović, A. Ilić, *Catalan matrix and related combinatorial identities*, Appl. Math. Comput. **215** (2009) 796–805.

[21] R. Stanley, *Enumerative Combinatorics, Volume 2*, Cambridge University Press, 1999.

Aleksandar Ilić:
University of Niš, Faculty of Sciences and Mathematics,
Višegradska 33, 18000 Niš, Serbia
E-mail: `aleksandari@gmail.com`

Andreja Ilić:
University of Niš, Faculty of Sciences and Mathematics,
Višegradska 33, 18000 Niš, Serbia
E-mail: `andrejko.ilic@gmail.com`

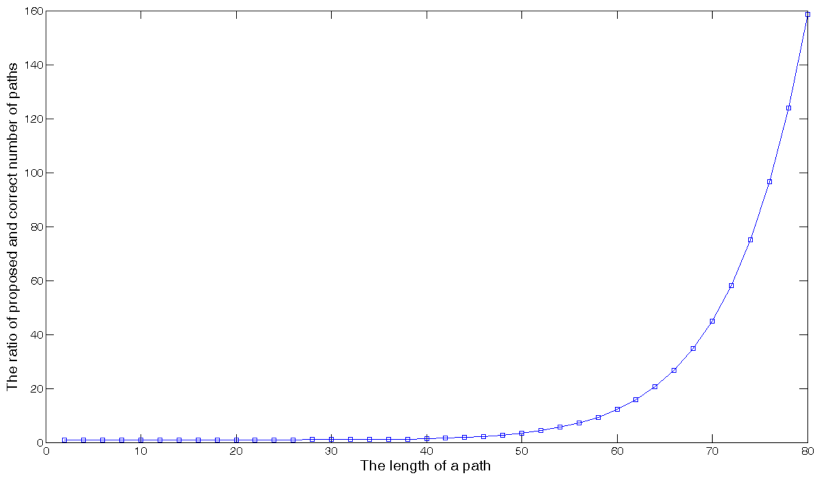

Figure 1: The ratio of proposed and correct number of paths for $k = 2$ and $m = 5$.

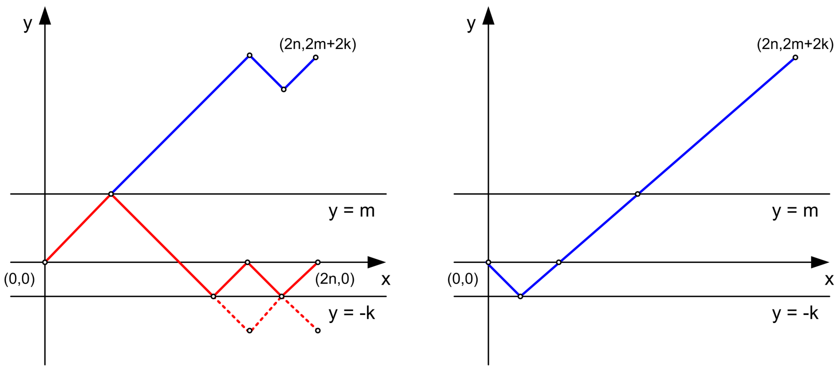

Figure 2: A counterexample to Theorem 1.