

Chapter 5

The Chung-Feller Theorem Revisited

Dyck paths are the most investigated objects related to the Catalan number C_n . For other families related to Catalan number, we refer to [25, 29, 45, 44, 74]. Surprisingly, the the number of the set of n -Dyck paths with k flaws is independent of k and equals to Catalan number C_n which is the *Chung-Feller Theorem*. In [19], the famous theorem was first proved by means of analytic method. The theorem was subsequently treated by more combinatorial methods in [51] (using cyclic permutation) and in [24] (using the cycle lemma). Recently, Eu, Fu, and Yeh [30] proved a refinement of this result by virtue of the research of Taylor expansions of generating functions. For more information of Chung-Feller Theorem, we refer to [4, 31, ?, 51, 68, 83].

In this chapter, our purpose is to reprove Chung-Feller Theorem, show three classes related to Chung-Feller Theorem, and study Chung-Feller Theorem for Motzkin paths with flaws and a labelled minimum. In Section 5.1, we reprove two well-known formulas: One is to provide a simple bijection between n -Dyck paths with k flaws and n -Dyck paths with $k + 1$ flaws for $i = 0, 1, \dots, n - 1$ (Theorem 5.1.1) and the other is to explain the Catalan identity appeared in Theorem 4.1.1 by n -Dyck paths with flaws (Theorem 5.1.2).

In Section 5.2, we are interested in studying *bi-color plane forests* (Definition

5.2.1) and obtain two main results. Theorem 5.2.2 provides a formula to count bi-color plane forests with n edges and k components. Theorem 5.2.3 is equivalent to the Chung-Feller Theorem.

In Section 5.3, we catch two results: One is to find the number of *semistandard tableaux* (Definition 5.3.1) of shape $2 \times n$ with k decreasing columns which is independent of k (Theorem 5.3.2), and the other is to find the number of *noncrossing semiordered pairs* (Definition 5.3.3) with n pairs and k d -arcs which only depends on n for $k = 0, 1, 2, \dots, n$ (Theorem 5.3.4).

In Section 5.4, we catch two main results. Theorem 5.3.2 provides a family related to Chung-Feller Theorem for Motzkin number. Theorem 5.3.4 provides two families related to Chung-Feller Theorem for the Riordan number.

5.1 A Simple Proof of Chung-Feller Theorem

In this section, we provide a simple bijection to reprove Chung-Feller Theorem. Furthermore, we explain the Catalan identity appeared in Theorem 4.1.1 by Chung-Feller Theorem.

Theorem 5.1.1 (Chung-Feller) *The number of n -Dyck paths with k flaws is the Catalan number C_n for $k = 0, 1, \dots, n$.*

Proof. Recall that $\mathbb{D}_{n,k}^*$ denote the set of n -Dyck paths with k flaws for $k = 0, 1, \dots, n$. We shall establish a bijection between $\mathbb{D}_{n,k}^*$ and $\mathbb{D}_{n,k+1}^*$ (see Figure 5.1).

On the one hand, for a given path D in $\mathbb{D}_{n,k}^*$, let $D = BuAdC$, where u is the first up step above the x -axis and d is the first down step touching the x -axis after u . It is easy to see B is a path all below the x -axis with k_1 flaws for some $k_1 \geq 0$, A is a path all above the x -axis, and C is the remaining path with $k - k_1$ flaws. Note that A and B may be empty. Switch Bu with Ad to obtain $D' = AdBuC$. Since A , dBu , and C have 0 , $k_1 + 1$, and $k - k_1$ flaws, respectively, then D' is a path in $\mathbb{D}_{n,k+1}^*$.

On the other hand, for a given path D' in $\mathbb{D}_{n,k+1}^*$, let $D' = AdBuC$, where d is the first down step below the x -axis and u is the first up step touching the x -axis after d . It is easy to see A is a path all above the x -axis, dBu is a path all above the x -axis with $k_1 + 1$ flaws for some $k_1 \geq 0$, and C is the remaining path with $k - k_1$ flaws. Switch Ad with Bu to obtain $D = BuAdC$. Since B , uAd , and C have k_1 , 0 , and $k - k_1$ flaws, respectively, then D is a path in $\mathbb{D}_{n,k}^*$. This completes the proof. \square

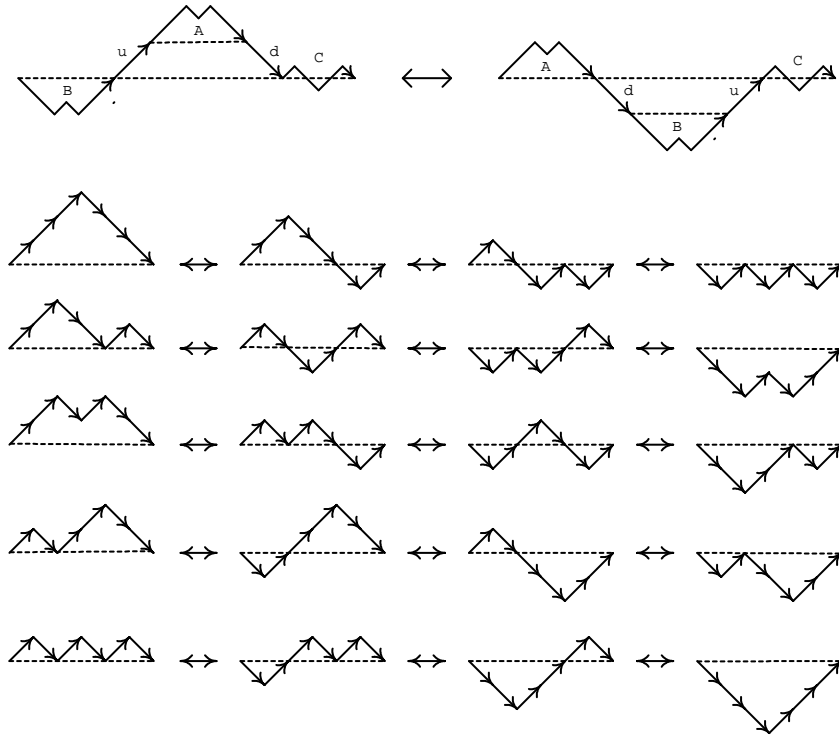


Figure 5.1: 3-Dyck paths with k flaws and their corresponding 3-Dyck paths with $k + 1$ flaws for $k = 0, 1, 2$

Next we reprove the Catalan identity appeared in Theorem 4.1.1. Using Equation (4.0.2) and two different methods to count n -Dyck paths with flaws, we have the following result.

Theorem 5.1.2 *The Catalan number $C_n = \frac{1}{n+1} \sum_{k=1}^n \frac{k2^k}{2n-k} \binom{2n-k}{n}$ for $n \geq 1$.*

Proof. Recall that \mathbb{D}_n^* be the set of n -Dyck paths with flaws. We use two different methods to calculate $|\mathbb{D}_n^*|$ and then derive this identity. One the one hand, by

Chung-Feller Theorem, $|\mathbb{D}_n^*| = (n+1)C_n$.

On the other hand, we evaluate $|\mathbb{D}_n^*|$ according to the number of intersections of n -Dyck paths with flaws with the x -axis except $(0,0)$. Let \mathbb{A}_k be the set of paths in \mathbb{D}_n^* such that each path intersects the x -axis k times except $(0,0)$ for $k = 1, 2, \dots, n$. Then $|\mathbb{D}_n^*| = \sum_{k=1}^n |\mathbb{A}_k|$. It suffices to prove that $|\mathbb{A}_k| = \frac{k2^k}{2n-k} \binom{2n-k}{n}$. Let $\mathbb{A}_{k,0}$ be the subset of \mathbb{A}_k where each path in $\mathbb{A}_{k,0}$ has no flaw. Then, by Equation (4.0.2), $|\mathbb{A}_{k,0}| = \frac{k}{2n-k} \binom{2n-k}{n}$. Denote $D \in \mathbb{A}_{k,0}$ as $D_1 D_2 \dots D_k$ where D_i is a minimal Dyck path with no flaw for $i = 1, 2, \dots, k$. Each path in \mathbb{A}_k can be constructed from some $D \in \mathbb{A}_{k,0}$ by remaining D_i or reflecting D_i to the x -axis. By multiplication principle, we obtain $|\mathbb{A}_k| = 2^k |\mathbb{A}_{k,0}| = \frac{k2^k}{2n-k} \binom{2n-k}{n}$. Hence, we complete the proof. \square

5.2 Bi-color Plane Forests

In this section we wish to study the relation between plane forests and Chung-Feller Theorem. For this purpose, we introduce the following technology.

Definition 5.2.1 *A bi-color plane tree is a plane tree with vertices either black or white such that the vertices of each subtree of the root have the same color. A bi-color plane forest is a plane forest where each component is a bi-color plane tree.*

Let \mathbb{BT}_n denote the set of bi-color plane trees with n edges and $\mathbb{BF}_{n,k}$ denote the set of bi-color plane forests with n edges and k components. Because \mathbb{BT}_n has a bijective correspondence to the set of n -Dyck paths with flaws (see Figure 5.2 for an example in \mathbb{BT}_5), we have $|\mathbb{BF}_{n,1}| = |\mathbb{BT}_n| = \binom{2n}{n}$.

Theorem 5.2.2 $|\mathbb{BF}_{n,k}| = \begin{cases} 2^{2n} \binom{n+\frac{k}{2}-1}{n}, & \text{if } k \text{ is even,} \\ \binom{2n+k-1}{2n} \binom{2n+\frac{k-1}{2}}{n} / \binom{2n+\frac{k-1}{2}}{2n}, & \text{if } k \text{ is odd.} \end{cases}$

Proof. Clearly, the generating function of \mathbb{BT}_n is $t(x) = \sum_{n \geq 0} \binom{2n}{n} x^n$. Hence, the generating function of $\mathbb{BF}_{n,k}$ is $f(x) = [t(x)]^k$. In [73], p. 52, we learned

$$\sum_{n \geq 0} \binom{2n}{n} x^n = (1-4x)^{-\frac{1}{2}}.$$

By algebraic calculation, $[x^n]f(x) = \frac{2^n}{n!} \prod_{i=0}^{n-1} (k+2i)$ and the proof is completed. \square

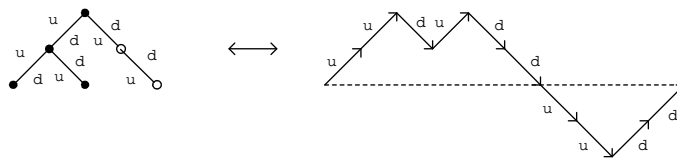


Figure 5.2: A bi-color plane tree with 5 edges and its corresponding 5-Dyck paths with 2 flaws

Let $\mathbb{B}_{n,k}^*$ be the set of bi-color plane forests with n edges and no trivial components such that vertices of each component have same colors, the colors of components are alternating either black or white, and there are k edges in white components for $k = 0, 1, \dots, n$. We shall discover the relation between bi-color plane forests $\mathbb{B}_{n,k}^*$ and the Chung-Feller Theorem.

Theorem 5.2.3 *The number of the bi-color plane forests in $\mathbb{B}_{n,k}^*$ is the Catalan number C_n for $k = 0, 1, \dots, n$.*

Proof. It is well-known that the Catalan number C_n counts n -plane trees. $\mathbb{B}_{n,0}^*$ is simply the set of n -plane trees, i.e. $|\mathbb{B}_{n,0}^*| = C_n$. We shall show that there is a bijection between $\mathbb{B}_{n,k}^*$ and $\mathbb{B}_{n,k+1}^*$ for $k = 0, 1, 2, \dots, n-1$ (see Figure 5.3). Then $|\mathbb{B}_{n,k}^*| = C_n$ for $k = 0, 1, \dots, n$.

Let A, B , and C be the first component, the second component, and the third component, respectively, of a given forest in $\mathbb{B}_{n,k}^*$. Without loss of generality, we may assume that B is a component with black vertices. Note that A may be empty. Partition B into three parts: The first part f is the subtree which is rooted by the first child of the root; the second part g is the root and its remaining subtree; the third part e is the edge adjacent to the root and its first child, and recolor the end points of e white which is denoted by e' . There are three cases:

Case 1: f and g are nontrivial. Then we exchange f and A , and identify the root of A and the first child of the root.

Case 2: f is trivial. Then we identify the root of A and the first child of the root.

Case 3: g is trivial. First, we exchange f and A , and identify the root of A and the first child of the root which is denoted by W . Secondly, we identify the root of

W and the root of C .

In any case, our construction will yield a forest in $\mathbb{B}_{n,k+1}^*$.

The converse holds if only assume that the above B is a component with white vertices for a given forest in $\mathbb{B}_{n,k+1}^*$ and proceed similar steps as the above. \square

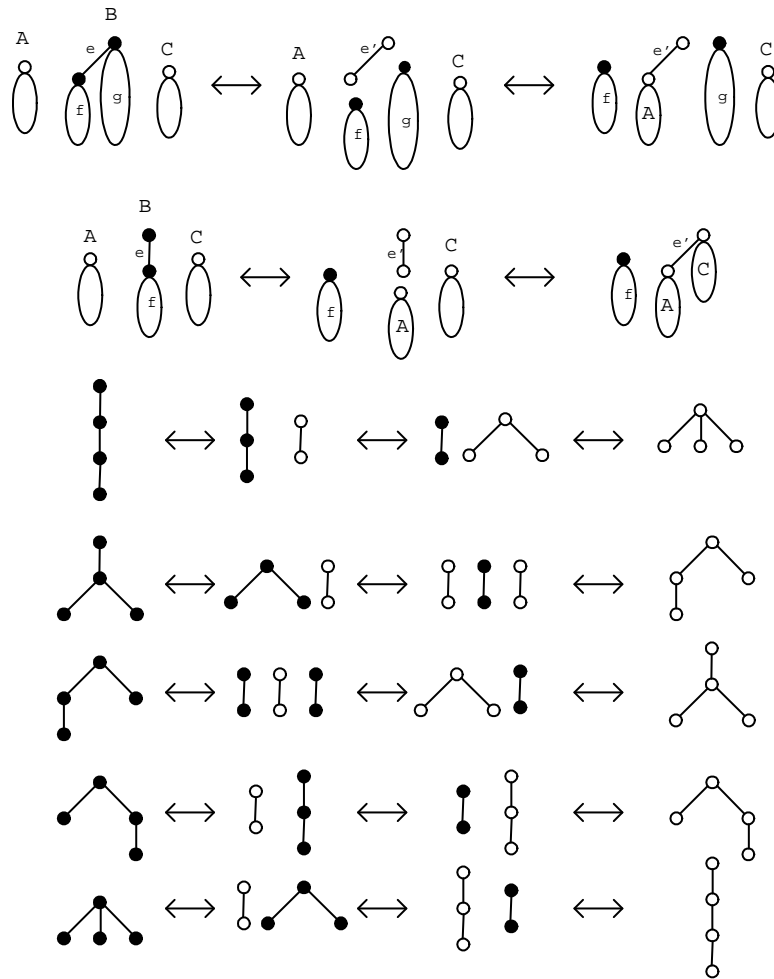


Figure 5.3: An illustration of the proof of Theorem 5.2.3.

Theorem 5.2.3 immediately yields Corollary 5.2.4.

Corollary 5.2.4 *The number of plane forests in \mathbb{F}_n^* , where there are k edges in all odd components or in all even components, equals to the Catalan number C_n , if $n \neq 2k$; $\frac{1}{2}C_n$, if $n = 2k$, for $k = 1, 2, \dots, n - 1$.*

5.3 Semistandard Tableaux and Noncrossing Semiordered Pairs

Recall that given a partition λ of n , a *standard tableau* is an arrangement of $[n]$ in the cells of the Ferrers diagram of λ which increases across rows and down columns. The Catalan number C_n counts standard tableaux in a $2 \times n$ rectangular Ferrers diagram (see [75]).

Definition 5.3.1 *Given a partition λ of n , a semistandard tableau is an arrangement of $[n]$ in the cells of the Ferrers diagram of λ which increases only across rows.*

We shall show that the number of semistandard tableaux of shape $2 \times n$ with k decreasing columns only depends on n .

Theorem 5.3.2 *The number of semistandard tableaux of shape $2 \times n$ with k decreasing columns is the Catalan number C_n for $k = 0, 1, 2, \dots, n$.*

Proof. Let $\mathbb{S}_{n,k}$ be the set of semistandard tableaux of shape $2 \times n$ with k decreasing columns for $k = 0, 1, 2, \dots, n$. The bijection ψ will be from $\mathbb{S}_{n,k}$ to $\mathbb{S}_{n,k+1}$ for $k = 0, 1, 2, \dots, n-1$. Let A be a semistandard in $\mathbb{S}_{n,k}$. Then $B = \psi(A)$ is obtained by the following steps.

1. Partition A into three parts: f , g , and h , where
 - (a) f contains a maximal tableau with decreasing columns, and the number x in the first row of the first increasing column,
 - (b) g contains the number y in the second row of the first increasing column, and a minimal tableau m (may be empty) with increasing columns such that each number in g is consecutively and the number in the first row and the first column of m is less than y , and
 - (c) h is the remaining tableau.
2. Switch f and g .

3. Replace s in g with $s - |f|$, where $|f|$ is the amount of numbers appeared in f .
4. Replace s in f with $s + |g|$, where $|g|$ is the amount of numbers appeared in g .
5. Let g' be a tableau obtained by removing the first row one position toward left in new g . Let f' be a tableau obtained by removing the second row one position toward right in new f .
6. Let B be the tableau composed of g' , f' , and h .

Clearly,

1. The set of numbers in g' is $[|g|]$ with increasing columns.
2. The set of numbers in f' is $\{|g|+1, |g|+2, \dots, |g|+|f|\}$ with decreasing columns.
3. The column composed of g' and f' is decreasing.

Hence $B \in \mathbb{S}_{n,k+1}$. In Figure 5.4 below, we have filled the numbers to help the reader trace what happens.

Conversely, $A = \phi(B)$ is obtained by the following steps.

1. Partition B into three parts: g' , f' , and h , where
 - (a) g' contains a maximal tableau with increasing columns, and the number y in the second row of the first decreasing column,
 - (b) f' contains the number x in the first row of the first decreasing column, and a minimal tableau m' (may be empty) with decreasing columns such that each number is consecutively and the number in the second row and the first column of m' is less than x , and
 - (c) h is the remaining tableau.
2. Remove the first row one position toward right in g' and remove the second row one position toward left in f' , respectively.

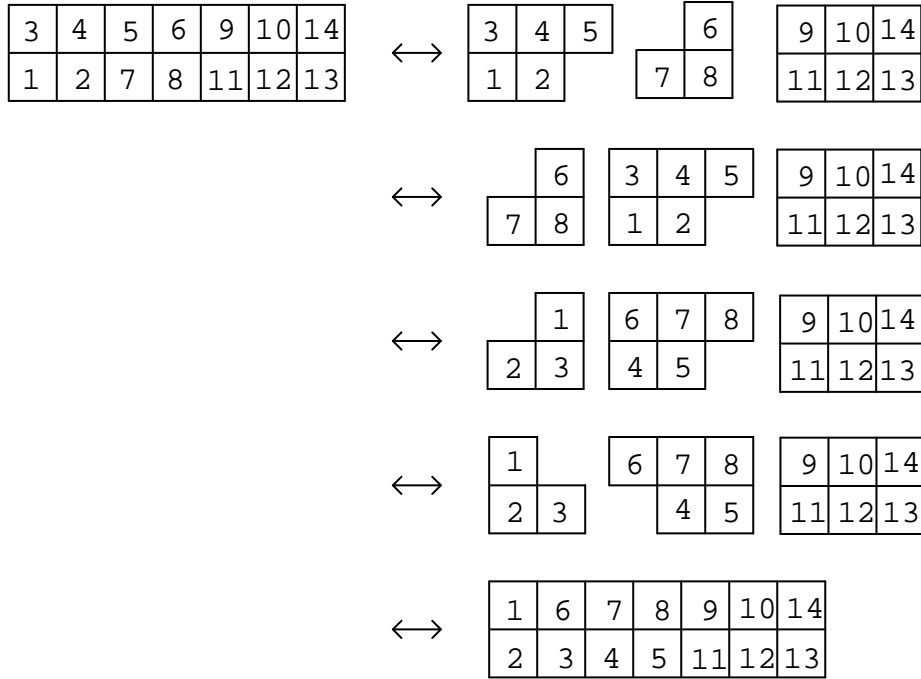


Figure 5.4: An illustration for the proof of Theorem 5.3.2

3. Replace s in g' with $s + |f'|$, where $|f'|$ is the amount of numbers appeared in f' . Denote this tableau by g .
4. Replace s in f' with $s - |g'|$, where $|g'|$ is the amount of numbers appeared in g' . Denote this tableau by f .
5. Let A be the tableau composed of f, g , and h .

Clearly,

1. The set of numbers in f is $[|f'|]$ with decreasing columns.
2. The set of numbers in g is $\{|f'| + 1, |f'| + 2, \dots, |f'| + |g'|\}$ with increasing columns.
3. The column composed of f and g is increasing.

Hence $A \in \mathbb{S}_{n,k}$. Obviously, for each $A \in \mathbb{S}_{n,k}$, $(\phi \circ \psi)(A) = A$ and for each $B \in \mathbb{S}_{n,k+1}$, $(\psi \circ \phi)(B) = B$. This implies ψ is a bijection. \square

Figure 5.5 is an example of semistandard tableaux with k decreasing columns and their corresponding semistandard tableaux with $k + 1$ decreasing columns for $k = 0, 1, 2$.

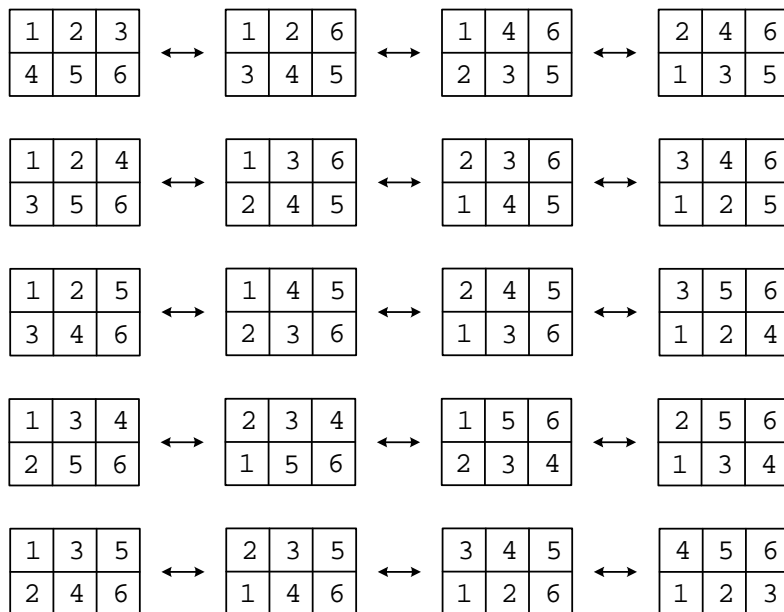


Figure 5.5: Semistandard tableaux of shape 2×3 .

Recall that a *noncrossing pair* is an arrangement of $2n$ points on a circle where we join them pairwise with n noncrossing chords. The number of noncrossing pairs of $2n$ points on a circle is the Catalan number C_n (see [75]).

Definition 5.3.3 *A noncrossing semiordered pair is an arrangement of $2n$ points on a circle with n noncrossing arcs satisfying the following conditions: The arcs of $v_i v_j$ and $v_m v_n$ has the same direction, if $i < m < n < j$, where v_k is the k^{th} point clockwise for $k = 1, 2, \dots, 2n$. We define $\overrightarrow{v_i v_j}$ to be an i -arc if $i < j$; a d -arc if $i > j$.*

In what follows, we shall show that the number of noncrossing semi-ordered pairs with n pairs and k d -arcs only depends on n .

Theorem 5.3.4 *The number of noncrossing semi-ordered pairs with n pairs and k d -arcs is the Catalan number C_n for $k = 0, 1, 2, \dots, n$.*

Proof. Let $\mathbb{P}_{n,k}$ be the set of noncrossing semi-ordered pairs with n pairs and k d -arcs. The bijection ψ will be from $\mathbb{P}_{n,k}$ to $\mathbb{P}_{n,k+1}$ for $k = 0, 1, 2, \dots, n-1$. Let P be a noncrossing semi-ordered pair in $\mathbb{P}_{n,k}$. Then $P' = \psi(P)$ is obtained by the following steps. Figure 5.6 is an example for illustrating this bijection.

Denote P as $BuAdC$ clockwise, where \overrightarrow{ud} is the first i -arc of P . It is easy to see B only contains d -arcs, say k_1 d -arcs for some $k_1 \geq 0$, A only contains i -arcs, and C is the remaining part with $k - k_1$ d -arcs. Note that A and B may be empty. Switch Bu with Ad to obtain $P' = AdBuC$. Obviously, \overrightarrow{ud} is the first d -arc of P' . Since A , dBu , and C have 0 , $k_1 + 1$, and $k - k_1$ d -arcs, respectively, then P' is a noncrossing semi-ordered pair in $\mathbb{P}_{n,k+1}$.

Conversely, for $P' \in \mathbb{P}_{n,k+1}$, then $P = \phi(P')$ is obtained by the following steps. Denote P' as $AdBuC$ clockwise, where \overrightarrow{ud} is the first d -arc of P' . It is easy to see A only contains i -arcs, dBu only contains d -arcs, say $k_1 + 1$ d -arcs for some $k_1 \geq 0$, and C is the remaining part with $k - k_1$ d -arcs. Switch Ad with Bu to obtain $P = BuAdC$. Obviously, \overrightarrow{ud} is the first i -arc of P . Since B , uAd , and C have k_1 , 0 , and $k - k_1$ d -arcs, respectively, then P is a noncrossing semi-ordered pair in $\mathbb{P}_{n,k}$. Obviously, for each $P \in \mathbb{P}_{n,k}$, $(\phi \circ \psi)(P) = P$ and for each $P' \in \mathbb{P}_{n,k+1}$, $(\psi \circ \phi)(P') = P'$. This implies ψ is a bijection. \square

5.4 Motzkin Paths with Flaws and a Labelled Minimum

In this section, we wish to introduce Chung-Feller Theorem for Motzkin number and Riordan number. For convenience, let $\mathbb{M}_n^{(k)}$ be the set of paths from $(0, k)$ to $(n, 0)$ allowing rise step $u = (1, 1)$, fall step $d = (1, -1)$, and level step $l = (1, 0)$. Let $\mathbb{R}_n^{(k)}$ be the set of paths with n steps starting from $(0, k)$, ending on x -axis and allowing rise steps (i, i) for $i \geq 1$ and fall step $(1, -1)$. In particular, $\mathbb{M}_n^{(0)}$ is the set of n -Motzkin paths with flaws.

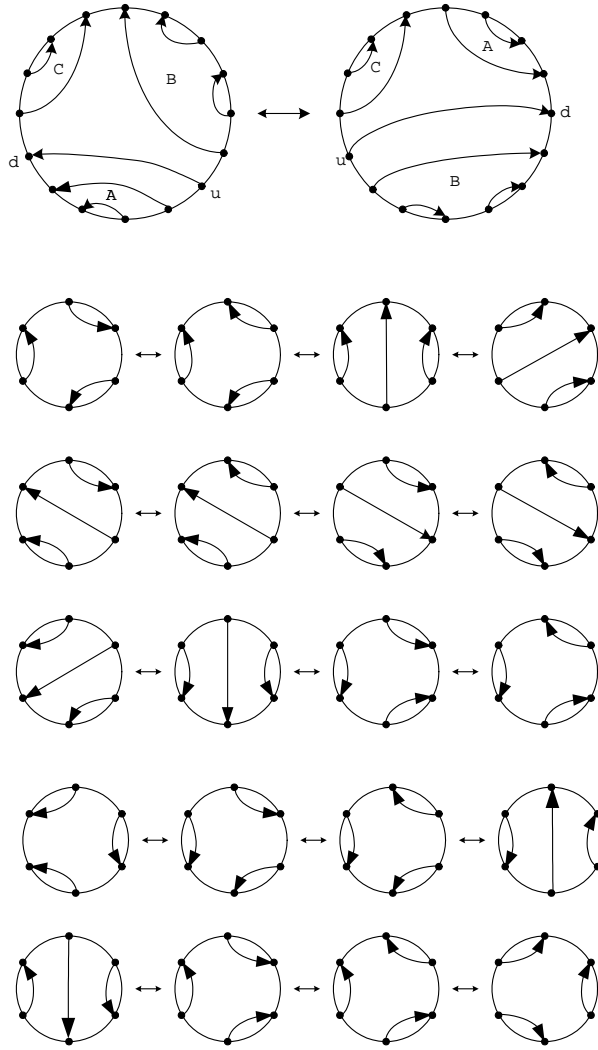


Figure 5.6: An illustration for the proof of Theorem 5.3.4 .

Theorem 5.4.1 *The number of paths in $\mathbb{M}_n^{(0)}$ with a labelled minimum after k steps for $k = 0, 1, 2, \dots, n$ is the Motzkin number M_n .*

Proof. Let \mathbb{S}_k be the subset of $\mathbb{M}_n^{(0)}$ with a labelled minimum after k steps for $k = 0, 1, \dots, n$. We shall provide a bijection between \mathbb{S}_k and \mathbb{S}_{k+1} for $k = 1, 2, \dots, n$. Figure 5.7 is an example for illustrating this bijection.

The bijection ψ will be from \mathbb{S}_k to \mathbb{S}_{k+1} . Let D be a path in \mathbb{S}_k . Then a path $D' = \psi(D)$ is obtained by the following step. We write $D = Ap$ and let $D' = pA$ where p is the last step of D . It is easy to check that D' is a path in \mathbb{S}_{k+1} and ψ is a bijection. Hence $|\mathbb{S}_k| = |\mathbb{S}_0|$ for $k = 0, 1, \dots, n$. By the definition of \mathbb{S}_0 , we obtain

$|\mathbb{S}_0| = M_n$. Hence, $|\mathbb{S}_k| = M_n$ for $k = 0, 1, 2, \dots, n$. □

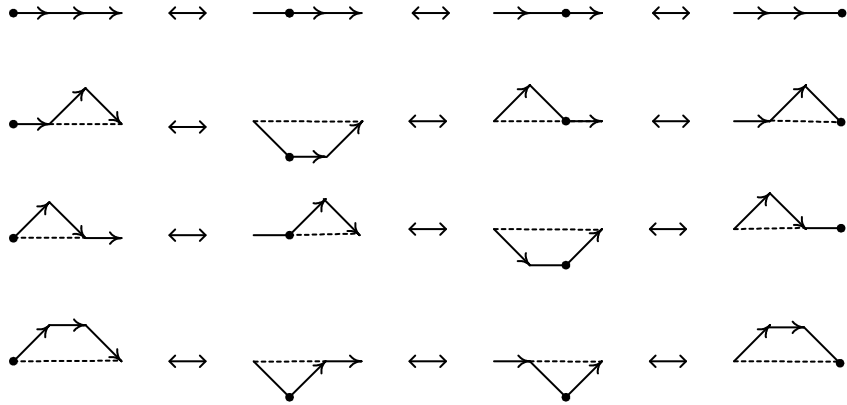


Figure 5.7: An illustration for the proof of Theorem 5.4.1

If we change the labelled minimum into the step u or the step d , then we obtain the following results. Figure 5.8 is an illustration for these correspondences.

- Corollary 5.4.2**
1. The number of paths in $\mathbb{M}_{n+1}^{(-1)}$ with a rightmost minimum after k steps for $k = 0, 1, 2, \dots, n$ is the Motzkin number M_n .
 2. The number of paths in $\mathbb{M}_{n+1}^{(1)}$ with a leftmost minimum after k steps for $k = 1, 2, \dots, n + 1$ is the Motzkin number M_n .

For Corollary 5.4.2-1, Shapiro [68] noted that it can be proven either by a version of the cycle lemma (see [74], p. 67 or [83]) or by generating functions. In [30], Eu-Fu-Yeh proved a refinement of this result.

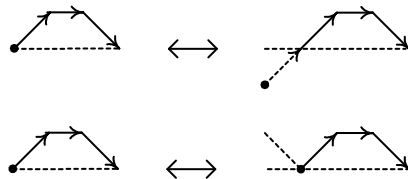


Figure 5.8: An illustration for Corollary 5.4.2

The following is an application of Corollary 5.4.2-1 in probability theory. Suppose there is a game of rolling one die. The player must pay one coin to the banker for entry fee. Each turn, if the player gets a number larger than the banker, then he

wins one coin from the banker. If the player gets a number smaller than the banker, then he loses one coin to the banker. If there is a tie, then he neither wins nor loses any coin. Let S_{n+1} be the number of coins the player wins after $n + 1$ turns containing one coin he pays to the banker for the entry fee. Clearly, $S_0 = -1$. Let B_k be the event that the player's last minimum occurs at the k^{th} turn for $k = 0, 1, \dots, n$. By Corollary 5.4.2-1, we have the following result in conditional probability.

Corollary 5.4.3 *Let S_{n+1} and B_k be defined as above. Then $|B_k| = M_n$ and the conditional probability $P(B_k | S_{n+1} = 0) = \frac{1}{n+1}$ for $k = 0, 1, \dots, n$.*

In what follows, we shall study Chung-Feller Theorem for Riordan number. Recall that the Riordan number counts n -Motzkin paths without level steps on the x -axis (see [33], p.456) and the short bushes with n edges (see [7], p. 85). Using the technique analogous to Proposition 6.2.1 – (v) in [74], p. 169, and the proof of Theorem 5.4.1, we shall obtain the following results.

Theorem 5.4.4 1. *The number of paths in $\mathbb{M}_n^{(0)}$ with a labelled minimum after k steps which is not an endpoint of a level step for $k = 0, 1, 2, \dots, n$ is the Riordan number R_n .*

2. *The number of paths in $\mathbb{R}_n^{(0)}$ with a labelled minimum after k steps for $k = 0, 1, 2, \dots, n$ is the Riordan number R_n .*

Figure 5.9-(a) is an illustration for Theorem 5.4.4-1 and Figure 5.9-(b) is an illustration for Theorem 5.4.4-2.

Again, if we change the labelled minimum into the step u or the step d , then we obtain the following results.

Corollary 5.4.5 1. *The number of paths in $\mathbb{M}_{n+1}^{(-1)}$ with a rightmost minimum after k steps for $k = 0, 1, 2, \dots, n$ which is not an endpoint of a level step is the Riordan number R_n .*

2. *The number of paths in $\mathbb{M}_{n+1}^{(1)}$ with a leftmost minimum after k steps for $k = 1, 2, \dots, n + 1$ which is not an endpoint of a level step is the Riordan number R_n .*

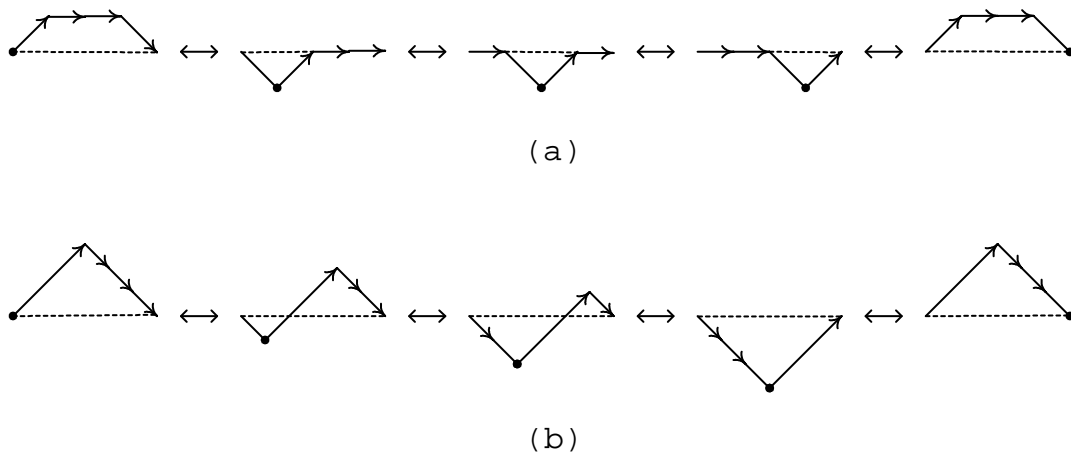


Figure 5.9: An illustration for the proof of Theorem 5.4.4

3. The number of paths in $\mathbb{R}_{n+1}^{(1)}$ with a leftmost minimum after k steps for $k = 1, 2, \dots, n + 1$ is the Riordan number R_n .