Total number of connected bipartite graphs with given Betti numbers and their applications

蓮井, 太朗

出版情報：Kyushu University, 2022, 博士（数理学）, 課程博士
バージョン：
権利関係：

Total number of connected bipartite graphs with given Betti numbers and their applications

Taro HASUI[1]

2023/2/16

[1]Kyushu Univerity Faculty of Mathematics. Email:hasui.taro.206@s.kyushu-u.ac.jp

Contents

Abstract

In the Erdős-Rényi random graph process, the expected of the number of connected components has been computed such that it depends on the increase or decrease of the Betti number. The computation of this expectation requires the total number of connected graphs and the asymptotic behavior of the total number of connected graphs for a given set of edges and vertices. On the other hand, in the Erdős-Rényi random bipartite graph process, the expectation regarding the number of connected components involved in this increase or decrease in the Betti number has not been computed. This is because the total number of bipartite connected graphs and their asymptotic behavior are not known. Therefore, to solve the problem of the total number of connected graphs of this bipartite Erdős-Rényi random graph process and their asymptotic behavior, this paper derives the number of connected bipartite graphs for a given number of edges and vertices and a given Betti number. For this reason, we can obtain a first-order linear partial differential equation satisfied by the exponential generating function of two variables. By solving this first-order linear partial differential equation inductively, we obtain the explicit form of the generating function and derive the asymptotic behavior of its coefficients. The cases of Betti numbers 1 and 2 were mainly discussed, but the same method can be used to derive Betti numbers 3 and beyond. We also introduce a set of basic graphs for classifying connected bipartite graphs and give another representation of the generating function as the sum over basic graphs of rational functions over the number of labeled bipartite rooted trees. As an application, we also discuss its relationship with Cuckoo hashing, a recently established data search algorithm.

1 Introduction

1.1 Previous work

In this section, we discuss the basics of graph theory and prior research.

Definition 1.1.1 (Graph)**.** For a set V, the structure $G = G(V, E)$ determined from its binary subset family $E \subset V \times V$ is called **graph**. The set V is called **vertex set** and the set E is called **edge set**. A graph G' taken from a graph G by selecting an arbitrary vertex and an arbitrary edge is called **subgraph of** G. Furthermore, a graph whose edges are oriented is called **directed graph** and a graph whose edges are not oriented is called **undirected graph**. A simple graph with edges at all vertices is called **complete graph**.

Definition 1.1.2 (Path, Cycle, Loop)**.** A sequence of connected edges is called **path**. A path whose start and end points are the same is called **cycle**. An edge with the same pair of vertices is called **multiple edge** and is called **loop** if both endpoints of an edge are equal.

A graph with no multiple edges and no loops is called a **simple graph**. In this paper, we use the term "graph" to mean a simple, undirected graph. Let us examine how well the vertices of a simple undirected graph G are connected to each other by edges. For this purpose, we introduce the following **connected graph**.

Definition 1.1.3 (Connected graph)**.** The existence of a path between two vertices is called a **connection**, and a graph whose any two vertices are connected is called a **connected graph**.

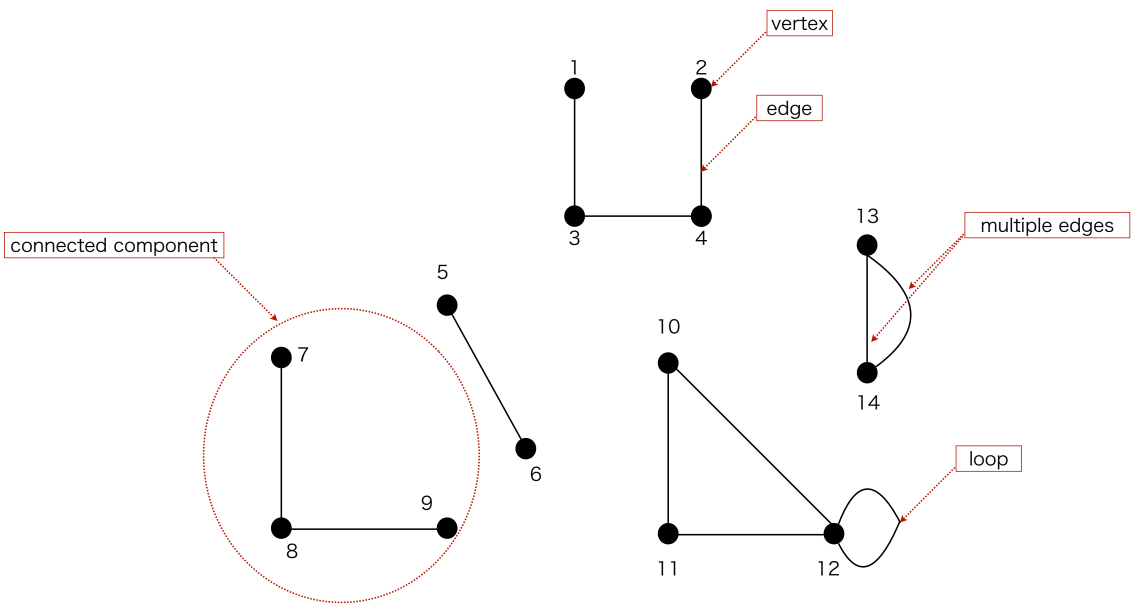


Figure 1: Disconnected and unsimple graph.

We apply the feature of connection not only to graphs, but also to **components of a graph**, i.e., the subset of vertices and edges that make up a graph. Furthermore, the components that are connected in a graph are called **connected components**. For example, in Figure 1, The vertex set is $V = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14\}$, the edge set is $E = \{1-3, 3-4, 4-2, 5-6, 7-8, 8-9, 10-11, 11-12, 10-12, 12-12, 13-14, 13-14\}$. Furthermore, it has $\{7, 8, 9\}$ as a connected component as shown in Figure 1. Obviously, there are other connected components, such as $\{5, 6\}$ and $\{13, 14\}$. A component with only one vertex is also considered

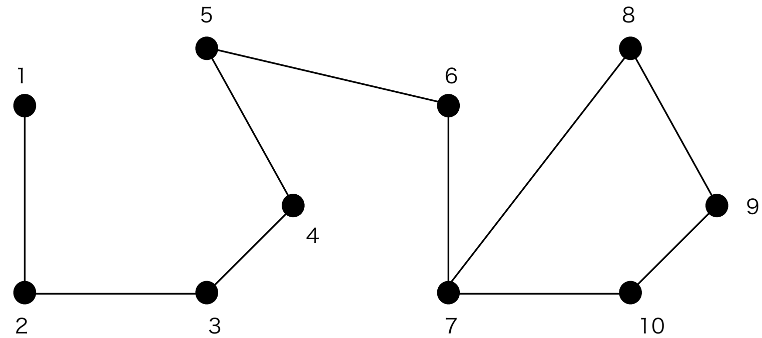


Figure 2: Connected simple graph.

a connected component. As an example, in Figure 1, The graph has five connected components but is a disconnected graph. On the other hand, in Figure 2, The graph is connected graph. It is obvious that there is only one connected component of Figure 2. It is important to emphasize that there are graphs that are disconnected but connected in the components that make up the graph.

Then, we will classify connected graphs or connected components into several types.

Definition 1.1.4 (Trees, Cycles, and Complexes of simple graphs)**.** We consider $|E| - |V| = k$ in a simple graph $G = (V, E)$. A graph with $k = -1$ is called **tree**. A graph with $k = 0$ is called **cycle**. Also, when $k \geq 1$, it is called $k-$**complex**.

In this paper, a cycle is also referred to as an **unicycle** to emphasize the fact that there is only one cycle. In some literature, it is also referred to as "unicyclic" instead of "unicycle". Again, Definition 1.1.4 can be applied to connected components. In particular, a tree consisting of all the vertices of a graph and a subset of the edges that make up the graph is called a **spanning tree**. We define (n, q)**-graph** to be the graph $G(V, E)$ such that $|V| = n, |E| = q$. Let $f(n, q)$ denote the number of connected (n, q)-graphs.

Suppose that there exists more than one spanning tree on the complete graph K_n. The total number of spanning trees of a graph is known by [Cay89] as follows.

Theorem 1.1.1 (Cayley's formula)**.** *For the number of spanning trees on K_n, we have*

$$f(n, n-1) = n^{n-2}.$$

For example, The graph in Figure 3(a) shows a spanning tree of a graph with $|V| = 9$ vertices. On the other hand, The graph in Figure 3(b) has a cycle formed by the red edges, which deviates from definition of a spanning tree as having no cycle. There are obviously other trees besides Figure 3, and from Theorem 1.1.1, the total number of trees in K_9 is 9^7.

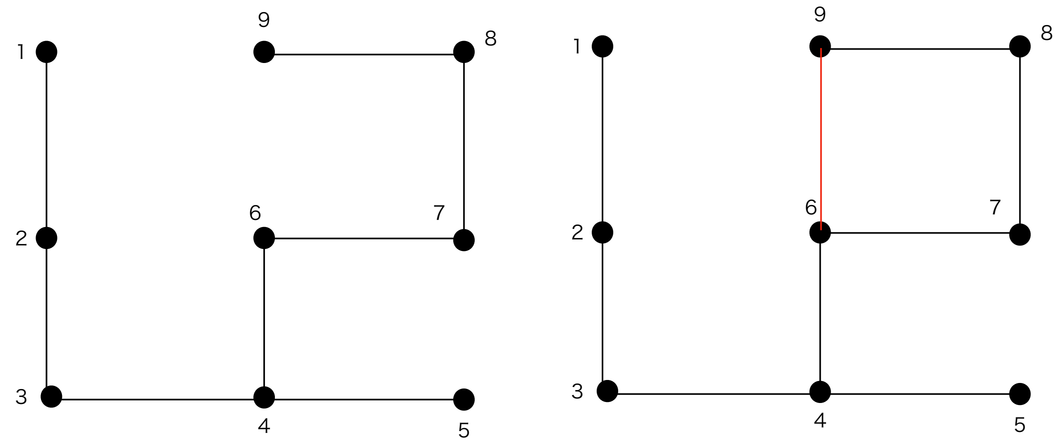
On the other hand, in the case $|V| = n, |E| = n$, i.e., the total number $f(n, n)$ of connected (n, n)-graphs is known by [Rén59] as follows.

Theorem 1.1.2. *In the total number $f(n, n)$ of cycle graphs on the n-vertex complete graph K_n, we have*

$$f(n, n) = \frac{1}{2}\left(\frac{h(n)}{n} - n^{n-2}(n-1)\right) \sim \sqrt{\frac{\pi}{8}}n^{n-1/2} \quad (n \to \infty),$$

where,

$$h(n) = \sum_{s=1}^{n-1} \binom{n}{s} s^s (n-s)^{n-s}.$$



(a) spanning tree **(b)** not spanning tree

Figure 3: An example of a spanning tree with $|V| = 9$.

We then define **bipartite graph** formed by partitioning the vertex set of a simple graph into two parts.

Definition 1.1.5 (Bipartite graph)**.** The vertex set V can be partitioned into mutually disjoint sets V_1, V_2, and if the edge set E is $E \subseteq \{(e, e') \mid e \in V_1, e' \in V_2\}$ then, this graph is called **bipartite graph** and is written as $BG(V_1, V_2, E)$. Also, a bipartite graph whose edges exist from each vertex of V_1 to all vertices of V_2 is called a **complete bipartite graph**, and is denoted by $K_{r,s}$ when $|V_1| = r$ and $|V_2| = s$.

Definition 1.1.6. Let $k \in \mathbb{Z}_{\geq 0}$ and the number of edges $|E| = |V_1| + |V_2| - 1 + k$ of a simple bipartite graph $BG(V_1, V_2, E)$, where a graph with $k = 0$ is called a **tree** and a graph with $k \geq 1$ is called a $k-$**cycle**.

Note that k is the number of cycles in a bipartite graph. This paper mainly discusses this connected bipartite graph. Obviously, every simple graph tree and every spanning tree is a bipartite graph. In this paper, a bipartite graph $BG(V_1, V_2, E)$ is called a (r, s, q)-graph if $|V_1| = r, |V_2| = s, |E| = q$. Let $f(r, s, q)$ denote the total number of connected (r, s, q)-graphs. According to [Sco62], the following holds.

Theorem 1.1.3 (Total number of spanning trees in the bipartite graph)**.** *For the number of spanning trees on $K_{r,s}$ of $|V_1| = r, |V_2| = s$, we have*

$$f(r, s, r + s - 1) = r^{s-1}s^{r-1}.$$

The total number of spanning trees of a bipartite graph can be computed as well as the total number of spanning trees of a simple graph. In Figure 4, an example of $(6, 6, 10)$-graph is given. There is no edge inside the two vertex sets V_1, V_2. Moreover, although Definition 1.1.5 is a spanning tree, the total number of such trees is the same as the total number of spanning trees in the bipartite graph, from Theorem 1.1.3 $f(6, 6, 10) = 6^{6-1} \cdot 6^{6-1}$

1.2 Main theorem

Let $F_k(x, y)$ be the exponential type generating function of $f(r, s, r+s-1+k)$. That is, $F_k(x, y)$ is a function of

$$F_k(x, y) := \sum_{r,s=0}^{\infty} \frac{f(r, s, r + s - 1 + k)}{r!s!} x^r y^s.$$

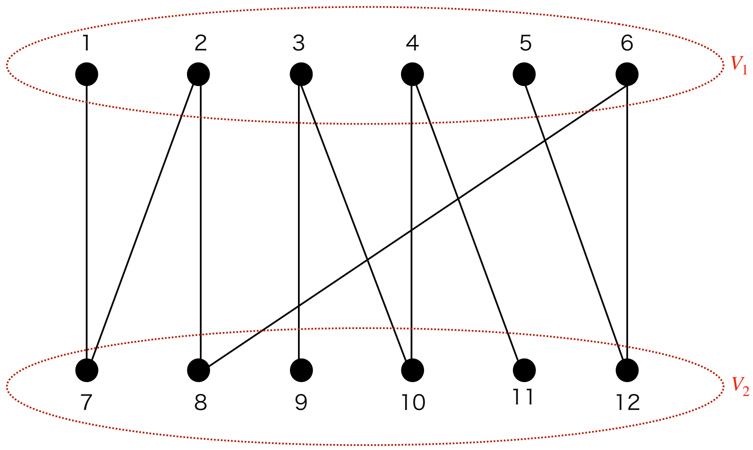


Figure 4: Bipartite connected graph of $|V_1| = |V_2| = 6$.

Since we are dealing with connected graphs, we note that k corresponds to the Betti number, the rank of the first homology group, of each $(r, s, r+s-1+k)$-graph. Note that $k-1$ is often called **excess** since such a connected graph has $k-1$ more edges than vertices. The excess is defined in Definition (6.2.3) below. Also, let $T := F_0(x, y)$. That is, for $k = 0$, from Theorem 1.1.3, we have

$$\begin{aligned} T(x,y) &:= F_0(x,y) \\ &= x + y + \sum_{r,s=1}^{\infty} \frac{r^{s-1}s^{r-1}}{r!s!} x^r y^s. \end{aligned} \tag{1.2.1}$$

Let us further set up the notation. Let D_x, D_y be the Eulerian derivatives of x, y, respectively. That is, D_x, D_y are the Euler derivatives of

$$\begin{aligned} D_x &:= x\frac{\partial}{\partial x}, \\ D_y &:= y\frac{\partial}{\partial y}. \end{aligned}$$

In addition, the notations for D_x, D_y are set as follows.

$$\begin{aligned} T_x &:= D_x T \\ T_y &:= D_y T \\ Z &:= T_x + T_y \\ W &:= T_x T_y \end{aligned}$$

In this case, the following holds.

Proposition 1.2.1. *The function $F_1(x, y)$ is expressed as $F_1 = f_1(W)$ with $f_1(w) = -\frac{1}{2}\big(\log(1-w) + w\big)$, i.e.,*

$$F_1(x,y) = -\frac{1}{2}\left(\log\left(1 - T_x T_y\right) + T_x T_y\right).$$

The following also holds for $F_2(x, y)$.

Proposition 1.2.2. *The function $F_2(x,y)$ is expressed as $F_2 = f_2(Z,W)$ with*

$$f_2(z,w) = \frac{w^2}{24(1-w)^3}\{(2+3w)z + 2w(6-w)\}. \tag{1.2.2}$$

Similarly, the coefficients of $F_1(x,x), F_2(x,x)$ are important for Proposition 1.2.1 and Proposition 1.2.2. For example, when $k=1$, $r+s=n$, Proposition 1.2.1 can be easily computed to be

$$\begin{aligned} F_1(x,x) &= \sum_{r,s=0}^{\infty} \frac{f(r,s,r+s)}{r!s!} x^n \\ &= \sum_{n=0}^{\infty} \left(\sum_{r+s=n} \frac{f(r,s,r+s)}{r!s!} \right) x^n. \end{aligned} \tag{1.2.3}$$

The coefficient of x^n in Equation (1.2.3) means that the number of connected components with Betti number corresponds to k. Also, the exponential generating function $A(x) = \sum_{n=0}^{\infty} a_n \frac{x^n}{n!}$ where the notation $\langle \cdot \rangle$ is defined as

$$\langle x^n \rangle A(x) := a_n.$$

The coefficient $\langle x^n \rangle F_k(x,x)$ is the total number of connected bipartite graphs of Betti number k with n vertices. Or, from $r+s=n$, it means to count the total number of connected $(r,s,n-1+k)$-graphs. Thus, when $k=0$, the total number of $(r,s,n-1+k)$-graphs that are connected is

$$F_0(x,x) = 2\left(x + \sum_{n=2}^{\infty} n^{n-2} \frac{x^n}{n!} \right). \tag{1.2.4}$$

Note that in Equation (1.2.4), the total number of spanning trees n^{n-2} on the complete graph, mentioned in the Theorem 1.1.1, appears in the coefficients. The correspondence between the spanning tree on the n-vertex complete bipartite graph and the spanning tree on K_n appears in the factor 2 on the right side of Equation (1.2.4). The details are given in Lemma 4.1.1. From Proposition 1.2.1 and Proposition 1.2.2, the following holds for the asymptotic behavior of the coefficients of $F_1(x,x)$ and $F_2(x,x)$.

Theorem 1.2.1. *For $n = 4, 5, \cdots$,*

$$\langle x^n \rangle F_1(x,x) = n^{n-1} \sum_{2 \le k \le n/2} \frac{n!}{(n-2k)! n^{2k}} \sim \sqrt{\frac{\pi}{8}} n^{n-1/2} \quad (n \to \infty).$$

Comparing with the previous work, Theorem 1.1.2, we see that the main terms of the asymptotic behavior of the total number of cycle graphs and the total number of bipartite cycle graphs coincide. We see that the main term of the asymptotic behavior of a bipartite graph of an $r+s=n$-vertex with two cycles coincides with the main term of the asymptotic behavior of a bipartite graph of an n-vertex with two cycles. They differ for finite n. This is clear from Table 1. Note that $u_n := \langle x^n \rangle F_1(x,x)$. Then, we give the reason why $f(n,n) > u_n$ for finite n. This is due to the way unicycle graphs are constructed. As mentioned in Definition 1.1.4, a spanning tree was a tree without cycles using all vertices of a certain simple graph. An unicycle graph on K_n is a simple graph that has one cycle. That is, a unicycle graph on K_n is completed by adding an edge to the spanning tree on K_n. Every spanning tree is a bipartite graph. However, whether or not the unicycle graph of an edge added to the spanning tree is bipartite depends on the way the edge is added. Some unicycle graphs are bipartite and some are not depending on the way edges are added, so the overall result is $f(n,n) > u_n$.

The following also holds for the coefficients of $F_2(x,x)$.

n	3	4	5	6	7	8	9	10	11
$f(n,n)$	1	15	222	3660	68295	1436568	33779340	880107840	25201854045
u_n	0	6	120	2280	46200	1026840	25102224	673706880	19745850960

Table 1: Comparison of $f(n,n)$ and u_n $(n = 3, 4, \cdots, 11)$.

n	4	5	6	7	8	9	10	11
$f(n,n+1)$	6	205	5700	156555	4483360	136368414	4432075200	154060613850
b_n	0	20	960	33600	1111040	37202760	1295884800	47478243120

Table 2: Comparison of $f(n,n)$ and b_n $(n = 4, 5, \cdots, 11)$

Theorem 1.2.2. *For $n \to \infty$,*

$$\langle x^n \rangle F_2(x,x) \sim \frac{5}{48} n^{n+1}.$$

Note that in [Wri77] The main term of the asymptotic behavior of the total number of connected graphs such that there are two cycles on K_n is $\frac{5}{24} n^{n+1}$. It is known that the asymptotic behavior of the $F_2(x,x)$ coefficients is twice as large as the expression in Theorem 1.2.2.

This also differs for finite n, as in the case of $F_1(x,x)$ coefficients. It should be clear from Figure 2. Note that $b_n := \langle x^n \rangle F_2(x,x)$. In principle, $F_k(x,y)$ can be calculated inductively. For general k, when $n \to \infty$, we can expect the following.

Theorem 1.2.3. *For $k \geq 2$, the function $F_k(x,y)$ is expressed as $F_k = f_k(Z,W)$ with*

$$f_k(z,w) = \frac{1}{(1-w)^{3(k-1)}} \sum_{j=0}^{k-1} q_{k,j}(w) z^j.$$

If $k \geq 3$, the form of the generating function becomes complicated and is difficult to express as in Proposition 1.2.2. Theorem 1.2.3 is the generalized form of the generating function. Thus, Theorem 1.2.3 corresponds to an extension of Theorem 1.2.2. In addition, we can conjecture that the generating function will take the following form.

Conjecture 1.2.1 (The value of F_k)**.** *For $k \geq 2$, the function $F_k(x,y)$ is expressed as $F_k = f_k(Z,W)$ with*

$$f_k(z,w) = \frac{w^2}{(1-w)^{3(k-1)}} \sum_{j=0}^{k-1} q_{k,j}(w) z^j$$

holds, where $q_{k,j}$ is a polynomial in w. Furthermore,

$$\langle x^n \rangle F_k(x,x) \sim \frac{1}{2^{k-1}} \rho_{k-1} n^{n+(3(k-1)-1)/2}$$

also holds. In addition,

$$f(n,n+k) \sim \rho_k n^{n+(3k-1)/2} \quad (n \to \infty)$$

and the explicit value of ρ_k is given by [Wri77].

The following results have also been obtained for the expression of $F_k(x,y)$.

Theorem 1.2.4. *Let $k \geq 2$. Then F_k is decomposed into a sum of rational functions of T_x, T_y using the basic graph.*

$$F_k(x, y) = \sum_{\mathcal{B} \in BG_k} J_{\mathcal{B}}(x, y),$$

where,

$$J_{\mathcal{B}}(x, y) = \frac{T_x^{r_{\text{sp}}+a_1+2a_2+b_2+c_2} T_y^{s_{\text{sp}}+2a_1+a_2+b_1+c_2}}{g_{\mathcal{B}} \left(1 - T_x T_y\right)^{a_1+a_2+b_1+b_2+c_1+c_2}}.$$

Note that BG_k is the set of all basic graphs. The $G_{\mathcal{B}}, r_{\text{sp}}, s_{\text{sp}}, a_i, b_j, c_k (i, j, k = 1, 2)$ are information about the basic graph $\mathcal{B}$.

The term "basic graph" first appeared in Theorem 1.2.4. This will be explained in detail later. That is, F_k is a rational function of T_x and T_y for $k \geq 2$.

2 Recurrence equations

In this section, we will introduce the various theorems that are the basis of the asymptotic formulas described in subsection 1.2. Again, $f(r, s, q)$ denotes the total number of connected (r, s, q)-graphs, and all graphs are assumed simple graphs. In addition, we assume that

$$f(r, s, q) = 0 \text{ if } q < r + s - 1 \text{ or } q > rs.$$

For example, $f(1, 0, 0) = f(0, 1, 0) = 1$ and $f(0, 0, -1) = 0$. Let $Q(r, s, q)$ be the total number of connected (r, s, q)-graphs completed by connecting the $(q + 1)$-th edge between the (r_1, s_1, t)- and $(r - r_1, s - s_1, q - t)$-components. This means that the two connected components are connected by the $(q + 1)$-th edge to form a single connected component. The $f(r, s, q)$ is simply counts how many connected graphs can be formed under the condition that (r, s)-vertexes and q-edges are given. The following corollary holds.

Lemma 2.0.1. *For $(r, s) \neq (0, 0)$, $q = -1, 0, 1, \cdots,$*

$$(q + 1) f(r, s, q + 1) = (rs - q) f(r, s, q) + Q(r, s, q), \tag{2.0.1}$$

where,

$$\begin{aligned} Q(r, s, q) =& \frac{1}{2} \sum_{r_1=0}^{r} \sum_{s_1=0}^{s} \sum_{t=0}^{q} \binom{r}{r_1} \binom{s}{s_1} \left\{ (r - r_1) s_1 + r_1 (s - s_1) \right\} \\ & \times f(r_1, s_1, t) f(r - r_1, s - s_1, q - t). \end{aligned} \tag{2.0.2}$$

Also, $Q(r, s, -1) = 0$.

Proof. Let us give an overview only. Let $G = (V_1, V_2, E)$ be a (r, s, q)-graph. Let us suppose the case where the $q + 1$-th edge is added to G. The following two cases are possible.

1. If G is a connected graph, i.e., one connected component, and we add the $q + 1$th edge to it.

2. If G is a disconnected graph consisting of two connected components, add a $q + 1$th edge between the connected components, and G is a connected graph.

In the case 1, the vertices of the $q+1$-th edge belong to V_1, V_2. On the other hand, in the case 2, $V_j = V_{j,1} \sqcup V_{j,2}$ $(j = 1, 2)$, there are four possible ways to add the $q+1$th edge.
$V_{1,1}$ and $V_{2,1}, V_{1,1}$ and $V_{2,2}, V_{1,2}$ and $V_{2,1}$, or $V_{1,2}$ and $V_{2,2}$. $\square$

The expression (2.0.1) means that the total number $f(r, s, q+1)$ of connected $(r, s, q+1)$-graphs is recounted in the proof in case 1 and case 2. The total number of additional possibilities $rs - q$ on the $(q+1)$-th edge of case 1 appears in the coefficient of $f(r, s, q)$ on the right side of Equation (2.0.1). As for the coefficient $q+1$ on the left-hand side of Equation (2.0.1) We will now discuss the coefficient $q+1$ on the left-hand side of Equation (2.0.1). In the right-hand counting, some connected $(r, s, q+1)$-graphs are counted up in duplicate in the case 1 and case 2. The number of overlaps is $q+1$.

As for Equation (2.0.1), the case $r = 12, s = 10, q = 18$ is noted in Figure 5 as an example. This shows the case 2 in the proof. The blue edge is the edge corresponding to the $(q+1)$-th edge. The graph G consists of connected components H_1, H_2 and 19-th edge is added between H_1 and H_2. The numbers of the vertices of H_1 and H_2 are $\binom{12}{6}, \binom{10}{5}$ and the number of potential 19-th edge is $(12-6)5 + 6(10-5) = 60$.

Also, regard H_1 as a graph with $(6, 5)$ vertices and 9 edges, and take the total number of connected graphs as $f(6, 5, 9)$ and H_2 as $f(6, 5, 9)$, and sum each of them. This is the graphical image of Lemma 2.0.1.

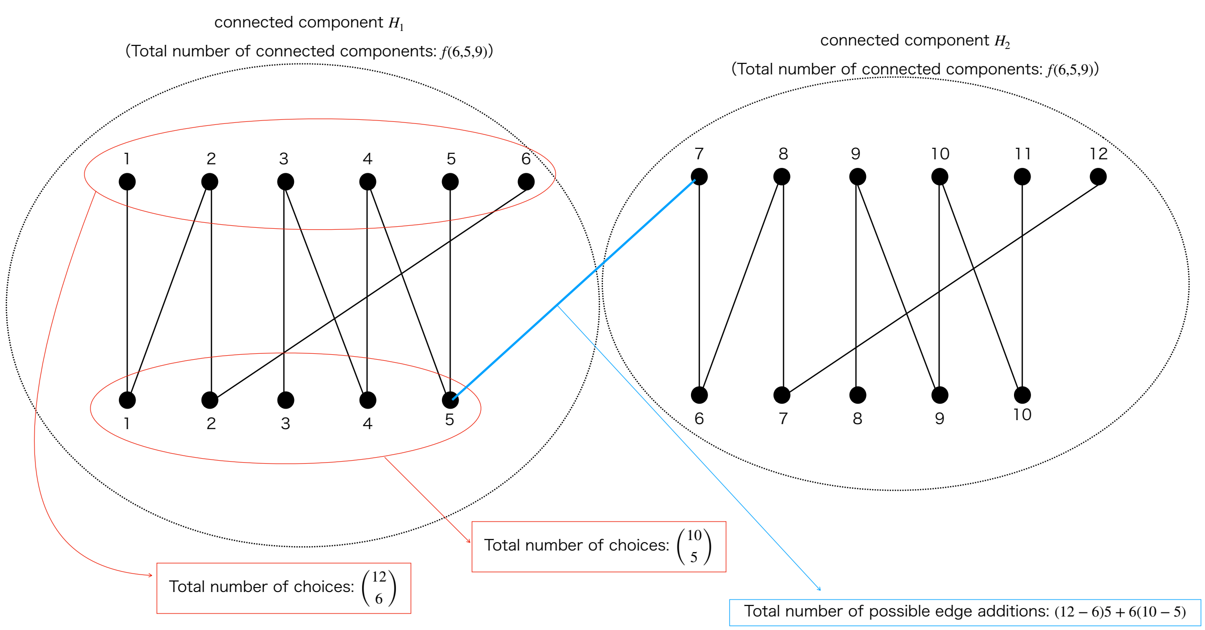


Figure 5: Image of Equation (2.0.1).

Then, we can construct the following linear partial differential Equation using Lemma 2.0.1. For convenience, F_{-1} is also assumed.

Proposition 2.0.1. *Let $D_x := x\partial_x, D_y := y\partial_y$. For $k = -1, 0, 1, 2, \cdots$, we have*

$$\begin{aligned}
&(D_x + D_y + k)\, F_{k+1} \\
&= (D_x D_y - D_x - D_y + 1 - k)\, F_k + \sum_{l=0}^{k+1} D_x F_l \cdot D_y F_{k+1-l}.
\end{aligned}$$

Proof. From Lemma 2.0.1,

$$\begin{aligned}
&(r+s+k)f(r,s,r+s+k)\\
&=(rs-r-s+1-k)f(r,s,r+s-1+k)+Q(r,s,r+s-1+k).
\end{aligned}$$

Furthermore,

$$\begin{aligned}
Q(r,s,r+s-1+k)=&\frac{1}{2}\sum_{r_1=0}^{r}\sum_{s_1=0}^{s}\sum_{t=r_1+s_1-1}^{r_1+s_1+k}\binom{r}{r_1}\binom{s}{s_1}\{(r-r_1)\,s_1+r_1\,(s-s_1)\}\\
&\times f\,(r_1,s_1,t)\,f\,(r-r_1,s-s_1,r+s-1+k-t)\,. \qquad (2.0.3)
\end{aligned}$$

Let us multiply both sides of Equation (2.0.3) by $\frac{x^r y^s}{r!s!}$ and take the infinite sum with respect to r,s. On the left-hand side of Equation (2.0.1),

$$\begin{aligned}
\sum_{r,s=0}^{\infty}(r+s+k)\frac{f(r,s,r+s+k)}{r!s!}x^r y^s=&\sum_{r,s=0}^{\infty}r\frac{f(r,s,r+s+k)}{r!s!}x^r y^s+\sum_{r,s=0}^{\infty}s\frac{f(r,s,r+s+k)}{r!s!}x^r y^s\\
&+\sum_{r,s=0}^{\infty}k\frac{f(r,s,r+s+k)}{r!s!}x^r y^s\\
=&D_xF_{k+1}+D_yF_{k+1}+kF_{k+1}.
\end{aligned}$$

Similarly for the right hand side of Equation (2.0.1), we obtain

$$(D_x+D_y+k)F_{k+1}=(D_xD_y-D_x-D_y+1-k)F_k+\sum_{r,s=0}^{\infty}\frac{Q(r,s,r+s-1+k)}{r!s!}x^r y^s$$

The second term on the right side is important here. Again, both sides of Equation (2.0.2) are multiplied by $\frac{x^r y^s}{r!s!}$ and summed to infinity with respect to r,s.

$$\begin{aligned}
\sum_{r,s=0}^{\infty}\frac{Q(r,s,r+s-1+k)}{r!s!}x^r y^s=&\sum_{r=0}^{\infty}\sum_{s=0}^{\infty}\sum_{r_1=0}^{r}\sum_{s_1=0}^{s}\sum_{t=0}^{r+s-1+k}\binom{r}{r_1}\binom{s}{s_1}\{(r-r_1)s_1+r_1(s-s_1)\}\\
&\times f(r_1,s_1,t)f(r-r_1,s-s_1,r+s-1+k-t)\\
=&\sum_{r=0}^{\infty}\sum_{s=0}^{\infty}\sum_{r_1=0}^{r}\sum_{s_1=0}^{s}\sum_{t=0}^{r+s-1+k}\{(r-r_1)s_1+r_1(s-s_1)\}\\
&\times f(r_1,s_1,t)f(r-r_1,s-s_1,r+s-1+k-t)\\
&\times\frac{x^{r-r_1}y^{s-s_1}}{(r-r_1)!(s-s_1)!}\frac{x^{r_1}y^{s_1}}{r_1!s_1!}.
\end{aligned}$$

Let us examine,

$$\begin{aligned}
&\sum_{r=0}^{\infty}\sum_{s=0}^{\infty}\sum_{r_1=0}^{r}\sum_{s_1=0}^{s}\sum_{t=0}^{r+s-1+k}r_1(s-s_1)f(r_1,s_1,t)f(r-r_1,s-s_1,r+s-1+k-t)\\
&\qquad\times\frac{x^{r-r_1}y^{s-s_1}}{(r-r_1)!(s-s_1)!}\frac{x^{r_1}y^{s_1}}{r_1!s_1!}.
\end{aligned}$$

With respect to the index of the sum, $\sum_{r=0}^{\infty}\sum_{r_1=0}^{r}=\sum_{r_1=0}^{\infty}\sum_{r=r_1}^{\infty}$. Furthermore, consider $f(r_1,s_1,t)f(r-r_1,s-s_1,r+s-1+k-t)$. The graph of $f(r,s,q)$ is a spanning tree when

$q = r + s - 1$. Therefore, $f(r, s, q) = 0$ when $q < r + s - 1$. That is, when $q \geq r + s - 1$, $f(r, s, q) \neq 0$. Thus, if $t \geq r_1 + s_1 - 1$ and $r + s + k - t \geq (r - r_1) + (s - s_1) - 1$, that is $r_1 + s_1 - 1 \leq t \leq r_1 + s_1 + k$, then

$$f(r_1, s_1, t) f(r - r_1, s - s_1, r + s - 1 + k - t) \neq 0.$$

Therefore,

$$\sum_{t=r_1+s_1-1}^{r_1+s_1+k} \sum_{r_1=0}^{\infty} \sum_{s_1=0}^{\infty} r_1 \frac{f(r_s, s_s, t)}{r_1! s_1!} x^{r_1} y^{s_1} \sum_{r=r_1}^{\infty} \sum_{s=s_1}^{\infty} (s - s_1) \frac{f(r - r_1, s - s_1, r + s - 1 + k - t)}{(r - r_1)!(s - s_1)!} x^{r-r_1} y^{s-s_1},$$

where $l := t - (r_1 + s_1 - 1)$, then

$$\sum_{l=0}^{k+1} \sum_{r_1=0}^{\infty} \sum_{s_1=0}^{\infty} r_1 \frac{f(r_s, s_s, t)}{r_1! s_1!} x^{r_1} y^{s_1} \sum_{r=r_1}^{\infty} \sum_{s=s_1}^{\infty} (s - s_1) \frac{f(r - r_1, s - s_1, r - r_1 + s - s_1 + k - l)}{(r - r_1)!(s - s_1)!} x^{r-r_1} y^{s-s_1}$$
$$= \sum_{l=0}^{k+1} \{D_x F_l \cdot D_y F_{k+1-l}\}.$$

Similarly,

$$\sum_{r=0}^{\infty} \sum_{s=0}^{\infty} \sum_{r_1=0}^{r} \sum_{s_1=0}^{s} \sum_{t=0}^{r+s-1+k} s_1 (r - r_1) f(r_1, s_1, t) f(r - r_1, s - s_1, q - t) \frac{x^{r-r_1} y^{s-s_1}}{(r - r_1)!(s - s_1)!} \frac{x^{r_1} y^{s_1}}{r_1! s_1!}$$
$$= \sum_{l=0}^{k+1} \{D_y F_l \cdot D_x F_{k+1-l}\}$$

is also valid. That is,

$$\begin{aligned} (D_x + D_y + k) F_{k+1} &= (D_x D_y - D_x - D_y + 1 - k) F_k \\ &\quad + \frac{1}{2} \sum_{l=0}^{k+1} \{D_x F_l \cdot D_y F_{k+1-l} + D_y F_l \cdot D_x F_{k+1-l}\} \\ &= (D_x D_y - D_x - D_y + 1 - k) F_k + \sum_{l=0}^{k+1} D_x F_l \cdot D_y F_{k+1-l}. \end{aligned}$$

This completes our proof. □

Let us discuss the case $k = -1$. Let $T := F_0$ and T_x, T_y, Z, W, then Proposition 2.0.1 can be expressed as follows. Note that $F_{-1} = 0$.

$$(D_x + D_y - 1) F_0 = D_x F_0 \cdot D_y F_0.$$

Also, since $T := F_0, T_x := D_x T, T_y := D_y T$, the following holds.

$$T_x T_y = T_x + T_y - T. \tag{2.0.4}$$

Let $\mathcal{L}_k := (1 - T_y) D_x + (1 - T_x) D_y + k$. Then Proposition 2.0.1is rewritten as follows. where $k = 1, 2, \cdots$.

$$\mathcal{L}_k F_{k+1} = (D_x D_y - D_x - D_y + 1 - k) F_k + \sum_{l=1}^{k} D_x F_l \cdot D_y F_{k+1-l}. \tag{2.0.5}$$

$T = F_0$ was known. Therefore, F_k can be expressed in terms of the known function T by recursively solving Equation (2.0.5) with $k = 1, 2, \cdots$. Before solving in this way, let us first state some algebraic properties about T.

Lemma 2.0.2 (Properties of T)**.**

$$T_{xx} = T_x\left(T_{xy}+1\right), \tag{2.0.6}$$
$$T_{xy} = T_{yx} = T_x T_{yy} = T_y T_{xx}, \tag{2.0.7}$$
$$T_{yy} = T_y\left(T_{xy}+1\right).$$

In particular, the following also hold

$$T_{xy} = \frac{T_x T_y}{1 - T_x T_y}. \tag{2.0.8}$$

Proof. From [Kut09], for T_x, T_y, we have

$$T_x = xe^{T_y}, \quad T_y = ye^{T_x}.$$

Therefore,

$$\begin{aligned}
T_{xx} =& D_x T_x \\
=& x\frac{\partial}{\partial x}(xe^{T_y}) \\
=& x(e^{T_y} + xe^{T_y}\frac{\partial}{\partial x}T_y) \\
=& x(e^{T_y} + e^{T_y}x\frac{\partial}{\partial x}T_y) \\
=& x(e^{T_y} + e^{T_y}D_x T_y) \\
=& x(e^{T_y} + e^{T_y}T_{yx}) \\
=& T_x(T_{xy}+1).
\end{aligned}$$

The same argument holds for $T_{yy} = T_y\left(T_{xy}+1\right)$. Furthermore, $T_{xy} = T_{yx}$ is obvious. In addition, we obtain

$$\begin{aligned}
T_x T_{yy} =& T_x T_y (T_{xy}+1) \\
=& T_y T_x (T_{xy}+1) \\
=& T_y T_{xx}.
\end{aligned}$$

Also, substituting Equation (2.0.6) into Equation (2.0.7), we obtain

$$T_{xy} = T_y T_x (T_{xy}+1).$$

Thus, we obtain Equation (2.0.8). $\square$

By Equation (2.0.4) and Equation (2.0.8)

$$T = Z - W,$$
$$T_{xy} = \frac{W}{1-W}.$$

In this case, the following holds.

Lemma 2.0.3. *Assume that $F(x,y)$ and $G(x,y)$ are differentiable functions $f(z), g(w)$ and can be expressed as $F(x,y) = f(Z(x,y))$, $G(x,y) = g(W(x,y))$. In this case, we have*

$$\mathcal{L}_0 F = \left(D_z f\right)(Z),$$
$$\mathcal{L}_0 G = 2\left(D_w g\right)(W). \tag{2.0.9}$$

Note that $\left(D_u f\right)(u) = uf'(u)$. Also, for a differentiable function $h(z,w)$, if $H = h(Z,W)$, then

$$\mathcal{L}_0 H = \left(D_z h\right)(Z,W) + 2\left(D_w h\right)(Z,W). \tag{2.0.10}$$

Proof. From Lemma 2.0.2,

$$\begin{aligned} D_x Z &= T_{xx} + T_{yx} = T_x + (T_x + 1)\, T_{xy}, \\ D_y Z &= T_{xy} + T_{yy} = T_y + (T_y + 1)\, T_{xy}, \end{aligned}$$

where from Equation (2.0.8), we have

$$\begin{aligned} \mathcal{L}_0 Z &= \{(1 - T_y) D_x + (1 - T_x) D_y\} Z \\ &= (1 - T_y) D_x Z + (1 - T_x) D_y Z \\ &= (1 - T_y) \{T_x + (T_x + 1)\, T_{xy}\} + (1 - T_x) \{T_y + (T_y + 1)\, T_{xy}\} \\ &= Z - 2W + \{(1 - T_y)(T_x + 1) + (1 - T_x)(T_y + 1)\}\, T_{xy} \\ &= Z. \end{aligned}$$

Since $\mathcal{L}$ is a linear operator,

$$\mathcal{L}_0 f(Z) = f'(Z) \mathcal{L}_0 Z = Z f'(Z) = (Df)(Z).$$

Note that Definition of $\mathcal{L}_0$ is

$$\begin{aligned} \mathcal{L}_0 T &= (1 - T_y)\, T_x + (1 - T_x)\, T_y \\ &= Z - 2W. \end{aligned}$$

Since $W = Z - T$, we have

$$\mathcal{L}_0 W = \mathcal{L}_0 Z - \mathcal{L}_0 T = 2W.$$

Therefore,

$$\mathcal{L}_0 g(W) = g'(W) \mathcal{L}_0 W = 2W g'(W) = 2(Dg)(W).$$

The same argument holds for $h(Z, W)$. □

3 Explicit expression of F_1, F_2, F_3 and F_4

In this section, we give an explicit expression of F_1, F_2 by solving PDE (2.0.5). Lemma (2.0.2) is essential.

3.1 For F_1

Let us discuss the case F_1, i.e., $k = 1$. The k corresponds to the number of cycles in the bipartite graph. In this case, we are discussing the total number of connected bipartite graphs such that there is one cycle. Also, PDE (2.0.5) is as follows when $k = 0$.

$$\mathcal{L}_0 F_1 = (D_x D_y - D_x - D_y + 1)\, F_0. \tag{3.1.1}$$

Rewriting it using $T := F_0$ yields the following.

$$\mathcal{L}_0 F_1 = T_{xy} - T_x - T_y + T.$$

Furthermore, using W, the left-hand side becomes

$$T_{xy} - T_x - T_y + T = \frac{W}{1 - W} - W.$$

Let us prove Proposition 1.2.1. Suppose there exists a function $f_1 = f_1(w)$ such that $F_1 = f_1(W)$. This f_1 has no constant term. That is, $f_1(0) = 0$. From Epuation (2.0.9),

$$\mathcal{L}_0 F_1 = 2\,(Df_1)\,.$$

In addition, from Equation (3.1.1),

$$2\,(Df_1)\,(w) = \frac{w}{1-w} - w.$$

i.e.,

$$f_1'(w) = \frac{1}{2}\left(\frac{1}{1-w} - 1\right).$$

Since $f_1(0) = 0$, we have

$$f_1(w) = -\frac{1}{2}(\log(1-w) + w).$$

3.2 For F_2

The case F_2, i.e., $k = 2$, is described. We discuss connected bipartite graphs such that there can be two cycles. That is, the case $k = 1$ in PDE (2.0.5). In other words, we have

$$\mathcal{L}_1 F_2 = (D_x D_y - D_x - D_y)\,F_1 + D_x F_1 \cdot D_y F_1. \tag{3.2.2}$$

Note that $\mathcal{L}_1 = \mathcal{L}_0 + 1$. Furthermore, F_1 is known by using Proposition 1.2.1. Let us prepare some complementary propositions for Z_x, Z_y.

Lemma 3.2.1 (Properties of Z_x, Z_y, Z_{xy})**.** *The following holds for Z_x, Z_y, Z_{xy}.*

$$\begin{aligned} Z_x &= \frac{W + T_x}{1-W}, \\ Z_y &= \frac{W + T_y}{1-W}, \\ Z_{xy} &= \frac{2+Z}{(1-W)^2} T_{xy}. \end{aligned}$$

Furthermore,

$$\begin{aligned} Z_x + Z_y &= \frac{Z + 2W}{1-W}, \\ Z_x Z_y &= \frac{W(Z+W+1)}{(1-W)^2}. \end{aligned}$$

Lemma 3.2.2 (Properties of W_x, W_y, W_{xy})**.** *The following holds for W_x, W_y, W_{xy}.*

$$\begin{aligned} W_x &= (1+T_x)\,T_{xy}, \\ W_y &= (1+T_y)\,T_{xy}, \\ W_{xy} &= T_{xy}^2 + \frac{1+Z+W}{(1-W)^2} T_{xy}. \end{aligned}$$

Furthermore,

$$\begin{aligned} W_x + W_y &= \frac{W(Z+2)}{1-W}, \\ W_x W_y &= \frac{W^2(Z+W+1)}{(1-W)^2}, \\ Z_x W_y + Z_y W_x &= \frac{W(ZW + Z + 4W)}{(1-W)^2}. \end{aligned}$$

Proof. From Lemma 2.0.2,

$$\begin{aligned}
W_x &= (T_x T_y)_x\,, \\
&= T_{xx} T_y + T_x T_{xy}, \\
&= (1+T_x)\, T_{xy}.
\end{aligned}$$

The same is true for W_y. In addition,

$$\begin{aligned}
(T_{xy})_y &= \left(\frac{W}{1-W}\right)_y, \\
&= \frac{1}{(1-W)^2} W_y, \\
&= \frac{1}{(1-W)^2}\,(1+T_y)\, T_{xy},
\end{aligned}$$

Thus,

$$\begin{aligned}
W_{xy} &= ((1+T_x)\, T_{xy})_y\,, \\
&= T_{xy}^2 + (T_x+1)\,(T_{xy})_y\,, \\
&= T_{xy}^2 + (1+T_x)\,(1+T_y)\,\frac{1}{(1-W)^2} T_{xy}.
\end{aligned}$$

We have completed the proof. □

Let us prove Proposition 1.2.2. Let $G = -2F_1$, then

$$G(W) = \log(1-W) + W.$$

Furthermore, the following holds.

$$G_x = \left(\frac{-1}{1-W} + 1\right) W_x = \frac{-W}{1-W} W_x = -T_{xy} W_x.$$

Similarly, $G_y = -T_{xy} W_y$. Therefore,

$$G_x + G_y = -T_{xy}\,(W_x + W_y) = -(2+Z) T_{xy}^2.$$

From Lemma 3.2.2,

$$\begin{aligned}
G_{xy} &= -\left(T_{xy} W_x\right)_y \\
&= -\left(T_{xy}\right)_y W_x - T_{xy} W_{xy} \\
&= -\frac{1}{(1-W)^2}\,(1+T_y)\, T_{xy} \cdot (1+T_x)\, T_{xy} - T_{xy}\left(T_{xy}^2 + (1+T_x)\,(1+T_y)\,\frac{1}{(1-W)^2} T_{xy}\right) \\
&= -T_{xy}^2 \left\{\frac{2}{(1-W)^2}\,(1+T_x)\,(1+T_y) + T_{xy}\right\} \\
&= -T_{xy}^2 \left\{\frac{2}{(1-W)^2} (1+Z+W) + T_{xy}\right\}.
\end{aligned}$$

Furtheremore,

$$G_x G_y = T_{xy}^2 W_x W_y = T_{xy}^4\,(1+T_x)\,(1+T_y)\,.$$

Thus, from Equation (3.2.2),

$$\begin{aligned}
4\mathcal{L}_1 F_2 &= 2\left(G_x + G_y\right) + G_x G_y - 2G_{xy} \\
&= -2(Z+2)T_{xy}^2 + T_{xy}^4\left(1+T_x\right)\left(1+T_y\right) + 2T_{xy}^2\left\{\frac{2}{(1-W)^2}\left(1+T_x\right)\left(1+T_y\right) + T_{xy}\right\} \\
&= \frac{T_{xy}^2}{(1-W)^2}\left\{-2(Z+2)(1-W)^2 + W^2(1+Z+W) + 4(1+Z+W) + 2W(1-W)\right\} \\
&= \frac{W^2}{(1-W)^4}\left\{\left(-W^2+4W+2\right)Z + \left(W^2-5W+14\right)W\right\}. \qquad (3.2.3)
\end{aligned}$$

Let us assume that there exists a function $a_0(w), a_1(w)$ such that $f_2(z,w) = a_1(w)z + a_0(w)$ in $F_2(Z,W)$. From $\mathcal{L}_1 = \mathcal{L}_0 + 1$, Equation (2.0.10), Equation (3.2.3) becomes

$$4\left(D_z f_2 + 2D_w f_2 + f_2\right) = \frac{w^2}{(1-w)^4}\left\{\left(-w^2+4w+2\right)z + \left(w^2-5w+14\right)w\right\}. \qquad (3.2.4)$$

In other words, $f_2(z,w) = a_1(w)z + a_0(w)$. Therefore,

$$\begin{aligned}
D_z f_2 + 2D_w f_2 + f_2 &= a_1(w)z + 2\left\{\left(Da_1\right)(w)z + \left(Da_0\right)(w)\right\} + a_1(w)z + a_0(w) \\
&= \left\{2a_1(w) + 2\left(Da_1\right)(w)\right\}z + \left\{2\left(Da_0\right)(w) + a_0(w)\right\}. \qquad (3.2.5)
\end{aligned}$$

We compare the coefficients of Equations (3.2.4) and (3.2.5).

$$\begin{aligned}
a_0(w) + 2\left(Da_0\right)(w) &= \frac{w^5}{4(1-w)^4}\left(w^2-5w+14\right) \\
2a_1(w) + 2\left(Da_1\right)(w) &= \frac{w^2}{4(1-w)^4}\left(-w^2+4w+2\right).
\end{aligned}$$

Since $a_0(0) = a_1(0) = 0$ holds, we have

$$a_0(w) = \frac{w^3(6-w)}{12(1-w)^3}, \quad a_1(w) = \frac{w^2(2+3w)}{24(1-w)^3}.$$

Thus,

$$f_2(z,w) = \frac{w^2(2+3w)}{24(1-w)^3}z + \frac{w^3(6-w)}{12(1-w)^3}.$$

Note that for the orders of f_1 and f_2, the lowest order when solving the recurrence formula of F_k came up. When solving for this recurrence formula, what is the lowest order of its coefficients?

The generating function of bipartite graph with number of vertex pairs (r,s) and number of edges is $r+s-1+k$ is

$$F_k(x,y) = \sum_{r,s=0}^{\infty} \frac{f(r,s,r+s-1+k)}{r!\cdot s!}x^r y^s. \qquad (3.2.6)$$

Note that k is the number of cycles. Let us consider the case $k=2$. In other words, bipartite graphs in which two cycles appear. Obviously, the generating function F_2 is

$$F_2(x,y) = \sum_{r,s=0}^{\infty} \frac{f(r,s,r+s+1)}{r!\cdot s!}x^r y^s.$$

The lowest order term of F_2 is when $(r,s)=(2,3),(3,2)$. For example, when $(r,s)=(2,2)$, we cannot construct two cycles. In this case, a simple calculation yields the generating function F_2 to

$$F_2(x,y)=\frac{f(3,2,6)}{3!2!}x^3y^2+\frac{f(2,3,6)}{2!3!}x^2y^3+\sum_{r}^{\infty}\sum_{s}^{\infty}\frac{f\left(r,s,r+s+1\right)}{r!\cdot s!}x^ry^s. \tag{3.2.7}$$

It is important to note the range of r,s in the infinite series of the third term. This is the shaded area in Figure 6. Let us further rearrange eq (3.2.7).

$$\begin{aligned}
F_2(x,y)&=\frac{f(3,2,6)}{3!2!}x^3y^2+\frac{f(2,3,6)}{2!3!}x^2y^3+\sum_{r}^{\infty}\sum_{s}^{\infty}\frac{f\left(r,s,r+s+1\right)}{r!\cdot s!}x^ry^s\\
&=\frac{1}{24}x^3y^2+\frac{1}{24}x^2y^3+\sum_{r,s}^{\infty}\frac{f(r,s,r+s+1)}{r!\cdot s!}x^ry^s\\
&=\frac{1}{24}(xy)^2(x+y)+\sum_{r,s}^{\infty}\frac{f(r,s,r+s+1)}{r!\cdot s!}x^ry^s\\
&=\frac{1}{24}W^2Z+\sum_{r,s}^{\infty}\frac{f(r,s,r+s+1)}{r!\cdot s!}x^ry^s.
\end{aligned}$$

Where the last Equation transformation used the relationship between T_x and T_y. That is, Definition of T,T_x,T_y is

$$\begin{aligned}
T&:=F_0,\\
T_x&:=D_xT=x+r\sum_{r,s=1}^{\infty}\frac{r^{s-1}\cdot s^{r-1}}{r!\cdot s!}x^ry^s,\\
T_y&:=D_yT=y+s\sum_{r,s=1}^{\infty}\frac{r^{s-1}\cdot s^{r-1}}{r!\cdot s!}x^ry^s.
\end{aligned}$$

Thus, with repect to $W:=T_xT_y$ and $Z:=T_x+T_y$,

$$\begin{aligned}
W&=xy+(\textit{higher order terms than } xy),\\
Z&=x+y+(\textit{higher order terms than } x+y).
\end{aligned}\tag{3.2.8}$$

Let us assume that $F_2(x,y)=\nu(Z,W)$, where

$$\nu(z,w):=\sum_{i_1=0,i_2=0}^{\infty}a_{i_1,i_2}z^{i_1}w^{i_2}.$$

However, for the index i_1,i_2 of an infinite series, $i_1=1,i_2=2$ is minimal.

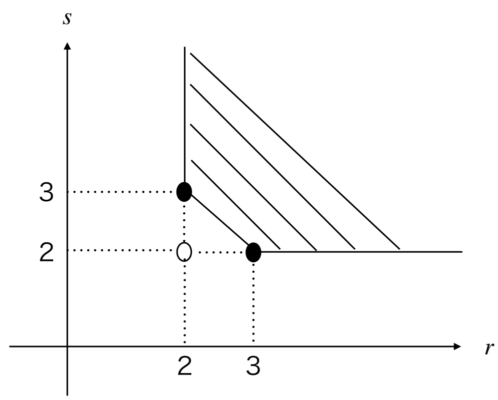


Figure 6: Range of r, s for Equation (3.2.7).

In this situation, what would be the lowest order term in a_{i_1,i_2}? To conclude, the lowest order term is zw^2. Therefore, for x and y, this means that fifth order is the lowest order. In addition, no z-only term, i.e., no constant term in w, appears. This is easy to understand if we consider the correspondence between the order of x, y and the bipartite graphs. It is clear from Equation (3.2.6) that the degree of x, y coincides with the number of vertices (r, s) in the bipartite graphs. For example, we consider the case zw. In this case, the numbers of corresponding vertices are $(r, s) = (1, 2), (2, 1)$, and two cycles cannot be constructed. For the same reason, it can be seen that the constant term of w does not appear either. From Equation (3.2.8), x^n and y^n appear when the z-only term is expands. The vertex corresponding to x^n or y^n is $(r, s) = (n, 0), (0, n)$. In other words, only one of the vertices V_1 or V_2 exists. Therefore, two cycles cannot be constructed. On the other hand, the constant term in z, i.e., the w-only term, appears.

3.3 For F_3

First, in Equation (2.0.5), when $k = 2$, we have

$$
\begin{aligned}
\mathcal{L}_2 F_3 &= (D_x D_y - D_x - D_y + 1 - 2) F_2 + \sum_{l=1}^{2} D_x F_l \cdot D_y F_{3-l} \\
&= (D_x D_y - D_x - D_y - 1) F_2 + \sum_{l=1}^{2} D_x F_l \cdot D_y F_{3-l} \\
&= (D_x D_y - D_x - D_y - 1) F_2 + D_x F_1 \cdot D_y F_2 + D_x F_2 \cdot D_y F_1 \\
&= V_{xy} - V_x - V_y - V + U_x V_y + V_x U_y. \qquad (3.3.9)
\end{aligned}
$$

Note $\mathcal{L}_2 := (1 - T_y) D_x + (1 - T_x) D_y + 2$. Then,

$$V_x = \partial_z v(z, w) \cdot Z_x + \partial_w v(z, w) \cdot W_x, \qquad (3.3.10)$$

$$V_y = \partial_z v(z, w) \cdot Z_y + \partial_w v(z, w) \cdot W_y. \qquad (3.3.11)$$

Similarly,

$$V_{xy} = D_y[\partial_z v(z,w)] \cdot Z_x + \partial_z v(z,w) \cdot Z_{xy} + D_y[\partial_w v(z,w)] \cdot W_x + \partial_w v(z,w) \cdot W_{xy} \tag{3.3.12}$$

From Lemma 3.2.1 and Lemma 3.2.2, Using Equation (3.3.10) (3.3.11) (3.3.12) and expressing Equation (3.3.9) in Z, W,

$$\begin{aligned}
\mathcal{L}_2 F_3 = &-\frac{(w(w(2w((w-8)w+28)-97)+215)+20)w^3}{12(w-1)^7} \\
&+\frac{(w(w(9(w-6)w+83)-470)-288)w^3}{48(w-1)^7} z \\
&-\frac{(w(22w+21)+2)w^2}{12(w-1)^7} z^2 \\
= &-\frac{(2w^5-16w^4+56w^3-97w^2+215w+20)w^3}{12(w-1)^7} \\
&+\frac{(9w^4-54w^3+83w^2-470w-288)w^3}{48(w-1)^7} z \\
&-\frac{(22w^2+21w+2)w^2}{12(w-1)^7} z^2.
\end{aligned}$$

On the other hand, since $f_3(z,w) = a_0(w) + a_1(w)z + a_2(w)z^2$ we can express

$$\begin{aligned}
\mathcal{L}_2 F_3(z,w) &= \left\{a_1(w)z + 2a_2(w)z^2\right\} + 2\left\{D_w a_0(w) + D_w a_1(w) z + D_w a_2(w) z^2\right\} \\
&\quad + 2\left\{a_0(w) + a_1(w)z + a_2(w)z^2\right\} \\
&= \{2a_0(w) + 2D_w a_0(w)\} + \{3a_0(w) + 2D_w a_1(w)\} z + \{4a_0(w) + 2D_w a_2(w)\} z^2.
\end{aligned}$$

Thus,

$$\begin{aligned}
2a_0(w) + 2D_w a_0(w) &= \frac{w^3\left(2w^5-16w^4+56w^3-97w^2+215w+20\right)}{12(1-w)^7} \\
3a_0(w) + 2D_w a_1(w) &= -\frac{w^3\left(9w^4-54w^3+83w^2-470w-288\right)}{48(1-w)^7} \\
4a_0(w) + 2D_w a_2(w) &= \frac{w^2\left(22w^2+21w+2\right)}{12(1-w)^7}
\end{aligned}$$

Solving this differential Equation by $a_0(0) = a_1(0) = a_2(0) = 0$, we conclude

$$\begin{aligned}
F_3(z,w) = &\frac{w^3\left(5+41w-23w^2+8w^3-w^4\right)}{24(1-w)^6} + \frac{w^3\left(32+34w-9w^2+3w^3\right)}{48(1-w)^6} z \\
&+\frac{w^2\left(1+8w+6w^2\right)}{48(1-w)^6} z^2.
\end{aligned}$$

3.4 For F_4

In Equation (2.0.5), note that $F_3 := X$, when $k = 3$,

$$\begin{aligned}
\mathcal{L}_3 F_4 &= (D_x D_y - D_x - D_y + 1 - 3) F_3 + \sum_{l=1}^{3} D_x F_l \cdot D_y F_{4-l} \\
&= X_{xy} - X_x - X_y - 2X + U_x \cdot X_y + V_x \cdot V_y + X_x \cdot U_y.
\end{aligned}$$

Note $\mathcal{L}_3 := (1-T_y)\,D_x + (1-T_x)\,D_y + 3 = \mathcal{L}_0 + 3$. Then,

$$X_x = \partial_z x(z,w)\cdot Z_x + \partial_w x(z,w)\cdot W_x$$
$$X_y = \partial_z x(z,w)\cdot Z_y + \partial_w x(z,w)\cdot W_y.$$

Similarly, the same is true for the

$$X_{xy} = D_y[\partial_z x(z,w)]\cdot Z_x + \partial_z x(z,w)\cdot Z_{xy} + D_y[\partial_w x(z,w)]\cdot W_x + \partial_w x(z,w)\cdot W_{xy}.$$

Thus,

$$\begin{aligned}
\mathcal{L}_3 F_4 =& \frac{w^2\left(841w^4+4120w^3+4096w^2+864w+24\right)}{576(w-1)^{10}}z^3\\
&+\frac{w^2\left(-61w^6+1223w^5+4646w^4+26250w^3+24180w^2+3432w\right)}{576(w-1)^{10}}z^2\\
&+\frac{w^2\left(-92w^8+791w^7-2850w^6+8699w^5+1600w^4+64248w^3+44856w^2+2088w\right)}{576(w-1)^{10}}z\\
&+\frac{w^2\left(76w^9-835w^8+4147w^7-12257w^6+24409w^5-23248w^4+70084w^3+17040w^2+144w\right)}{576(w-1)^{10}}.
\end{aligned} \tag{3.4.13}$$

Also, since $\mathcal{L}_3 = \mathcal{L}_0 + 3$, let

$$F_4(x,y) = a_0(w) + a_1(w)z + a_2(w)z^2 + a_3(w)z^3,$$

from Equation (2.0.10), the following holds.

$$\begin{aligned}
\mathcal{L}_3 F_4 =&(\mathcal{L}_0+3)F_4\\
=&\mathcal{L}_0 F_4 + 3F_4\\
=&a_1(w)z + 2a_2(w)z^2 + 3a_3(w)z^3\\
&+2\left\{D_w a_0(w) + D_w a_1(w)z + D_w a_2(w)z^2 + D_w a_3(w)z^3\right\}\\
&+3\left\{a_0(w) + a_1(w)z + a_2(w)z^2 + a_3(w)z^3\right\}\\
=&\{2D_w a_3(w) + 6a_3(w)\}z^3 + \{2D_w a_2(w) + 5a_2(w)\}z^2\\
&+\{2D_w a_1(w) + 4a_1(w)\}z + 2D_w a_0(w) + 3a_0(w).
\end{aligned}$$

Compared to Equation (3.4.13)

$$2D_w a_3(w) + 6a_3(w) = \frac{w^2\left(841w^4+4120w^3+4096w^2+864w+24\right)}{576(w-1)^{10}},$$
$$2D_w a_2(w) + 5a_2(w) = \frac{w^2\left(-61w^6+1223w^5+4646w^4+26250w^3+24180w^2+3432w\right)}{576(w-1)^{10}},$$

$$\begin{aligned}
&2D_w a_1(w) + 4a_1(w)\\
&=\frac{w^2\left(-92w^8+791w^7-2850w^6+8699w^5+1600w^4+64248w^3+44856w^2+2088w\right)}{576(w-1)^{10}},
\end{aligned}$$

$$\begin{aligned}
&2D_w a_0(w) + 3a_0(w)\\
&=\frac{w^2\left(76w^9-835w^8+4147w^7-12257w^6+24409w^5-23248w^4+70084w^3+17040w^2+144w\right)}{576(w-1)^{10}}.
\end{aligned}$$

Solving this partial differential Equation under the initial conditions $a_1(0) = a_2(0) = a_3(0) = 0$ yields

$$\begin{aligned}
a_0(w) &= \frac{w^3\left(76w^7 - 809w^6 + 3746w^5 - 9889w^4 + 15356w^3 - 22820w^2 - 7680w - 80\right)}{2880(w-1)^9},\\
a_1(w) &= -\frac{w^3\left(230w^6 - 1425w^5 + 5568w^4 - 6617w^3 + 30468w^2 + 35988w + 2088\right)}{5760(w-1)^9},\\
a_2(w) &= -\frac{w^3\left(61w^4 + 64w^3 + 1186w^2 + 1692w + 312\right)}{576(w-1)^9},\\
a_3(w) &= -\frac{w^2\left(254w^4 + 1919w^3 + 2624w^2 + 704w + 24\right)}{5760(w-1)^9}.
\end{aligned}$$

Therefore,

$$\begin{aligned}
F_4(z,w) =& \frac{w^3\left(76w^7 - 809w^6 + 3746w^5 - 9889w^4 + 15356w^3 - 22820w^2 - 7680w - 80\right)}{2880(w-1)^9}\\
&- \frac{w^3\left(230w^6 - 1425w^5 + 5568w^4 - 6617w^3 + 30468w^2 + 35988w + 2088\right)}{5760(w-1)^9} z\\
&- \frac{w^3\left(61w^4 + 64w^3 + 1186w^2 + 1692w + 312\right)}{576(w-1)^9} z^2\\
&- \frac{w^2\left(254w^4 + 1919w^3 + 2624w^2 + 704w + 24\right)}{5760(w-1)^9} z^3.
\end{aligned}$$

4 Asymptotic behavior for the coefficients of $F_1(x,x)$ and $F_2(x,x)$

In this section, we derive the asymptotic behavior of the coefficients of F_1 and F_2 obtained in Section 3.

4.1 For $F_1(x,x)$

From Definition of convolution, with respect to $\langle x^n \rangle A(x) = a_n$ and $\langle x^n \rangle B(x) = b_n$, we see that

$$\langle x^n \rangle A(x)B(x) = \sum_{k=0}^{n} \binom{n}{k} a_k b_{n-k}. \tag{4.1.1}$$

Then, introduce the following symbols for the two variables x, y in exponential generating function

$$C(x,y) := \sum_{r,s=0}^{\infty} c_{rs} \frac{x^r y^s}{r!s!}.$$

That is as follows.

$$\langle x^r y^s \rangle C(x,y) = c_{rs}$$

Also, note that for $C(x,x)$ it is given by

$$\langle x^n \rangle C(x,x) = \sum_{r+s=n} \binom{n}{r} c_{rs}.$$

Focusing on the coefficients of $F_1(x, x)$, we see that

$$u_n := \langle x^n \rangle F_1(x, x) = \sum_{r+s=n} \binom{n}{r} f(r, s, r+s).$$

This corresponds to the number of unicycles of the complete graph $K_{r,s}$ on $r + s = n$. Let us check the asymptotic behavior of this u_n at $n \to \infty$. From 1.2.1,

$$F_1(x, x) = \frac{1}{2} \sum_{k=2}^{\infty} \frac{W(x, x)^k}{k}. \tag{4.1.2}$$

We first consider the coefficients of $W(x, x)$. Since $W = T_x + T_y - T$, we have

$$W(x, y) = \sum_{r,s=1}^{\infty} \frac{w(r, s)}{r!s!} x^r y^s.$$

Note that $w(r, s) = r^{s-1} s^{r-1}(r + s - 1)$. Thus, we have

$$W(x, x) = \sum_{n=2}^{\infty} \left(\sum_{r+s=n} \frac{w(r, s)n!}{r!s!} \right) \frac{x^n}{n!} =: \sum_{n=2}^{\infty} w_n \frac{x^n}{n!}.$$

Then,

$$\begin{aligned} w_n &= \sum_{r+s=n} \frac{r^{s-1} s^{r-1}(r + s - 1)n!}{r!s!} \\ &= (n - 1) \sum_{r=1}^{n-1} \binom{n}{r} r^{n-r-1}(n - r)^{r-1}. \end{aligned} \tag{4.1.3}$$

The sum of Equation (4.1.3), which is known as [KP09], is as follows

Lemma 4.1.1 ([KP09])**.** *For $n = 2, 3, \cdots$,*

$$\sum_{r=1}^{n-1} \binom{n}{r} r^{n-r-1}(n - r)^{r-1} = 2n^{n-2}.$$

Proof. The proof is done as in combinatorics. Let S_n be the set of all spanning trees on a vertex-labeled complete graph K_n, and let S_n^b be the set of spanning trees on the vertex-labeled complete bipartite graph K_n. Let $S_{r,s}^b$ be the set of all trees on the complete bipartite graph $K_{r,s}$ such that $r + s = n$. In this case, we have

$$S_n^b = \bigsqcup_{1 \le r \le n-1} S_{r,n-r}^b.$$

where $(V_1 \sqcup V_2, E_{r,n-r}) \in S_{r,n-r}^b$, for $|V_1| = r, |V_2| = n - r$, let us define the map $\phi : S_n^b \to S_n$ as follows.

$$\phi\left((V_1 \sqcup V_2, E_{r,n-r})\right) := (V, E_{r,n-r}).$$

That is, ϕ is a map that performs an operation that removes the restriction of being bipartite. Since any spanning tree on K_n is a bipartite graph, ϕ is a surjection. More specifically, ϕ is a

2-to-1 mapping. Moreover, for $1 \leq r \leq n-1$ and $(V_1 \sqcup V_2, E_{r,n-r}) \in S^b_{r,n-r}$, there exists only one $(V_1' \sqcup V_2', E_{n-r,r}) \in S^b_{n-r,r}$ such that

$$\begin{aligned}
&V_1' = V_2,\\
&V_2' = V_1,\\
&E_{n-r,r} = \left\{(i,j) \in V_1' \times V_2' : (j,i) \in E_{r,n-r}\right\},
\end{aligned}$$

is satisfied. Also, for $1 \leq r \leq n-1$, from the choice of r vertices labeled V_1 and Theorem 1.1.3

$$\begin{aligned}
|S^b_{r,n-r}| &= \binom{n}{r} f(r, n-r, n-1)\\
&= \binom{n}{r} r^{n-r-1}(n-r)^{r-1}.
\end{aligned}$$

Therefore,

$$|S^b_n| = \sum_{r=1}^{n-1} \binom{n}{r} r^{n-r-1}(n-r)^{r-1}.$$

On the other hand, from Theorem 1.1.1,

$$|S_n| = n^{n-2}.$$

Thus, we have

$$S^b_n = \bigsqcup_{1 \leq r \leq n-1} S^b_{r,n-r}.$$

□

The following also holds.

Corollary 4.1.1. *For $n = 1, 2, \ldots,$*

$$w_n = \langle x^n \rangle W(x,x) = 2(n-1)n^{n-2}.$$

Let us now discuss the case of powers of $W(x,x)$. For $k = 1, 2, \cdots$, let

$$w_n^{*k} := \langle x^n \rangle W(x,x)^k.$$

From Corollary 4.1.1, we see that $w_n^{*1} = w_n$. Note that the minimum order of the term $W(x,x)$ is 2, and $w_n^{*k} = 0$ for $n = 1, 2, \ldots, 2k-1$. From Equation (4.1.1), w_n^{*k} is the k-time convolution of $(w_n)_{n=2,3}, \ldots$. Therefore, it can be defined as follows.

$$w_n^{*(k+1)} = \sum_{r=2k}^{n-2} \binom{n}{r} w_r^{*k} w_{n-r}, \tag{4.1.4}$$

where from Equation (4.1.2), the coefficient u_n of F_1 is given by

$$u_n = \frac{1}{2} \sum_{2 \leq k \leq n/2} \frac{w_n^{*k}}{k}. \tag{4.1.5}$$

Corollary 4.1.2. *For $k = 1, 2, \ldots, \lfloor n/2 \rfloor$, we have*

$$w_n^{*k} = 2k \cdot (2k)! n^{n-2k-1} \binom{n}{2k}. \tag{4.1.6}$$

Proof. Fix n to be one. Let us prove Equation (4.1.6) by induction on k. First, $w_n^{*1} = w_n$ for $k = 1$. Next, if Equation (4.1.6) holds for k, then from the corollary 4.1.1 and Equation (4.1.4), we obtain the following

$$\begin{aligned}
w_n^{*(k+1)} &= \sum_{r=1}^{n} \binom{n}{r} w_r^{*k} w_{n-r} \\
&= \sum_{r=2k}^{n-2} \binom{n}{r} 2k \cdot (2k)! r^{r-2k-1} \binom{r}{2k} \cdot 2(n-r-1)(n-r)^{n-r-2} \\
&= 4k \sum_{r=2k}^{n-2} \binom{n}{r} (r-1) \cdots (r-(2k-1)) r^{r-1-(2k-1)} (n-r-1)(n-r)^{n-r-2}.
\end{aligned}$$

Now we introduce a class of polynomials which appears in Abel's generalization of the binomial formula [Rio68]:

$$A_n(x, y; p, q) := \sum_{r=0}^{n} \binom{n}{r} (x+r)^{r+p} (y+n-r)^{n-r+q}.$$

In particular $p = q = -1$,

$$A_n(x, y; -1, -1) = \left(x^{-1} + y^{-1}\right)(x+y+n)^{n-1}. \tag{4.1.7}$$

Multiplying both sides by xy yields

$$\begin{aligned}
(x+y) Q(x,y) &= xy \sum_{r=0}^{n} \binom{n}{r} (x+r)^{r-1} (y+n-r)^{n-r-1} \\
&= x(x+n)^{n-1} + y(y+n)^{n-1} + xyS(x,y),
\end{aligned} \tag{4.1.8}$$

where $Q(x,y) := (x+y+n)^{n-1}$. In addition, we define

$$S(x,y) := \sum_{r=1}^{n-1} \binom{n}{r} (x+r)^{r-1} (y+n-r)^{n-r-1}.$$

For $p \in \mathbb{N}$, from Leibniz's chain rule for partial differentiation, we have

$$\begin{aligned}
\partial_x^p (xS(x,y)) &= p\partial_x^{p-1} S(x,y) + x\partial_x^p S(x,y) \\
\partial_x^p ((x+y)Q(x,y)) &= p\partial_x^{p-1} Q(x,y) + (x+y)\partial_x^p Q(x,y).
\end{aligned}$$

Then,

$$\partial_x^p \partial_y^2 (xyS(x,y))\big|_{x=y=0} = 2pS^{(p-1,1)}(0,0), \tag{4.1.9}$$

$$\partial_x^p \partial_y^2 ((x+y)Q(x,y))\big|_{x=y=0} = (p+2)Q^{(p-1,2)}(0,0). \tag{4.1.10}$$

Note that

$$\begin{aligned}
S^{(p,q)}(x,y) &:= \partial_x^p \partial_y^q S(x,y), \\
Q^{(p,q)}(x,y) &:= \partial_x^p \partial_y^q Q(x,y).
\end{aligned}$$

When $k = 1, 2, \ldots, \lfloor n/2 \rfloor$, differentiate both sides of Equation (4.1.8) twice with respect to x and 2 times with respect to y. Furthermore, with $p = 2k$, we use the partial derivative of Equation (4.1.9) and the partial derivative of Equation (4.1.10).

$$\begin{aligned}
\partial_x^{2k}\partial_y^2(\text{RHS of } (4.10))\Big|_{x=y=0} &= \partial_x^{2k}\partial_y^2(xyS(x,y))\Big|_{x=y=0} \\
&= 4kS^{(2k-1,1)}(0,0) = w_n^{*(k+1)} \\
\partial_x^{2k}\partial_y^2(\text{LHS of } (4.10))\Big|_{x=y=0} &= (2k+2)Q^{(2k-1,2)}(0,0) \\
&= 2(k+1)\cdot(n-1)\cdots(n-(2k+1))n^{n-1-(2k+1)} \\
&= 2(k+1)\cdot(2(k+1))!n^{n-2(k+1)-1}\binom{n}{2(k+1)}.
\end{aligned}$$

We have now completed our proof. □

Next, we show Theorem 1.2.1. From Equation (4.1.5) and Corollary 4.1.2,

$$\begin{aligned}
u_n &= \sum_{2\le k\le n/2}(2k)!n^{n-2k-1}\binom{n}{2k} \\
&= n^{n-1}\sum_{2\le k\le n/2}\frac{n!}{(n-2k)!n^{2k}},
\end{aligned}$$

where the last sum is similar to the Ramanujan Q-function. Therefore, let us refer to the following from [FS09]. Suppose $k_0 \in \mathbb{Z}_{\ge 0}$ satisfies $k_0 = o\left(n^{2/3}\right)$. In this case, the sum is divided into two parts as follows.

$$\sum_{2\le k\le n/2}\frac{n!}{(n-2k)!n^{2k}} = \sum_{2\le k\le k_0}\frac{n!}{(n-2k)!n^{2k}} + \sum_{k_0<k\le n/2}\frac{n!}{(n-2k)!n^{2k}}.$$

Since $k = o\left(n^{2/3}\right)$, from [FS09], we have

$$\frac{n!}{(n-2k)!n^{2k}} = e^{-2k^2/n}\left(1 + O\left(\frac{k}{n}\right) + O\left(\frac{k^3}{n^2}\right)\right).$$

Since the sum term decreases with k and $e^{-2k^2/n}$ becomes exponentially small for $k > k_0$, the sum of the two terms can be ignored. Therefore,

$$\begin{aligned}
\sum_{2\le k\le n/2}\frac{n!}{(n-2k)!n^{2k}} &= \sum_{2\le k\le k_0} e^{-2k^2/n}\left(1 + O\left(\frac{k}{n}\right) + O\left(\frac{k^3}{n^2}\right)\right) + o(1) \\
&= \sum_{2\le k\le k_0} e^{-2k^2/n} + O(1).
\end{aligned}$$

Again, $e^{-2k^2/n}$ is exponentially small for $k > k_0$. Hence, it can be summed by $2 \le k \le n/2$. Then, using Euler-Maclaurin's formula, we can obtain the following equation

$$\sum_{2\le k\le n/2} e^{-2k^2/n} = \sqrt{n}\int_0^\infty e^{-2x^2}dx + O(1) = \sqrt{\frac{\pi}{8}}\sqrt{n} + O(1).$$

4.2 For $F_2(x, x)$

From Equation (1.2.1),

$$Z(x, y) = T_x + T_y = \sum_{r,s=0}^{\infty} \frac{(r+s)r^{s-1}s^{r-1}}{r!s!} x^r y^s.$$

From Lemma 4.1.1,

$$\begin{aligned} Z(x, x) &= \sum_{r,s=0}^{\infty} \frac{(r+s)r^{s-1}s^{r-1}}{r!s!} x^{r+s} \\ &= \sum_{n=1}^{\infty} n \left(\sum_{r+s=n} \frac{n!r^{s-1}s^{r-1}}{r!s!} \right) \frac{x^n}{n!} = \sum_{n=1}^{\infty} 2n^{n-1} \frac{x^n}{n!}. \end{aligned} \tag{4.2.11}$$

Let $Y(x)$ be the exponential generating function for the number of labeled rooted spanning trees in K_n:

$$Y(x) := \sum_{n=1}^{\infty} n^{n-1} \frac{x^n}{n!}.$$

First, we see the powers of $Y(x)$.

Lemma 4.2.1 (Power of $Y(x)$). *For $k = 1, 2, \dots,$*

$$Y(x)^k = \sum_{n=1}^{\infty} k(n-1)(n-2)\cdots(n-(k-1))n^{n-k} \frac{x^n}{n!}.$$

Proof. We show by induction on k. Suppose that the complement 4.2.1 holds for k. That is,

$$\begin{aligned} Y(x)^{k+1} &= \left(\sum_{n=1}^{\infty} k(n-1)(n-2)\cdots(n-(k-1))n^{n-k} \frac{x^n}{n!} \right) \left(\sum_{n=1}^{\infty} n^{n-1} \frac{x^n}{n!} \right) \\ &= kx^{k+1} \left(\sum_{n=0}^{\infty} (n+k)^{n-1} \frac{x^n}{n!} \right) \left(\sum_{n=0}^{\infty} (n+1)^{n-1} \frac{x^n}{n!} \right) \\ &= kx^{k+1} \sum_{n=0}^{\infty} \left(\sum_{r=0}^{n} \binom{n}{r} (k+r)^{r-1} (1+n-r)^{n-r-1} \right) \frac{x^n}{n!}. \end{aligned}$$

By Equation (4.1.7), we have

$$\begin{aligned} \sum_{r=0}^{n} \binom{n}{r} (k+r)^{r-1} (1+n-r)^{n-r-1} &= A_n(k, 1; -1, -1) \\ &= \left(\frac{1}{k} + 1 \right) (k+1+n)^{n-1}. \end{aligned}$$

Therefore,

$$\begin{aligned} Y(x)^{k+1} &= x^{k+1} \sum_{n=0}^{\infty} (k+1)(n+k+1)^{n-1} \frac{x^n}{n!} \\ &= (k+1) \sum_{n=0}^{\infty} (n+1)(n+2)\cdots(n+k)(n+k+1)^n \frac{x^{n+k+1}}{(n+k+1)!} \\ &= (k+1) \sum_{n=k+1}^{\infty} (n-1)(n-2)\cdots(n-k)n^{n-(k+1)} \frac{x^n}{n!}. \end{aligned}$$

Now it holds for $k+1$, so it is proved. □

Lemma 4.2.1 for $a_k \in \mathbb{R}, k = 1, 2, \ldots,$

$$\sum_{k=1}^{\infty} a_k Y(x)^k = \sum_{n=1}^{\infty} n^{n-1} \left(\sum_{k=1}^{\infty} a_k k \frac{(n-1)(n-2)\cdots(n-(k-1))}{n^{k-1}} \right) \frac{x^n}{n!}, \tag{4.2.12}$$

where the sum over k is finite. From Corollary 4.1.1, Equation (4.2.11), and Lemma 4.2.1, we obtain

$$Z(x,x) = \sum_{n=1}^{\infty} 2n^{n-1} \frac{x^n}{n!} = 2Y(x), \tag{4.2.13}$$

$$W(x,x) = \sum_{n=1}^{\infty} 2(n-1)n^{n-2} \frac{x^n}{n!} = Y(x)^2. \tag{4.2.14}$$

Therefore, instead of $Z(x,x)$ and $W(x,x)$, only $Y(x)$ can be used to represent $F_2(x,x)$. Expressing Equation (4.2.13) and Equation (4.2.14) in terms of Y in Proposition 1.2.2, we have

$$\begin{aligned} F_2(x,x) =& f_2\left(2Y, Y^2\right) = \frac{Y^5\left(2+4Y-Y^2\right)}{12(1-Y)^3(1+Y)^2} \\ =& \frac{Y^2-3Y-3}{12} - \frac{11}{64(1+Y)} + \frac{1}{32(1+Y)^2} \\ &+ \frac{143}{192(1-Y)} - \frac{11}{24(1-Y)^2} + \frac{5}{48(1-Y)^3}. \end{aligned} \tag{4.2.15}$$

In K_n, a similar equation appears in the [Wri77]. As we will see later, Equation (4.2.15) determines the asymptotic behavior of the coefficients in Equation (1.2.2). To obtain the asymptotic behavior of the coefficients of F_2, we can estimate the coefficients of $\frac{1}{(1-Y)^p}, \frac{1}{(1+Y)^p}, p \in \mathbb{N}$ from Equation (4.2.15). Next, fix one $p \in \mathbb{N}$. Also define the polynomial $\{t_n(p)\}_{n\geq 0}$ as follows.

$$\frac{1}{(1-Y(x))^p} = \sum_{n=0}^{\infty} t_n(p) \frac{x^n}{n!}.$$

This polynomial and its asymptotic behavior are detailed in[KP89].

Lemma 4.2.2 ([KP89])**.** *Fix $p \in \mathbb{N}$. For $n \to \infty$,*

$$t_n(p) = \frac{\sqrt{2\pi} n^{n-1}}{2^{p/2}} \left(\frac{n^{(p+1)/2}}{\Gamma(p/2)} + \frac{\sqrt{2}p}{3} \frac{n^{p/2}}{\Gamma((p-1)/2)} + O\left(n^{(p-1)/2}\right) + O(1) \right)$$

Therefore, the asymptotic behavior for $\frac{1}{(1-Y)^p}, p \in \mathbb{N}$ is obtained. On the other hand, for $\frac{1}{(1+Y)^p}, p \in \mathbb{N}$, we give a loose evaluation of the coefficient of $\frac{1}{(1+Y)^p}$. From the binomial expansion and Equation (4.2.12), we obtain

$$\begin{aligned} \frac{1}{(1+Y(x))^p} &= \sum_{k=0}^{\infty} \binom{p+k-1}{k} (-1)^k Y(x)^k \\ &= 1 + \sum_{n=1}^{\infty} \left(\frac{n^{n-1}}{\Gamma(p)} \sum_{k=0}^{\infty} \binom{n-1}{k} (-1)^{k+1} \frac{\Gamma(p+k+1)}{n^k} \right) \frac{x^n}{n!}. \end{aligned}$$

Thus, for $n \to \infty$,

$$\begin{aligned}\langle x^n \rangle \frac{1}{(1+Y(x))^p} &= \frac{n^{n-1}}{\Gamma(p)} \sum_{k=0}^{\infty} \binom{n-1}{k} (-1)^{k+1} \frac{\Gamma(p+k+1)}{n^k} \\ &\leq t_n(p) = O\left(n^{n+(p-1)/2}\right). \end{aligned} \tag{4.2.16}$$

Now we are in a position to prove Theorem 1.2.2. By Equation (4.2.13), Equation (4.2.14), Lemma 4.2.2 and Equation (4.2.16), we obtain the leading asymptotic behavior of coefficients of $F_2(x,x)$ appeared in Equation (4.2.15). We obtain the following.

$$\begin{aligned}\langle x^n \rangle F_2(x,x) =& -\frac{1}{12} n^{n-1} + O\left(n^n\right) + O\left(n^{n+1/2}\right) + \frac{143}{192}\left(n^n + O\left(n^{n-1/2}\right)\right) \\ & - \frac{11}{24}\left(\sqrt{\frac{\pi}{2}} n^{n+1/2} + O\left(n^n\right)\right) + \frac{5}{48}\left(n^{n+1} + O\left(n^{n+1/2}\right)\right) \\ =& \frac{5}{48} n^{n+1} + O\left(n^{n+1/2}\right).\end{aligned}$$

We have now completed our proof.

5 Another expression of $F_k(x,y)$

In this section, we prove Theorem 1.2.4.The proof is based on the combinatorial method introduced in [Wri77].

First, we discuss how to construct a basic graph for a connected bipartite graph. Fix $k \geq 2$ and consider a labeled connected $(r,s,r+s-1+k)$-graph G. Note that the vertex set is $V = (V_1, V_2)$ and $|V_1| = r, |V_2| = s$. Delete the leaf and adjacent edges from this G and repeat this operation until all the leaves in G are gone. Each operation removes only one vertex and one edge. Thus, for some $t \leq r, u \leq s$, we obtain a connected $(r,s,r+s-1+k)$-graph G labeled with no leaves. Let G' be the graph thus obtained, and let G' be independent of the order of leaf elimination. Let $V' = (V_1', V_2')$ be a vertex of G'. If V_1 is $\deg(v) \geq 3$, then we call this vertex V a **special point**. On the other hand, let us call V the vertex V such that $\deg(v) \leq 2$, and let us call it **normal point**. Let r_{sp} be the number of special points belonging to V_1', and let S_{sp} be the number of special points belonging to V_2'. Applying the handshaking lemma to G', we get

$$\sum_{v \in V'} (\deg(v) - 2) = 2(k-1).$$

Therefore,

$$r_{\mathrm{sp}} + s_{\mathrm{sp}} \leq 2(k-1). \tag{5.0.1}$$

In a graph G', an edge of a special point with different endpoints is called a **special path**, and a cycle containing a special point is called a **special cycle**. A cycle containing one special point is called a **special cycle**. Since G' is connected and $\deg(v) \geq 2$, it is clear that G' consists of a special path and a special cycle, which are disjoint except for the special point. Then, special cycles and special paths are classified into seven categories and minimally contracted to obtain a basic graph $\mathcal{B}(G)$. The special paths and special cycles are described in Figure 7. The vertex of the white circle is the special point.

These seven types of special paths and cycles are described below.

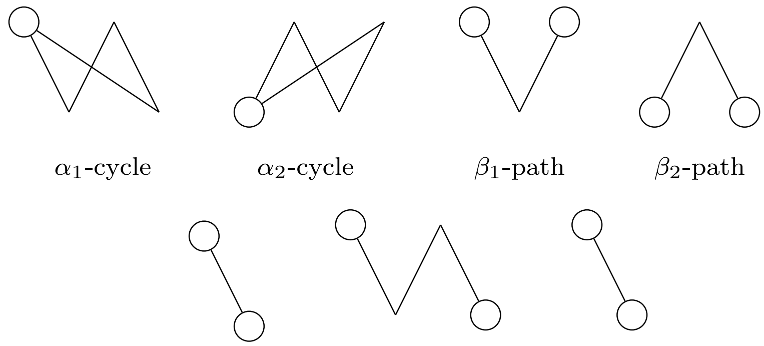


Figure 7: Special path and special cycle.

- The α_i**-cycle** is a special cycle such that $V_i'(i = 1, 2)$ has exactly one special point. By the structure of the bipartite graph, α_i-cycle has exactly three normal points. The smallest α_i-cycle has three normal points as shown in Figure 7.
- β_j**-path** is a special path whose endpoints are two different special points belonging to $V_j'(j = 1, 2)$. By the structure of the bipartite graph, β_j-path contains at least one normal point. The smallest β_j-path has only one normal point, as shown in Figure 7.
- Special paths with V_1' and V_2' special points as endpoints are classified according to the following situations.
  - There is only one special path connecting V_1' and V_2': Such a special path is called a γ_1**-path**. The length of the smallest γ_1-path is 1 as shown in Figure 7.
  - There are multiple special paths connecting V_1' and V_2': Due to the structure of simple graphs, there is only one such special path of length 1. Therefore, a path of length 3 or more is called a γ_2**-path** and a path of length 1 is called a δ**-path**. The minimum length of γ_2-path is 3.

Let us then decompose G' into the union of the sets α_i-path, β_j-path, γ_k-path and δ-path. The basic graph $\mathcal{B}(G)$ is reduced to α_i-path, β_j-path, γ_k-path and δ-path as in Figure 7. In the degenerate procedure, the vertices are unlabeled. The following is a summary of the reduction procedure.

- If each α_i-cycle$(i = 1, 2)$ contains more than 5 normal points, let the normal points reduce to the 3 smallest α_i-cycles.
- If β_i-path$(i = 1, 2)$ contains more than three normal points, it is reduced to the smallest β_i-path that contains only one normal point.
- If each γ_1-path contains a normal point, it is reduced to the smallest γ_1-path that does not contain a normal point.
- If each γ_2-path contains more than four normal points, it is reduced to the smallest γ_2-path that contains two normal points.

We have then summarized how to construct a basic graph $\mathcal{B}(G)$ from the given labeled connected $(r, s, r + s - 1 + k)$-graph. Note that the number k of cycles in the graph is universal by contraction. Therefore, the basic graph $\mathcal{B}(G)$ also has k cycles. We reconstruct a labeled bipartite graph from each basic graph $\mathcal{B}(G)$ and introduce $J_{\mathcal{B}}(x,y)$ to express F_k as the sum of $J_{\mathcal{B}}(x,y)$. In the following, we will prove Theorem 1.2.4. Let G be a labeled connected $(r,s,r+s-1+k)$-graph G. Let $V''=(V_1'',V_2'')$ be the vertex set of $\mathcal{B}(G)$. Let a_i,b_j,c_k and d be α_i-path, β_j-path, γ_k-path and δ-path respectively. In this case, from the number of vertices in $\mathcal{B}(G)$, we have

$$\left|V_1''\right| = r_{\mathrm{sp}} + a_1 + 2a_2 + b_2 + c_2 \leq t \leq r, \tag{5.0.2}$$

$$\left|V_2''\right| = s_{\mathrm{sp}} + 2a_1 + a_2 + b_1 + c_2 \leq u \leq s. \tag{5.0.3}$$

Also, since the number of vertices and edges removed by the contraction is the same for the number of edges in $\mathcal{B}(G)$, we have

$$4a_1 + 4a_2 + 2b_1 + 2b_2 + c_1 + 3c_2 + d = \left|V_1''\right| + \left|V_2''\right| + k - 1. \tag{5.0.4}$$

In addition, from Equation (5.0.2), Equation (5.0.4) and Equation (5.0.1),

$$\begin{aligned} a_1 + a_2 + b_1 + b_2 + c_1 + c_2 + d &= r_{\mathrm{sp}} + s_{\mathrm{sp}} + k - 1 \\ &\leq 3(k-1). \end{aligned} \tag{5.0.5}$$

Thus, if G is a labeled connected $(r,s,r+s-1+k)$-graph, then $\mathcal{B}(G)$ should satisfy the case from Equation (5.0.1) to Equation (5.0.5). Let BG_k be the set of all basic graphs with k cycles. That is,

$$BG_k := \{\mathcal{B}(G) : G \text{ is a labeled bipartite } (r,s,r+s-1+k)\text{-graph for some } r,s.\}.$$

From Equation (5.0.5), BG_k is a finite set. Then let us now fix one $\mathcal{B} \in BG_k$. Let $J_{\mathcal{B}}(r,s)$ be the number of labeled connected $(r,s,r+s-1+k)$-graphs G such that $\mathcal{B}(G)=\mathcal{B}$. Also define the exponential type generating function of $J_{\mathcal{B}}(r,s)$ by

$$J_{\mathcal{B}} = J_{\mathcal{B}}(x,y) := \sum_{r,s=0}^{\infty} j_{\mathcal{B}}(r,s)\frac{x^r y^s}{r!s!}.$$

5.1 Proof of Theorem 1.2.4

In the following, we show that $J_{\mathcal{B}}(x,y)$ is a rational function of T_x and T_y. At first, a bipartite graph $(r,s,r+s-1+k)$- is constructed from $\mathcal{B}$ in the following two steps.

step1：

Select one $\mathcal{B} \in BG_k$. Let $V''=(V_1'',V_2'')$ be the vertex set of $\mathcal{B}$. Also,let

$$M := a_1 + a_2 + b_1 + b_2 + c_1 + c_2$$

be the number of all minimal special paths and special cycles in $\mathcal{B}$. It also includes δ-path. Then, let t and u be taken such that

$$\begin{aligned} \left|V_1''\right| &\leq t \leq r, \\ \left|V_2''\right| &\leq u \leq s. \end{aligned}$$

Also, the smallest α_i-path, β_j-path, γ_k-path, and δ-path that make up $\mathcal{B}$ are labeled as follows $\mathbf{S}_1, \mathbf{S}_2, \ldots, \mathbf{S}_M$. We also add pairs of normal points to these special paths and special cycles. From the structure of the bipartite graph, the number of pairs to be added to each $\mathbf{S}_j$ for each

$j = 1, 2, \ldots, M$ is even and the number of normal points added to V_1'', V_2'' is equal. Let m be this number. Therefore, the necessary condition for the vertices added to V_1'', V_2'' is

$$t - \left|V_1''\right| = u - \left|V_2''\right| = \sum_{j=1}^{M} m_j.$$

Combining Equations (5.0.2) and Equation (5.0.3) with this necessary condition, a non-negative integer $\{m_j\}_{j=1}^{M}$ satisfies

$$m_1 + m_2 + \cdots + m_M = t - (r_{\mathrm{sp}} + a_1 + 2a_2 + b_2 + c_2)\,, \tag{5.1.6}$$
$$m_1 + m_2 + \cdots + m_M = u - (s_{\mathrm{sp}} + 2a_1 + a_2 + b_1 + c_2)\,. \tag{5.1.7}$$

Then, Let $y_{\mathcal{B}}(t,u) = y_{\mathcal{B}}(t, u, r_{\mathrm{sp}}, s_{\mathrm{sp}}, a_1, a_2, b_1, b_2, c_1, c_2)$ be the number of the solutions of $\{m_j\}_{j=1}^{M}$ Equation (5.1.6) and Equation (5.1.7). For each $\{m_j\}_{j=1}^{M}$, we obtain an unlabeled connected bipartite graph $(t, u, t+u-1+k)$- graph. As a result, $Y_{\mathcal{B}}(t,u)$ is obtained from $\mathcal{B}$.

step2:

Take one unlabeled connected bipartite graph $(t, u, t+u-1+k)$-$y_{\mathcal{B}}(t,u)$ of the graph. Let $T_1, \ldots, T_t$ and $U_1, \ldots, U_u$ be its vertices. Let $\mathcal{I}_{t,u}$ be a set $\{(r_{1i}, s_{1i})\}_{i=1}^{t}$ and a set $\{(r_{2j}, s_{2j})\}_{j=1}^{u}$ such that

$$\begin{aligned}
&r_{1i} \geq 1, r_{2j} \geq 0, s_{1i} \geq 0, s_{2j} \geq 1,\\
&\sum_{i=1}^{n} r_{1i} + \sum_{j=1}^{n} r_{2j} = r,\\
&\sum_{i=1}^{t} s_{1i} + \sum_{j=1}^{u} s_{2j} = s.
\end{aligned}$$

Then, for any $\mathcal{I}_{t,u}$, $\{(r_{1i}, s_{1i})\}_{i=1}^{t}$ and $\{(r_{2j}, s_{2j})\}_{j=1}^{u}$, where for $i = 1, 2, \ldots, t$, attach a rooted tree of size (r_{1i}, s_{1i}) to T_i. Let J_j be a rooted tree of size (r_{2j}, s_{2j}) for $j = 1, 2, \ldots, u$ as well. Let $\phi(r, s, t, u)$ be the number of labeled connected $(r, s, r+s-1+k)$-graphs. Let t rooted trees with roots in V_1 and $(r, s, r+s-1+k)$-graphs with roots in V_2. By counting u rooted trees whose roots are in V_2, we obtain the following.

$$\begin{aligned}
\phi(r, s, t, u) = &\sum{}' \binom{r}{r_{11}, \ldots, r_{1t}, r_{21}, \ldots, r_{2u}} \binom{s}{s_{11}, \ldots, s_{1t}, s_{21}, \ldots, s_{2u}}\\
&\times \prod_{i=1}^{t} r_{1i}^{s_{1i}} s_{1i}^{r_{1i}-1} \prod_{j=1}^{u} r_{2j}^{s_{2j}-1} s_{2j}^{r_{2j}}.
\end{aligned} \tag{5.1.8}$$

Note that the symbol $\sum'$ is the sum over the set $\mathcal{I}_{t,u}$. From the above two steps, a labeled connected $(r, s, r+s-1+k)$-graph can be obtained from the basic graph $\mathcal{B}$. However, after we label all the vertices with a rooted tree, we forget the labels $\mathbf{S}_1, \mathbf{S}_2, \ldots, \mathbf{S}_M$. Therefore, not all of them are different. If $G_{\mathcal{B}}$ is a graph automorphism of $\mathcal{B}$, then every graph appears exactly $G_{\mathcal{B}}$ times. Therefore,

$$j_{\mathcal{B}}(r,s) = \sum_{\substack{\left|V_1''\right| \leq t \leq r \\ \left|V_2''\right| \leq u \leq s}} \frac{y_{\mathcal{B}}(t,u)\phi(r,s,t,u)}{g_{\mathcal{B}}}.$$

That is,

$$J_{\mathcal{B}}(x,y) = \frac{1}{g_{\mathcal{B}}} \sum_{\substack{|V_1''|\le t \\ |V_2''|\le u}} y_{\mathcal{B}}(t,u) \sum_{\substack{t\le r \\ u\le s}} \phi(r,s,t,u) \frac{x^r y^s}{r!s!}. \tag{5.1.9}$$

Summing both sides of Equation (5.1.8) with respect to r, s, we obtain

$$\begin{aligned}\sum_{\substack{t\le r \\ u\le s}} \phi(r,s,t,u) \frac{x^r y^s}{r!s!} &= \sum_{\substack{t\le r \\ u\le s}} \sum{}' \prod_{i=1}^{t} r_{1i}^{s_{1i}i} s_{1i}^{r_{1i}-1} \frac{x^{r_{1i}} y^{s_{1i}}}{r_{1i}! s_{1i}!} \prod_{j=1}^{u} r_{2j}^{s_{2j}-1} s_{2j}^{r_{2j}} \frac{x^{r_{2j}} y^{s_{2j}}}{r_{2j}! s_{2j}!} \\ &= T_x^t T_y^u.\end{aligned}$$

On the other hand, a straightforward calculation shows that

$$\begin{aligned}\sum_{\substack{|V_1''|\le t \\ |V_2''|\le u}} y_{\mathcal{B}}(t,u) T_x^t T_y^u &= \sum_{\substack{m_1,\ldots,m_M \\ |V_1''|\le t \\ |V_2''|\le u}} T_x^t T_y^u \\
&= T_x^{|V_1''|} T_y^{|V_2''|} \sum_{\substack{|V_1''|\le t \\ |V_2''|\le u}} \sum_{\substack{m_1,\ldots,m_M \\ \sum -|V_1''| = u-|V_2''|}} (T_x T_y)^{\sum_{j=1}^M m_j} \\
&= T_x^{|V_1''|} T_y^{|V_2''|} \sum_{n\ge 0} \sum_{\substack{m_1,\ldots,m_M \\ \sum m_j = n}} (T_x T_y)^{\sum_{j=1}^M m_j} \\
&= T_x^{|V_1''|} T_y^{|V_2''|} \prod_{j=1}^{M} \left(\sum_{m_j \ge 0} (T_x T_y)^{m_j} \right) \\
&= T_x^{|V_1''|} T_y^{|V_2''|} (1 - T_x T_y)^{-M}. \end{aligned} \tag{5.1.10}$$

Then, combining Equations (5.0.2), (5.0.3), (5.1.9) and (5.1.10), we have

$$J_{\mathcal{B}}(x,y) = \frac{T_x^{r_{\mathrm{sp}}+a_1+2a_2+b_2+c_2} T_y^{s_{\mathrm{sp}}+2a_1+a_2+b_1+c_2}}{g_{\mathcal{B}} (1 - T_x T_y)^{a_1+a_2+b_1+b_2+c_1+c_2}}.$$

In k-cycle graph automorphism, the basic graph leads to a labeled bipartite $(r, s, r+s-1+k)$-graph that is not graph automorphism. Thus, taking the sum $J_{\mathcal{B}}$ with respect to $\mathcal{B} \in BG_k$, we obtain

$$F_k(x,y) = \sum_{\mathcal{B}\in BG_k} J_{\mathcal{B}}(x,y).$$

This completes the proof.

5.2 Basic graph with $k = 2$

When $k = 2$, Equation (5.0.1) is

$$r_{sp} + s_{sp} \le 2.$$

Furthermore, Equation (5.0.4) is

$$\begin{aligned} a_1 + a_2 + b_1 + b_2 + c_1 + c_2 + d &= r_{\text{sp}} + s_{\text{sp}} + 1 \\ &\leq 3. \end{aligned}$$

Therefore, the possible special point combination $(r_{\text{sp}}, s_{\text{sp}})$ is $(r_{\text{sp}}, s_{\text{sp}}) = (1, 0), (0, 1), (1, 1), (2, 0), (0, 2)$. Write down the basic graph for each of these cases. Note that the basic graph is unlabeled for both vertices and edges.

The case $(r_{\text{sp}}, s_{\text{sp}}) = (1, 0)$

Let us consider the number of cycles and paths that make up a basic graph. If

$$a_1 + a_2 + b_1 + b_2 + c_1 + c_2 + d = 2$$

is used, the number of $a_1, a_2, b_1, b_2, c_1, c_2, d$ can be expressed in the Table 3. That is, α_1-cycle is

variable	a_1	a_2	b_1	b_2	c_1	c_2	d
quantity	2	0	0	0	0	0	0

Table 3: The number of each part of the basic graph of $(r_{\text{sp}}, s_{\text{sp}}) = (1, 0)$

a combination of the two. Therefore, the basic graph is shown in Figure 8. The graph that is graph automorphism to the basic graph is shown in Figure 9. This gives $g_{\mathcal{B}} = 8$.

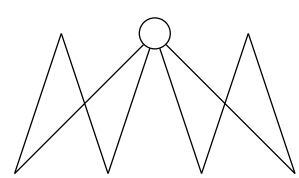


Figure 8: Basic graph of $(r_{\text{sp}}, s_{\text{sp}}) = (1, 0)$.

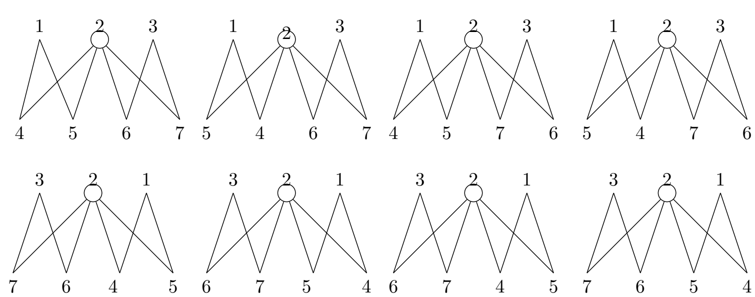


Figure 9: Graph automorphism of $(r_{\text{sp}}, s_{\text{sp}}) = (1, 0)$ in the basic graph.

The number of graph automorphism are all counted up as in Figure 9, but it is also possible to compute them by looking at Figure 8. First, the two α_1-cycles are arranged in 2! ways. There are two ways to label the vertices of each α_1-cycle: $1 \to 2 \to 3 \to 4 \to 1$ with special point as 1, or in reverse order ($1 \to 4 \to 3 \to 2 \to 1$). Thus $G_{\mathcal{B}} = 2! \times 2^2 = 8$. Therefore $J_{\mathcal{B}}(x, y)$ yields

$$J_{\mathcal{B}}(x, y) = \frac{T_x^3 T_y^4}{8(1 - T_x T_y)^2}.$$

The case $(r_{\mathrm{sp}}, s_{\mathrm{sp}}) = (0, 1)$

In this case, it is equivalent to Section 5.2. Using

$$a_1 + a_2 + b_1 + b_2 + c_1 + c_2 + d = 2$$

the number of $a_1, a_2, b_1, b_2, c_1, c_2, d$ can be calculated from Table 4.

variable	a_1	a_2	b_1	b_2	c_1	c_2	d
quantity	0	2	0	0	0	0	0

Table 4: The number of each part of the basic graph of $(r_{\mathrm{sp}}, s_{\mathrm{sp}}) = (0, 1)$

In other words, it is a basic graph combining two α_2-cycles. Therefore, the basic graph is shown in Figure 10. The graph automorphism is also omitted, since it is just a transposition of the upper and lower parts of Figure 9. That is, $G_{\mathcal{B}} = 8$. Therefore, $J_{\mathcal{B}}(x, y)$ is

$$J_{\mathcal{B}}(x, y) = \frac{T_x^4 T_y^3}{8(1 - T_x T_y)^2}.$$

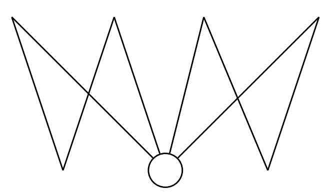


Figure 10: Basic graph of $(r_{\mathrm{sp}}, s_{\mathrm{sp}}) = (0, 1)$.

The case $(r_{\mathrm{sp}}, s_{\mathrm{sp}}) = (1, 1)$

Let us consider the number of cycles and paths that make up a basic graph. Using

$$a_1 + a_2 + b_1 + b_2 + c_1 + c_2 + d = 3,$$

the number of $a_1, a_2, b_1, b_2, c_1, c_2, d$ is obtained as shown in Table 5.

variable	a_1	a_2	b_1	b_2	c_1	c_2	d
pattern 1	0	0	0	0	0	3	0
pattern 2	0	0	0	0	0	2	1
pattern 3	1	1	0	0	1	0	0

Table 5: The number of each part of the basic graph of $(r_{\mathrm{sp}}, s_{\mathrm{sp}}) = (1, 1)$

Pattern 1 is a basic graph combining three γ_2-paths. Pattern 2 is a basic graph combining two γ_2-paths and one δ-path. Pattern 3 is a basic graph combining one α_1-cycle, one α_2-cycle and one γ_1-cycle.

Therefore, the basic graph is Figure 11.

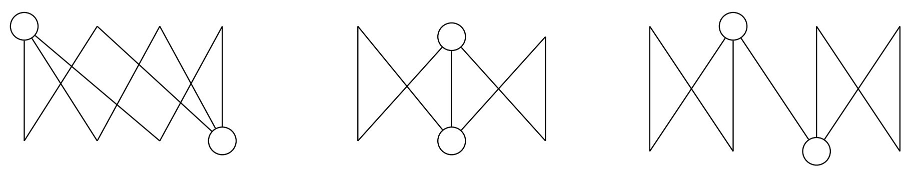


Figure 11: Basic graph of $(r_{\mathrm{sp}}, s_{\mathrm{sp}}) = (1, 1)$.(From left to right: Pattern 1, Pattern 2, Pattern 3)

Let us calculate $g_{\mathcal{B}}$ and $J_{\mathcal{B}}$ for each basic graph. The basic graph of Pattern 1 has $G_{\mathcal{B}} = 3! = 6$ ways in the total number of γ_2-path 3 sequences. Therefore, we have

$$J_{\mathcal{B}} = \frac{T_x^4 T_y^4}{6(1 - T_x T_y)^3}.$$

The basic graph of pattern 2 is $G_{\mathcal{B}} = 2$ due to symmetry around δ-path. It is not a combination of α_1-cycle and α_2-cycle, although it is easy to misunderstand from the shape of the graph. Therefore,

$$J_{\mathcal{B}} = \frac{T_x^3 T_y^3}{2(1 - T_x T_y)^2}.$$

The basic graph of pattern 3 is $g_{\mathcal{B}} = 1! \times 1! \times 2^2 = 4$ based on the total number of sequences in α_1-cycle and α_2-cycle and the labeled orientation of its vertices. Therefore, we have

$$J_{\mathcal{B}} = \frac{T_x^4 T_y^4}{4(1 - T_x T_y)^3}.$$

The case$(r_{\mathrm{sp}}, s_{\mathrm{sp}}) = (2, 0)$

Let us consider the number of cycles and paths that make up a basic graph. Using

$$a_1 + a_2 + b_1 + b_2 + c_1 + c_2 + d = 3$$

the number of $a_1, a_2, b_1, b_2, c_1, c_2, d$ is obtained as shown in Table 6.

variable	a_1	a_2	b_1	b_2	c_1	c_2	d
pattern 1	0	0	3	0	0	0	0
pattern 2	2	0	1	0	0	0	0

Table 6: The number of each part of the basic graph of $(r_{\mathrm{sp}}, s_{\mathrm{sp}}) = (1, 1)$

That is, β_1-path is a combination of the three. Therefore, the basic graph is shown in Figure 12.

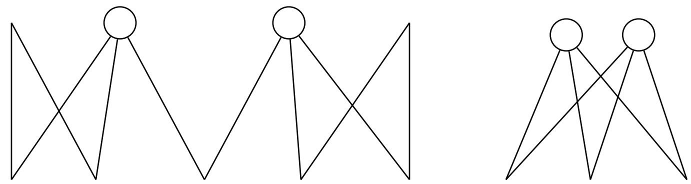


Figure 12: Basic graph of $(r_{\mathrm{sp}}, s_{\mathrm{sp}}) = (2, 0)$.(pattern 1 (left) and pattern 2 (right))

Let us compute $g_{\mathcal{B}}$ and $J_{\mathcal{B}}$ for each basic graph. Pattern 1 is $g_{\mathcal{B}} = 1! \times 1! \times 2^2 \times 2 = 8$, considering the total number of permutations of α_1-cycle, the orientation of the labels of its vertices, and its symmetry. Therefore, we have

$$J_{\mathcal{B}} = \frac{T_x^4 T_y^5}{8(1 - T_x T_y)^3}.$$

Pattern 2 considers the total number of special point alignments, γ_2-path alignments, $g_{\mathcal{B}} = 2! \times 3! = 12$. Therefore, we have

$$J_{\mathcal{B}} = \frac{T_x^2 T_y^3}{12(1 - T_x T_y)^3}.$$

The case$(r_{\mathrm{sp}}, s_{\mathrm{sp}}) = (0, 2)$

In this case, it is equivalent to Section 5.2. Therefore, the basic graph is Figure 13.

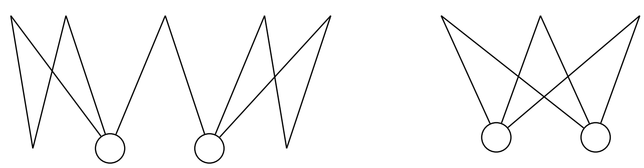


Figure 13: Basic graph of $(r_{\mathrm{sp}}, s_{\mathrm{sp}}) = (0, 2)$.(From left to right:Pattern1, Pattern2)

Thus $J_{\mathcal{B}}$ for pattern 1 is

$$J_{\mathcal{B}} = \frac{T_x^5 T_y^4}{8(1 - T_x T_y)^3}.$$

On other hand, $J_{\mathcal{B}}$ for pattern 2 is

$$J_{\mathcal{B}} = \frac{T_x^3 T_y^2}{12(1 - T_x T_y)^3}.$$

Summary

The results of Subsection 5.2 are summarized in Figure 14.

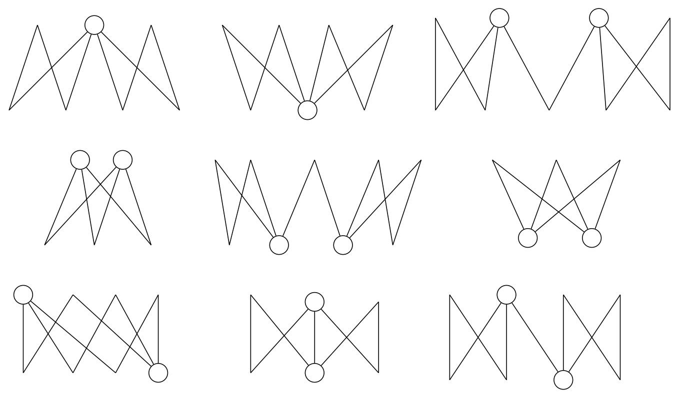


Figure 14: Basic graph with $k = 2$.

We also computed $J_{\mathcal{B}}$ for each basic graph. Let us now calculate F_2 using Theorem 1.2.4. In addition, using $W := T_xT_y, Z := T_x + T_y$, we can calculate F_2 using the following equation. We have

$$\begin{aligned}
F_2(x,y) = &\sum_{\mathcal{B}\in BG_2} J_{\mathcal{B}}(x,y) \\
=&\frac{T_x^3T_y^4}{8(1-T_xT_y)^2} + \frac{T_x^4T_y^3}{8(1-T_xT_y)^2} + \frac{T_x^4T_y^5}{8(1-T_xT_y)^3} \\
&+ \frac{T_x^2T_y^3}{12(1-T_xT_y)^3} + \frac{T_x^5T_y^4}{8(1-T_xT_y)^3} + \frac{T_x^3T_y^2}{12(1-T_xT_y)^3} \\
&+ \frac{T_x^4T_y^4}{6(1-T_xT_y)^3} + \frac{T_x^3T_y^3}{2(1-T_xT_y)^2} + \frac{T_x^4T_y^4}{4(1-T_xT_y)^3} \\
=&\frac{W^2(2+3W)}{24(1-W)^3}Z + \frac{W^3(6-W)}{12(1-W)^3}.
\end{aligned}$$

It is indeed consistent with Equation (1.2.2).

5.3 Details of calculation methods for graph automorphism

In obtaining $J_{\mathcal{B}}$ for a basic graph with $k = 2$, the number of graphs $G_{\mathcal{B}}$ that are graph automorphism to each basic graph is calculated. This section describes the calculation method in more detail. First, the basic graph itself is unlabeled in terms of vertices and edges. The vertices are labeled appropriately, and $g_{\mathcal{B}}$ is calculated based on the labeled basic graph. An image of this process is shown in Figure 1.

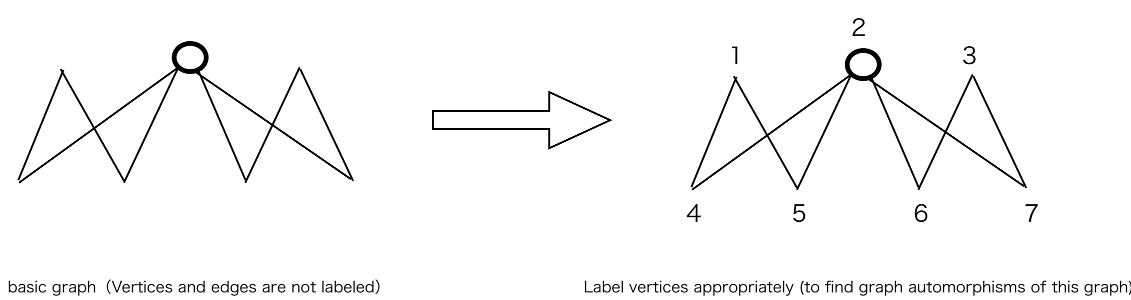


Figure 15: How to find graph automorphism.

When $G_{\mathcal{B}}$ is derived by calculation instead of counting, the calculation procedure depends on the shape of the labeled basic graph. What is important is the symmetry of the basic graph and the existence of special points to be fixed. Below is an example of a complicated calculation of $g_{\mathcal{B}}$ from a basic graph with $k = 2$, and a detailed description of the calculation procedure.

Then, in Section 5.2, if the graph is symmetric around the special point when calculating the graph automorphism, the symmetry is taken into account and the calculation is completed at the end. If the graph is symmetric with respect to the special point, $\times 2$ is used at the end of the calculation to account for the symmetry. However, this is not the case, for example, in the case of 8. In the case of 8, there is only one special point. As a result, the total number of permutations of two α_1-cycles is not $1! \times 1!$ with one permutation each on the left and the right, but α_1-cycles with a total of two permutations. See Figure 16 for an image. If we consider α_1-cycles to be interchangeable because of the symmetry, the number of graph automorphisms is $1! \times 1! \times 2^2 \times 2 = 8$.

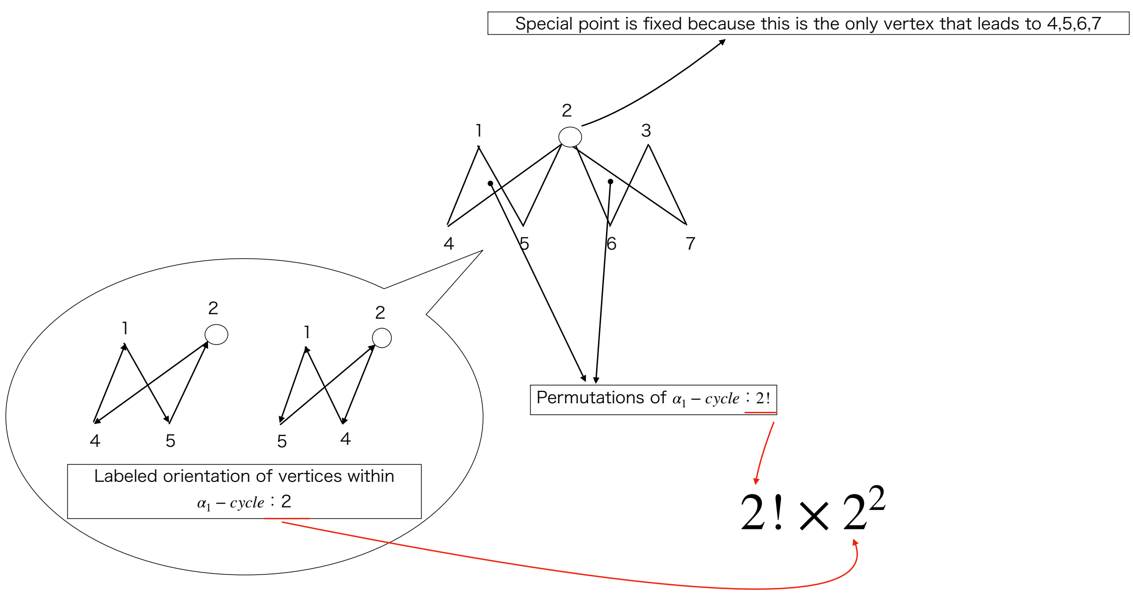


Figure 16: Computational methods one of graph automophism.

Here is another example. In the case of Pattern 2 in Figure 2, the exchange of labeled special points does not change the connection relationships of the entire graph. Therefore, it is still symmetric and does not need to be considered as $\times 2$. Therefore, it can be considered as shown in Figure 17.

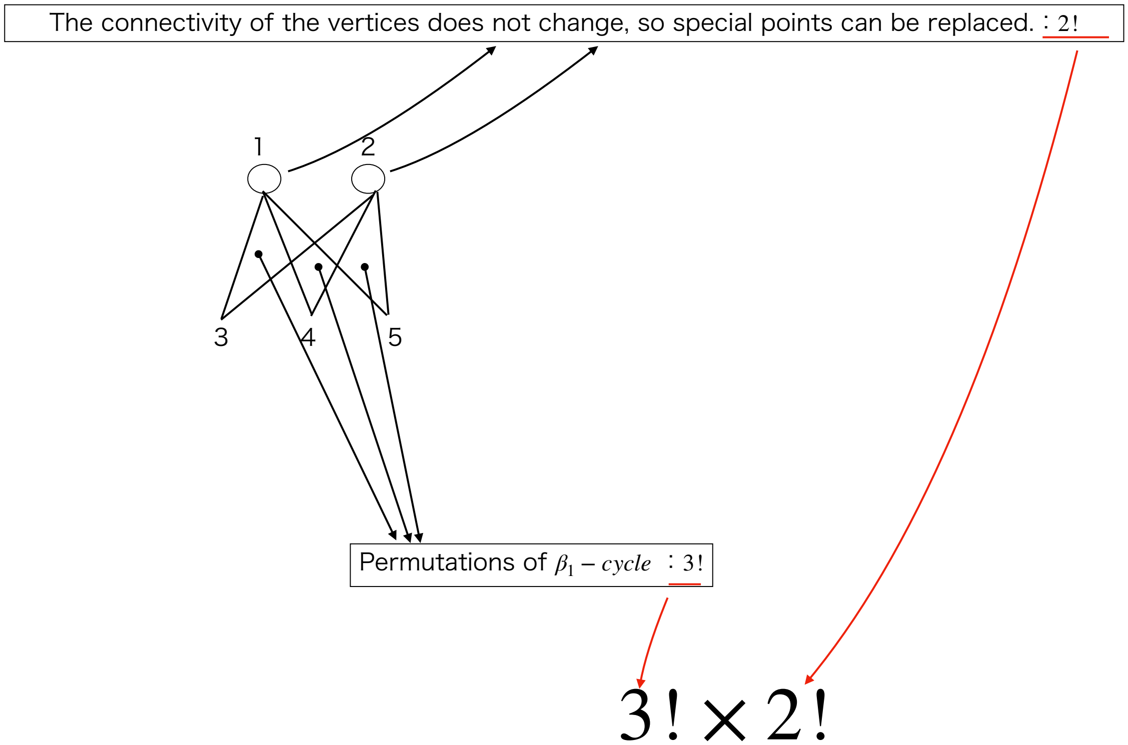


Figure 17: Computational methods two of graph automophism.

Let us compute the graph automophism of Pattern 1 in Figure 12. Unlike the previous examples, this one must be symmetric and then $\times 2$ is needed at the end of the computation. First, we note the degree of each vertex of the special point. In the case of labeling as shown in Figure 18, vertex 2, which is a special point, is connected to vertices 5, 6, and 7, and vertex 3, which is also a special point, is connected to vertices 7, 8, and 9. If vertex 2 and vertex 3 are swapped, the new vertex 3 will be connected to vertices 5, 6, and 7, but not to vertices 8 and 9, thus changing the structure of the graph. Therefore, the graph with vertex 2 and vertex 3 swapped is not graphically isomorphic to the graph in Figure 18. Hence, we cannot think of it in the same way as in Figure 16 or Figure 17.

Therefore, it is necessary to first consider the permutations of cycles and paths on each side of the blue dotted line in Figure 18. The fact that the permutations are considered "always on the left and right sides" is a clear difference between Figure 16 and Figure 17. After considering the permutations of *cycle* and *path*, the blue dotted line is considered as the center, and the left and right sides are swapped. This is the reason why $G_{\mathcal{B}}$ is (or must be) calculated with $\times 2$.

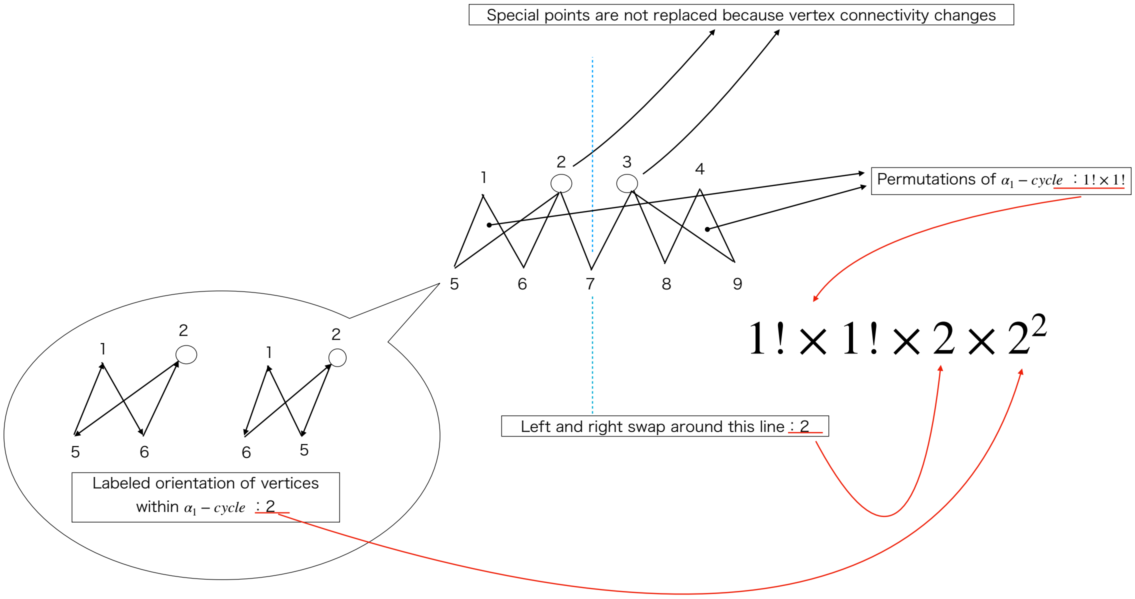


Figure 18: Computational methods three of graph automophism.

5.4 Proof of Theorem 1.2.3

Next, let us prove Theorem 1.2.3. From Theorem 1.2.4 and Equation (5.0.5), we have

$$F_k(x,y) = \frac{1}{(1-T_xT_y)^{3(k-1)}} \sum_{\mathcal{B}\in BG_k} \frac{1}{g_{\mathcal{B}}} T_x^{\alpha_{\mathcal{B}}} T_y^{\beta_{\mathcal{B}}} (1-T_xT_y)^{p_{\mathcal{B}}}\,. \tag{5.4.11}$$

Note that,

$$\alpha_{\mathcal{B}} = r_{\mathrm{sp}} + a_1 + 2a_2 + b_2 + c_2,$$
$$\beta_{\mathcal{B}} = s_{\mathrm{sp}} + 2a_1 + a_2 + b_1 + c_2.$$

Also,

$$\begin{aligned} p_{\mathcal{B}} &= 3(k-1) - (a_1+a_2+b_1+b_2+c_1+c_2)\,, \\ &= 2(k-1) - (r_{\mathrm{sp}} + s_{\mathrm{sp}}) + d, \\ &= \sum_{v\in \text{ special points}} (\deg(v)-3) + d. \\ &\geq 0 \end{aligned} \tag{5.4.12}$$

Note that there exists $\mathcal{B}\in BG_k$ such that $p_{\mathcal{B}}=0$. For example, if we construct a basic graph $\mathcal{B}^*$ such that $r_{\mathrm{sp}} = 2(k-1), a_1 = 2, b_1 = 3k-5$ and no other constants exist, we label the $2(k-1)$ special points of V_1'' as $r_1, r_2, \ldots, r_{2(k-1)}$. Next, attach α_1-cycle to each of r_1 and $r_{2(k-1)}$, and connect r_{2j-1} and r_{2j} by β_1-path. Also, connect R_{2j} and R_{2j+1} by a second β_1-path. Thus we obtain a basic graph $\mathcal{B}^*$.

In the case$k=2$, Figure 5.2 corresponds. Clearly, $s_{\mathrm{sp}} = d = 0$ for all $k\geq 2$. Also, computing Equation (5.4.12), we see that $p_{\mathcal{B}}=0$ also holds. From this, the numerator of the right-hand side of Equation (5.4.11) is

$$Q(x,y) = \sum_{i=1}^{m} C_i x^{a_i} y^{b_i} + \sum_{i=m+1}^{m+n} C_i x^{a_i} y^{b_i} (1-xy)^{p_i}$$

for some integer m, n. Note that for any i, a_i, b_i are non-negative integers, p_i is an integer, and $C_i > 0$. If $Q(x, y)$ has coefficients $1 - xy$, then plugging $y = x^{-1}$ into both sides yields ,

$$0 = \sum_{i=1}^{m} C_i > 0,$$

which is a contradiction. Therefore, $Q(x, y)$ does not factor the coefficient $1 - xy$. This means that the numerator of the right-hand side of Equation (5.4.11) does not have $1 - T_x T_y$. Note that since we are dealing with bipartite graphs, there exist graphs that are symmetric in their bipartite structure. Therefore, $F_k(x, y)$ is symmetric about T_x and T_y. That is, the numerator of (5.4.11) is expressed by using $Z = T_x + T_y$ and $W = T_x T_y$, and Theorem 1.2.3 is proved.

5.5 Another proof of Proposition 1.2.1

Finally, we give another proof of Proposition 1.2.1. This is an argument similar to the proof of Theorem 1.2.4, and is a bipartite version of the combinatorial argument of [Wri77]. The notation will be the same as before. In a preliminary step, the leaf and adjacent edges are repeatedly deleted. In this case, this step yields the only cycle with length $2t$. Then fix $r, s \geq 2$ and let $V'' = (V_1'', V_2'')$ be the vertex set. Consider an unlabeled bipartite graph with a cycle of length $2t$, taking t such that $2 \leq t \leq \min\{r, s\}$. First, it is clear that $|V_1''| = |V_2''| = t$. For the bipartite graph thus set up, rooted trees are attached in the same way as in Step 2 of Theorem 1.2.4. To create $2t$ rooted trees, we partition (r, s) vertices into $2t$ vertex sets. All these vertex sets are contained in $\mathcal{I}_{t,t}$. This procedure yields a connected $(r, s, r+s)$-graph labeled $\phi(r, s, t, t)$. Note that $\phi(r, s, t, t)$ is defined by Equation (5.1.8). For each graph obtained, there exists a $2t$ graph automorphism due to the cycle and the root of the rooted tree. Let $j(r, s)$ denote the number of labeled connected $(r, s, r+s)$-graphs. In this case,

$$j(r, s) = \sum_{2 \leq t \leq \min\{r,s\}} \frac{\phi(r, s, t, t)}{2t}.$$

Also, let $J(x, y)$ be the exponential-type generating function of $J(r, s)$.

$$\begin{aligned} J(x, y) &= \sum_{r,s=0}^{\infty} j(r, s) \frac{x^r y^s}{r!s!} = \sum_{t=2}^{\infty} \frac{1}{2t} \sum_{\substack{t \leq r \\ t \leq s}} \frac{\phi(r, s, t, t)}{r!s!} x^r y^s \\ &= \frac{1}{2} \sum_{t=2}^{\infty} \frac{(T_x T_y)^t}{t} = -\frac{1}{2} \left(\log(1 - T_x T_y) + T_x T_y\right) = F_1(x, y), \end{aligned}$$

This completes the proof.

6 Application

In this section, we describe two applications that are closely related to the main results of this paper. For more details, see [Kut08].

6.1 Relationship to cuckoo hashing

6.1.1 What is hashing ?

Before describing Cuckoo hashing, let's start with hashing (in a very simple way). Hashing is a data searching algorithm that replaces a specific string of letters or numbers with another value (hash value) using a calculation procedure based on a certain set of rules (hash function). The substituted values are stored in a simple one-dimensional array called a hash table, and the values are searched to find the desired source data. The image in Figure 19 shows the atmosphere of this process.

For example, a student of a certain university logs into the university's student-only website. To log in, an e-mail address and a password are required. If the password is "suugaku1234", it can be replaced with the hash value "Oidw98-aonlz" using a hash function. It is important to note that this is only a replacement. In other words, the student's password "suugaku1234" and the hash value "Oidw98-aonlz" are the same in meaning, and have not been changed to different values.

As a side note, there are other encryption methods that transform passwords so that they cannot be read. Encryption is the process of making information inaccessible to the outside world by following specific rules (algorithms).

Both encryption and hashing are very similar, but the biggest difference is whether the information can be undone or not. Encrypted information can be restored by using a "key" that can only be known by those who are authorized to possess it. In contrast, hashed information cannot be undone by anyone, including the person who hashed it. Encryption is a reversible replacement, while hashing is an irreversible replacement.

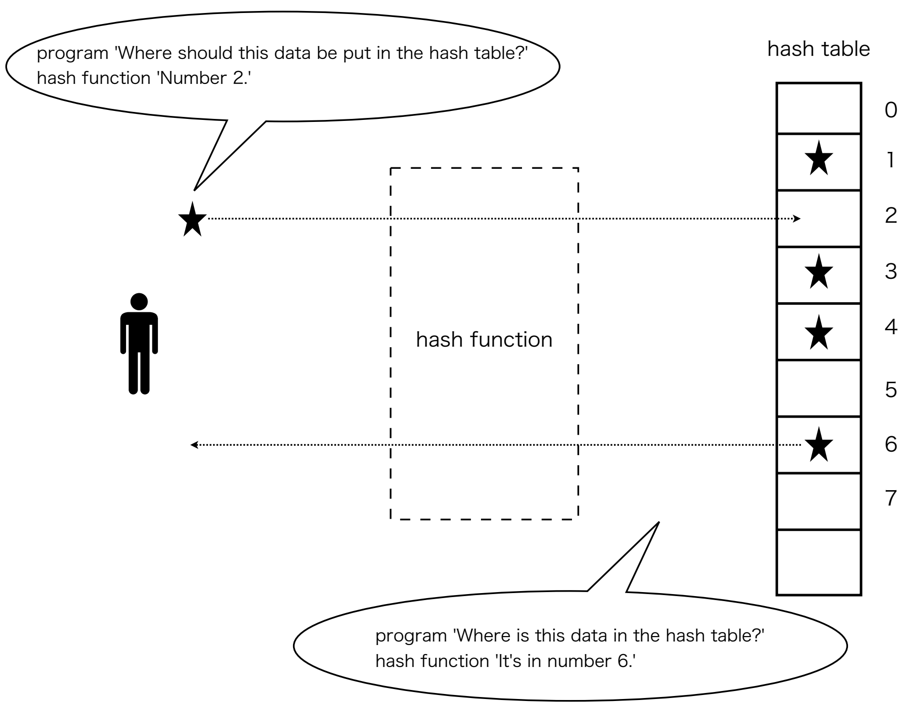


Figure 19: Image of hashing.

The most important feature of hashing is that the time required for searching, inserting, and deleting data is "theoretically" only $O(1)$. Cleary, the $O(1)$ search time is independent of the number N of data, which is ideal as a search technique. This is ideal as a search technique. On the other hand, however, there are some problems such as the impossibility of approximate search (searching for data close to the target data) and time-consuming expansion and contraction of the hash table.

Let the data structure under consideration be an associative array. That is, each data record is uniquely defined by a key. For example, the Austrian personal information database corresponds to this data structure. A national insurance number is assigned to each citizen as a key, and the owner's personal information is obtained by querying the database for this number. There is also hashing for data formats that are not associative arrays. See, for example, [CLRS09] and [SF13].

The number of data records is usually much smaller than the number of keys. To take the example of the personal information database, the number of national insurance numbers that qualify as keys is 10 digits (the total number of combinations is 10^{10}), whereas the number of data records, the total population, is only 8 million people. Therefore, it is completely inefficient to store 8 million or so data using an array with 10^{10} cells. Therefore, hashing is used. By implementing a hash table using a hash function, only memory proportional to the maximum number of keys to be stored is required. (See Figure 20).

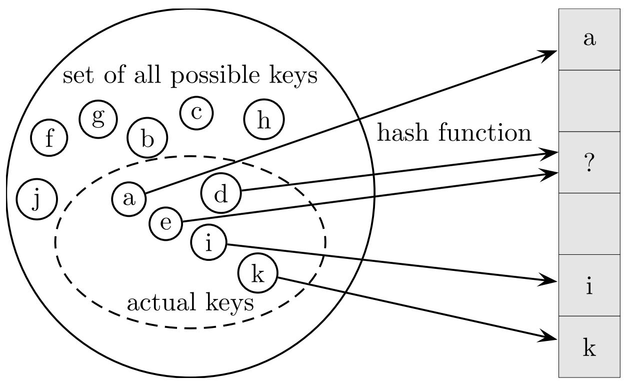


Figure 20: Hash function and hash table.

However, there is one major problem. This is that the converted values using the hash function are the same, i.e., the destination data is the same even though the original data is different. In Figure 20, d and e correspond to this case. When data are stored in the same destination of a hash table by a hash function even though they are different, it is called a collision. Although it depends on how the hash function is constructed, collisions cannot be avoided in many cases. The important point is not how to avoid collisions, but how to handle them when they occur. In the next section, we discuss countermeasures against collisions.

6.1.2 Open address method

The most natural approach is to computationally re-find a new storage location in the event of a collision. Finding a valid storage location in this way is called rehashing, and hashing that avoids collisions by rehashing is called the open address method.

Let x be the key and $h(x)$ be the hash function. In this case, the hash function only depends on the key information. Then, add $i = 0, 1, 2, \cdots$ as a new value and set the hash function to $h(x, i)$. Thus, if a collision occurs, we store the ith slot counting from that slot. Initially, the calculation is performed with $i = 0$, and if it has already been stored, $i = 1, \cdots, m-1$, and the calculation is repeated until a free slot is reached. The m denotes the number of slots in the hash table.

The diagram in Figure 21 shows an image of such an open address method. Here, d and e collide and e is moved to a new storage position. In this case, data larger than the hash table cannot be handled, and is also called closed hashing. See [Bin96] for more details. The sequence of searched slot locations is called a prob sequence. A prob sequence is obviously a permutation of $\{0, 1, 2, \cdots\}$, and is detailed in the general selection method [GBY91] and [Knu73].

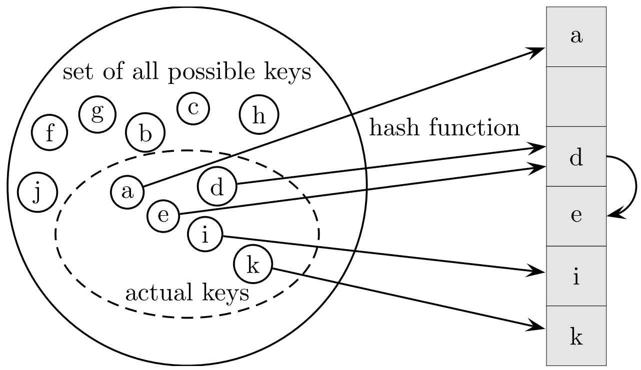


Figure 21: Image of open address method.(by [Kut08])

6.1.3 Chain method

In the case of a collision, multiple data are stored in the colliding slot without looking for a vacant slot. This method is called the chain method because data are chained together by creating a chained list in the collided slots. The diagram in Figure 22 shows an image of this method.

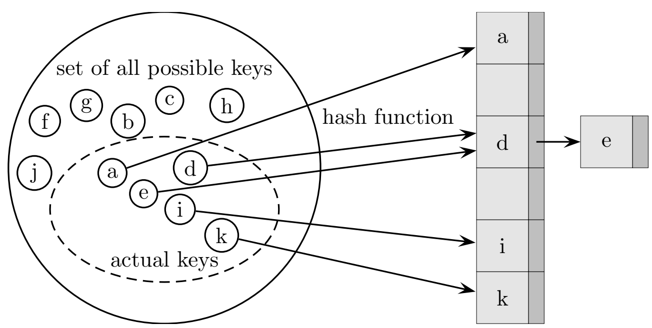


Figure 22: Image of chain method.(by [Kut08])

Although d and e still collide, we do not search for a vacant slot as in the open hashing method. A new linked list, called a linked list, is created in the same slot, and the data is searched by traversing the linked list. See Cormen([CLRS09]) for more details.

The greatest advantage of the chain method is that the number of keys stored in the hash table can exceed the size of the hash table. Note that unlike the open-address method, the longer the linked list, the more time is required to search the linked list for the source data. When using short keys such as 32-bit integers, the open address method is preferable, while hashing by the chain method is recommended when using large keys such as strings.

6.1.4 Cuckoo hashing

Cuckoo hashing is a relatively newer method than the open hashing and chaining methods. It was first introduced by [PR04] and further analyzed by [DM03]. As mentioned earlier, the hashing method theoretically requires only $O(1)$ search time. In practice, however, collisions occur, and the more collisions that occur, the further away from a constant time the search time becomes. Cuckoo hasing, on the other hand, ensures that the search time is constant even when collisions are taken into account.

Then, let us explain how Cuckoo hasing works. Let x be a key and h_1, h_2 be hash functions. First, a new key x is inserted into a new storage location $h_1(x)$ by the hash function h_1. If $h_1(x)$ is empty before storage, the operation is terminated. If it is not empty, there exists a key y such that $h_1(x) = h_1(y)$. This y is stored in another storage location $h_2(y)$. If this slot is empty, the operation ends there. If this slot is also non-empty, i.e., there exists a key z such that $h_2(y) = h_2(z)$, then this z is stored in another storage location $h_1(z)$. This operation is repeated until all keys are stored. The algorithm of "sequentially storing the keys that come in first and then the keys that come out later" is similar to the cuckoo's act of mending its eggs, hence the name "cuckoo hashing".

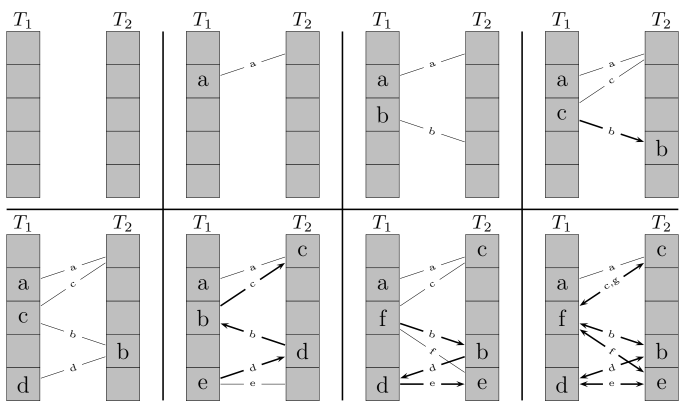


Figure 23: Image of cuckoo hashing.

For example, at what position is the key x_1 inserted in a hash table T_1 of size m? It can be inserted at any position with probability $\frac{1}{m}$ by the hash function h_1. At what position in the hash table T_2 of size m is $h_1(x_1)$ inserted in this way? It can be inserted at any position with probability $\frac{1}{m}$ by the hash function h_2. The images are shown in Figure 23. Each figure shows the state of the key after it has been inserted. The lines indicate the two possible locations for each key, and the arrows indicate the movement of the key. Finally, we try to store the key g at position f, which is impossible because it would cause an infinite loop.

As can be seen in Figure 23, Cukoo hashing uses two hash tables T_1, T_2 of size m. Naturally, two hash functions are required, but after all operations are completed, the position of a key x, for example, is guaranteed to be in either $T_1[h_1(x)]$ or $T_2[h_2(x)]$. In other words, the target key x can be retrieved by simply looking for these two keys.

Note, as with the open address method and the chain method, the hash value of the key $\{x_1, x_2, \cdots, x_n\}$ is assumed to form an independent uniform random sequence of numbers $\{1, 2, \cdots, m\}$. Furthermore, if a re-hash is necessary, the new hash value is assumed to be independent of the previous attempts.

Then, as can be easily seen in Figure 23, the key to modeling Cuckoo hashing is the bipartite graph (the basis of bipartite graphs and graph theory). See [Die05] for the basics of bipartite graphs and graph theory. In a bipartite graph $BG(V_1, V_2, E)$, the two vertex sets V_1, V_2 such that $V = V_1 \sqcup V_2$ correspond to the hash tables T_1, T_2 in Cuckoo hashing. Furthermore, the insertion of the key x is encoded as the edge $(h_1(x), h_2(x)) \in T_1 \times T_2$. The evolution of hash tables such as the one shown in Figure 23 is handled by labeling each edge of the corresponding bipartite graph. That is, an edge with label j is the jth key inserted into the hash table.

The structure of the modeled bipartite graph determines whether the hash table can be constructed successfully, i.e., whether the Cuckoo hashing algorithm works well. First, the structure of Cuckoo hashing makes it impossible to store more than k keys in k slots. Therefore, every component of the bipartite graph must have an edge less than or equal to a vertex. This implies that the bipartite graph is a spanning tree or unicycle. In this case, infinite loops such as the one seen at the end of Figure 23 in the corresponding hash table cannot occur. On the other hand, a bipartite graph with $|E| - |V| \geq 1$, i.e., a complex graph, will have infinite loops. In Figure 23, the hash table T_2 of key g overlaps with the location of c. This means that g and c can form a multiple edge. Note that an infinite loop will always occur even if g is inserted at a and d instead of f and T_2 is stored at any of $c, b, ande$. This means that if the modeled graph is a bipartite graph with a unicycle structure, $|V_1| + |V_2| = n$ keys are stored in $|E| = n + 1$ locations. However, when adding a new key from there, $n + 1$ keys are added to $|V_1| + |V_2| = n$ storage locations.

6.1.5 Previous work

In this section, we present several major previous works on cuckoo hashing. Let T_1 and T_2 be two hash tables and m be their size. Let $n = (1 - \epsilon)m$ be the number of keys. See also $\epsilon \in (0, 1)$. Then, the following theorem on the probability of success of cuckoo hashing is given by [Kal91].

Theorem 6.1.1. *The probability that a Cuckoo hash of $n = (1 - \epsilon)m$ data points into two tables of size m succeeds (i.e.the corresponding bipartite graph contains no complex component), provided that $\epsilon \in (0, 1)$, is equal to*

$$1 - \frac{\left(2\varepsilon^2 - 5\varepsilon + 5\right)(1 - \varepsilon)^3}{12(2 - \varepsilon)^2\varepsilon^3}\frac{1}{m} + O\left(\frac{1}{m^2}\right).$$

A complex component is a connected component that has two cycles. The following theorem also holds for connected components that have a single cycle, i.e., unicycle components.

Theorem 6.1.2 ([KP09])**.** *The number of unicycle components with cycle length $2c$ has in limit a Poisson distribution $Po(\lambda_c)$ with parameter*

$$\lambda_c = \frac{1}{2c}(1 - \varepsilon)^{2c}$$

Further, the number of tree components with l vertices converges in distribution to a Gaussian random variable with asymptotic mean

$$\mu = 2m\frac{l^{l-2}(1 - \varepsilon)^{l-1}e^{l(\varepsilon-1)}}{l!},$$

and variance

$$\sigma^2 = \mu - 2m\frac{e^{2l(\varepsilon-1)}l^{2l-4}(1 - \varepsilon)^{2l-3}\left(l^2\varepsilon^2 + l^2\varepsilon - 4l\varepsilon + 2\right)}{(l!)^2}.$$

Also, if cuckoo hashing is successful, the following holds with respect to the upper bound of its running time.

Theorem 6.1.3 ([Kut08])**.** *The expected construction time of a successful Cuckoo fash attempt is bounded above by*

$$2n\frac{1 - e^{\varepsilon-1}}{1 + \varepsilon - e^{\varepsilon-1}}.$$

6.2 Relationship to random graph theory

6.2.1 What is random graph ?

The graph is called a **random graph** if it is completed by adding edges with certain probability to a set of vertices. The most classical random graph is the Erdős–Rényi graph defined below.

Definition 6.2.1 (Erdős–Rényi graph of type $G(n, M)$)**.** A random graph that chooses one graph uniformly at random from all graphs of n-vertices and M-edges is called an Erdős–Rényi graph of type $G(n, M)$.

For example, suppose $G(3, 2)$. The graph in Figure 24 is all possible graphs with three vertices and two edges. Assume that only one of these graphs is chosen, each with probability 1/3. The graph thus chosen is called $G(3, 2)$.

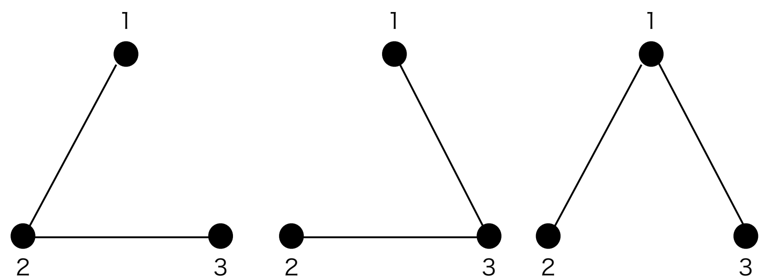


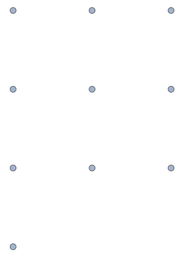
Figure 24: All possible graphs with $n = 3, M = 2$.

Definition 6.2.2 (Erdős–Rényi graph of type $G(n, p)$)**.** For an isolated point on n vertices with no edges, we denote by Erdős–Rényi graph of type $G(n, p)$ the random graph to which edges are added independently with probability p.

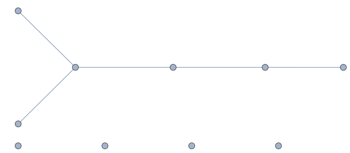
In the case of time expansion, the process ends when an n-vertex complete graph K_n is created. Note that for the degree $\deg(\nu)$ of the vertex ν, we have

$$P[\deg(\nu) = k] = \binom{n-1}{k} p^k (1-p)^{n-1-k}.$$

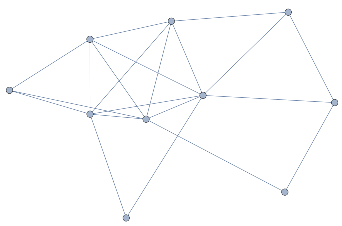
Also, by a short calculation, the expected value of the degree is $(n-1)p$ and the expected value of the number of edges is $\binom{n}{2}p$. As an example of $G(n, p)$, the case $n = 10$ is shown in Figure 25. The isolated point is $n = 0$. From there, the probability of introducing an edge is gradually increased, and Figure 25(e) has all edges. Hence, we have a complete graph K_{10}.



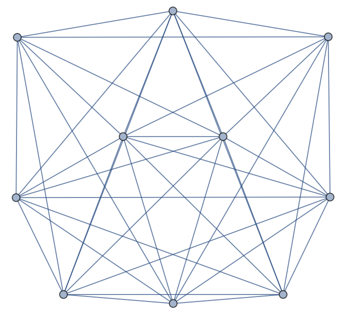
(a) $p = 0$



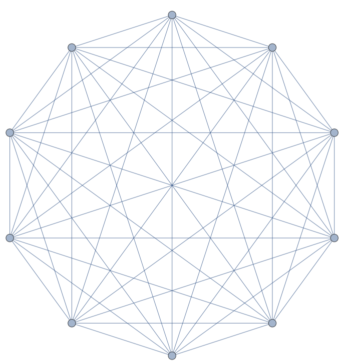
(b) $p = 0.1$



(c) $p = 0.5$



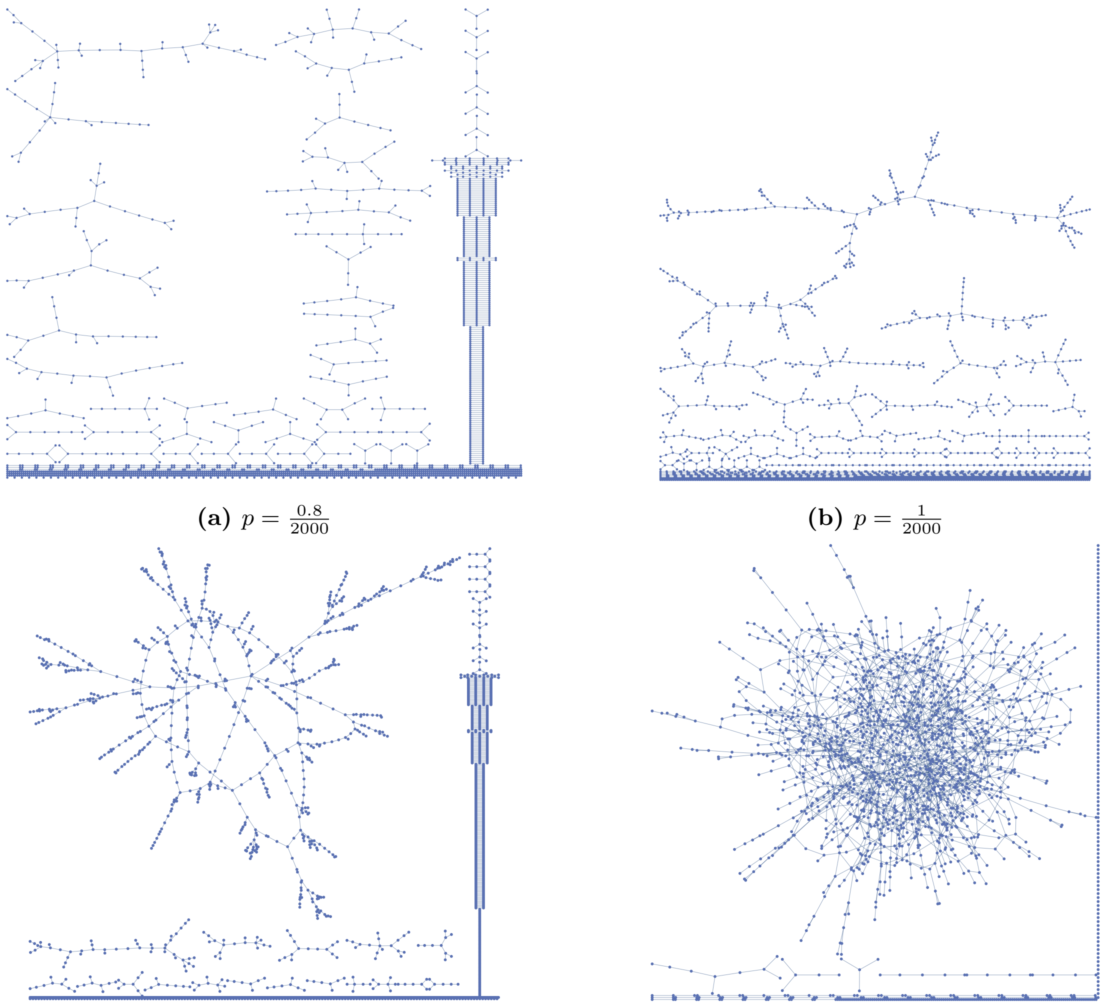
(d) $p = 0.9$



(e) $p = 1$

Figure 25: Examples of $G(10, p)$ for $p = 0, 0.1, 0.5, 0.9, 1$.

As can be seen in Figure 25, the connection relationship of the graph changes when the probability p is varied. For $G(n, p)$-type Erdős–Rényi graphs, varying the probability p of introducing an edge as $p(n) = \frac{c}{n}$ depends on the number of vertices n, the the connection relationship of each component of the graph changes drastically. Such a change is called **phase transition**. Note that by [ER$^+$60] Erdős–Rényi graph is also known to have an asymptotic threshold of $\frac{\log n}{n}$. The following figure shows an example of a phase transition. By gradually increasing the probability of introducing the edges from Figure 26(a) to Figure 26(d), the graph indeed changes.



(a) $p = \frac{0.8}{2000}$ **(b)** $p = \frac{1}{2000}$ **(c)** $p = \frac{1.2}{2000}$ **(d)** $p = \frac{2}{2000}$

Figure 26: Phase transition of $G(2000, p)$.

Erdős–Rényi graph is one of the most classical random graphs, but it has been actively studied in recent years. For example, in [ST22], in an Erdős–Rényi graph of type $G(n, p)$, if $X_{n,p}$ is a random variable that counts cliques (the largest complete subgraph of a graph), the following is shown to hold.

Theorem 6.2.1 ([ST22])**.**

$$\mathbb{E}\left[X_{n,p}\right] = n^{\frac{1}{-2\log p}(\log n - 2\log\log n + O(1))}$$

6.2.2 Previous work

Let $G(n, m)$ be a random graph chosen uniformly from a simple graph of n vertices and m edges, and Let $\{G(n, m)\}_{0 \leq m \leq \binom{n}{2}}$ be its stochastic process. For a simple graph G, we define an **excess** as follows.

Definition 6.2.3 (excess)**.** For a simple graph G with V as the vertex set and E as the edge set,

$$excess(G) := |E| - |V|.$$

For a d-dimensional hypergraph H, we define $excess(H) := (d-1)|E|-|V|$. Let l-component be a connected **component** (not a connected **graph**) whose excess is l. Note that excess is the Betti number -1. For k in [HSY22], $l = k-1$. We denote the -1-component of the l-component by **tree component**, the 0-component by **tree component**, and the k-component by **tree component**. 0-component is converted to **Unicycle component**, $l \geq 1$-component The l-component of $l \geq 1$ is called **complex component**.

In the random graph process $\{G(n,m)\}_{0 \leq m \leq \binom{n}{2}}$, the l-component develops in the following two situations.

1. Add a new edge between two vertices in the same connected component.
2. Add new edges between different connected components.

Therefore, the following holds for l-component **generation** and **growth**.

Corollary 6.2.1 (Generation and growth of l-component)**.** *The l-component is generated from the $l-1$-component by the circumstance 1. Furthermore, l_1+l_2+1-component is generated from l_1-component and l_2-component by circumstance 2. Also, if l-component and -1-component are situation 2, this grows to l-component.*

We denote $l_1 \oplus l_2$ as the case where l_1-component and l_2-component are in the situation 2in the system 6.2.1. The growth of l-component is the case of $l_1 \oplus (-1)$, i.e., when an edge is added between l-component and the tree. Examples are shown in Figure 27 and Figure 28. The red lines indicate newly added edges. In Figure 27, a red edge is added to -1-component by the condition 1, and 0-component is generated from -1-component. On the other hand, in the -1-component in Figure 28, -1-component swallowed the tree component, and the tree component grew as a -1-component. In other words, the tree component swallows the tree component and grows further.

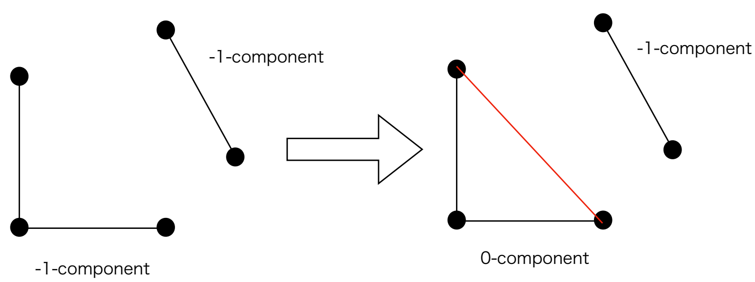


Figure 27: Generation of 0-component.

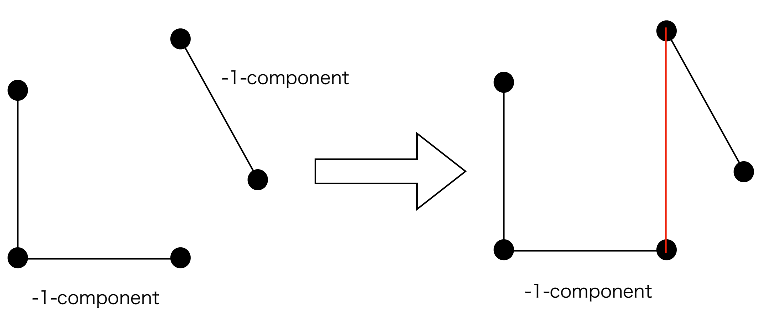


Figure 28: Growth of -1-component.

Let us introduce the symbol for expected value. Let $\alpha(l;k)$ be the expected value of "the number of times that l-component at k-point becomes $(l+1)$-component by adding an edge".

Also, let $\beta(l_1, l_2; k_1, k_2)$ be the expected value of "the number of times that l_1+l_2+1-component is obtained by adding an edge between l_1-component H_1 at k_1 and l_2-component H_2 at k_2".

We want to discuss how the expected value $\alpha(l;k), \beta(l_1, l_2; k_1, k_2)$ varies with the value of l, k. Here, the following is known by [Wri77].

Theorem 6.2.2 ([Wri77])**.** *The total number $C(k, k+l)$ of connected graphs of a k-point, $k+l$-edge simple graph, when fixing $l \leq -1$, is*

$$C(k, k+l) = \rho_l k^{k-1/2+3l/2}\left(1 + O\left(k^{-1/2}\right)\right).$$

In particular, $\rho_{-1} = 1, \rho_0 = \sqrt{\pi/8}, \rho_1 = 5/24$.

The well-known one is the total number of spanning trees $C(k, k-1) = k^{k-1}$. The following results are quoted from the [Jan00]. First, as a corollary, the following holds.

Lemma 6.2.1 ([Jan00])**.**

$$\begin{aligned}
\alpha(l;k) =& \binom{n}{k} C(k, k+l)\left(\binom{k}{2} - k - l\right)\int_0^1 t^{k+l}(1-t)^{(n-k)k+\binom{k}{2}-k-l-1}dt \\
\beta\left(l_1, l_2; k_1, k_2\right) =& \binom{k}{2}\binom{k}{k_1} C\left(k_1, k_1+l_1\right) C\left(k_2, k_2+l_2\right) k_1 k_2 \\
&\times \int_0^1 t^{k+l}(1-t)^{(n-k)k+\binom{k}{2}-k-l-1}dt
\end{aligned}$$

Given the Lemma 6.2.1, for example, the following holds for the complex component. This describes the case where a 0-component generates a 1-component.

Theorem 6.2.3 ([Jan00])**.**

$$\begin{aligned}
\sum_{k=1}^{n} \alpha(0;k) &= \frac{1}{2}\rho_0 n^{-1}\sum_{k=1}^{n}\left(k^{1/2} + O\left(\frac{k^{3/2}}{n} + \frac{k^{9/2}}{n^3} + 1\right)\right)e^{-k^3/24n^2} \\
&= \frac{1}{2}\rho_0 2^{3/2}3^{-1/2}\Gamma(1/2)(1+o(1)) + O\left(n^{5/3-2}\right) + O\left(n^{11/3-4}\right) + O\left(n^{2/3-1}\right) \\
&= \frac{\pi}{2\sqrt{3}} + o(1)
\end{aligned}$$

The following is also valid. This is the case where 0-component and 0-component are generated, i.e., $0 \oplus 0$.

Theorem 6.2.4 ([Jan00])**.**

$$\frac{1}{2}\sum_{k_1,k_2=1}^{n}\beta\left(0,0;k_1,k_2\right) \sim \frac{1}{2}\left(\frac{1}{2\pi}\right)^{1/2}\frac{\pi}{8}\sum_{k=2}^{n} k^{1/2}n^{-1}e^{-k^3/24n^2} \sim \frac{\pi}{8\sqrt{3}}$$

Theorem 6.2.3 and Theorem 6.2.4 is one of the main results of [Jan00]. [Jan00]also exhibits asymptotic behavior for the unicycle and tree components.

6.2.3 Application to bipartite graph

Let $BG(V_1, V_2, E)$ be a simple bipartite graph with V_1, V_2 as the vertex set and E as the edge set. Note that $V = V_1 \sqcup V_2$. When $|V_1| = r, |V_2| = s$, let $K_{r,s}$ be a complete simple bipartite graph. Note that $|V| = n, |V_1| + |V_2| = |V|$. Then, on all edges of $K_{r,s}$, we put random variables that follow a uniform distribution $U[(0,1)]$ on $(0,1)$ independently and identically. Let $0 \leq t \leq 1$ evolve in time and let $BG(n,r,s,t)$ be the graph completed by edges under t. Let $\{BG(n,r,s,t)\}_{0 \leq t \leq 1}$ be a random graph process. The following corollary holds for the expected value $\alpha_{r,s}(l;k_1,k_2)$ of the number of times that an edge is added to the interior of an l-component consisting of k points to become an $(l+1)$-component.

Lemma 6.2.2. *Let $C(k_1,k_2,l+k)$ be the total number of connected bipartite graphs on k points consisting of $(l+k)$ edges. Suppose $k = k_1 + k_2$, then*

$$\alpha_{r,s}(l;k_1,k_2) = \binom{r}{k_1}\binom{s}{k_2} C(k_1,k_2,l+k)(k_1k_2 - (l+k)) \int_0^1 t^{k+l}(1-t)^{k_1s+k_2r-k_1k_2-k-l-1}dt$$

Also, by adding an edge between l_1-componentH_1 consisting of k_1 points and l_2-componentH_2 consisting of k_2 points Let $\beta_{r_1,s_1,r_1,r_1,r_2}(l;k_1,k_2)$ be the expected number of times (l_1+l_2+1)-component is reached. The vertices of the connected component $H_i (i = 1,2)$ of the bipartite graph are r_i from V_1 and s_i from V_2. It is obvious that $r_i + s_i = k_i$. In this case, the following holds.

Lemma 6.2.3. *Let $C(r_i,s_i,l_i+k_i)$ be the total number of connected bipartite graphs on (r_i,s_i) points consisting of (l_i+k_i) edges.*

$$\begin{aligned}\beta_{r_1,s_1,r_1,r_2}(l;k_1,k_2) = &\binom{r}{r_1}\binom{s}{s_1}\binom{r-r_1}{r_2}\binom{s-s_1}{s_2} C(r_1,s_1,l_1+k_1)C(r_2,s_2,l_2+k_2) \\ &\times A^n_{r_1,r_2,s_1,s_2} \int_0^1 t^{k+l}(1-t)^{A^n_{r_1,r_2,s_1,s_2} - B^n_{r_1,r_2,s_1,s_2} - (l+k) - 1}dt\end{aligned}$$

where,

$$\begin{aligned}A^n_{r_1,r_2,s_1,s_2} =& k_1k_2 - \binom{r_1}{2} - \binom{r_2}{2} - \binom{s_1}{2} - \binom{s_2}{2} \\ B^n_{r_1,r_2,s_1,s_2} =& k(n-k) - r_1(r-r_1) - \{r_2(r-r_2) - r_1r_2\} \\ &- s_1(s-s_1) - \{s_2(s-s_2) - s_1s_2\} - (l+k).\end{aligned}$$

Summary of notations used repeatedly in this paper

- $G(V, E)$: Simple graph with vertex set V and edge set E
- (n, q)-graph : Simple graph such that $|V| = n, |E| = q$
- $f(n, q)$: Total number of connected (n, q)-graphs
- $BG(V_1, V_2, E)$: Simple bipartite graph with vertex sets V_1, V_2 and edge set E
- (r, s, q)-graph : Simple graph such that $|V_1| = r, |V_2| = r, |E| = q$
- $f(r, s, q)$: Total number of connected (r, s, q)-graphs
- $Q(r, s, q)$: Total number of connected (r, s, q)-graphs completed by connecting the $q+1$-th edge between the (r_1, s_1, t)-connected component and the $(r - r_1, s - s_1, q - t)$-connected component
- D_x, D_y : Euler operator. $D_x := x\partial_x, D_y := y\partial_y$
- $F_k(x, y) := \sum_{r,s=0}^{\infty} \frac{f(r,s,r+s-1+k)}{r!s!} x^r y^s$
- $T(x, y) := F_0(x, y)$
- $T_x := D_x F_0, T_y := D_y F_0$
- $Z = Z(x, y) := T_x + T_y$
- $W = W(x, y) := T_x T_y$
- $U(x, y) := F_1(x, y)$
- $V(x, y) := F_2(x, y)$
- $\mathcal{L}_k := (1 - T_y) D_x + (1 - T_x) D_y + k$

Acknowledgments

I would like to thank Professor Tomoyuki Shirai and Professor Satoshi Yabuoku for their enthusiastic guidance throughout the preparation of this paper.

References

[Bin96] Andrew Binstock. Hashing rehashed: Is ram speed making your hashing less effcient. *Dr. Dobb's Journal*, 4(2), 1996.

[Cay89] Arthur Cayley. A theorem on trees. *Quart. J. Math.*, 23:376–378, 1889.

[CLRS09] Thomas H. Cormen, Charles E. Leiserson, Ronald L. Rivest, and Clifford Stein. *Introduction to algorithms*. MIT press, 2009.

[Die05] Reinhard Diestel. *Graph Theory*. Springer, Berlin,, 2005.

[DM03] Luc Devroye and Pat Morin. Cuckoo hashing: further analysis. *Information Processing Letters*, 86(4):215–219, 2003.

[ER+60] Paul Erdős, Alfréd Rényi, et al. On the evolution of random graphs. *Publ. Math. Inst. Hung. Acad. Sci*, 5(1):17–60, 1960.

[FS09] Philippe Flajolet and Robert Sedgewick. *Analytic combinatorics*. cambridge University press, 2009.

[GBY91] Gaston H. Gonnet and Ricardo Baeza-Yates. *Handbook of algorithms and data structures: in Pascal and C*. Addison-Wesley Longman Publishing Co., Inc., 1991.

[HSY22] Taro Hasui, Tomoyuki Shirai, and Satoshi Yabuoku. Enumeration of connected bipartite graphs with given betti number. *arXiv preprint arXiv:2208.03996*, 2022.

[Jan00] Svante Janson. Growth of components in random graphs. *Random Structures & Algorithms*, 17(3-4):343–356, 2000.

[Kal91] IB Kalugin. The number of components of a random bipartite graph. 1991.

[Knu73] Donald E. Knuth. The art of computer programming. vol. 3, sorting and searching. 1973.

[KP89] Donald E. Knuth and Boris Pittel. A recurrence related to trees. *Proceedings of the American Mathematical Society*, 105(2):335–349, 1989.

[KP09] P. Mark Kayll and David Perkins. Combinatorial proof of an abel-type identity. *J. Combin. Math. Combin. Comput*, 70:33–40, 2009.

[Kut08] Reinhard Kutzelnigg. *Random bipartite graphs and their application to Cuckoo Hashing*. PhD thesis, 2008.

[Kut09] Reinhard Kutzelnigg. *Random Graphs and Cuckoo Hashing*. Südwestdeutscher Verlag für Hochschulschriften, 2009.

[PR04] Rasmus Pagh and Flemming Friche Rodler. Cuckoo hashing. *Journal of Algorithms*, 51(2):122–144, 2004.

[Rén59] Alfred Rényi. *On connected graphs*. Akadémiai Kiadó, 1959.

[Rio68] John Riordan. *Combinatorial identities*. Wiley, 1968.

[Sco62] H.I. Scoins. The number of trees with nodes of alternate parity. In *Mathematical Proceedings of the Cambridge Philosophical Society*, volume 58, pages 12–16. Cambridge University Press, 1962.

[SF13] Robert Sedgewick and Philippe Flajolet. *An introduction to the analysis of algorithms.* Pearson Education India, 2013.

[ST22] Taro Sakurai and Norihide Tokushige. Counting cliques in a random graph. *arXiv preprint arXiv:2208.07492*, 2022.

[Wri77] Edward M. Wright. The number of connected sparsely edged graphs. *Journal of Graph Theory*, 1(4):317–330, 1977.