Texas College Mathematics Journal
Volume 1, Number 1, 2005, Pages 5–14 (January 14, 2005)
S applied for(XX)0000-0

A SHORT (LATTICE) PATH FROM BEATING YOUR KIDS TO PERMUTATION STATISTICS

TOMISLAV DOŠLIĆ

ABSTRACT. Motivated by a simple card guessing game, we consider the problem of enumerating grand Dyck lattice paths with respect to the number of turns. An explicit solution to this problem is given and subsequently applied to enumeration of permutations of $\{1, \ldots, n\}$ according to the number of parity changes. Some properties of the enumerating sequences, such as their symmetry, unimodality, and the position of the maximal value are briefly investigated, and some further possibilities of research have been suggested.

1. RED OR BLACK

A couple of years ago, while I was still a young and relatively inexperienced father, I came across an article with a very intriguing title: "How often should you beat your kids?" [1]. As it happened, I had already formed a rather strong and clear opinion on this question, but I decided to read the article nevertheless. Although it started with a shocking conclusion that one should beat one's kids every day except Sunday, it soon became clear that the article does not advocate regular physical abuse of children. Instead, it was concerned with establishing the probability of winning in a simple card guessing game; the beating from the title was only metaphorical. I have not regretted reading the article. Its delayed time effects are still felt (and appreciated) whenever my daughters are occupied by playing the card guessing game described there. Besides, the game provided me with the motivation for the present note.

The game considered here is so simple that it was undoubtedly reinvented many times. It was introduced into the mathematical literature in the paper "How to beat your kids in their own game" ([2]), that prompted the article mentioned in the introduction. It is quite possible, and even likely, that some of its mathematical aspects were studied even earlier, but I am not aware of any such references.

The game consists in guessing the color (red or black) of the top card in a deck of playing cards. After each guess the top card is removed, and the game ends when the deck is empty.

There are various strategies for playing the game, but they are not the concern in this note. Their performances were discussed and compared in the already mentioned references, and the analysis was further refined in [3]. Instead, this note is

Received by the editors June 9, 2004 and, in revised form, December 15, 2004.
1991 *Mathematics Subject Classification.* Primary 54C40; Secondary 46E25.
Key words and phrases. lattice path, grand Dyck path, permutation statistics.
The author was supported in part by the Welch Foundation of Houston, Texas, via Grant # BD-0894.

concerned not with the game, but with the deck. More precisely, we will try to find the probability that in a well-shuffled deck there are exactly k color changes. Equivalently, we ask how many ways there are to have a deck with k color changes. In fact, we consider the following more general problem.

Problem: What is the number of decks with m red and m black cards which have exactly k color changes?

Let us denote the number of such decks by $p(2m, k)$. If we disregard for the moment the differences in suits and values between the cards of the same color, we get an auxilliary problem that can be formulated (and subsequently solved) in the terms of lattice paths.

2. Lattice paths

There are several advantages in casting our problem in terms of lattice paths. First, the lattice paths provide a convenient way of visualizing the problem. Second, the lattice paths are a well-researched area, with lot of connections to other areas of combinatorics, statistics and probability ([4], [5], [6]). Finally, the lattice paths were already employed in the references [2] and [1] to analyze our game.

A **lattice path** of length n between the points P_0 and P_n in the (x, y) coordinate plane is any sequence P of n segments $\left(\overline{P_{j-1}P_j}\right)_{j=1}^{n}$ both of whose endpoints have integer coordinates. The segment $\overline{P_{j-1}P_j}$ is called the j-th **step** of the path P. By imposing various restrictions on the size and orientation of steps, on the initial and final points, and on the areas of the lattice that must be visited or avoided by the path, we obtain different classes of lattice paths. All lattice paths in this paper will be self-avoiding, i.e. no point in the plane will be visited more than once. Further, they will be directed, in the sense that for any step $\overline{P_{j-1}P_j}$ the coordinates of P_j are not smaller than the corresponding coordinates of P_{j-1}. Finally, the lattice paths considered here will use only two types of steps: the R steps, whose endpoints' abscissas differ by one, while the ordinates are the same, and U steps, whose endpoints have the same abscissas, and their ordinates differ by one. A sequence of consecutive steps of the same type is called a **block**. A place where a step of one type is immediately followed by a step of another type is called a **turn**; if an R step is followed by a U step, the turn is of RU **type**; otherwise, it is of **type** UR.

As noted in [2], each game of Red or Black corresponds to a lattice path in the (x, y)-coordinate plane that connects the points $(0, 0)$ and (m, m) using only the steps R and U. For each red card removed from the top of the deck we make a step to the right (an R step); for each black card, a step up (a U step) is made. The number of color changes in a deck is the same as the number of turns in the corresponding lattice path. Hence, we are left with the task of enumerating lattice paths from $(0, 0)$ to (m, m), that use only the steps R and U, with respect to the number of turns. A lattice path of this type with 4 turns is shown in Fig. 1.

The directed self-avoiding lattice paths from $(0, 0)$ to (m, m) that use only the steps R and U are known in combinatorial literature as **grand Dyck paths**. There are $\binom{2m}{m}$ such paths, since each of them is uniquely determined by deciding which m of its $2m$ steps are of the type R. If an additional constraint is imposed, namely that such a path is not allowed to wander above the line $y = x$, we get the class of **Dyck paths**. The Dyck paths are among the best researched combinatorial families; we refer the reader to [7] for a recent survey. In particular, it is well known that they

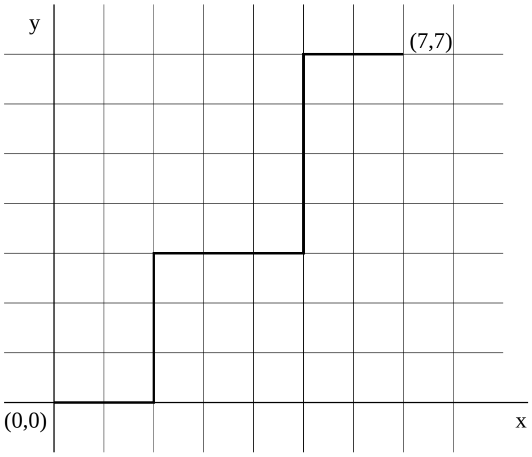

FIGURE 1. A grand Dyck path of 14 steps with 4 turns.

are enumerated by the Catalan numbers, $C_m = \frac{1}{m+1}\binom{2m}{m}$, and with respect to the number of RU turns, they are enumerated by the Narayana numbers. For given positive integers m and k, the Narayana number $N(m,k) = \frac{1}{m}\binom{m}{k}\binom{m}{k-1}$ is the number of Dyck paths on $2m$ steps with exactly k turns of the RU type. Hence, $\sum_k N(m,k) = C_m$. As a Dyck path with k RU turns has exactly $k-1$ UR turns, the Narayana numbers also enumerate Dyck paths with respect to the total number of turns.

For the grand Dyck paths, the situation is a bit different, since a grand Dyck path with k RU turns can have $k-1$, k, or $k+1$ UR turns. That was, probably, the reason that a search of the literature failed to produce an explicit formula for the number of grand Dyck paths on $2m$ steps with exactly k turns. We denote the number of such paths by $q(2m,k)$. Fortunately, the formula for $q(2m,k)$ can be easily derived as follows.

Let us first consider the case of an even k, $k = 2l$. Any grand Dyck path with such a number of turns must start and end with a step of the same type. Let it be a step of the type R. Now, the m steps of the type R must be partitioned in $l+1$ blocks of consecutive steps, and the m steps of the type U must be partitioned in l blocks. The first partition is completely determined by the positions of l breaks in the sequence of m R steps, and these positions can be selected in $\binom{m-1}{l}$ ways. As the same reasoning can be applied to the steps of the type U, it follows that there are $\binom{m-1}{l}\binom{m-1}{l-1}$ grand Dyck paths that start and end by an R step with exactly $2l$ turns. By switching the roles of R and U steps, one obtains the following formula for the number of grand Dyck paths on $2m$ steps with $2l$ turns:

$$q(2m, 2l) = 2\binom{m-1}{l}\binom{m-1}{l-1}.$$

When the number of turns is odd, $k = 2l+1$, the same reasoning gives the following expression:

$$q(2m, 2l+1) = 2\binom{m-1}{l}^2.$$

The above two results can be expressed in a single formula, using the floor and ceiling functions $\lfloor x \rfloor$ and $\lceil x \rceil$. (For a given real number x, $\lfloor x \rfloor$ is the greatest integer that does not exceed x. Similarly, the ceiling function $\lceil x \rceil$ is defined as the smallest integer that is not smaller than x. The reader unfamiliar with the $\lfloor x \rfloor$ and $\lceil x \rceil$ functions should consult the book [8] for an excellent introduction.)

$$q(2m,k) = 2\binom{m-1}{\lceil \frac{k-1}{2} \rceil}\binom{m-1}{\lfloor \frac{k-1}{2} \rfloor}. \tag{2.1}$$

Namely, for an even $k = 2l$, we have $l = \lceil \frac{k-1}{2} \rceil$, $l-1 = \lfloor \frac{k-1}{2} \rfloor$, while for an odd $k = 2l+1$, we have $l = \lceil \frac{k-1}{2} \rceil = \lfloor \frac{k-1}{2} \rfloor$.

When dealing with an enumeration problem, it is often very useful to check the arising enumerating sequences against those listed in the *On-Line Encyclopedia of Integer Sequences* ([9]). This reference contains a wealth of information on all kind of integer sequences and on relations and connections between them. Curiously enough, the sequence $q(2m,k)$ does not appear among more than 100000 sequences collected there. Apparently, either nobody ever bothered to enumerate grand Dyck paths with respect to the number of turns, or, if the enumeration was done, the sequence was not submitted to the *Encyclopedia*. Hence, we propose to call the numbers $q(2m,k)$ **grand Narayana numbers**, since they decompose the central binomial coefficients in a similar way as the ordinary Narayana numbers decompose Catalan numbers.

Before we turn back to our deck of cards, let us make a digression and count all paths of the grand Dyck type from $(0,0)$ to $(m+1,m)$ with respect to the number of turns. The total number of such paths is $\binom{2m+1}{m}$. There are $m+1$ R steps and m U steps. Let us first consider the case of an even $k = 2l$. Again, a path with an even number of turns must start and end with a step of the same type, say R. The $m+1$ R steps must be divided in $l+1$ blocks of consecutive steps, and the m U steps must be partitioned in l such blocks. This can be effected in $\binom{m}{l}\binom{m-1}{l-1}$ different ways. By switching the roles of R and U steps, the number of paths with an even number of turns is given by

$$q(2m+1,2l) = \binom{m}{l}\binom{m-1}{l-1} + \binom{m}{l-1}\binom{m-1}{l}.$$

Let us now consider the case of an odd number of turns, $k = 2l+1$. If a path starts with an R step, then it must end with a U step. There are l breaks in the sequence of $m+1$ R steps, and l breaks in the sequence of m U steps. These breaks can be made in $\binom{m}{l}\binom{m-1}{l}$ different ways. By reversing the roles of R and U steps, we obtain

$$q(2m+1,2l+1) = 2\binom{m}{l}\binom{m}{l-1}.$$

As before, both results can be written in a single formula, using the floor and ceiling functions.

$$q(2m+1,k) = \binom{m}{\lfloor \frac{k-1}{2} \rfloor}\binom{m-1}{\lceil \frac{k-1}{2} \rceil} + \binom{m}{\lceil \frac{k-1}{2} \rceil}\binom{m-1}{\lfloor \frac{k-1}{2} \rfloor}. \tag{2.2}$$

By denoting the total number of steps in the considered path by n and assuming $n \geq 2$, formulas (2.1) and (2.2) can be united in the following expression:

$$q(n,k) = \binom{\lceil \frac{n-2}{2} \rceil}{\lceil \frac{k-1}{2} \rceil}\binom{\lfloor \frac{n-2}{2} \rfloor}{\lfloor \frac{k-1}{2} \rfloor} + \binom{\lfloor \frac{n-2}{2} \rfloor}{\lceil \frac{k-1}{2} \rceil}\binom{\lceil \frac{n-2}{2} \rceil}{\lfloor \frac{k-1}{2} \rfloor}. \tag{2.3}$$

Formula (2.3) is, in a sense, the central result of this note. Besides being of certain interest on its own in the context of lattice path enumeration (the sequence $q(n,k)$ not being in [9]), it readily yields the solution to our initial problem, and this, in turn, leads to a result in the context of permutation statistics. Before turning our attention to these questions, we list some properties of the sequence $q(n,k)$ that can be easily verified, either by simple combinatorial arguments, or by routine (although somewhat tedious) direct calculations.

Proposition 2.1. Let $n \geq 2$, $m \geq 1$, and $k \geq 1$ be integers, and $q(n,k)$ the sequence defined by formula (2.3). Then the following are valid.

(i) $q(n,k) = 0$ for $k \geq n$.

(ii) $q(2m,k)$ is even, for all m,k.

(iii) $q(2m,k) = q(2m, 2m-k)$, for all m,k.

(iv) The sequence $q(n,k)$ is unimodal in k, i.e., there is an index k_m such that $q(n,1) \leq \ldots \leq q(n,k_m) \geq \ldots \geq q(n,n-1)$.

(v) The index k_m is unique if n is not divisible by 4, and its value is equal to $2t+1$ for n of the form $4t+1$, $4t+2$, or $4t+3$. If n is a multiple of 4, then $k_m \in \{n/2-1, n/2, n/2+1\}$.

(vi) $\sum_k q(n,k) = \binom{n}{\lfloor \frac{n}{2} \rfloor}$.

Proof. Claim (i) is obvious, since the maximal number of turns in a path cannot be equal to or exceed the number of steps. Claim (ii) follows from the explicit formula (2.1), and (iii) is a consequence of the symmetry property for the binomial coefficients. In order to prove (iv), we first note that, for a fixed n, the quotients of consecutive elements of $q(n,k)$ form a (weakly) decreasing sequence in k. For example, $\frac{q(2m,2l)}{q(2m,2l-1)} = \frac{q(2m,2l+1)}{q(2m,2l)} = \frac{m-l}{l}$, and similarly for an odd n. Further, for all n we have $\frac{q(n,2)}{q(n,1)} \geq 1$ and $\frac{q(n,n-1)}{q(n,n-2)} \leq 1$. Hence, there must exist a k such that $\frac{q(n,k)}{q(n,k-1)} \geq 1$ and $\frac{q(n,k+1)}{q(n,k)} \leq 1$. This, in turn, implies (iv). The claim (v) can be verified by finding the value j such that the expression $\frac{q(n,j)}{q(n,j-1)} - 1$ changes sign, and showing that this value cannot be an integer unless n is divisible by 4. An easier way is to check that the indicated values are greater than their immediate neighbors, and then the rest follows by unimodality. Finally, (vi) follows from the fact that summing over all possible number of turns gives the total number of paths. $\square$

The numbers $q(n,k)$ can be ordered in a triangular array. The first few rows are given in Table 1.

3. Decks and permutations

Let us now come back to our initial problem: How many decks of m red and m black cards have exactly k color changes? If all cards of the same color were identical, the answer would have been given by formula (2.1). However, they are not identical - they differ by suits and values, and there are no two of them that

$n\backslash k$	1	2	3	4	5	6	7	8	9
2	2								
3	2	1							
4	2	2	2						
5	2	3	4	1					
6	2	4	8	4	2				
7	2	5	12	9	6	1			
8	2	6	18	18	18	6	2		
9	2	7	24	30	36	18	8	1	
10	2	8	32	48	72	48	32	8	2

TABLE 1. The first few rows of the array $q(n, k)$.

have the same suit and value. Hence, we have to allow for different ways of ordering cards of each color in their positions. Doing so, we obtain the following answer to our initial question.

Proposition 3.1.

$$p(2m, k) = 2\binom{m-1}{\lfloor \frac{k-1}{2} \rfloor}\binom{m-1}{\lceil \frac{k-1}{2} \rceil}(m!)^2.$$

Proof. In each of $q(2m, k)$ decks counted by formula (2.1) there are $m!$ ways to order red cards in their positions, and $m!$ ways to order black cards in their positions. Therefore, the number $q(2m, k)$ must be multiplied by $(m!)^2$. □

The following two results are direct consequences of Propositions 2.1 and 3.1.

Corollary 3.2. The most probable number of color changes in a deck of m red and m black cards is m, for an odd m. If m is even, then the decks with $k = m - 1$, m, and $m + 1$ color changes are all equally likely, and no other number of changes is more probable.

Proof. The number of all possible decks is $(2m)!$, and hence the probability that a given deck has exactly k color changes is given by $p(2m, k)/(2m)!$. As the denominator does not depend on k, this expression is maximized for the same value(s) of k that maximize(s) $p(2m, k)$. The claim now follows by applying Proposition 2.1, (iv) and (v). □

Corollary 3.3. The expected number of color changes in a deck of m red and m black cards is equal to m.

Proof. The expected number of color changes is defined [6] by

$$E(k) = \frac{\sum_{k=1}^{2m-1} kp(2m, k)}{\sum_{k=1}^{2m-1} p(2m, k)}.$$

By grouping together the terms with k and $2m - k$ in the numerator, and by using the symmetry property (iii) from Proposition 2.1, we obtain

$$E(k) = \frac{m}{\sum_{k=1}^{2m-1} p(2m, k)}\left[2\sum_{k=1}^{m-1} p(2m, k) + p(2m, m)\right].$$

But again by the symmetry, the term in the square brackets is equal to $\sum_{k=1}^{2m-1} p(2m, k)$, and the claim of the Corollary follows. □

Hence, in the standard case of $m = 26$, one can expect 26 changes of color in

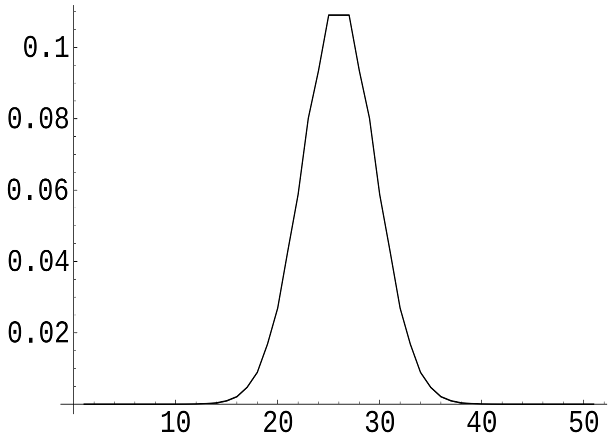

FIGURE 2. Probability distribution of the number of color changes in the standard deck.

a well-shuffled deck. Since 26 is even, it follows from Corollary 3.2 that the decks with 25, 26, and 27 changes are all equally likely. The probability for any of these number of color changes is given by $p(52, 26)/52!$, and this is close to 0.109. So, the three most likely cases together account for about one third of all possible decks. On the other hand, the probability of only one color change is approximately $4 \cdot 10^{-15}$, the same as the probability of a deck in which no two cards of the same color come together. The distribution of the probability of k color changes for the standard deck of 52 cards is shown in Fig. 2.

Let us now consider a deck of n cards of two colors such that the numbers of cards of each color differ by at most one. By a suitable labeling of cards with the first n natural numbers, we can use formula (2.3) to enumerate permutations of the set $[n] = \{1, \dots, n\}$ with respect to the number of parity changes.

Theorem 3.4. *The number $r(n, k)$ of permutations of $[n]$ with exactly k changes of parity is given by the formula*

$$r(n,k) = \left[\binom{\lceil \frac{n-2}{2} \rceil}{\lceil \frac{k-1}{2} \rceil} \binom{\lfloor \frac{n-2}{2} \rfloor}{\lfloor \frac{k-1}{2} \rfloor} + \binom{\lfloor \frac{n-2}{2} \rfloor}{\lceil \frac{k-1}{2} \rceil} \binom{\lceil \frac{n-2}{2} \rceil}{\lfloor \frac{k-1}{2} \rfloor} \right] \left\lfloor \frac{n}{2} \right\rfloor! \left\lceil \frac{n}{2} \right\rceil! . \tag{3.1}$$

Proof. Let us take a deck of m red and m black cards, and label the red cards by the first m odd numbers, and the black cards by the first m even numbers. Every such deck gives rise to a permutation of the set $\{1, \dots, 2m\}$, and the color changes in the deck correspond to the changes of parity between neighboring elements of the permutation. By labeling a deck of $m + 1$ red and m black cards in the same manner, one can also get the mapping between such decks and the permutations of the first $2m+1$ natural numbers. In both cases, the correspondence is bijective. □

Written in a more compact way, $r(n, k) = q(n, k)\lfloor \frac{n}{2} \rfloor! \lceil \frac{n}{2} \rceil!$. Again, the sequence $r(n, k)$ is new, in the sense that it is not listed in [9]. This comes as a bit of surprise, since permutation statistics have a long history; the first results were given by Euler, and the area was extensively studied by MacMahon at the begining of the

20th century [10]. Among the most famous combinatorially interesting sequences originating in this historical material are certainly the Eulerian numbers and the Stirling numbers of the first kind. We refer the reader to Ch. 6 of [8] for more information on these important sequences. The reader interested in making his own permutation statistic may consult the recent reference [11].

For a fixed n, the behavior of the numbers $r(n,k)$ is very much the same as for the sequence $q(n,k)$. Hence, for example, in a typical permutation of $[n]$ one can expect $n/2$ parity changes. The first few rows of the array $r(n,k)$ are given in Table 2.

$n\backslash k$	1	2	3	4	5	6	7	8	9
2	2								
3	4	2							
4	8	8	8						
5	24	36	48	12					
6	72	144	288	144	72				
7	288	720	1728	1296	864	144			
8	1152	3456	10368	10368	10368	3456	1152		
9	5760	20160	69120	86400	103680	51840	23040	2880	
10	28800	115200	460800	691200	1036800	691200	460800	115200	28800

TABLE 2. The first few rows of the array $r(n,k)$.

4. Further directions

The path that leads one away from beating one's kids to counting permutations can be followed even further. In what follows, we indicate several trails that an interested reader might wish to explore.

First, it should be straightforward to extend our results to the case of unbalanced decks, i.e., to the decks with m_1 red and m_2 black cards; the case $m_1 = m_2 + 1$ has been already solved. Further, it might be interesting to consider the case when some of the cards are identical. Such a situation arises, for example, when one combines two (or more) decks for certain games such as rummy. This case should not be much harder than the solved one.

More difficulties are to be expected if instead of colors one considers changes of suits. This problem can further be generalized to counting the permutations of $[n]$ with respect to the number of changes of congruences $\pmod{s}$ for some $s \geq 3$. It would be interesting to see how much of the symmetry of formula (2.3) (or (3.1)) would be preserved in such a case.

It was mentioned above that the sequences $q(n,k)$ and $r(n,k)$ do not appear in [9]. However, there is a sequence there (A088855) that corresponds to the numbers $q(2m,k)/2$. As a description, it is stated that this sequence enumerates symmetric Dyck paths of semilength m with k peaks. From there it follows that there should exist a two-to-one correspondence between the set of all grand Dyck paths on $2m$ steps with k turns and the set of all Dyck paths on $4m-2$ steps that are symmetric with respect to the line $x+y=m$ with k RU turns. (Equivalently, there should be a bijection between the set of all grand Dyck paths on $2m$ steps that start with an R step and the set of all symmetric Dyck paths on $4m-2$ steps with k RU turns.)

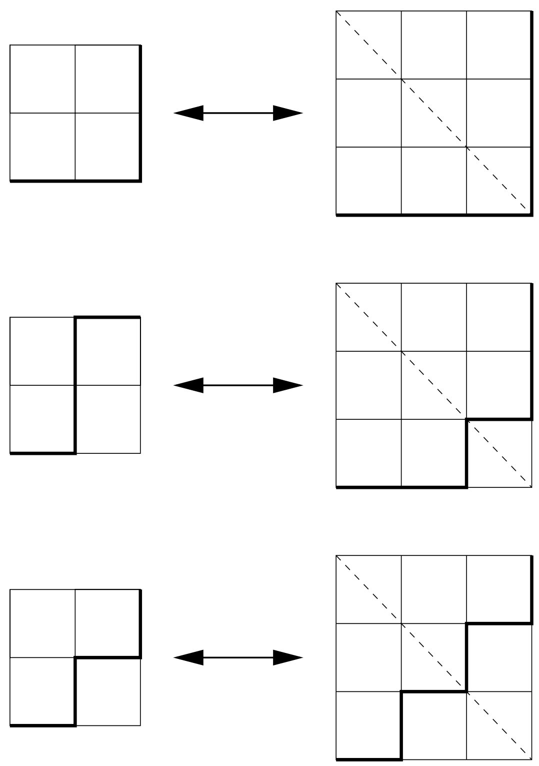

FIGURE 3. Correspondence between grand Dyck paths and symmetric Dyck paths.

The corresponding sets for $m = 2$ are shown in Fig. 3. The reader is invited to describe the correspondence for a general m and to verify that it is a bijection. The author is indebted to Daniel Soll of Marburg, Germany, who brought to his attention a mistake in earlier formulation of the bijection.

We conclude by mentioning that the results presented here may be also interpreted in terms of random walks on nonnegative integers. For more information on this topic, we refer the reader to the monograph [6].

References

1. D. Zagier, *How often should you beat your kids*, Math. Mag. **63** (1990) 89–92.
2. K.M. Levasseur, *How to beat your kids in their own game*, Math. Mag. **61** (1988) 301–305.
3. A. Knopfmacher, H. Prodinger, *A simple card guessing game revisited*, Elec. J. Comb. **8** (2001) #R13 (9 pp.).
4. S.G. Mohanty, *Lattice Path Counting and Applications*, Academic Press, New York, 1979.
5. T.V. Narayana, *Lattice Path Combinatorics with Statistical Applications*, Toronto Univ. Press, Toronto, 1979.
6. W. Feller, *An Introduction to Probability Theory and its Applications*, Vol 1, (3rd ed.), Wiley, New York, 1968.
7. E. Deutsch, *Dyck path enumeration*, Discr. Math. **204** (1999) 167–202.
8. R.L. Graham, D.E. Knuth, O. Patashnik, *Concrete Mathematics*, (2nd ed.), Addison Wesley, Reading, MA, 1994.
9. N.J.A. Sloane, *The On-Line Encyclopedia of Integer Sequences*, published electronically at http://www.research.att.com/∼njas/sequences/index.html

10. P.A. MacMahon, *Combinatory Analysis*, 2 vols, Cambridge University Press, Cambridge, 1915 (reprinted by Chelsea, New York, 1955).
11. E. Deutsch, W.P. Johnson, *Create your own permutation statistics*, Math. Mag. **77** (2004) 130–134.

Texas A&M University at Galveston, Galveston, Texas, 77551, U.S.A. and Faculty of Agriculture, University of Zagreb, Svetošimunska c. 25, 10000 Zagreb, CROATIA
E-mail address: `doslict@tamug.tamu.edu`