

# Koroljuk's formula for counting lattice paths revisited

James J.Y. Zhao

Center for Applied Mathematics  
Tianjin University, Tianjin 300072, P. R. China

Email: jjyzhao@tju.edu.cn

**Abstract.** Koroljuk gave a summation formula for counting the number of lattice paths from  $(0, 0)$  to  $(m, n)$  with  $(1, 0), (0, 1)$ -steps in the plane that stay strictly above the line  $y = k(x - d)$ , where  $k$  and  $d$  are positive integers. In this paper we obtain an explicit formula for the number of lattice paths from  $(a, b)$  to  $(m, n)$  above the diagonal  $y = kx - r$ , where  $r$  is a rational number. Our result slightly generalizes Koroljuk's formula, while the former can be essentially derived from the latter. However, our proof uses a recurrence with respect to the starting points, and hereby presents a new approach to Koroljuk's formula.

*Keywords:* Koroljuk's formula, forward recursion, lattice path, ballot problem.

*Mathematics Subject Classification:* 05A15, 05A19

## 1 Introduction

Lattice paths are simple and elementary combinatorial structures, while they have many interesting applications in combinatorics, probability and statistics, see for instance Mohanty [15] and Krattenthaler and Mohanty [13]. In the study of the discrepancy of empiric distributions, Koroljuk [12] obtained a formula for counting the number of lattice paths in the plane that stay strictly above certain line with slope and  $y$ -intercept corresponding to a pair of integers. The main objective of this paper is to give a new approach to Koroljuk's formula. The original proof given by Koroljuk used a technique of decomposing the paths, and our proof relies on a recurrence with respect to the starting points.

Let us first give an overview of Koroljuk's formula. It should be mentioned that Koroljuk's formula is originally concerned with the enumeration of lattice paths from  $(0, 0)$  to  $(m - pn, m + n)$  using the steps  $(1, 1)$  and  $(-p, 1)$  in the plane that intersect the line  $x = c$ , for some positive integers  $m, n, p$  and  $c$ . Koroljuk [12, eq. (4)]<sup>1</sup> found that such lattice paths are counted by

$$c + \lfloor \frac{m+n-c}{p+1} \rfloor (p+1) \sum_{s \equiv c \pmod{p+1}} \frac{c}{s} \binom{s}{\frac{s-c}{p+1}} \binom{m+n-s}{n - \frac{s-c}{p+1}}, \quad (1.1)$$

where  $\lfloor x \rfloor$  denotes the largest integer not greater than  $x$ .

Koroljuk's formula also counts the number of certain lattice paths using  $(1, 0), (0, 1)$ -steps in the plane with some restrictions, see for instance Niederhausen [17, p. 9]. For some positive integers  $m, k$  and rational number  $d \geq \frac{k-1}{k}m$ , Niederhausen [17, p. 9]<sup>2</sup> showed that the number of lattice paths from  $(0, 0)$  to  $(m, n)$  with  $(1, 0), (0, 1)$ -steps that stay strictly above the line  $y = k(x - d)$  is equal to

$$\binom{m+n}{m} - \sum_{i=\lfloor \frac{kd-1}{k+1} \rfloor + 1}^m \frac{n-k(m-d)}{n-k(i-d)} \binom{i+k(i-d)}{i} \binom{m-i-1+n-k(i-d)}{m-i}. \quad (1.2)$$

Substituting  $p = k$ ,  $m = n$ ,  $n = m$ ,  $c = n - k(m - d)$  and  $s = n - k(m - d) + (k + 1)i$  in Koroljuk's formula (1.1), we obtain

$$\sum_{i=0}^{m - \lfloor \frac{kd}{k+1} \rfloor} \frac{n-k(m-d)}{n-k(m-d)+(k+1)i} \binom{n-k(m-d)+(k+1)i}{i} \binom{m+k(m-d)-(k+1)i}{m-i},$$

<sup>1</sup>Here the original upper bound  $m + n - \lfloor \frac{c}{p} \rfloor$  of summation is changed to  $c + \lfloor \frac{m+n-c}{p+1} \rfloor (p+1)$  without affecting the value of the summation.

<sup>2</sup>Here the original lower bound  $\lfloor d \rfloor + 1$  of summation is changed to  $\lfloor \frac{kd-1}{k+1} \rfloor + 1$ , which is more accurate. The details are shown in Section 3.

where  $\lceil x \rceil$  is the least integer not less than  $x$ . By further substituting  $m - i$  for  $i$ , the above formula becomes

$$\sum_{i=\lceil \frac{kd}{k+1} \rceil}^m \frac{n-k(m-d)}{m+n+kd-(k+1)i} \binom{m+n+kd-(k+1)i}{m-i} \binom{(k+1)i-kd}{i}.$$

It is easy to verify that the above sum is equal to the subtrahend term of (1.2). Therefore, Koroljuk's formula also gives the number of lattice paths from  $(0, 0)$  to  $(m, n)$  with  $(1, 0), (0, 1)$ -steps that touch or cross the line  $y = k(x - d)$ . The equivalence between these two kinds of lattice paths counted by Koroljuk's formula will be illustrated in the proof of Theorem 3.1.

In this paper, we are concerned with the enumeration of lattice paths from  $(a, b)$  to  $(m, n)$ , using  $(1, 0), (0, 1)$ -steps in the plane, that stay above or strictly above a boundary line  $y = kx - r$ , where  $a, b, m, n, k$  are nonnegative integers and  $r$  is a rational number. Although the enumeration of lattice paths starting from a non-origin point can be transformed to that from the origin, it is convenient to directly work with the former. Based on this consideration, we present another approach to Koroljuk's formula. The main feature of the present approach is to use the forward recursion with respect to the starting points of lattice paths. By using some reflections on lattice paths, we derive the enumeration formula for the lattice paths with respect to the boundary line  $y = x/k - r$  from that with respect to  $y = kx - r$ .

We would like to point out that the enumeration problem considered here has attracted much attention over the past several decades. We shall give an overview of some main methods which have been used in this field, and these methods vary significantly depending upon the boundary lines and the starting points. For the convenience, let us first introduce some notations. Given rational numbers  $q > 0$  and  $r$ , and integers  $a, b, m, n$  with  $0 \leq a \leq m$  and  $0 \leq b \leq n$ , let  $L_{q,r}(a, b; m, n)$  denote the set of lattice paths from  $(a, b)$  to  $(m, n)$  that stay above the line  $y = qx - r$ , and let  $\mathfrak{L}_{q,r}(a, b; m, n)$  denote the set of lattice paths from  $(a, b)$  to  $(m, n)$  that stay strictly above the line  $y = qx - r$ . With these notations, Koroljuk's formula is related to the evaluation of  $|\mathfrak{L}_{k,kd}(0, 0; m, n)|$ .

The first lattice path enumeration problem is to count  $|L_{1,0}(0, 0; m, n)|$ . It is often thought that this problem originated from Bertrand's ballot problem [5], with the solution

$$|L_{1,0}(0, 0; m, n)| = \frac{n-m+1}{n+1} \binom{m+n}{m},$$

which, known as the ballot number, is a generalization of the well-known Catalan number  $\frac{1}{n+1} \binom{2n}{n}$  (see [20, A000108]). One of the most classical techniques for solving Bertrand's ballot problem is the renowned *reflection principle*, which has been widely contributed to André [2]. Renault [18] pointed out that the reflection principle is in fact a variation due to Aebly [1] and Mirimanoff [14], see also [10, §2.3].

Barbier [4] further studied the generalized ballot problem, which is equivalent to that of counting  $|L_{k,0}(0, 0; m, n)|$ . The solution

$$|L_{k,0}(0, 0; m, n)| = \binom{m+n}{m} - k \binom{m+n}{m-1} \quad (1.3)$$

is known as the generalized ballot number. In the spirit of the reflection principle, Goulden and Serrano [8] gave a bijective proof for the generalized ballot problem by using rotation instead of reflection.

It is clear that the boundary line for the ballot problem or its generalization has zero  $y$ -intercept. If the  $y$ -intercept of the boundary line is nonzero, the corresponding lattice path enumeration problem will become a little complicated. As remarked before, Koroljuk [12] essentially studied the problem of counting  $|\mathfrak{L}_{k,kd}(0, 0; m, n)|$ , where  $k, d$  are positive integers. Niederhausen [16, Example 2.2] gave a formula of evaluating  $|\mathfrak{L}_{k,r}(0, 0; m, n)|$  for positive integers  $k, r$ . In fact, Niederhausen considered a multi-dimensional lattice path enumeration problem, and his method is based on the finite operator calculus developed in Rota et al. [19]. It should be mentioned that Böhm [6, (10)] studied an enumeration problem which is equivalent to the evaluation of  $|\mathfrak{L}_{k,r}(0, 0; m, n)|$ . This equivalence will be illustrated in Section 3. A formula of  $|\mathfrak{L}_{\frac{1}{k}, \frac{c}{k}}(0, 0; m, n)|$  with positive integers  $k, c$  was given in Niederhausen [17, Proposition 4].

Among the above lattice path enumeration problems, the starting points are fixed to be the origin of the plane. For this reason, all manipulations on the lattice paths are naturally required to fix the starting points. In fact, all known methods obeyed this spirit. Since we are considering the enumeration of lattice paths starting from a point not necessarily being fixed (usually the origin is chosen as the starting point), we could use induction on the starting points. This suggests a forward recursion approach to the enumeration problem considered here.

This paper is organized as follows. In Section 2, by using the forward recursion, we obtain a formula of  $|L_{k,r}(a, b; m, n)|$  for integers  $k, r, a, b, m, n$ , as well as that of  $|\mathfrak{L}_{k,r}(a, b; m, n)|$ ,  $|L_{\frac{1}{k},r}(a, b; m, n)|$  and  $|\mathfrak{L}_{\frac{1}{k},r}(a, b; m, n)|$ . In Section 3, we shall show how to derive Koroljuk's formula from the results in Section 2. In particular, we shall show the relations among the lattice paths with unit steps  $(1, 1)$  and  $(-p, 1)$  studied by Koroljuk, the lattice paths with unit steps  $(1, t)$  and  $(1, -1)$  studied by Böhm, and the lattice paths with unit steps  $(1, 0)$  and  $(0, 1)$  studied by Niederhausen, where  $p, t$  are some fixed positive integers.

## 2 The Lattice Path Enumeration Formulas

In this section we give an enumeration formula for  $|L_{k,r}(a, b; m, n)|$ . By using some simple bijections, we then derive the enumeration formulas for  $|\mathfrak{L}_{k,r}(a, b; m, n)|$ ,  $|L_{\frac{1}{k},r}(a, b; m, n)|$  and  $|\mathfrak{L}_{\frac{1}{k},r}(a, b; m, n)|$  from that for  $|L_{k,r}(a, b; m, n)|$ .

Our first main result is as follows.

**Theorem 2.1** *Suppose that  $k, r, a, b, m$  and  $n$  are integers satisfying*

$$k \geq 1, \quad 0 \leq a \leq m, \quad n \geq km - r, \quad \max\{0, ka - r\} \leq b \leq n. \quad (2.1)$$

*Then we have*

$$|L_{k,r}(a, b; m, n)| = \sum_{i=0}^{\lfloor \frac{b+r-ka}{k+1} \rfloor} (-1)^i \binom{b+r-k(a+i)}{i} h(a, m, n, k, r, i), \quad (2.2)$$

where

$$h(a, m, n, k, r, i) = \frac{n+r+1-km}{n+r+1-k(a+i)} \binom{m+n+r-(k+1)(a+i)}{m-a-i}. \quad (2.3)$$

**Remark.** Note that when  $a = b = r = 0$ ,  $k \rightarrow k - 1$  and  $n = (k - 1)m$ , (2.2) becomes

$$|L_{k-1,0}(0, 0; m, (k-1)m)| = \frac{1}{(k-1)m+1} \binom{km}{m},$$

which are known as the order- $k$  Fuss-Catalan numbers, see [9, Eq. (7.67)]. They also enumerate  $(k - 1)$ -Dyck paths, as well as  $k$ -ary trees. In particular, when  $k = 2$  they are the well-known Catalan numbers, and when  $k = 3$  they form the sequence A001764 in [20] which count 2-Dyck paths, as well as ternary trees. For more details, see [3].

We shall prove Theorem 2.1 by induction. Before doing this, let us first give two lemmas, one of which is concerned with a recursion relation on  $|L_{k,r}(a, b; m, n)|$ , and the other some initial values of  $|L_{k,r}(a, b; m, n)|$ . For the recursion relation on  $|L_{k,r}(a, b; m, n)|$ , we have the following result.

**Lemma 2.1** *Suppose that  $k, r, a, b, m$  and  $n$  are integers satisfying*

$$k, m \geq 1, \quad 0 \leq a \leq m, \quad n \geq km - r, \quad k(a+1) - r \leq b \leq n. \quad (2.4)$$

*Then we have*

$$|L_{k,r}(a, b+1; m, n)| = |L_{k,r}(a, b; m, n)| - |L_{k,r}(a, b-k; m-1, n-k)|.$$

*Proof.* Note that, for each lattice path of  $L_{k,r}(a, b; m, n)$ , the first step is either from  $(a, b)$  to  $(a+1, b)$  or from  $(a, b)$  to  $(a, b+1)$ . Moreover, if  $k, r, a, b, m$  and  $n$  satisfy (2.4), then  $(a, b)$  and  $(m, n)$  must stay above the boundary line  $y = kx - r$ . It is easy to see that

$$|L_{k,r}(a, b; m, n)| = |L_{k,r}(a+1, b; m, n)| + |L_{k,r}(a, b+1; m, n)|.$$

Now it suffices to show that

$$|L_{k,r}(a+1, b; m, n)| = |L_{k,r}(a, b-k; m-1, n-k)|. \quad (2.5)$$

To this end, define a map

$$f : L_{k,r}(a+1, b; m, n) \longrightarrow L_{k,r}(a, b-k; m-1, n-k)$$

as follows: for each lattice path  $\xi \in L_{k,r}(a+1, b; m, n)$ , let  $f(\xi)$  be the lattice path obtained from  $\xi$  by translating  $\xi$  horizontally left by one unit and vertically down by  $k$  units. Thus, the resulting path  $f(\xi)$  is from  $(a, b-k)$  to  $(m-1, n-k)$ . We claim that  $f(\xi) \in L_{k,r}(a, b-k; m-1, n-k)$ , since each lattice point  $(x, y) \in \xi$  satisfies  $y \geq kx - r$ , and its image  $(x-1, y-k) \in f(\xi)$  naturally satisfies  $y-k \geq k(x-1) - r$ , i.e.,  $f(\xi)$  stays above the line  $y = kx - r$ . Clearly, the map  $f$  is reversible and unique. Therefore, we obtain the equality (2.5). This completes the proof.  $\square$

Next we give a formula for  $|L_{k,r}(a, b; m, n)|$  when  $r = 0$  and  $0 \leq b - ka \leq k$ . For convenience, we abbreviate  $L_{k,0}(a, b; m, n)$  by  $L_k(a, b; m, n)$ . The following result serves as the base case of the induction proof of Theorem 2.1.

**Lemma 2.2** *Suppose that  $k, a, b, m$  and  $n$  are integers satisfying*

$$k, m \geq 1, \quad 0 \leq a \leq m, \quad 0 \leq b \leq n, \quad n \geq km, \quad 0 \leq b - ka \leq k.$$

*Then we have*

$$|L_k(a, b; m, n)| = \frac{n+1-km}{n+1-ka} \binom{m+n-(k+1)a}{m-a}.$$

*Proof.* Given the conditions satisfied by  $a, b, m, n$  and  $k$ , each path in  $L_k(a, b; m, n)$  must pass through the point  $(a, ka+k)$ . Thus, we have

$$|L_k(a, b; m, n)| = |L_k(a, ka; m, n)|,$$

since the path from  $(a, b)$  to  $(a, ka+k)$  is unique due to the restriction of unit steps.

Now we are to compute  $|L_k(a, ka; m, n)|$ . Let  $\mathcal{C}_k(a, b; m, n)$  denote the set of lattice paths from  $(a, b)$  to  $(m, n)$  which cross the boundary line  $y = kx$ . Since the number of lattice paths from  $(a, ka)$  to  $(m, n)$  with unit steps  $(1, 0)$  and  $(0, 1)$  is  $\binom{m+n-(k+1)a}{m-a}$ , we get that

$$|L_k(a, ka; m, n)| = \binom{m+n-(k+1)a}{m-a} - |\mathcal{C}_k(a, ka; m, n)|.$$

By translating each lattice path in  $\mathcal{C}_k(a, ka; m, n)$  horizontally left by  $a$  units and vertically down by  $ka$  units, we shall obtain a lattice path in  $\mathcal{C}_k(0, 0; m-a, n-ka)$ . Conversely, each path in  $\mathcal{C}_k(0, 0; m-a, n-ka)$  can be obtained in such a way. Therefore, we have

$$|\mathcal{C}_k(a, ka; m, n)| = |\mathcal{C}_k(0, 0; m-a, n-ka)|.$$

By the solution of the generalized ballot problem [8], we see that

$$|\mathcal{C}_k(0, 0; m, n)| = k \binom{m+n}{m-1},$$

and hence

$$|\mathcal{C}_k(0, 0; m-a, n-ka)| = k \binom{m+n-(k+1)a}{m-a-1}.$$

Therefore,

$$|L_k(a, b; m, n)| = \binom{m+n-(k+1)a}{m-a} - k \binom{m+n-(k+1)a}{m-a-1} = \frac{n+1-km}{n+1-ka} \binom{m+n-(k+1)a}{m-a}.$$

This completes the proof.  $\square$

We now proceed to prove Theorem 2.1.

*Proof of Theorem 2.1.* There is a one-to-one correspondence between the sets  $L_{k,r}(a, b; m, n)$  and  $L_k(a, b+r; m, n+r)$ , since a lattice path in the set  $L_{k,r}(a, b; m, n)$  will be transformed into a lattice path in  $L_k(a, b+r; m, n+r)$  by simply shifting the  $x$ -axis vertically down or up by  $|r|$  units according to the sign of  $r$ . This means that

$$|L_{k,r}(a, b; m, n)| = |L_k(a, b+r; m, n+r)|. \quad (2.6)$$

Now it is sufficient to prove the theorem for the case of  $r = 0$ , namely

$$|L_k(a, b; m, n)| = \sum_{i=0}^{\lfloor \frac{b-ka}{k+1} \rfloor} (-1)^i \binom{b-k(a+i)}{i} h(a, m, n, k, i), \quad (2.7)$$

where

$$h(a, m, n, k, i) = \frac{n+1-km}{n+1-k(a+i)} \binom{m+n-(k+1)(a+i)}{m-a-i}$$

is the abbreviation of  $h(a, m, n, k, 0, i)$ . For  $r \neq 0$ , one can easily derive (2.2) from (2.7) together with (2.6).

In the following we shall prove (2.7) by induction on the value  $\ell(a, b, k) = \lfloor \frac{b-ka}{k+1} \rfloor$ . For  $\ell(a, b, k) = 0$ , namely,  $0 \leq b-ka \leq k$ , it is easy to see that (2.7) holds true by Lemma 2.2. Assume that (2.7) holds for any  $a, b, k$  satisfying  $\lfloor \frac{b-ka}{k+1} \rfloor < \ell$ . We are to prove that it is also valid for  $b = \ell(k+1) + ka + j$  for some  $0 \leq j \leq k$ , namely

$$|L_k(a, \ell(k+1) + ka + j; m, n)| = \sum_{i=0}^{\ell} (-1)^i \binom{\ell(k+1)+j-ki}{i} h(a, m, n, k, i). \quad (2.8)$$

By Lemma 2.1, we have

$$|L_k(a, b; m, n)| = |L_k(a, b-1; m, n)| - |L_k(a, b-1-k; m-1, n-k)|.$$

Since for  $L_k(a, b-1-k; m-1, n-k)$  it holds

$$\ell(a, b-1-k, k) = \left\lfloor \frac{(b-1-k)-ka}{k+1} \right\rfloor = \left\lfloor \frac{\ell(k+1)+ka+j-1-k-ka}{k+1} \right\rfloor = \ell - 1,$$

we can determine the value of  $|L_k(a, b-1-k; m-1, n-k)|$  by the induction hypothesis. However, for  $L_k(a, b-1; m, n)$  it holds

$$\ell(a, b-1, k) = \left\lfloor \frac{(b-1)-ka}{k+1} \right\rfloor = \left\lfloor \frac{\ell(k+1)+ka+j-1-ka}{k+1} \right\rfloor = \ell$$

for  $j \geq 1$ . In order to use the induction hypothesis, we need to apply Lemma 2.1 to  $L_k(a, b-1; m, n)$ , and we get

$$|L_k(a, b-1; m, n)| = |L_k(a, b-2; m, n)| - |L_k(a, b-2-k; m-1, n-k)|.$$

Now we have

$$|L_k(a, b; m, n)| = |L_k(a, b-2; m, n)| - \sum_{t=0}^1 |L_k(a, b-2+t-k; m-1, n-k)|.$$

Again, we check that  $L_k(a, b-2-k; m-1, n-k)$  satisfies the induction hypothesis, while  $L_k(a, b-2; m, n)$  does not for  $j \geq 2$ . We now need to apply Lemma 2.1 to  $L_k(a, b-2; m, n)$ . Iterate this procedure until reaching the following formula

$$|L_k(a, b; m, n)| = |L_k(a, b-j-1; m, n)| - \sum_{t=0}^j |L_k(a, b-j-1+t-k; m-1, n-k)|.$$

By using the induction hypothesis and substituting  $\ell(k+1) + ka + j$  for  $b$ , we find that

$$|L_k(a, b-j-1; m, n)| = \sum_{i=0}^{\ell-1} (-1)^i \binom{\ell(k+1)-1-ki}{i} h(a, m, n, k, i),$$

and

$$\sum_{t=0}^j |L_k(a, b-j-1+t-k; m-1, n-k)|$$

$$\begin{aligned}
&= \sum_{t=0}^j \sum_{i=0}^{\ell-1} (-1)^i \binom{(\ell-1)(k+1) + t - ki}{i} h(a, m-1, n-k, k, i) \\
&= \sum_{t=0}^j \sum_{i=1}^{\ell} (-1)^{i-1} \binom{\ell(k+1) + t - 1 - ki}{i-1} h(a, m, n, k, i),
\end{aligned}$$

where the last equality follows from the fact that  $h(a, m-1, n-k, k, i-1) = h(a, m, n, k, i)$ . Therefore,

$$\begin{aligned}
|L_k(a, b; m, n)| &= \sum_{i=0}^{\ell-1} (-1)^i \binom{\ell(k+1) - 1 - ki}{i} h(a, m, n, k, i) \\
&\quad - \sum_{t=0}^j \sum_{i=1}^{\ell} (-1)^{i-1} \binom{\ell(k+1) + t - 1 - ki}{i-1} h(a, m, n, k, i) \\
&= \sum_{i=0}^{\ell-1} (-1)^i \binom{\ell(k+1) - 1 - ki}{i} h(a, m, n, k, i) \\
&\quad - \sum_{i=1}^{\ell} (-1)^{i-1} \left( \sum_{t=0}^j \binom{\ell(k+1) + t - 1 - ki}{i-1} \right) h(a, m, n, k, i).
\end{aligned}$$

Note that

$$\begin{aligned}
\sum_{t=0}^j \binom{\ell(k+1) + t - 1 - ki}{i-1} &= \sum_{t=0}^j \left( \binom{\ell(k+1) + t - ki}{i} - \binom{\ell(k+1) + t - 1 - ki}{i} \right) \\
&= \binom{\ell(k+1) + j - ki}{i} - \binom{\ell(k+1) - 1 - ki}{i}.
\end{aligned}$$

Thus,

$$\begin{aligned}
|L_k(a, b; m, n)| &= \sum_{i=0}^{\ell-1} (-1)^i \binom{\ell(k+1) - 1 - ki}{i} h(a, m, n, k, i) \\
&\quad - \sum_{i=1}^{\ell} (-1)^i \binom{\ell(k+1) - 1 - ki}{i} h(a, m, n, k, i) \\
&\quad - \sum_{i=1}^{\ell} (-1)^{i-1} \binom{\ell(k+1) + j - ki}{i} h(a, m, n, k, i) \\
&= h(a, m, n, k, 0) + \sum_{i=1}^{\ell} (-1)^i \binom{\ell(k+1) + j - ki}{i} h(a, m, n, k, i) \\
&= \sum_{i=0}^{\ell} (-1)^i \binom{\ell(k+1) + j - ki}{i} h(a, m, n, k, i),
\end{aligned}$$

as desired. This completes the proof.  $\square$

Now we are able to give a formula to compute  $|\mathfrak{L}_{k,r}(a, b; m, n)|$ .

**Corollary 2.1** *Suppose that  $k, r, a, b, m$  and  $n$  are integers satisfying*

$$k \geq 1, \quad 0 \leq a \leq m, \quad \max\{0, ka - r\} < b \leq n, \quad n > km - r.$$

Then we have

$$|\mathfrak{L}_{k,r}(a, b; m, n)| = \sum_{i=0}^{\lfloor \frac{b+r-1-ka}{k+1} \rfloor} (-1)^i \binom{b+r-1-ki}{i} h(a, m, n-1, k, r, i), \quad (2.9)$$

where  $h(a, m, n, k, r, i)$  is defined as in (2.3).

*Proof.* There is an obvious bijection between  $\mathfrak{L}_{k,r}(a, b; m, n)$  and  $L_{k,r}(a, b-1; m, n-1)$ . Shift any path  $\xi \in \mathfrak{L}_{k,r}(a, b; m, n)$  by one vertical down unit, and the resulting path obviously belongs to  $L_{k,r}(a, b-1; m, n-1)$ . Clearly, this map is reversible and unique. Thus

$$|\mathfrak{L}_{k,r}(a, b; m, n)| = |L_{k,r}(a, b-1; m, n-1)|.$$

By (2.2), we immediately obtain the desired result. This completes the proof.  $\square$

Theorem 2.1 can also be used to derive the number of lattice paths above the boundary line  $y = \frac{x}{k} - r$  for some positive integer  $k$ .

**Corollary 2.2** *Suppose that  $k, r, a, b, m$  and  $n$  are integers satisfying*

$$k \geq 1, \quad 0 \leq a \leq m, \quad \max\{0, \frac{a}{k} - r\} \leq b \leq n, \quad n \geq \frac{m}{k} - r.$$

*Then we have*

$$|L_{\frac{1}{k},r}(a, b; m, n)| = \sum_{i=0}^{\lfloor \frac{k(n+r)-m}{k+1} \rfloor} (-1)^i \binom{k(n+r-i)-m}{i} s(a, b, n, k, r, i), \quad (2.10)$$

*where*

$$s(a, b, n, k, r, i) = \frac{k(b+r) - a + 1}{k(n+r-i) - a + 1} \binom{(k+1)(n-i) - a - b + kr}{n-b-i}. \quad (2.11)$$

*Proof.* We shall show that there is a one-to-one correspondence between  $L_{\frac{1}{k},r}(a, b; m, n)$  and  $L_k(0, k(n+r) - m; n - b, k(n+r) - a)$ . Define a map  $f : L_{\frac{1}{k},r}(a, b; m, n) \rightarrow L_k(0, k(n+r) - m; n - b, k(n+r) - a)$  in the following two steps.

1. First, shift the  $x$ -axis vertically up by  $n$  units and the  $y$ -axis horizontally right by  $k(n+r)$  units such that the new origin is located at  $(k(n+r), n)$ , and then reverse the direction of each axis. Now, a lattice path  $\xi \in L_{\frac{1}{k},r}(a, b; m, n)$  in the original coordinate system becomes a lattice path from  $(k(n+r) - m, 0)$  to  $(k(n+r) - a, n - b)$  that stays below the line  $y = \frac{x}{k}$  in the new coordinate system, denoted by  $\xi_1$ .
2. Second, reflect  $\xi_1$  along the diagonal  $y = x$ . Denote the resulting path by  $f(\xi)$ . It is clear that  $f(\xi)$  is a lattice path from  $(0, k(n+r) - m)$  to  $(n - b, k(n+r) - a)$  that stays above the line  $y = kx$ , namely,  $f(\xi) \in L_k(0, k(n+r) - m; n - b, k(n+r) - a)$ .

Clearly, each step is uniquely reversible. So,  $f$  is bijective. We immediately derive the desired result from (2.7). This completes the proof.  $\square$

With the aid of Corollary 2.2, we obtain the following result for counting the lattice paths of  $\mathfrak{L}_{\frac{1}{k},r}(a, b; m, n)$ . The proof is similar to that of Corollary 2.1 as derived from Theorem 2.1 and is omitted.

**Corollary 2.3** *Suppose that  $k, r, a, b, m$  and  $n$  are integers satisfying*

$$k \geq 1, \quad 0 \leq a \leq m, \quad \max\{0, \frac{a}{k} - r\} < b \leq n, \quad n > \frac{m}{k} - r.$$

*Then we have*

$$|\mathfrak{L}_{\frac{1}{k},r}(a, b; m, n)| = \sum_{i=0}^{\lfloor \frac{k(n+r)-m-1}{k+1} \rfloor} (-1)^i \binom{k(n+r-i)-m-1}{i} s(a+1, b, n, k, r, i), \quad (2.12)$$

*where  $s(a, b, n, k, r, i)$  is defined as in (2.11).*

It should be mentioned that Niederhausen [17, Proposition 4] gave a formula for the number of lattice paths from  $(0, 0)$  to  $(n, m)$  staying strictly above the line  $y = (x - c)/k$  for positive integers  $k$  and  $c$ , by using Koroljuk's method of decomposing lattice paths. This enumeration formula is equivalent to  $|\mathfrak{L}_{\frac{1}{k}, \frac{c}{k}}(0, 0; n, m)|$ .

**Remark.** Note that when  $r$  is a real number but not an integer, the restriction of the boundary  $y = kx - r$  on the lattice paths is equivalent to that of the line  $y = kx - \lfloor r \rfloor$  for (2.2). Now we have  $|L_{k,r}(a, b; m, n)| = |\mathfrak{L}_{k,r}(a, b; m, n)| = |L_{k, \lfloor r \rfloor}(a, b; m, n)|$ . In addition, as long as  $kr$  is an integer, the formulae (2.10) and (2.12) still work. If  $kr$  is not an integer, it is easy to see that  $L_{\frac{1}{k},r}(a, b; m, n) = \mathfrak{L}_{\frac{1}{k},r}(a, b; m, n) = L_{\frac{1}{k}, \lfloor kr \rfloor}(a, b; m, n)$ . Substituting  $\lfloor kr \rfloor$  for  $kr$  in (2.10), the formula still works.

### 3 Koroljuk's formula

The aim of this section is to give a new proof of Koroljuk's formula (1.1) by using the above enumeration results. To this end, we first prove the following result, a counterpart of Koroljuk's formula.

**Theorem 3.1** *For some positive integers  $m, n, p$  and  $c$ , the number of lattice paths from  $(0, 0)$  to  $(m - pn, m + n)$ , using the steps  $(1, 1)$  and  $(-p, 1)$  in the plane and not intersecting the line  $x = c$ , is enumerated by  $|\mathfrak{L}_{p, c+pn-m}(0, 0; n, m)|$ .*

*Proof.* Given a lattice path  $\xi$  from  $(0, 0)$  to  $(m - pn, m + n)$ , using the steps  $(1, 1)$  and  $(-p, 1)$  in the plane and not intersecting the line  $x = c$ , we shall map it into a lattice path in  $\mathfrak{L}_{p, c+pn-m}(0, 0; n, m)$  in the following way:

1. First, rotate  $\xi$  by  $90^\circ$  clockwise about the origin and then shift the  $x$ -axis by  $c$  units downwards. Denote the image by  $\xi_0$  (see Figure 3.1), which is a path from  $(0, c)$  to  $(m + n, c + pn - m)$  with  $(1, p), (1, -1)$ -steps that stays strictly above the  $x$ -axis. Let  $v = c + pn - m$ , which is obviously a positive integer.
2. Second, rotate  $\xi_0$  by  $180^\circ$  about the origin, and then shift the image horizontally right such that the leftmost lattice point is located on the  $y$ -axis. The resulting image, denoted by  $\xi_1$  (see Figure 3.2), is a path from  $(0, -v)$  to  $(m + n, -c)$  with  $(1, p), (1, -1)$ -steps that stays strictly below the line  $y = 0$ .
3. Thirdly, shift the  $y$ -axis horizontally left by  $v$  units. Then rotate the axes by  $45^\circ$  counterclockwise about the origin. Next reverse the direction of the  $y$ -axis and then exchange the two axes. Now scale the Cartesian coordinate system with a scale factor  $\cos \frac{\pi}{4}$ . The resulting path, denoted by  $\xi_2$  (see Figure 3.3), is from  $(v, 0)$  to  $(c + \frac{\sqrt{2+2p^2}}{2}n \sin \theta, \frac{\sqrt{2+2p^2}}{2}n \sin \theta)$ , where  $\theta = \frac{3\pi}{4} - \arctan p$  is the angle between the non-horizontal step and the horizontal step. Moreover,  $\xi_2$  stays strictly below the line  $y = x$ .
4. Next, keep the portion of  $\xi_2$  on the  $x$ -axis fixed and shear the remaining portion of  $\xi_2$  rightward parallel to the  $x$ -axis such that all the non-horizontal steps become vertical, which means there is a rotation by  $\beta_1 = \frac{\pi}{2} - \theta$  degrees clockwise for each non-horizontal step. Then scale the length of each vertical step to 1. At the same time, rotate the boundary line  $y = x$  by  $\beta_2 = \arctan p - \frac{\pi}{4}$  degrees clockwise about the origin. The boundary line becomes  $y = x/p$ . Denote the resulting path by  $\xi_3$  (see Figure 3.4). Note that  $\beta_1 = \beta_2$ . It is easy to see that  $\xi_3$  goes from  $(v, 0)$  to  $(c + pn, n)$  and stays strictly below the line  $y = x/p$ .
5. Finally, shift the  $y$ -axis horizontally right by  $v$  units, and then reflect  $\xi_3$  along the diagonal  $y = x$ . Denote the image by  $\xi'$ . It is clear that  $\xi' \in \mathfrak{L}_{p, c+pn-m}(0, 0; n, m)$ .

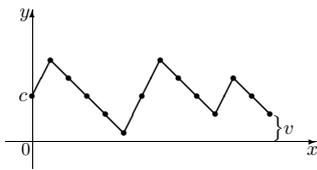


Figure 3.1  $\xi_0$  with  $p = 2$  and  $n = 4$ .

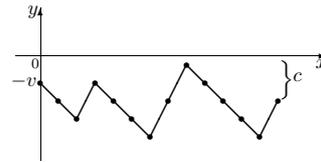


Figure 3.2  $\xi_1$  with  $p = 2$  and  $n = 4$ .

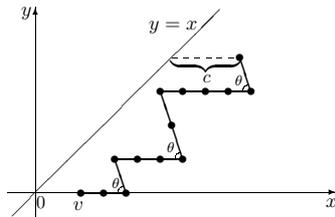


Figure 3.3  $\xi_2$  with  $p = 2$  and  $n = 4$ .

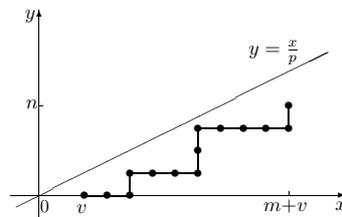


Figure 3.4  $\xi_3$  with  $p = 2$  and  $n = 4$ .

It is easy to verify that each step is reversible and unique. Therefore the map is a bijection. This completes the proof.  $\square$

**Remark.** The lattice paths obtained in the first step of the proof have been studied by Böhm [6]. As pointed out by Böhm [6, Eq. (10)], the number of paths, going from initial altitude  $m > 0$  to terminal altitude  $k > 0$  with  $(1, r), (1, -1)$ -steps such that the paths do not touch or cross the  $x$ -axis and the number of  $(1, r)$ -steps is equal to  $a$ , can be given by

$$N_{m,k}(a) = \sum_{\ell \geq 0} (-1)^\ell \frac{m}{m+(r+1)(a-\ell)} \binom{m+(r+1)(a-\ell)}{a-\ell} \binom{k-r\ell-1}{\ell}. \quad (3.1)$$

Böhm [6] attributes the above formula to Niederhausen. By the proof of the theorem, we see that  $N_{m,k}(a)$  also counts the lattice paths from  $(0, 0)$  to  $(a, m + ra - k)$  that strictly stay above the line  $y = rx - k$ . Thus,  $N_{m,k}(a) = |\mathfrak{L}_{r,k}(0, 0; a, m + ra - k)|$ . By (2.9), we can also obtain Niederhausen's formula (3.1).

Now we can prove Koroljuk's formula.

**Corollary 3.1** *For some positive integers  $m, n, p$  and  $c$ , the number of lattice paths from  $(0, 0)$  to  $(m - pn, m + n)$ , using the steps  $(1, 1)$  and  $(-p, 1)$  in the plane and intersecting the line  $x = c$ , is enumerated by (1.1).*

*Proof.* Note that there are  $\binom{m+n}{n}$  lattice paths from  $(0, 0)$  to  $(m - pn, m + n)$  by using the steps  $(1, 1)$  and  $(-p, 1)$ . By the above theorem, the number of lattice paths from  $(0, 0)$  to  $(m - pn, m + n)$ , using the steps  $(1, 1)$  and  $(-p, 1)$  in the plane and intersecting the line  $x = c$ , is counted by

$$\binom{m+n}{n} - |\mathfrak{L}_{p,c+pn-m}(0, 0; n, m)|.$$

By (2.9), it is equal to

$$\binom{m+n}{n} - \sum_{i=0}^{\lfloor \frac{c+pn-m-1}{p+1} \rfloor} (-1)^i \frac{c}{c+(p+1)(n-i)} \binom{c+(p+1)(n-i)}{n-i} \binom{c-m-1+p(n-i)}{i}. \quad (3.2)$$

It is well known that

$$\binom{x}{k} = (-1)^k \binom{k-x-1}{k}, \quad \text{for } x \in \mathbb{R}, k \in \mathbb{N},$$

see [11, Eq. (17), p. 58]. Thus, we have

$$\binom{c-m-1+p(n-i)}{i} = (-1)^i \binom{m+n-c-(p+1)(n-i)}{i},$$

and (3.2) is equal to

$$\begin{aligned} \binom{m+n}{n} - \sum_{i=0}^{\lfloor \frac{c+pn-m-1}{p+1} \rfloor} \frac{c}{c+(p+1)(n-i)} \binom{c+(p+1)(n-i)}{n-i} \binom{m+n-c-(p+1)(n-i)}{i} \\ = \binom{m+n}{n} - \sum_{i=n-\lfloor \frac{c+pn-m-1}{p+1} \rfloor}^n \frac{c}{c+(p+1)i} \binom{c+(p+1)i}{i} \binom{m+n-c-(p+1)i}{n-i}. \end{aligned}$$

By the following Hagen-Rothe identity [7, Eq. (11)],

$$\sum_{i=0}^n \frac{\gamma}{\gamma + \beta i} \binom{\gamma + \beta i}{i} \binom{\alpha + \beta(n-i)}{n-i} = \binom{\alpha + \gamma + \beta n}{n}, \quad \alpha, \beta, \gamma \in \mathbb{R}, n \in \mathbb{N},$$

we find that

$$\sum_{i=0}^n \frac{c}{c+(p+1)i} \binom{c+(p+1)i}{i} \binom{m+n-c-(p+1)i}{n-i} = \binom{m+n}{n}.$$

Therefore, (3.2) is equal to

$$\sum_{i=0}^{n-\lfloor \frac{c+pn-m-1}{p+1} \rfloor-1} \frac{c}{c+(p+1)i} \binom{c+(p+1)i}{i} \binom{m+n-c-(p+1)i}{n-i},$$

that is

$$\sum_{i=0}^{\lfloor \frac{m+n-c}{p+1} \rfloor} \frac{c}{c+(p+1)i} \binom{c+(p+1)i}{i} \binom{m+n-c-(p+1)i}{n-i}, \quad (3.3)$$

since, for positive integers  $m, n, p$  and  $c$ , we have

$$n - \left\lfloor \frac{c + pn - m - 1}{p + 1} \right\rfloor - 1 = n + \left\lfloor \frac{m - pn - c + 1}{p + 1} \right\rfloor - 1 = \left\lfloor \frac{m + n - c + 1}{p + 1} \right\rfloor - 1 = \left\lfloor \frac{m + n - c}{p + 1} \right\rfloor.$$

By the substitution of  $i = \frac{s-c}{p+1}$  for the summation index, one can obtain Koroljuk's formula (1.1) from (3.3) immediately. This completes the proof.  $\square$

As noted in the introduction, Niederhausen [17, p. 9] studied the enumeration of lattice paths from  $(0, 0)$  to  $(m, n)$  with  $(1, 0), (0, 1)$ -steps that stay strictly above the line  $y = k(x - d)$ , an equivalent problem to Koroljuk's formula. By (2.9), such lattice paths are counted by

$$\begin{aligned} & |\mathfrak{L}_{k, kd}(0, 0; m, n)| \\ &= \sum_{i=0}^{\lfloor \frac{kd-1}{k+1} \rfloor} (-1)^i \frac{n + kd - km}{m + n + kd - (k+1)i} \binom{m + n + kd - (k+1)i}{m-i} \binom{kd-1-ki}{i} \\ &= \sum_{i=0}^{\lfloor \frac{kd-1}{k+1} \rfloor} \frac{n + kd - km}{m + n + kd - (k+1)i} \binom{m + n + kd - (k+1)i}{m-i} \binom{ki - kd + i}{i} \\ &= \sum_{i=0}^m g(m, n, k, d, i) - \sum_{i=\lfloor \frac{kd-1}{k+1} \rfloor + 1}^m g(m, n, k, d, i), \end{aligned}$$

where

$$g(m, n, k, d, i) = \frac{n + kd - km}{m + n + kd - (k+1)i} \binom{m + n + kd - (k+1)i}{m-i} \binom{ki - kd + i}{i}.$$

By the Hagen-Rothe identity, it is clear that

$$\sum_{i=0}^m g(m, n, k, d, i) = \binom{m+n}{m}.$$

So, it follows that

$$|\mathfrak{L}_{k, kd}(0, 0; m, n)| = \binom{m+n}{m} - \sum_{i=\lfloor \frac{kd-1}{k+1} \rfloor + 1}^m g(m, n, k, d, i),$$

which is just (1.2) by collecting some terms.

**Acknowledgments.** I am grateful to Arthur L.B. Yang and Bing-Qing Li for inspiring discussions and valuable suggestions. I would also like to thank the referee for helpful comments. This work was supported by the 863 Program and the National Science Foundation of China.

## References

- [1] J. Aebly, Démonstration du problème du scrutin par des considérations géométriques, L'enseignement mathématique, 23 (1923) 185–186.
- [2] D. André, Solution directe du problème résolu par M. Bertrand, Comptes Rendus de l'Académie des Sciences, Paris, 105 (1887) 436–437.
- [3] J.-C. Aval, Multivariate Fuss-Catalan numbers, Discrete Math. 308 (2008) 4660–4669.
- [4] E. Barbier, Note: Calcul des probabilités. Generalisation du problème résolu par M. J. Bertrand, Co. Re. Acad. Sci. Paris, 105 (1887) 407.

- [5] J. Bertrand, Solution d'un problème, *Comptes Rendus de l'Académie des Sciences, Paris*, 105 (1887) 369.
- [6] W. Böhm, Lattice path counting and the theory of queues, *Journal of Statistical Planning and Inference*, 140 (2010) 2168–2183.
- [7] H.W. Gould, Some generalizations of Vandermonde's convolution, *Amer. Math. Month.*, 63 (1956) 84–91.
- [8] I.P. Goulden and L.G. Serrano, Maintaining the spirit of the reflection principle when the boundary has arbitrary integer slope, *J. Combin. Theory A*, 104 (2003) 317–326.
- [9] R.L. Graham, D.E. Knuth, and O. Patashnik, *Concrete Mathematics: A Foundation for Computer Science*, 2nd ed., Reading, MA: Addison-Wesley, 1994.
- [10] K. Humphreys, A history and a survey of lattice path enumeration, *Journal of Statistical Planning and Inference*, 140 (2010) 2237–2254.
- [11] D.E. Knuth, *The Art of Computer Programming, Vol. 1: Fundamental Algorithms, Third Edition*, Reading, Massachusetts: Addison-Wesley, 1997.
- [12] V.S. Koroljuk, On the discrepancy of empiric distributions for the case of two independent samples. *Izvestiya Akad. Nauk SSSR. Ser. Mat.* 19 (1955) 81–96. (Selected Translations in Mathematical Statistics and Probability, 4:105–121, Institute of Mathematical Statistics, American Mathematical Society, 1963.)
- [13] C. Krattenthaler and S.G. Mohanty, Lattice path combinatorics – applications to probability and statistics, 14 pages; written for the *Encyclopedia of Statistical Sciences, Second Edition*, Samuel Kotz, Norman L. Johnson, Campbell B. Read, N. Balakrishnan, Brani Vidakovic, eds., Wiley, New York, 2003.
- [14] D. Mirimanoff, A propos de l'interprétation géométrique du problème du scrutin, *L'enseignement mathématique*, 23 (1923) 187–189.
- [15] S.G. Mohanty, *Lattice Path Counting and Applications*, Academic Press, New York, 1979.
- [16] H. Niederhausen, Linear recurrences under side conditions, *European J. Combin.* 1 (1980) 353–368.
- [17] H. Niederhausen, Catalan traffic at the beach. *Electron. J. Combin.* 9(1)(2002) #R33.
- [18] M. Renault, Lost (and found) in translation: André's actual method and its application to the generalized ballot problem, *Amer. Math. Monthly*, 115 (2008) 358–363.
- [19] G.-C. Rota, D. Kahaner and A. Odlyzko, On the foundations of combinatorial theory. VIII. Finite operator calculus, *J. Math. Anal. Appl.* 42 (1973) 684–760.
- [20] N.J.A. Sloane, *The On-Line Encyclopedia of Integer Sequences*, published electronically at <http://oeis.org/>.